Distributions:
Topology and Sequential Compactness

Acknowledgement

First of all I want to thank Professor Mouhot for setting this essay topic, which I found very interesting and enjoyable. I also want to thank him for the fruitful discussion we had during the process of writing it.
Furthermore I want to thank Professor Escher, who has been a good mentor throughout my entire undergraduate studies, I am deeply grateful for his great support whenever it was needed. His lectures and seminars on various topics in Analysis and differential equations clearly formed the basis of my knowledge in Analysis. Furthermore I want to thank him for pointing out the book of Hörmander to me, which I found to be a valuable reference.
Similarly I want to thank Dr. Zsak and Dr. Garling for their interesting Part III lectures on Functional Analysis and Analysis on Polish Spaces respectively. I enjoyed both lectures and some of the covered material certainly helped in understanding various aspects within this essay.
Finally I am also much obliged to Grégoire Sergeant-Perthuis for his help with translating Shirai's paper, which helped me a lot to understand the material presented in the paper. I also want to thank Sandra Marschke for reading the draft version of the essay and giving me some helpful feedback.

Contents

1 Introduction

In this essay I will introduce the space of distributions and a selection of it's properties. The space of distributions arose by difficulties in the study of differential equations when having to deal with non-differentiable functions. To tackle these problem we will construct a space which in some sense extends the space of continuous and therefore in particular differentiable functions. Our first aim will be to construct this space, the space of distributions. It will turn out to be the dual space of a function space of very well-behaved functions, which we will call the space of test functions. Well-behaved here means that the function space is closed under derivation. This will allow us to quasi "push" the derivative from the newly defined distribution to the well-behaved function. As the dual space, its behaviour and structure strongly depend on the topology on the function space, we will start by having a close look at the topologies we can put on the space of test functions. Having found a suitable topology we will define the distributions to be the dual space and look at its properties and induced topology.
Having constructed the space, the next step has to be to define the generalised derivative for this larger space. To be consistent with existing results, we need to make sure that the generalised derivative coincides with the classical derivative wherever it is defined.

With the space of distributions, it is now possible to study a much wider range of problems which include the necessity to derive an object. But to tackle these problems one might also want to apply other well known tools from Analysis such as the Fourier transform. Thus our second goal will be to extend the constructions in a way that allows us to to do this. Following the same idea, we will choose a function space which is closed under derivation and Fourier transform and consider it's dual space. This space will be the space of tempered distributions with a generalised derivative and Fourier transformation. It will turn out that it can be identified with a subspace of distributions and that our results from the first part generally also apply to tempered distributions.

Being familiar with these spaces and their basic properties we will establish a connection to operator theory. This will shed light on a property helping with a problem which will have arisen in the first section. Namely we will see in the first section that it is in general not possible to multiply distributions. However, if the distributions are defined on independent variables, it is always possible to define a multiplication. Following the idea of considering test functions on independent variables, we will see another strong result. It will turn out that every bounded linear operator on a space of test functions into a space of distributions is given by an integral transform if the test function space and the space of distributions are defined on independent variables. This is very remarkable as it is far from true in general.

Finally, as many methods in studying differential equations make use of compactness arguments, we will examine compactness properties in the space of distributions. The Banach Alaoglu Bourbaki provides us with a large amount of compact sets. But in practise it is often desirable to use sequences to characterise compactness properties. We will equip the space of distributions with the weak* topology, which is not metrisable since the space of test functions is not finite dimensional. For non-metrisable topologies compactness and sequential compactness are not equivalent in general, which we will illustrate with an example. Thus at this point we cannot characterise the compact sets we get from the Banach Alaoglu Bourbaki theorem by sequences, so we will finish the essay by examining the relation between compactness and sequences in the space of distributions. It will turn out that we have enough structure that compact sets are actually also sequentially compact.

Overall we see that the spaces of distributions and tempered distributions allow us to apply a lot of very powerful machinery. Furthermore, despite the fact that the involved topologies often lack much seemingly important structure, the spaces behave much better than one might think on first sight. This good structure makes them very useful for application.

1.1 General definitions and conventions

Unless otherwise specified, throughout the essay Ω will denote an open subset of $\mathbb{R}^n$ and K will be a compact subset of Ω.
When dealing with dual spaces we will denote the duality pairing by $\langle x, f\rangle$ with $x \in X$ and $f \in X'$ the dual space, i.e. the element of the dual space on the right hand side acts on the element of the vector space on the left hand side, $\langle x, f\rangle = f(x)$.

2 The Space of Distributions

In this section we will introduce a space of well-behaved functions, the Schwartz space. It is named in honour of Laurent Schwartz, who was a pioneer in the theory of distributions and proved many fundamental results. Equipping this space with a suitable topology will yield the space of distributions as it's dual space. We will then examine the properties of the space of distributions with it's induced topology. In this chapter we will mainly follow the constructions in Rudin's book [10].

2.1 The Schwartz space

2.1.1 Definition and the first natural topology

We want to consider very well behaved functions, namely smooth functions with compact support.

Definition 2.1 *(The Schwartz space)*
For fixed $K \subset \Omega$ compact, let

$$\mathscr{D}_K := \{f \in C^\infty(\Omega) \,:\, \text{supp}(f) \subseteq K\}$$

Then the Schwartz space on Ω is the union over all such $\mathscr{D}_K$:

$$\mathscr{D}(\Omega) := \bigcup_{K \subset \Omega} \mathscr{D}_K = \{f \in C^\infty(\Omega) \,:\, \text{supp}(f) \text{ compact}\}$$

So the Schwartz space is the space of all smooth functions on Ω which have compact support. It is also often called the space of test functions.

For this space we can find a topology in a very natural way. As we consider smooth functions with compact support, the functions and all their derivatives are bounded. So we can construct a family of norms by examining the derivatives of our functions up to a certain degree. This idea is made precise in the following lemma.

Lemma 2.2
Define a family of norms on $\mathscr{D}(\Omega)$ by

$$\|\varphi\|_n := \max_{|\alpha| \leq n} \sup_{x \in \Omega} |D^\alpha \varphi(x)|$$

Then these norms induce a locally convex, metrizable topology on $\mathscr{D}(\Omega)$.
A local base is given by

$$V_n := \left\{\varphi \in \mathscr{D}(\Omega) \,:\, \|\varphi\|_n \leq \frac{1}{n}\right\}$$

However, this topology is not complete.

Notation 2.1:
We will denote the topology on $\mathscr{D}(\Omega)$ induced by $(\|\cdot\|_n)$ by $\overline{\tau}$.

Proof:

Part 1: *($\overline{\tau}$ locally convex and metrizable)*
The topology is clearly locally convex and as the V_n form a countable local base it is also metrisable.
So it remains to show that the topology it is not complete.

Part 2: *($\overline{\tau}$ not complete)*
To see this let us consider an easy example of a Cauchy sequence in $\overline{\tau}$ that does not converge in this space (c.f. [10, p.151]).
Let $\Omega = \mathbb{R}, \varphi \in \mathcal{D}(\mathbb{R}), \operatorname{supp}(\varphi) = [0, 1]$
Now define a sequence (ψ_k) by

$$\psi_k(x) := \varphi(x-1) + \frac{1}{2}\varphi(x-2) + \ldots + \frac{1}{k}\varphi(x-k)$$

(ψ_k) is clearly a $\overline{\tau}$-Cauchy sequence in $\mathcal{D}(\mathbb{R})$ as

$$\|\psi_k - \psi_{k-1}\|_n = \|\frac{1}{k}\varphi(x-k)\|_n = \frac{1}{k}\|\varphi\|_n$$

But $\psi := \lim \psi_k \notin \mathcal{D}(\mathbb{R})$ as $\operatorname{supp}(\psi_k) = [1, k+1]$ and hence $\operatorname{supp}(\psi) = [1, \infty)$ so it does not have compact support.

❑

The lack of completeness is rather unfortunate. In what follows, we will consider the dual space and equip it with the weak* topology. If we can construct a topology on $\mathscr{D}(\Omega)$ which is complete, the weak* topology on the respective dual space will be well behaved; e.g. in the dual of a Banach space w*-boundedness implies norm boundedness, so completeness allows us to deduce information about the behaviour in norm from the weak* topology. Therefore we shall now construct a different topology for the Schwartz space which is complete. As the topology we have just seen arose in a very natural way, we want the new topology of to be very close to it.
In above example we can see that the sequence failed to converge because of the limit not having compact support. Hence our aim has to be to include a control over the support of the limit function into the topology.

2.1.2 A better topology

To construct a topology which suits our needs better than the one we have seen in the last part we shall first slightly weaken our norms.

Lemma 2.3
Let $(K_n),\ K_n \nearrow \Omega$ a strictly increasing sequence of compact sets. Then

$$\rho_n(\varphi) := \max_{|\alpha| \leq n} \sup_{x \in K_n} |D^\alpha \varphi(x)|$$

is a family of seminorms on $C^\infty(\Omega)$. The induced topology on $C^\infty(\Omega)$ is locally convex, metrisable and complete.
Furthermore, for any fixed K the induced subspace topology coincides with the respective subspace topology of $\overline{\tau}$.

Notation 2.2:

We will denote the topology on $C^\infty(\Omega)$ by τ_Ω and the subspace topology on $\mathscr{D}_K$ for any fixed K by τ_K.

Proof:
For the properties of the topology on $C^\infty(\Omega)$, c.f. [10, p.35].
As for the furthermore part, since, for fixed K, the norms from lemma 2.2 and the seminorms ρ coincide from a certain n onwards (where $K \subset K_n$) and both are monotone increasing one can "chop off" the first finitely many terms. Hence both topologies coincide.

❑

Using this topology on the space of smooth functions, or rather it's subspace topologies on the $\mathscr{D}_K$, we now want to construct the desired topology on $\mathscr{D}(\Omega)$. We have just seen, that the topology induced by the family of seminorm is very close to the one we obtained before following the very natural approach. The following construction will preserve this but will add a certain control over bounded sets. This will help to control Cauchy sequences as they are bounded.

Proposition 2.4

Let

$$\beta := \{U \subset \mathscr{D}(\Omega) \,:\, U \text{ absolutely convex}, U \cap \mathscr{D}_K \in \tau_K \,\forall\, K\}$$

The family of all unions of the sets

$$\{\varphi + U \,:\, \varphi \in \mathscr{D}(\Omega), U \in \beta\}$$

is a topology for $\mathscr{D}(\Omega)$ with local base β.

Notation 2.3:

We will denote this topology on $\mathscr{D}(\Omega)$ by τ. A topology constructed this way are called an inductive limit topology. If we restricting attention to an increasing sequence of compact sets as in lemma 2.3, which we can certainly do, it is called a strict inductive limit topology.
For τ-neighbourhoods we will use one of two notations, depending on the situation. A τ-neighbourhood of some $\varphi \in \mathscr{D}(\Omega)$ will either simply be written as $U \in \tau$ or as $\varphi + U$, where $U \in \beta$ is a neighbourhood of 0.

Proof:
If τ is a topology, β is by definition a local base.
Clearly both the empty set (empty union) and $\mathscr{D}(\Omega)$ (union over all $\varphi \in \mathscr{D}(\Omega)$) are elements of τ. Furthermore, by definition, any union of elements of τ is an element of τ. So to proof that τ is a topology it remains to check that it is closed under finite intersections.
Let $U_1, U_2 \in \tau$ with $U_1 \cap U_2 \neq \emptyset$. Fix $\varphi \in U_1 \cap U_2$. We need to find a τ-neighbourhood of φ contained in $U_1 \cap U_2$, which is equivalent to finding a 0-neighbourhood $V \in \beta$ such

that $\varphi + V \subset U_1 \cap U_2$. As the argument is symmetric for U_1 and U_2, in the following we will use the index $i = 1, 2$.
As $U_i \in \tau$ we can find $\varphi_i \in \mathscr{D}(\Omega), V_i \in \beta$ such that $\varphi \in \varphi_i + V_i \subset U_i$.
As the V_i are open we can shrink them a bit while keeping φ_i in the set, i.e.
$\varphi \in \varphi_i + (1 - \delta_i) \cdot V_i$ for some $\delta_i > 0$. But this implies that

$$\varphi + \delta_i \cdot V_i \subset \varphi_i + (1 - \delta_i) \cdot V_i + \delta_i \cdot V_i$$

By convexity of V_i the right hand side is just $\varphi_i + V_i$. Hence we have

$$\varphi + (\delta_1 \cdot V_1 \cap \delta_2 \cdot V_2) \subset (\varphi_1 + V_1) \cap (\varphi_2 + V_2) \subset U_1 \cap U_2$$

So τ is closed under finite intersection and hence a topology with local base β.

❑

After having assured that τ is a topology on $\mathscr{D}(\Omega)$, we have to check, that it suits our needs, i.e. that it is reasonably close to out first topology and complete. But firstly we should check that the topology is conform with the vector space structure, i.e. that addition and scalar multiplication are continuous with respect to τ.

Theorem 2.5
$(\mathscr{D}(\Omega), \tau)$ is a locally convex topological vector space.

Proof:
As β is a local base of convex sets, $(\mathscr{D}(\Omega), \tau)$ is clearly locally convex. We have to check that addition and scalar multiplication are τ-continuous.

Part 1: *(Addition is τ-continuous)*
Let $\varphi_1, \varphi_2 \in \mathscr{D}(\Omega)$ and $(\varphi_1 + \varphi_2) + U$ an arbitrary τ-neighbourhood of $\varphi_1 + \varphi_2$.
We want τ-neighbourhoods $\varphi_1 + U_1$ and $\varphi_2 + U_2$, such that

$$\psi_1 + \psi_2 \in (\varphi_1 + \varphi_2) + U \quad \forall\ \psi_1 \in \varphi_1 + U_1,\ \psi_2 \in \varphi_2 + U_2$$

But if we choose $U_1 = U_2 = \frac{1}{2}U$, by the convexity of U we have
$(\varphi_1 + \frac{1}{2}U) + (\varphi_2 + \frac{1}{2}U) = (\varphi_1 + \varphi_2) + U$. Hence the addition is continuous.

Part 2: *(Multiplication is τ-continuous)*
Let $\varphi \in \mathscr{D}(\Omega), \lambda \in \mathbb{R}$ and $\lambda \cdot \varphi + U$ an arbitrary τ-neighbourhood of $\lambda \cdot \varphi$.
We need to find a τ-neighbourhood $\varphi + V$ of φ and an open ball around λ, $B(\lambda, \delta)$ say, such that

$$\mu \cdot \psi \in \lambda \cdot \varphi + U \quad \forall\ \psi \in \varphi + V, \mu \in B_\delta \tag{1}$$

But this is equivalent to saying

$$\begin{aligned} &\mu \cdot \psi \in \lambda \cdot \varphi + U \\ \Leftrightarrow &\mu \cdot \psi - \lambda \cdot \varphi \in U \\ \Leftrightarrow &\mu\,(\psi - \varphi) + (\mu - \lambda)\,\varphi \in U \end{aligned} \tag{2}$$

A way to make sure this equation holds is to ensure both summands are in $\frac{1}{2}U$, say. Convexity of U then again implies continuity.

As U is a neighbourhood of 0, it is absorbing and hence we can find a $t \in \mathbb{R}$ such that $t \cdot \varphi \in \frac{1}{2}U$. If $|\mu - \lambda| < t$ then

$$|\mu| \leq |\lambda| + t \tag{3}$$

Next we need to scale U by a factor c such that $\mu(\psi - \varphi) \in \frac{1}{2}U$. So we want

$$\mu \cdot c \leq \frac{1}{2}$$
$$\overset{(3)}{\Leftrightarrow} c \leq \frac{1}{2(|\lambda| + t)}$$

Hence (2) and therefore (1) holds if we choose $V = c \cdot U$ and $\delta = t$ for t, c as above, so the multiplication is continuous.

❑

The topology $\overline{\tau}$ arose in a very natural way, so it would be desirable to have our new topology to be very close. In fact it coincides on every fixed $\mathscr{D}_K$ as can be seen in the next theorem.

Theorem 2.6

On any $\mathscr{D}_K$ the subspace topology inherited from $(\mathscr{D}(\Omega), \tau)$ coincides with τ_K and therefore by Lemma 2.3 also with the subspace topology of $\overline{\tau}$.

Proof:

Part 1: *($\tau|_{\mathscr{D}_K} \subseteq \tau_K$)*

Let $U \in \tau$. We need $U \cap \mathscr{D}_K \in \tau_K$. Thus we fix $\varphi \in U \cap \mathscr{D}_K$ and construct a τ_K-neighbourhood of it.

By definition of τ we can find $V \in \beta$ such that $\varphi + V \subset U$. But as φ is in $\mathscr{D}_K$ we have

$$\varphi + (V \cap \mathscr{D}_K) \subset U \cap \mathscr{D}_K$$

By definition of β we have $V \cap \mathscr{D}_K \in \tau_K$, hence we have a τ_K-neighbourhood of φ. As φ was arbitrary $U \cap \mathscr{D}_K \in \tau_K \, \forall \, K$

Part 2: *($\tau_K \subseteq \tau|_{\mathscr{D}_K}$)*

Let $E \in \tau_K$. We need to find $U \in \tau$ such that $E = U \cap \mathscr{D}_K$.

We have seen before, that τ_K coincides with the subspace topology of $\overline{\tau}$. Hence for all $\varphi \in E$ we can find $n \in \mathbb{N}$ and $\epsilon > 0$ such that

$$\{\psi \in \mathscr{D}_K \, : \, \|\varphi - \psi\|_n < \epsilon\} \subset E \tag{4}$$

To get a τ-neighbourhood we need to consider functions from the whole of $\mathscr{D}(\Omega)$, so for all $\varphi \in E$ let $U_\varphi := \{\psi \in \mathscr{D}(\Omega) \, : \, \|\psi\|_n < \epsilon\}$. This is a set in β and hence $\varphi + U_\varphi$ is a τ-neighbourhood of φ.

By (4) $\varphi + (U_\varphi \cap \mathscr{D}_K) = (\varphi + U_\varphi) \cap \mathscr{D}_K \subset E$. Let

$$V = \bigcup_{\varphi \in E} (\varphi + U_\varphi)$$

Certainly $E \subset V$ as every $E \ni \varphi \in \varphi + U_\varphi$ and $V \subset E$ as every $\varphi + U_\varphi \subset E$.

❑

2.1.3 The fundamental structure of the topology

The last theorems have shown that the topology τ is very reasonable. Thus we will now show that we achieved our goal of constructing a complete topology. For this we will show that we gained a certain control over bounded sets. Completeness will then be an instant corollary as Cauchy sequences are bounded.

Theorem 2.7

Let E be a bounded set in $\mathscr{D}(\Omega)$. Then

(i) There exists a common compact set containing the support of all functions in E, i.e. $\exists\, K$ such that $\operatorname{supp}(\varphi) \subset K \,\forall\, \varphi \in E$

(ii) There exists a uniform bound for every norm from Lemma 2.2, i.e. $\forall\, n \in \mathbb{N}\, \exists\, M_n$ such that $\|\varphi\|_n \leq M_n \,\forall\, \varphi \in E$

Proof:
We will show the contrapositive of (i) and (ii) will be an immediate consequence as the topologies coincide on any fixed $\mathscr{D}_K$.

Part 1: *(Construct a τ-neighbourhood that fails to absorb E)*
Let $E \subset \mathscr{D}(\Omega)$ be such that there exists no compact set K which contains the support of every function in E, that is

$$S := \bigcup_{\varphi \in E} \operatorname{supp}(\varphi)$$

is not compact.
Hence we can find a sequence (x_n) in S which does not have a limit point. As we are choosing the sequence in S we can make sure that for every x_n we can find a function φ_n in E that is supported around x_n, i.e. we choose $(x_n) \subset S$ and $(\varphi_n) \subset E$ such that $\varphi_n(x_n) \neq 0$ for every n.
We now have a, in a way, widely spread out set of points in Ω and functions in E which have an effect at these points. To deduce information about the whole of E from this we consider the functions in $\mathscr{D}(\Omega)$ which are sufficiently bounded at these points. Precisely let

$$W := \left\{ \varphi \in \mathscr{D}(\Omega) \,:\, |\varphi(x_n)| < \frac{1}{n} |\varphi_n(x_n)| \right\}$$

Then W belongs to β. Assuming this is true, W would absorb E if E is bounded, because the multiplication in $\mathscr{D}(\Omega)$ with respect to τ is continuous, so the sets in β are absorbent. But by construction $\varphi_n \notin n \cdot W$ as $\varphi_n(x_n) \not< \varphi_n(x_n)$. So there is no scalar such that W absorbs E, hence E is not bounded. This proves (i), every bounded set in $\mathscr{D}(\Omega)$ must lay completely in some $\mathscr{D}_K$.

Part 2: *(Show that W actually is a set in β)*
To finish the proof of (i) it remains to check this. By definition we need to proof that the intersection of W with any $\mathscr{D}_K$ is open in the respective topology, since W is certainly absolutely convex. Therefore fix K. As we have chosen the sequence (x_n) such that it doesn't have a limit point there are only finitely many elements of the sequence in any sufficiently small open set. But as K is compact finitely many of these sufficiently small open sets cover K. Therefore we have only finitely many x_n in K. Hence we know that all functions in $\mathscr{D}_K$ are 0 for all but finitely many x_n. Let

$$I := \{n \in \mathbb{N} : x_n \in K\}$$

the finite index set of our chosen points laying in K. This allows us to write the intersection of W and $\mathscr{D}_K$ as follows

$$\begin{aligned} W \cap \mathscr{D}_K &= \left\{ \varphi \in \mathscr{D}_K : |\varphi(x_n)| < \frac{1}{n} |\varphi_n(x_n)| \quad \forall n \in \mathbb{N} \right\} \\ &= \bigcap_{n \in I} \left\{ \varphi \in \mathscr{D}_K : |\varphi(x_n)| < \frac{1}{n} |\varphi_n(x_n)| \right\} \end{aligned} \tag{5}$$

For every $\psi \in \{\varphi \in \mathscr{D}_K : |\varphi(x_n)| < c_n\}$ the set $\psi + \{\varphi \in \mathscr{D}_K : \|\varphi\|_0 < c_n - \psi(x_n)\}$ is a τ_K-neighbourhood of ψ inside $\psi \in \{\varphi \in \mathscr{D}_K : |\varphi(x_n)| < c_n\}$. So (5) is a finite intersection of τ_K-open sets and therefore in τ_K. Hence W is in β and the proof in part 1 is valid.

Part 3: *(There exists a uniform bound for every norm)*
We have seen before that on $\mathscr{D}_K$ the topologies τ and τ_K coincide, so E is also a bounded set in $\mathscr{D}_K$. But being bounded in a locally convex space means exactly that every seminorm (or norm as the case may be) inducing the topology is bounded. This is exactly (ii).

❑

Corollary 2.8
$(\mathscr{D}(\Omega), \tau)$ is complete.

Proof:
Let (φ_n) be a τ-Cauchy sequence. As Cauchy sequences are bounded, by Theorem 2.7, there exists a common support K. Hence $(\varphi_n) \subset \mathscr{D}_K$ for this K. By Theorem 2.6 the subspace topology of τ coincides with τ_K on $\mathscr{D}_K$ so (φ_n) is a τ_K-Cauchy sequence. But $(\mathscr{D}_K, \tau_K)$ is complete, so the sequence converges. Hence $(\mathscr{D}(\Omega), \tau)$ is complete.

❑

So the topology τ seems very well behaved. It's structure gives us a strong control over bounded sets and thus over convergence of sequences. However it has a minor drawback, it is not metrisable. It will turn out that this won't be an issue for our intended constructions. We will see that the space of distributions behaves very nicely, which is due to the fact that we can usually prove things for an arbitrary $\mathscr{D}_K$ and than deduce that the result holds for $\mathscr{D}(\Omega)$. Nevertheless we will have a closer look at the metrisability and related structure in the following section to get a better understanding of the topology.

2.1.4 A deeper inspection of the structure of the topology

The fact that the space of test functions is not metrizable is the most commonly mentioned property and can easily be proved using a Baire argument (c.f. [10]). However even more is true, it actually is, in a manner of speaking, very far from being metrizable. Recall the following definition.

Definition 2.9
A space is called sequential if it satisfies either of the two equivalent conditions

(i) Every sequentially open subset is open

(ii) Every sequentially closed subset is closed

In fact the space of test functions is even far from being sequential. Klaus Floret mentions in [8] that it actually contains sequentially closed linear subspaces which are not closed. The fact that the space of test functions is not sequential can sometimes be found as an exercise to the reader, e.g. in the book by Wilansky [17, p. 224]. Here the reader is pointed to a paper by Webb [16] and to Shirai and Dudley. In fact amongst others Webb discusses the properties of sequential closures and sequential closedness and their relation to closed sets in some detail. He also refers to Dudley's paper [6], in which he proves an important necessary and sufficient criterion for a space to be sequential. Furthermore he points out that Shirai (in [13]) proofed that the space of test functions does not fulfil this condition and therefore fails to be sequential.

Proposition 2.10
A topological vector space which is Hausdorff is sequential if and only if there exists no strictly finer topology with the same convergent sequences.

This result was given by Dudley in [6]. We shall use it in this section to prove that the space of test functions is not sequential. We will do this by presenting the most relevant results from Shirai's paper [13], with a few adaptations to fill in gaps which were either left by Shirai, or which arose from not covering some of the results stated in the paper, as we will not consider all aspects of the paper. These results establish the existence of a topology on the space of test functions which is strictly finer than the topology τ.
At first we need to make some constructions about the topology τ, which go along with the way Schwartz introduced them in his book [12] and Shirai in his paper [13]. As a simplification we will consider test functions on the whole of $\mathbb{R}^n$. This allows us to choose $K_n = \mathcal{B}_n$ in Lemma 2.3 as the sequence of compact sets converging to $\Omega = \mathbb{R}^n$.

Remark 2.11
With the choice $K_n = \mathcal{B}_n$ in Lemma 2.3 we get basic open neighbourhoods of 0 for every $\mathscr{D}_{K_n}$ by taking the sets

$$\{\varphi \in \mathscr{D}_{\mathcal{B}_n} : |D^\alpha \varphi(x)| < \varepsilon, |\alpha| \leq m\}$$

We can extend these functions to the whole of $\mathbb{R}^n$ by setting them 0 outside $\mathcal{B}_n$. We will denote these sets by

$$V(m,\varepsilon,\mathcal{B}_n) := \{\varphi \in \mathscr{D}(\mathbb{R}^n) \,:\, |D^\alpha\varphi(x)| < \varepsilon, |\alpha| \le m, \operatorname{supp}(\varphi) \subseteq \mathcal{B}_n\}$$

Note that the way we constructed the topology τ means that if A is τ-open, then for every $\varphi \in A$ and every K such that $\operatorname{supp}(\varphi) \subseteq K$ there exists a τ_K-neighbourhood of 0 which is contained in A. This motivates the following construction.

Definition 2.12
Let

- $(m_n) \subset \mathbb{N}$ be an unbounded, increasing sequence of natural numbers
- $(\varepsilon_n) \subset \mathbb{R}^+$ be a decreasing null sequence of positive real numbers

For these sequences define inductively for $\lambda, \mu \in \mathbb{N}$

$$U_{\lambda,0} := V(m_\lambda, \varepsilon_\lambda, \mathcal{B}_\lambda)$$

and

$$U_{\lambda,\mu} := \bigcup_{0\neq\varphi\in U_{\lambda,\mu-1}} (\varphi + V(m_{\lambda+\mu}, \varepsilon_{\lambda+\mu}, \mathcal{B}_{\lambda+\mu}))$$

By taking the union over all these $U_{\lambda,\mu}$ we obtain a τ-open, absolutely convex set.

$$U = \bigcup_{\substack{0\le\mu<\infty \\ 1\le\lambda<\infty}} U_{\lambda,\mu}$$

Now every $\varphi \in U$ is contained in some U_{λ_0,μ_0} for some $\lambda_0, \mu_0 \in \mathbb{N}$. Thus it can be written as

$$\varphi = \varphi_{\lambda_0} + \varphi_{\lambda_0+1} + \ldots + \varphi_{\lambda_0+\mu_0}$$

such that $\varphi_{\lambda_0} \neq 0$ and

$$\varphi_{\lambda_0+\mu} \in V\left(m_{\lambda_0+\mu}, \varepsilon_{\lambda_0+\mu}, \mathcal{B}_{\lambda_0+\mu}\right) \quad \forall\, 0 \le \mu \le \mu_0$$

We obtain a collection of sets by taking all possible choices of sequences as above:

$$\mathcal{U} = \left\{U((m_n),(\varepsilon_n)) \,:\, (m_n) \subset \mathbb{N}, m_n \nearrow \infty, (\varepsilon_n) \subset \mathbb{R}^+, \varepsilon_n \searrow 0\right\}$$

As long as no confusion is to be expected we will omit the parameters indicating the corresponding sequences and will simply write $U \in \mathcal{U}$.

Lemma 2.13
$\mathcal{U}$ is a local base of neighbourhoods at 0 for the topology τ.

Proof:
It is clear that $\mathcal{U}$ is a local base of 0 for the topology it induces on $\mathscr{D}(\mathbb{R}^n)$. We need to proof that the induced topology is exactly τ. To do this within the prove we will denote the topology induced by $\mathcal{U}$ by $\hat{\tau}$, following the notation used by Shirai.

Part 1: *($\hat{\tau} \subseteq \tau \;\rightarrow\; \tau$ is finer than $\hat{\tau}$)*
Let $U \in \mathcal{U}$. U is an absolutely convex subset of $\mathscr{D}(\mathbb{R}^n)$, so we need to show that the intersection with any $\mathscr{D}_K$ is τ_K-open. As mentioned in proposition 2.4 and earlier in this chapter we can restrict our attention to the closed balls $\mathcal{B}_n$.
Thus fix a radius n and let $\varphi \in U \cap \mathscr{D}_{\mathcal{B}_n}$. We need to find a $\tau_{\mathcal{B}_n}$ neighbourhood of φ which is contained in U. We have

$$U \supset U_{\lambda_0,\mu_0} \ni \varphi = \varphi_{\lambda_0} + \ldots + \varphi_{\lambda_0+\mu_0}$$

for some $\lambda_0, \mu_0 \in \mathbb{N}$ and $\varphi_{\lambda_0+\mu} \in V(m_{\lambda_0+\mu}, \varepsilon_{\lambda_0+\mu}, \mathcal{B}_{\lambda_0+\mu})$ for $0 \le \mu \le \mu_0$.
We can now find a $\delta_\mu > 0$ for every $0 \le \mu \le \mu_0$ such that

$$|D^\alpha \varphi_{\lambda_0+\mu}(x)| + \delta_\mu < \varepsilon_{\lambda_0+\mu} \quad \forall\, |\alpha| \le m_{\lambda_0+\mu}$$

by the definition of the V's (as open sets). Now let $\delta = \min\{\delta_0, \ldots, \delta_{\mu_0}\}$ and define $\tau_{\mathcal{B}_n}$-neighbourhoods of the $\varphi_{\lambda_0+\mu}$ by

$$\widetilde{V}_\mu = \left\{ \psi \in \mathscr{D}_{\mathcal{B}_n} \,:\, |D^\alpha \left(\psi - \varphi_{\lambda_0+\mu}\right)(x)| < \frac{\delta}{\mu_0 + 2}, |\alpha_0| \le m_{\lambda_0+\mu_0}, x \in \mathcal{B}_n \right\}$$

for $0 \le \mu \le \mu_0$, where $\varphi_{\lambda_0+\mu}$ as regarded as functions in $\mathscr{D}_{\mathcal{B}_n}$, "chopped off" outside $\mathcal{B}_n$ as their support lays completely inside $\mathcal{B}_n$. Then for any $\psi = \psi_0 + \ldots + \psi_{\mu_0}$ with $\psi_\mu \in \widetilde{V}_\mu$

$$|D^\alpha \psi(x)| < \sum_{\mu=0}^{\mu_0} \left(|D^\alpha \varphi_{\lambda_0+\mu}(x)| + \frac{\delta}{\mu_0 + 2} \right) < \delta + \sum_{\mu=0}^{\mu_0} |D^\alpha \varphi_{\lambda_0+\mu}(x)| < \sum_{\mu=0}^{\mu_0} \varepsilon_{\lambda_0+\mu}$$

for every $x \in \mathcal{B}_n$, thus $\psi \in U_{\lambda_0,\mu_0}$, as the distance to φ is less then the difference contained in each of the V's. Furthermore $\varphi \in \widetilde{V}_0 + \ldots + \widetilde{V}_{\mu_0}$ which is a $\tau_{\mathcal{B}_n}$-open set. Thus $U \cap \mathscr{D}_{\mathcal{B}_n} \in \tau_{\mathcal{B}_n}$, which proves that U is τ-open.

Part 2: *($\tau \subseteq \hat{\tau} \;\rightarrow\; \hat{\tau}$ is finer than τ)*
Fix a τ-open set A and some $\varphi_0 \in A$. We need to find $U \in \mathcal{U}$ such that $\varphi_0 \in U \subset A$. Since A is τ-open it's intersection with any $\mathscr{D}_K$ is open. Thus we have $\varphi_0 \in V_n \in \tau_{\mathcal{B}_n}$ for every $\mathcal{B}_n$ such that the support of φ_0 is contained in it. Without loss of generality we can assume that the V_n are basic open sets, i.e.

$$V_n = \{\varphi \in \mathscr{D}_{\mathcal{B}_n} \,:\, |D^\alpha \varphi(x)| < \varepsilon_n, |\alpha| \le m_n, x \in \mathcal{B}_n\}$$

Let n_0 be the smallest integer such that $\operatorname{supp}(\varphi_0) \subseteq \mathcal{B}_{n_0}$. Now define sequences $(\widetilde{m}_n)$ and $(\widetilde{\varepsilon}_n)$ by setting

$$\widetilde{m}_\lambda = \begin{cases} m_{n_0} & \lambda \le n_0 \\ \max\{m_\lambda, \widetilde{m}_{\lambda-1}\} & \lambda > n_0 \end{cases} \qquad \widetilde{\varepsilon}_\lambda = \begin{cases} \frac{\varepsilon_{n_0}}{2^\lambda} & \lambda \le n_0 \\ \min\{\frac{\varepsilon_\lambda}{2^\lambda}, \frac{\widetilde{\varepsilon}_{\lambda-1}}{2}\} & \lambda > n_0 \end{cases}$$

Now $(\widetilde{m}_\lambda)$ is clearly increasing and $(\widetilde{\varepsilon}_\lambda)$ is decreasing, thus they define a $U \in \mathcal{U}$. The choice of $\widetilde{\varepsilon}_\lambda$ makes sure that any function in U, being the sum of finitely many functions from the $V(\widetilde{m}_\lambda, \widetilde{\varepsilon}_\lambda, \mathcal{B}_\lambda)$, will not exceed the bounds of the according V_n, up to at least the same degree of derivatives by the choice of the $\widetilde{m}_\lambda$. Thus we have that $U \subset A$ as desired.

❑

Remark 2.14

Note that this already gives us a hint towards the "size" of τ. The set of all pairs of sequences $(m_n), (\varepsilon_n)$ as in definition 2.12 is clearly not countable, so this local base is also not countable. Furthermore the set of pairs of sequences is not separable, so it is not to be expected that we can reduce it to a countable local base.

We can construct another topology in a very similar way.

Definition 2.15

Let $(m_n), (\varepsilon_n)$ as in definition 2.12 and similar to the constructions there define inductively for $\lambda, \mu \in \mathbb{N}$

$$W_{\lambda,0} := V(m_\lambda, \varepsilon_\lambda, \mathcal{B}_\lambda)$$

and

$$W_{\lambda,\mu} := \bigcup_{0 \neq \varphi \in W_{\lambda,\mu-1}} \left(\varphi + V \left(m_{\lambda+\mu}, \sup_{x\in\mathbb{R}^n} \frac{|\varphi(x)|}{5^\mu + 5^{\mu-1} + \ldots + 5}, \mathcal{B}_{\lambda+\mu} \right) \right)$$

As seen in definition 2.12 we take the union over all such $W_{\lambda,\mu}$.

$$W = \bigcup_{\substack{0\le\mu<\infty \\ 1\le\lambda<\infty}} W_{\lambda,\mu}$$

Again we find that every $\varphi \in W$ is contained in some W_{λ_0,μ_0} for some $\lambda_0, \mu_0 \in \mathbb{N}$. Thus it can be written as

$$\varphi = \varphi_{\lambda_0} + \varphi_{\lambda_0+1} + \ldots + \varphi_{\lambda_0+\mu_0}$$

such that $0 \neq \varphi_{\lambda_0} \in V(m_{\lambda_0}, \varepsilon_{\lambda_0}, \mathcal{B}_{\lambda_0})$ and

$$\varphi_{\lambda_0+\mu} \in V \left(m_{\lambda_0+\mu}, \sup_{x\in\mathbb{R}^n} \frac{|(\varphi_{\lambda_0} + \varphi_{\lambda_0+1} + \ldots + \varphi_{\lambda_0+\mu-1})(x)|}{5^\mu + 5^{\mu-1} + \ldots + 5}, \mathcal{B}_{\lambda_0+\mu} \right) \quad \forall\, 1 \le \mu \le \mu_0$$

We will denote the collection of neighbourhoods of 0 one obtains by again taking all sequences as described above by $\mathcal{W}$ and the resulting topology by $\widetilde{\tau}$.

Theorem 2.16

$\widetilde{\tau}$ is strictly finer than τ, i.e. $\tau \subsetneq \widetilde{\tau}$.

Proof:

First we need to show that $\tau \subseteq \widetilde{\tau}$ i.e. $\widetilde{\tau}$ is finer than τ. Then we will assume for the contrary that also τ is finer than $\widetilde{\tau}$, which would mean they are the same. We will then reach a contradiction, which implies that $\widetilde{\tau}$ must be strictly finer.

Part 1: *(Show that $\tau \subseteq \widetilde{\tau} \ \rightarrow \ \widetilde{\tau}$ is finer than τ)*
Let A be a τ-open set, we need to find $W \in \mathcal{W}$ such that $W \subset A$. By definition of τ we have $A \cap \mathscr{D}_{\mathcal{B}_n} \in \tau_{\mathcal{B}_n}$ for every $\mathcal{B}_n$. Thus we can find a basic $\tau_{\mathcal{B}_n}$-neighbourhood for each n, say

$$V_n = \{\varphi \in \mathscr{D}_{\mathcal{B}_n} \ : \ |D^\alpha \varphi(x)| < \varepsilon_n, |\alpha| \leq m_n, x \in \mathcal{B}_n\}$$

Consider a $W \in \mathcal{W}$ for some $(\widetilde{m}_n), (\widetilde{\varepsilon}_n)$ to be determined in the proof. If $\varphi \in W$, we know from definition 2.15 that $\varphi = \varphi_{\lambda_0} + \ldots + \varphi_{\lambda_0+\mu_0} \in W_{\lambda_0,\mu_0}$ for some $\lambda_0, \mu_0 \in \mathbb{N}$. By construction we know that $\varphi_{\lambda_0} \neq 0$ and we can assume that $\varphi_{\lambda_0+\mu_0} \neq 0$, otherwise we would have $\varphi \in W_{\lambda_0,\mu_0-1}$. A direct calculation inductively now shows that

$$\sup_{x\in\mathbb{R}^n} |D^\alpha \varphi_{\lambda_0+\mu}(x)| < \sup_{x\in\mathbb{R}^n} \frac{|(\varphi_{\lambda_0} + \ldots + \varphi_{\lambda_0+\mu-1})(x)|}{5^\mu + \ldots + 5} \leq \sup_{x\in\mathbb{R}^n} \frac{|\varphi_{\lambda_0}(x)|}{5^\mu} \tag{6}$$

for every $|\alpha| \leq \widetilde{m}_{\lambda_0}$ and every $0 \leq \mu \leq \mu_0$. Since $\sup\limits_{x\in\mathbb{R}^n} |\varphi_{\lambda_0}(x)| < \widetilde{\varepsilon_{\lambda_0}}$ we find

$$\begin{aligned} \sup_{x\in\mathbb{R}^n} |D^\alpha \varphi(x)| &\leq \sum_{\mu=0}^{\mu_0} \sup_{x\in\mathbb{R}^n} |D^\alpha \varphi_{\lambda_0+\mu}(x)| \\ &< \sum_{\mu=0}^{\infty} \frac{\widetilde{\varepsilon}_{\lambda_0}}{5^\mu} \\ &= \frac{5}{4}\widetilde{\varepsilon}_{\lambda_0} \end{aligned} \tag{7}$$

for every $|\alpha| \leq \widetilde{m}_{\lambda_0}$.
Now let $n = \lambda_0 + \mu_0$ and write $\varphi = \varphi_k + \ldots + \varphi_n$ for some $k \leq n$. Then clearly $\varphi \in \mathscr{D}_{\mathcal{B}_n}$ and by (7) we know that $\sup\limits_{x\in\mathbb{R}^n} |D^\alpha \varphi(x)| < \frac{5}{4}\widetilde{\varepsilon}_k$ for all $|\alpha| \leq \widetilde{m}_k$. So if we choose $\widetilde{\varepsilon}_k < \frac{4}{5}\varepsilon_n$ and $\widetilde{m}_k \geq m_n$ for all $1 \leq k \leq n$, where ε_n, m_n are the the variables from the respective V_n, then $\varphi \in V_n \subset A$. Iterating this process to choose all $\widetilde{m}_n$ and $\widetilde{\varepsilon}_n$ we obtain a $W \in \mathcal{W}$ which is contained in A as we can obviously without loss of generality make the sequences $\widetilde{m}_n$ increasing and $\widetilde{\varepsilon}_n$ decreasing. Therefore we see that $\widetilde{\tau}$ is finer than τ, as claimed.

Part 2: *(Examine the structure of the open sets $U \in \mathcal{U}$ and $W \in \mathcal{W}$)*
To show that $\widetilde{\tau}$ is actually strictly finer than τ, we will assume the contrary, i.e. the reverse inclusion also holds. This is, if we fix a $W \in \mathcal{W}$ then there must exist a $\mathcal{U} \ni U \subseteq W$.
Denote the sequences inducing U by (m_n) and (ε_n) and the sequences inducing W by $(\widetilde{m}_n)$ and $(\widetilde{\varepsilon}_n)$.
As $\widetilde{m}_n \to \infty$ there exists $n_0 \in \mathbb{N}$ such that $m_1 + 1 < \widetilde{m}_{n_1}$. Now let

$$\eta_{n_1} = \min\{\varepsilon_{n_1}, \widetilde{\varepsilon}_{n_1}\} \quad \text{and} \quad \rho = \sup_{\varphi\in V(m_{n_1},\eta_{n_1},\mathcal{B}_{\frac{1}{3}})} \sup_{x\in\mathcal{B}_{\frac{1}{3}}} |\varphi(x)|$$

Then there exists a function $\varphi_0 \in V(m_{n_1}, \eta_{n_1}, \mathcal{B}_{\frac{1}{3}})$ which is somewhat close to the supremum, say

$$\sup_{x\in\mathcal{B}_{\frac{1}{3}}} |\varphi_0(x)| > \frac{\rho}{2}$$

Since ρ is clearly bigger then 0 we can find an integer k such that $\rho > \frac{\widetilde{\varepsilon}_{n_1}}{k}$. But $\widetilde{\varepsilon}_{n_1}$ is at least as big as η_{n_1} by definition of η_{n_1}. Thus we find that

$$\sup_{x\in\mathcal{B}_{\frac{1}{3}}} |\varphi_0(x)| > \frac{\eta_{n_1}}{2k} \tag{8}$$

Remark: Shirai defines ρ in his paper, but it seems to be unnecessary. We will have $\rho = \eta_{n_1}$ as there will clearly be a function in $V(m_{n_1}, \eta_{n_1}, \mathcal{B}_{\frac{1}{3}})$ coming arbitrary close to η_{n_1} inside $\mathcal{B}_{\frac{1}{3}}$. Thus we can replace ρ by η_{n_1} in what we have just seen and all arguments remain true. In what follows we will only use (8) which is independent of ρ.

Continuing with the main argument we can also find a function $f \in V(m_1, \frac{\eta_{n_1}}{10k}, \mathcal{B}_{\frac{1}{3}})$ such that there exists a multi-index α_0 such that

$$|\alpha_0| = m_1 + 1 \quad \text{ and } \quad \sup_{x\in\mathcal{B}_{1/3}} |D^{\alpha_0} f(x)| > 2\widetilde{\varepsilon}_1 \tag{9}$$

since the set V only gives control over the derivatives up to degree m_1.

Part 3: *(Construct a function in U using above structure)*
Combining above constructions we now define a function which we will find to lay in U and thus by assumption also in W. First translate φ_0 by $n_1 - \frac{1}{2}$ in x_1 direction, i.e.

$$\phi_0(x) := \varphi_0\left(x_1 + \left(n_1 - \frac{1}{2}\right), x_2, \ldots, x_n\right)$$

Now define a new function as the sum of f as above and the ϕ we just defined

$$\psi(x) = f(x) + \phi_0(x)$$

By definition of η_{n_1} and since ε_n is decreasing we have $\frac{\eta_{n_1}}{10k} < \varepsilon_{n_1} < \varepsilon_1$. Thus we find that

$$f \in V(m_1, \frac{\eta_{n_1}}{10k}, \mathcal{B}_{\frac{1}{3}}) \subset V(m_1, \varepsilon_1, \mathcal{B}_1) \tag{10}$$

Furthermore φ_0 was chosen to be compactly supported in the closed ball with radius $\frac{1}{3}$. We translated this ball to be centred around $n_1 - \frac{1}{2}$. As n_1 is a natural number, the translated function ϕ_0 lays completely between n_1 and $n_1 - 1$, whence

$$\text{supp}(\phi_0) \subset \mathcal{B}_{n_1} \setminus \mathcal{B}^\circ_{n_1-1}$$

Thus the same specifications we had for φ_0 are true for the translated version if we enlarge the ball appropriately, i.e.

$$\phi_0 \in V(m_{n_1}, \eta_{n_1}, \mathcal{B}_{n_1}) \subset V(m_{n_1}, \varepsilon_{n_1}, \mathcal{B}_{n_1}) \tag{11}$$

From (10) and (11) we see (using the notation from definition 2.12)

$$\psi \in V(m_1, \varepsilon_1, \mathcal{B}_1) + V(m_{n_1}, \varepsilon_{n_1}, \mathcal{B}_{n_1}) \subset U_{1,n_1} \subset U$$

But U is contained in W by assumption, so $\psi \in W$.

Part 4: *(Show that such a function cannot exist in W)*
As $\psi \in W$ there exist $\lambda_0, \mu_0 \in \mathbb{B}$ such that $\psi \in W_{\lambda_0,\mu_0}$, using the notation from definition 2.15. We will have a look at the size of λ_0 compared to n_1 and doing so reach contradictions, which will tell us that such a λ_0 cannot exist.
Note first that the supports of ϕ_0 and f are disjoint

$$\underbrace{\mathrm{supp}(f)}_{\supset \mathcal{B}_{\frac{1}{3}}} \cap \underbrace{\mathrm{supp}(\phi_0)}_{\supset \mathcal{B}_{n_1} \setminus \mathcal{B}^\circ_{n_1-1}} = \emptyset \tag{12}$$

Now we consider two possible cases.

- *Assume $\lambda_0 < n_1$:*
  Recalling the choice of φ_0 (8) and since ϕ_0 and f have disjoint support we have

$$\frac{4}{5} \cdot \frac{\eta_{n_1}}{2k} < \frac{4}{5} \cdot \sup_{x \in \mathbb{R}^n} |\phi_0(x)| \leq \frac{4}{5} \cdot \sup_{x \in \mathbb{R}^n} |\psi(x)|$$

  As before we can write $\psi = \psi_{\lambda_0} + \ldots + \psi_{\lambda_0+\mu_0}$, and as in (6) a direct calculation shows that

$$\sup_{x \in \mathbb{R}^n} |\psi_{\lambda_0+\mu}(x)| < \sup_{x \in \mathbb{R}^n} \frac{|\psi_{\lambda_0}(x)|}{5^\mu} \qquad 0 \leq \mu \leq \mu_0$$

  Using this, similar to (7), we find

$$\begin{aligned} \frac{4}{5} \cdot \sup_{x \in \mathbb{R}^n} |\psi(x)| &\leq \frac{4}{5} \cdot \sum_{\mu=0}^{\mu_0} \sup_{x \in \mathbb{R}^n} |\psi_{\lambda_0+\mu}(x)| \\ &\leq \frac{4}{5} \cdot \sup_{x \in \mathbb{R}^n} |\psi_{\lambda_0}(x)| \cdot \sum_{\mu=0}^{\infty} \frac{1}{5^\mu} \\ &= \sup_{x \in \mathbb{R}^n} |\psi_{\lambda_0}(x)| \end{aligned}$$

  But ψ_{λ_0} is supported inside $\mathcal{B}_{\lambda_0}$ so we can restrict the last supremum to $x \in \mathcal{B}_{\lambda_0}$. By again examining the sum $\psi = \psi_{\lambda_0} + \ldots + \psi_{\lambda_0+\mu_0}$ we can make another estimation of the supremum

$$\begin{aligned} \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi(x)| &= \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}(x) + \ldots + \psi_{\lambda_0+\mu_0}(x)| \\ &\geq \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}(x)| - \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0+1}| - \ldots - \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0+\mu_0}(x)| \\ &\geq \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}(x)| - \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}| \sum_{\mu=1}^{\mu_0} \frac{1}{5^\mu} \\ &\geq \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}(x)| - \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}| \frac{1}{5} \sum_{\mu=0}^{\infty} \frac{1}{5^\mu} \\ &\geq \frac{3}{4} \cdot \sup_{x \in \mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}(x)| \end{aligned}$$

Thus we see that

$$\sup_{x\in\mathcal{B}_{\lambda_0}} |\psi_{\lambda_0}(x)| \leq \frac{4}{3}\cdot \sup_{x\in\mathcal{B}_{\lambda_0}} |\psi(x)|$$

By the assumption for λ_0 we have that $\mathcal{B}_{\lambda_0} \subseteq \mathcal{B}_{n_1-1}$ and thus $supp(\phi_0)\cap\mathcal{B}_{\lambda_0} = \emptyset$. So inside $\mathcal{B}_{\lambda_0}$ only f makes a contribution to ψ and we have

$$\frac{4}{5}\cdot\frac{\eta_{n_1}}{2k} < \frac{4}{3}\cdot \sup_{x\in\mathcal{B}_{\lambda_0}} |\psi(x)| = \frac{4}{3}\cdot \sup_{x\in\mathcal{B}_{\lambda_0}} |f(x)| \leq \frac{4}{3}\cdot\frac{\eta_{n_1}}{10k}$$

by the choice of f. But this is clearly a contradiction.

- *Assume* $\lambda_0 \geq n_1$*:*
  By construction of f (9) there exists a multi-index α_0 such that

  $$2\widetilde{\varepsilon}_1 < \sup_{x\in\mathbb{R}^n} |D^{\alpha_0} f(x)|$$

  and $|\alpha_0| = m_1 + 1 < \widetilde{m}_{n_1} \leq \widetilde{m}_{\lambda_0}$ as $\widetilde{m}_n$ is increasing.
  Furthermore by (12) we that

  $$\sup_{x\in\mathbb{R}^n} |D^{\alpha_0} f(x)| \leq \sup_{x\in\mathbb{R}^n} |D^{\alpha_0}\psi(x)|$$

  Finally writing $\psi = \psi_{\lambda_0} + \ldots + \psi_{\lambda_0+\mu_0}$ using an argument similar to (6) and (7) as before we find

  $$\begin{aligned}\sup_{x\in\mathbb{R}^n} |D^{\alpha}\psi(x)| &\leq \sum_{\mu=0}^{\mu_0} \sup_{x\in\mathbb{R}^n} |D^{\alpha}\psi_{\lambda_0+\mu}(x)| \\ &< \sum_{\mu=0}^{\infty} \frac{\widetilde{\varepsilon}_{\lambda_0}}{5^{\mu}} \\ &= \frac{5}{4}\widetilde{\varepsilon}_{\lambda_0}\end{aligned}$$

  for every α such that $|\alpha| \leq \widetilde{m}_{\lambda_0}$, which is true for α_0 as seen above. So putting all this together we obtain

  $$2\widetilde{\varepsilon}_1 < \sup_{x\in\mathbb{R}^n} |D^{\alpha_0} f(x)| \leq \sup_{x\in\mathbb{R}^n} |D^{\alpha_0}\psi(x)| \leq \frac{5}{4}\widetilde{\varepsilon}_{\lambda_0}$$

  But this is a contradiction, as $\widetilde{\varepsilon}_1 > \widetilde{\varepsilon}_{\lambda_0}$ as $\widetilde{\varepsilon}_n$ is decreasing.

Thus we have reached a contradiction and $\widetilde{\tau}$ must be finer than τ as claimed.

❑

Corollary 2.17

τ is not sequential.

Proof:
Since $\widetilde{\tau}$ is finer than τ we trivially have that a $\widetilde{\tau}$-convergent sequence is τ-convergent. For the reverse recall that a sequence (φ_n) converges to 0 in the τ-topology if and only if there exists a compact set K say, such that the support of every φ_n is contained in K and (φ_n) converges to 0 in $\mathscr{D}_K$. But this means

$$\forall\ \varepsilon > 0, m \in \mathbb{N}\ \exists\ N \in \mathbb{N} \text{ such that } \|D^\alpha \varphi_n(x)\| < \varepsilon \quad \forall\ |\alpha| \leq m, x \in K$$

Now fix a $W((m_n), (\varepsilon_n)) \in \mathcal{W}$, then $V(m_\lambda, \varepsilon_\lambda, \mathcal{B}_\lambda) \subset W$ for every λ. Let λ_0 be big enough, such that K is contained in $\mathcal{B}_{\lambda_0}$. Then, as above,

$$\exists N_0 \in \mathbb{N} \text{ such that } \|D^\alpha \varphi_n(x)\| < \varepsilon_{\lambda_0} \quad \forall\ |\alpha| \leq m_{\lambda_0}, x \in K \subset \mathcal{B}_{\lambda_0}$$

for every $n \geq N_0$. Thus $\varphi_n \in V(m_{\lambda_0}, \varepsilon_{\lambda_0}, \mathcal{B}_{\lambda_0}) \subset W$ and thus (φ_n) is also $\widetilde{\tau}$-convergent. So $\widetilde{\tau}$ is a strictly finer topology than τ with the same convergent sequences, thus by proposition 2.10 τ is not sequential.

❑

Since every first countable space is sequential this result tells us that there cannot exist a countable local base for this topology and thus it also proves that the topology is not metrisable. Throughout the essay we will encounter several cases where metrisability would have been very convenient to have, but, as mentioned before, usually we can prove a result by restricting attention to the metrisable spaces $\mathscr{D}_K$ and than prove that the result also holds for all test functions on Ω.

2.2 The space of Distributions

Now that we have equipped the Schwartz space with a suitable topology we can define the space of distributions, as mentioned before as the dual space of the Schwartz space. After the short "detour" in the last section, we are now following Rudin's constructions [10] again.

2.2.1 Definition and basic properties

Definition 2.18 *(The space of distributions)*
Let

$$\mathscr{D}'(\Omega) = \{T : \mathscr{D}(\Omega) \to \mathbb{R} \,:\, T \text{ continuous, linear}\}$$

$\mathscr{D}'(\Omega)$ is called the space of distributions.

Having this definition, we should have a look at which functionals belong to $\mathscr{D}'(\Omega)$. A general result from Functional Analysis characterises when linear functionals on a locally convex space are continuous. This provides us with a criterion to check if a given functional is a distribution.

Theorem 2.19
Let $T : \mathscr{D}(\Omega) \to \mathbb{R}$ be linear. The following are equivalent

(i) T is continuous (at 0)

(ii) T is bounded

(iii) $T|_{\mathscr{D}_K}$ is continuous for all K

(iv) For every K there exists $N_K \in \mathbb{N}$ and a constant $0 < C_K < \infty$ such that

$$|\langle \varphi, T \rangle| \leq C_K \|\varphi\|_{N_K} \quad \forall\, \varphi \in \mathscr{D}_K$$

Note that we will mostly use (iv) to verify the continuity of a given functional.

Proof:
The equivalence of (i) and (ii) is a well known fact from Functional Analysis, the same is true for (iii) and (iv), i.e. a set is bounded if and only if all (semi)norms are bounded (c.f. [10, pp.27-28]).
Trivially (i) implies (iii), so it remains to show that (iii) also implies (i).
That is for every open neighbourhood V of 0 in $\mathbb{R}$ we need to find a neighbourhood U of 0 in $\mathscr{D}(\Omega)$ such that $T(U) \subset V$. Without loss of generality we may assume that $V = \{x \in \mathbb{R} : |x| < \varepsilon\}$.
By (iv) for all K we have $N_K \in \mathbb{N}$ and a constant $0 < C_K < \infty$ such that

$$|\langle \varphi, T \rangle| \leq C_K \|\varphi\|_{N_K} \quad \forall\, \varphi \in \mathscr{D}_K$$

But then the set

$$U := \left\{ \varphi \in \mathscr{D}(\Omega) \,:\, \|\varphi\|_{N_K} < \frac{\varepsilon}{C_K} \text{ for } \operatorname{supp}(\varphi) \subset K \right\}$$

is a 0 neighbourhood in $\mathscr{D}(\Omega)$ which is as desired: For $\varphi \in U$ we have

$$|\langle \varphi, T\rangle| \leq C_K \|\varphi\|_{N_K} < C_K \cdot \frac{\varepsilon}{C_K} = \varepsilon$$

❑

Definition 2.20 *(Order of a distribution)*

If in theorem 2.19 (iv) we can choose N independently of K, i.e. such that for every K

$$|\langle \varphi, T\rangle| \leq C_K \|\varphi\|_N \quad \forall \varphi \in \mathscr{D}_K$$

we call the smallest possible choice of N the order of T. If no N works for every K we say that T has infinite order.

The following result concerned with sequential continuity of distributions is quite remarkable. We have seen that the topology on the space of test functions is not first countable, in which case sequential continuity and continuity are not equivalent in general. But as the following theorem states, despite the fact that the topology is not first countable, the structure in this case is good enough to get an equivalence again.

Lemma 2.21

Let $T : \mathscr{D}(\Omega) \to \mathbb{R}$ be linear. Then T is a distribution if and only if $\langle \varphi_n, T\rangle \underset{n\to\infty}{\longrightarrow} 0$ for every sequence (φ_n) of test functions such that $\varphi_n \underset{n\to\infty}{\longrightarrow} 0$ in $(\mathscr{D}(\Omega), \tau)$.

Proof:

Note that this is just sequential continuity at 0 and thus by linearity of T sequential continuity everywhere. As in general a map between topological spaces which is continuous is also sequentially continuous, the condition is certainly true if T is a distribution, i.e. continuous. The remarkable part is the reverse direction. In general sequential continuity may be strictly weaker than continuity when dealing with non first-countable spaces. The fact that we actually have equivalence for distributions is due to the fact that for every sequence we can always restrict to a fixed $\mathscr{D}_K$, which is metrisable.
More precisely let (φ_n) be a sequence in $\mathscr{D}(\Omega)$ that converges to 0. We saw in theorem 2.7 and corollary 2.8 that this means that the φ_n have a common compact support, i.e. $(\varphi_n) \subset \mathscr{D}_K$ and $\varphi_n \underset{n\to\infty}{\longrightarrow} 0$ in $\mathscr{D}_K$. As $\mathscr{D}_K$ is metrisable, sequential continuity implies continuity. But this is true for any sequence in $\mathscr{D}(\Omega)$ and thus for every K, so by theorem 2.19 (iii) T is continuous.

❑

Definition 2.22 *(Support of a distribution)*
Let $T \in \mathscr{D}'(\Omega)$. If $\langle \varphi, T \rangle = 0$ for every φ that is supported inside some open U, T is said to vanish in U. Let V be the union over all such U

$$V = \bigcup \{U \,:\, T \text{ vanishes in } U\}$$

Then the support of T is defined as

$$\operatorname{supp}(T) = \Omega \setminus V$$

Proposition 2.23
It can be checked that this definition has exactly the properties one might expect for this to be the support. For $T \in \mathscr{D}'(\Omega)$

a) T vanishes on $\Omega \setminus \operatorname{supp}(T)$

b) If $\varphi \in \mathscr{D}(\Omega)$ and $\operatorname{supp}(\varphi) \cap \operatorname{supp}(T) = \emptyset$ then $\langle \varphi, T \rangle = 0$

c) If $\operatorname{supp}(T) = \emptyset$ then $T = 0$

2.2.2 Some members of the space of distributions

Having this result to decide if any object is a distribution, lets find some objects belonging to $\mathscr{D}'(\Omega)$. Initially we said that we would like the continuous functions to be included in our larger space of objects. In the way we constructed the space it cannot contain the space of continuous functions in the classical way, but in what follows we will show that we can identify every continuous function with exactly one distribution. In fact something stronger is true, we can identify a much larger class of functions with distributions.

Theorem 2.24 *(Regular distributions)*
Every $f \in L_{1,loc}(\Omega)$ can be identified with a distribution T_f of degree 0 in $\mathscr{D}'(\Omega)$. This distribution is given by

$$\langle \varphi, T_f \rangle := \int_\Omega f(x) \cdot \varphi(x) dx \tag{13}$$

If $T_f = T_g$ for some $f, g \in L_{1,loc}(\Omega)$ then f and g agree almost everywhere.
The set $\{T_f \,:\, f \in L_{1,loc}(\Omega)\}$ is called the set of regular distributions.

Proof:

Part 1: *(T_f is a distribution)*
First we need to check that (13) really defines a distribution. By Theorem 2.19 (iv) we need to find $N_K \in \mathbb{N}$ and $0 < C_K < \infty$ such that $|\langle \varphi, T_f \rangle| \leq C_K \|\varphi\|_{N_K} \quad \forall\, \varphi \in \mathscr{D}_K$.
So fix K, then for any $\varphi \in \mathscr{D}_K$ the integrand is clearly 0 outside K so we have

$$\langle \varphi, T_f \rangle \leq \int_K |f(x)| dx \cdot \|\varphi\|_0$$

Hence we can choose $N_K = 0$ and $C_K = \int_K |f(x)| dx$ to find that T_f is indeed a distribution of degree 0.

Part 2: *($T_f = T_g \Rightarrow f = g$ a.e.)*
Clearly we have

$$\begin{aligned} T_f = T_g &\Leftrightarrow \int_\Omega f(x) \cdot \varphi(x) dx = \int_\Omega g(x) \cdot \varphi(x) dx \quad \forall \varphi \in \mathscr{D}(\Omega) \\ &\Leftrightarrow \int_\Omega (f(x) - g(x)) \cdot \varphi(x) dx = 0 \quad \forall \varphi \in \mathscr{D}(\Omega) \end{aligned}$$

But $\mathscr{D}(\Omega)$ is dense in $L_{1,loc}(\Omega)$, so we can choose φ arbitrary close to $f - g$ to find that

$$\begin{aligned} T_f = T_g &\Leftrightarrow \int_\Omega (f(x) - g(x))^2 \, dx = 0 \\ &\Leftrightarrow \int_\Omega |f(x) - g(x)| \, dx = 0 \end{aligned}$$

Thus f and g agree almost everywhere.

❑

Notation 2.4:
We will be abusing notation and write $f \in \mathscr{D}'(\Omega)$ instead of $T_f \in \mathscr{D}'(\Omega)$ as long as no confusion is to be expected.

So we indeed constructed a space which extends the space of continuous functions in a certain sense. But the space $\mathscr{D}'(\Omega)$ is much bigger then $L_{1,loc}$ as the following results show.

Lemma 2.25 *(Dirac delta distribution)*
For all $x_0 \in \Omega$ a distribution of degree 0 is given by

$$\langle \varphi, \delta_{x_0} \rangle := \varphi(x_0)$$

This is called the Dirac delta distribution. It is not a regular distribution.

Proof:
We have

$$|\langle \varphi, \delta_{x_0} \rangle| = |\varphi(x_0)| \leq \|\varphi\|_0$$

So choosing $N_K = 0$ and $C_K = 1$ for all K in Theorem 2.19 (iv) shows that δ_{x_0} is in $\mathscr{D}'(\Omega)$ and of degree 0.
It remains to proof that δ_{x_0} is not regular. To do this, assume the contrary, that is there exists $f \in L_{1,loc}$ such that for all $\varphi \in \mathscr{D}(\Omega)$ we have

$$\langle \varphi, \delta_{x_0} \rangle = \int_\Omega f(x) \cdot \varphi(x) dx \tag{14}$$

But this is not true as can be seen by considering the following parametrised test function.
Let

$$\varphi_\epsilon := \begin{cases} e^{-\frac{1}{1-\left|\frac{x-x_0}{\epsilon}\right|^2}} & x_0 - \epsilon < x < x_0 + \epsilon \\ 0 & \text{otherwise} \end{cases}$$

Then $\langle \varphi_\epsilon, \delta_{x_0} \rangle = \varphi_\epsilon(x_0) = e^{-1}$. But as φ_ϵ is bounded from above by e^{-1} for all ϵ, assuming (14) holds and estimating the right hand side we find

$$\left|e^{-1}\right| = \left| \int_\Omega f(x) \cdot \varphi_\epsilon(x) dx \right| \leq \int_{x_0-\epsilon}^{x_0+\epsilon} |f(x)| \cdot e^{-1} dx$$

Passing to the limit $\epsilon \to 0$ this integral clearly converges to 0 for every $L_{1,loc}$ function. But this is a contradiction to being bounded from below by e^{-1}. Hence δ_{x_0} cannot be a regular distribution.

Lemma 2.26

Similar to Theorem 2.24 one can identify certain measures with distributions. For μ a complex Borel measure or μ a positive measure which is finite for all K a distribution is defined by

$$\langle \varphi, T_\mu \rangle := \int_\Omega \varphi d\mu$$

Usually one will again simply write μ instead of T_μ as long as no confusion is to be expected.

Remark 2.27

We see that we indeed achieved our goal of enlarging the space of continuous functions, in fact, looking at above examples, it seems like we enlarged it quite a lot. However, it turns out that we enlarged it by a pretty reasonable amount. One can show that, at least locally, every distribution is given by a partial derivative of some continuous function. To provide every continuous functions with all possible partial derivatives, we therefore couldn't use any subspace of the distributions. More detailed information can be found e.g. in Rudin's book [10].

2.2.3 The Topology of $\mathscr{D}'(\Omega)$

We will equip the space of distributions with the weak* topology, i.e. the topology of pointwise convergence on $\mathscr{D}(\Omega)$.
We will start with a quick reminder how the weak* topology looks like.

Proposition 2.28 *(w^*-open sets)*
We say a set $V \subset \mathscr{D}'(\Omega)$ is w^*-open if and only if

$$\forall\, T_0 \in V \ \exists\, n \in \mathbb{N}, \varphi_1, \ldots, \varphi_n \in \mathscr{D}(\Omega), \varepsilon > 0 \text{ such that}$$
$$T_0 \in \big\{T \in \mathscr{D}'(\Omega) \,:\, |\langle \varphi_i, T_0 - T\rangle| < \varepsilon \text{ for } 1 \le i \le n\big\} \subseteq V$$

Proposition 2.29 *(w^*-convergence)*
Let $(T_n) \subset \mathscr{D}'(\Omega), T \in \mathscr{D}'(\Omega)$. We say T_n converges w^* to T and write $T_n \overset{w^*}{\to} T$ if and only if

$$\langle \varphi, T_n\rangle \to \langle \varphi, T\rangle \quad \forall\, \varphi \in \mathscr{D}(\Omega)$$

in $\mathbb{R}$. Note that this illustrates why we often say that the weak* topology is the topology of pointwise convergence.

Proposition 2.30 *(w^*-boundedness)*
A set $B \subset \mathscr{D}'(\Omega)$ is w^*-bounded if for all $\varphi \in \mathscr{D}(\Omega)$ the set

$$\{\langle \varphi, T\rangle \,:\, T \in B\}$$

is bounded in $\mathbb{R}$.

Having recalled these basic properties of the weak* topology it would be reasonable to examine compactness, as compactness is very desirable for many applications. However this will be done very detailed and thus in section 5. Instead we will now examine “how big” the weak* topology on the space of distributions is.

Theorem 2.31
The weak* topology on $\mathscr{D}'(\Omega)$ is not first countable, i.e. it does not have a countable local neighbourhood base.

Proof:
Assume there is a countable base of neighbourhoods of $0 \in \mathscr{D}'(\Omega)$, $\mathcal{B}$ say.

Part 1: *(Gain control by replacements with standard basic sets)*
Clearly we can obtain a base of the neighbourhoods of 0 in $(\mathscr{D}'(\Omega), w^*)$ by taking finite intersections of the sets

$$W_{\varphi,\varepsilon} := \big\{T \in \mathscr{D}'(\Omega) \,:\, |\langle \varphi, T\rangle| < \varepsilon\big\}$$

for all $\varepsilon > 0$ and $\varphi \in \mathscr{D}(\Omega)$. So for every fixed $V \in \mathcal{B}$ we can find $n \in \mathbb{N}, \varphi_1, \dots, \varphi_n \in \mathscr{D}(\Omega)$ and $k_1, \dots, k_n$ such that

$$W_{\varphi_1, \frac{1}{k_1}} \cap \dots \cap W_{\varphi_n, \frac{1}{k_n}} \subset V \tag{15}$$

Furthermore we can choose the φ_i to be linearly independent. To see this assume they were not linearly independent i.e. $\varphi_n = \lambda_1 \cdot \varphi_1 + \ldots + \lambda_{n-1} \cdot \varphi_{n-1}$. Now choose in (15) $k_1 = \ldots = k_{n-1} := k' > (|\lambda_1| + \ldots + |\lambda_{n-1}|) \cdot k_n$. But then

$$\begin{aligned}\bigcap_{i=1}^{n-1} W_{\varphi_i, \frac{1}{k'}} &\subset \bigcap_{i=1}^{n-1} \left\{ T \in \mathscr{D}'(\Omega) \, : \, |\langle \varphi_i, T \rangle| < \frac{1}{(|\lambda_1| + \ldots + |\lambda_{n-1}|) \cdot k_n} \right\} \\ &\subset \left\{ T \in \mathscr{D}'(\Omega) \, : \, |\lambda_1| \cdot |\langle \varphi_1, T \rangle| + \ldots + |\lambda_{n-1}| \cdot |\langle \varphi_{n-1}, T \rangle| < \frac{1}{k_n} \right\} \\ &= W_{\varphi_n, \frac{1}{k_n}}\end{aligned}$$

So if the φ_i were linearly dependent we could discard some of them by choosing the k_i big enough so that the remaining φ_i are linearly independent.
If we now replace every $V \in \mathcal{B}$ by the intersection (15) we obtain $\mathcal{B}'$, which is still a countable base of neighbourhoods of 0.

Part 2: *(Construct a basic open set not containing any set in the base)*
Now let $X := \bigcup_{V \in \mathcal{B}} \{\varphi_1, \dots, \varphi_n\}$ as above. Since we take a countable union of finite sets X is countable, so we can relabel so that $X = \{\varphi_n \, : \, n \in \mathbb{N}\}$.
But the space of test functions has uncountable dimension, e.g. for all $\alpha > 0$ the functions $\varphi_\alpha(x) = e^{\frac{-\alpha}{1-|x|^2}}$ for $x \in \mathcal{B}_1$ and $\varphi_\alpha(x) = 0$ outside this ball are all linearly independent. Thus we can find a test function which is not in the span of X.
Fix $\psi \in \mathscr{D}(\Omega) \setminus \operatorname{span}\{X\}$. Using the Hahn-Banach-Theorem for locally convex spaces we can find distributions T_j such that

$$T_j(\psi) = 1 \quad \text{and} \quad T_j(\varphi_i) = \frac{1}{j} \qquad \text{for } i = 1, \dots, j$$

By definition $T_j \in W_{\varphi_i, \frac{1}{k_i}}$ if $\langle \varphi_i, T \rangle < \frac{1}{k_i}$
But by construction we have control over $\langle \varphi_i, T \rangle$ if $j > i$ and it is less then $\frac{1}{k_i}$ if $j > k_i$. So we certainly have $T_j \in W_{\varphi_i, \frac{1}{k_i}}$ if $j > \max(i, k_i)$. Hence

$$T_j \in W_{\varphi_i, \frac{1}{k_i}} \quad \forall \; j > \max(i, k_i)$$

Since $\langle \psi, T_j \rangle = 1$ for all j, this means no T_j can be contained in $W_{\psi,1} = \{T \in \mathscr{D}'(\Omega) \, : \, |\langle \psi, T \rangle| < 1\}$. So by the construction in part 1 no $U \in \mathcal{B}$ can be contained in $W_{\psi,1}$. But this is a contradiction as $W_{\psi,1}$ is clearly an open set, hence it would have to contain a set of $\mathcal{B}$ if $\mathcal{B}$ was a local base.

❑

Actually even more is true, the weak* topology on $\mathscr{D}'(\Omega)$ is not even sequential. This is strongly related to the compactness properties and will thus also be discussed in chapter 5.

2.3 Operations on $\mathscr{D}'(\Omega)$

2.3.1 Derivation

We have now extended the space of continuous functions to a much larger class of objects, but the initial goal was to construct an extension with a generalised derivative. As we need the generalised derivative to agree with the classical one, where it is defined, lets first look at what we can find for differentiable functions. To simplify the notation also assume that $\Omega \subseteq \mathbb{R}$.
Let $f \in C^1(\Omega)$ and T_f the associated regular distribution. We want $T_f' = f'$ in the distributional sense, so $T_f' = T_{f'}$. But by investigating the right hand side we find

$$T_{f'} = \int_\Omega f'(x)\varphi(x)dx \overset{IBP}{=} -\int_\Omega f(x)\varphi'(x)dx$$

As φ is compactly supported, the boundary terms of the integration by parts vanish. For sufficiently differentiable functions one can iterate this argument to find

$$\int_\Omega f^{(n)}(x)\varphi(x)dx \overset{IBP}{=} (-1)^n \int_\Omega f(x)\varphi^{(n)}(x)dx$$

We see that we can, in a manner of speaking, move the derivative over to φ. This gives rise to the following definition.

Definition 2.32 *(Generalised derivative)*
Let $T \in \mathscr{D}'(\Omega)$. Then the distributional derivative of T is defined as

$$\langle \varphi, D^\alpha T\rangle := (-1)^{|\alpha|} \cdot \langle D^\alpha \varphi, T\rangle$$

Lemma 2.33
Let $T \in \mathscr{D}'(\Omega)$, then $D^\alpha T$ is also a distribution.

Proof:
By Theorem 2.19 (iv) we have that for any fixed K

$$|\langle \varphi, T\rangle| \leq C_K \|\varphi\|_{N_K} \quad \forall \varphi \in \mathscr{D}_K$$

for $N_K \in \mathbb{N}$ and $0 < C_K < \infty$.
But then we have

$$\begin{aligned} |\langle \varphi, D^\alpha T\rangle| &= |\langle D^\alpha \varphi, T\rangle| \\ &\leq C_K \|D^\alpha \varphi\|_{N_K} \\ &\leq C_K \|\varphi\|_{N_K + |\alpha|} \end{aligned}$$

Hence $D^\alpha T \in \mathscr{D}'(\Omega)$ by Theorem 2.19 (iv).

❑

Remark 2.34
Note that this shows that if T is of degree N, then $D^\alpha T$ is of degree $N + |\alpha|$.

Obviously we have all usual derivation laws for the generalised derivative. The observations before the definition of the generalised derivative make clear that for continuously differentiable functions the classical and the generalised derivative are coherent, i.e. $D^\alpha T_f = T_{D^\alpha f}$.

2.3.2 Multiplication

We extended the vector space of continuous functions to a much larger space equipped with a generalised derivative. However, the space of continuous functions also has an algebra structure. Hence one might wish to also extend the multiplication of continuous functions to the space of distributions. Unfortunately this is not possible in general. We will see in this section that a multiplication of two distributions can only be defined if a certain amount of structure is assumed for at least one of the factors. That is, the space of distributions with its generalised derivative cannot be given an algebra structure which is coherent with the multiplication of continuous functions.
Further reading can be found in the book by Colombeau [3] which has been the main reference for this section or in the original paper by Laurent Schwartz [11].

Definition 2.35
Let $C(\Omega)$ be the algebra of continuous functions on Ω.
Denote it's unit, the constant 1 function, by $\mathbb{1}$.

Theorem 2.36
Let A be an algebra extension of $C(\Omega)$ equipped with a generalised derivative $D : A \to A$, satisfying the product rule. Denote by x the element in A such that $D\mathrm{x} = \mathbb{1}$. Then $D^2(|\mathrm{x}|) = 0$.

Proof:

Part 1: *(Compare the classical and generalised derivative of a suitable function)*
Consider $\mathrm{x} \cdot |\mathrm{x}| \in C(\Omega) \subset A$. By derivating twice, using the generalised derivative and applying the Leibniz rule, one finds

$$\begin{aligned} D^2(\mathrm{x} \cdot |\mathrm{x}|) &= D\left(\underbrace{D(\mathrm{x})}_{=\mathbb{1}} \cdot |\mathrm{x}| + \mathrm{x} \cdot D(|\mathrm{x}|)\right) \\ &= D(|\mathrm{x}|) + \underbrace{D(\mathrm{x})}_{=\mathbb{1}} \cdot D(|\mathrm{x}|) + \mathrm{x} \cdot D^2(|\mathrm{x}|) \\ &= 2 \cdot D(|\mathrm{x}|) + \mathrm{x} \cdot D^2(|\mathrm{x}|) \end{aligned}$$

But $\mathrm{x} \cdot |\mathrm{x}|$ is also classically differentiable, so we have

$$\begin{aligned} D^2(\mathrm{x} \cdot |\mathrm{x}|) &= D\left(2 \cdot |\mathrm{x}|\right) \\ &= 2 \cdot D(|\mathrm{x}|) \end{aligned}$$

This tells us that $\mathrm{x} \cdot D^2(|\mathrm{x}|) = 0$.
But x is invertible in A so that

$$\begin{aligned} 0 &= \mathrm{x}^{-1} \cdot \big(\mathrm{x} \cdot D^2(|\mathrm{x}|)\big) \\ &= \big(\mathrm{x}^{-1} \cdot \mathrm{x}\big) \cdot D^2(|\mathrm{x}|) \\ &= \mathbb{1} \cdot D^2(|\mathrm{x}|) \end{aligned}$$

So $D^2(|\mathrm{x}|) = 0$ as claimed.

Part 2: *(x is invertible in A)*
To see that x is invertible, consider $\mathrm{x}^2 \cdot (\log|\mathrm{x}| - 1)$. This is a continuous function on $\Omega \setminus \{0\}$ and can by continuously extended by giving it the value 0 at $\mathrm{x} = 0$. Then the continuous extension lays in A. As the extension is unique we can identify it with the original function. Thus in what follows we will regard $\mathrm{x}^2 \cdot (\log|\mathrm{x}| - 1)$ as being continuous everywhere. The extension is even differentiable, hence we can apply the same idea as before and compare the generalised and classical derivative.
For the generalized derivative, applying the product rule, one finds

$$\begin{aligned} &D^2\left[\mathrm{x}^2 \cdot (\log|\mathrm{x}| - 1)\right] \\ &= D^2\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1) \cdot \mathrm{x}\right] \\ &= D\left[D\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \cdot \mathrm{x} + \mathrm{x} \cdot (\log|\mathrm{x}| - 1) \cdot \underbrace{D[\mathrm{x}]}_{=\mathbb{1}}\right] \\ &= D^2\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \cdot \mathrm{x} + D\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \cdot \underbrace{D[\mathrm{x}]}_{=\mathbb{1}} + D\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \\ &= D^2\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \cdot \mathrm{x} + 2 \cdot D\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \end{aligned}$$

Looking at the classical derivative we have

$$\begin{aligned} D^2\left[\mathrm{x}^2 \cdot (\log|\mathrm{x}| - 1)\right] &= D\left[D\left[\mathrm{x}^2\right] \cdot (\log|\mathrm{x}| - 1) + \mathrm{x}^2 \cdot D\left[(\log|\mathrm{x}| - 1)\right]\right] \\ &= D\left[2 \cdot \mathrm{x} \cdot (\log|\mathrm{x}| - 1) + \mathrm{x}^2 \cdot \frac{1}{\mathrm{x}}\right] \\ &= 2 \cdot D\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] + \mathbb{1} \end{aligned}$$

Plugging both together and rearranging we get

$$\begin{aligned} D^2\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \cdot \mathrm{x} + 2 \cdot D\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] &= 2 \cdot D\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] + \mathbb{1} \\ \Leftrightarrow D^2\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] \cdot \mathrm{x} &= \mathbb{1} \end{aligned}$$

So we have $D^2\left[\mathrm{x} \cdot (\log|\mathrm{x}| - 1)\right] = \mathrm{x}^{-1}$ which completes the proof.

❑

Corollary 2.37

The algebra structure of the space of continuous functions cannot be extended to the space of distributions.

Proof:
Assume the algebra structure can be extended to $\mathscr{D}'(\Omega)$. Then by Theorem 2.36 we know that $D^2(|\mathrm{x}|) = 0$.
We can view $D^2(|\mathrm{x}|)$ as a distribution. To understand a distribution we have to study its effect on an arbitrary test function.

$$\langle \varphi, D^2(|\mathrm{x}|) \rangle = \langle D^2\varphi, |\mathrm{x}| \rangle = \int_\Omega |\mathrm{x}|(x) \cdot D^2\varphi(x) dx$$

We can split the integral in two parts by considering

$$\Omega^- := \{x \in \Omega \,:\, \mathrm{x}(x) < 0\}$$

and

$$\Omega^+ := \{x \in \Omega \,:\, \mathrm{x}(x) > 0\}$$

as $\{x \in \Omega \,:\, \mathrm{x}(x) = 0\}$ is a zero set since by continuity and the choice of x ($D\mathrm{x} \equiv 1$) we have $\mathrm{x}(x) = 0 \Leftrightarrow x = 0$.

$$\begin{aligned}
& \int_\Omega |\mathrm{x}|(x) \cdot D^2\varphi(x) dx \\
= & \int_{\Omega^+} |\mathrm{x}|(x) \cdot D^2\varphi(x) dx + \int_{\Omega^-} |\mathrm{x}|(x) \cdot D^2\varphi(x) dx \\
\overset{\text{IBP}}{=} & \int_{\partial\Omega^+} |\mathrm{x}|(x) D\varphi(x) dx + \int_{\partial\Omega^-} |\mathrm{x}|(x) D\varphi(x) dx - \left(\int_{\Omega^+} \mathbb{1} \cdot D^2\varphi(x) dx + \int_{\Omega^-} -\mathbb{1} \cdot D^2\varphi(x) dx \right) \\
= & \int_{\Omega^+} -\mathbb{1} \cdot D\varphi(x) dx + \int_{\Omega^-} \mathbb{1} \cdot D\varphi(x) dx
\end{aligned}$$

The boundary terms vanish as $\partial\Omega^+$ and $\partial\Omega^-$ can be split into a part of $\partial\Omega$ and a part where $\mathrm{x} = 0$. So on either part of the boundary one of the factors is 0. Applying integration by parts again yields

$$\begin{aligned}
& \int_{\Omega^+} -\mathbb{1} \cdot D\varphi(x) dx + \int_{\Omega^-} \mathbb{1} \cdot D\varphi(x) dx \\
\overset{\text{IBP}}{=} & \int_{\partial\Omega^+} -\mathbb{1} \cdot \varphi(x) dx + \int_{\partial\Omega^-} \mathbb{1} \cdot \varphi(x) dx - \int_{\Omega^+} 0 \cdot D\varphi(x) dx - \int_{\Omega^-} 0 \cdot D\varphi(x) dx \\
= & \; 2 \cdot \varphi(0)
\end{aligned}$$

This shows that $D^2(|\mathrm{x}|) = 2\delta_0$ in the space of distributions. But this is a contradiction to Theorem 2.36 if the algebra structure of the continuous functions could be extended to the space of distributions.

❑

Remark 2.38

We actually proved something stronger. Corollary 2.37 tells us that there cannot be an Algebra extension of the continuous functions, which contains the Dirac δ distribution.

However the multiplication can be extended to some cases by imposing structure on at least on of the factors. To see one example, assume the first factor to be a regular distribution given by a smooth function. In this case a sensible multiplication can be defined with the second factor being an arbitrary distribution. One can say that the good structure of the first factor, being smooth, evens out the bad structure of the second, being completely arbitrary.

Proposition 2.39

Let $T \in \mathscr{D}'(\Omega), f \in C^\infty(\Omega)$. Define the product $f \cdot T$ by

$$\langle \varphi, f \cdot T \rangle = \langle f \cdot \varphi, T \rangle$$

Then $f \cdot T \in \mathscr{D}'(\Omega)$.

We will see another example when a sensible multiplication can be defined in chapter 4.

3 The space of Tempered Distributions

Recall that the basic idea to construct the generalised derivative on the space of distributions was to push the derivative to the test function. This was a sensible approach because the space of test functions is closed under derivation.
Being able to derive a very large class of objects, one might wish to extend other powerful tools of Analysis to a larger class of objects. Many problems arising in PDEs and elsewhere can be tackled using Fourier transformation. Hence we shall shrink the space of distributions such that we can extend the Fourier transformation to this space. This subspace of the distributions will be called the space of tempered distributions. To shrink the space of distributions appropriately we will enlarge the Schwartz space to a space which is closed under Fourier transformation.
There exist further generalisations, but it is custom to regard these spaces on the whole of $\mathbb{R}^n$. We will follow this practise and hence also compare the new spaces to the special cases $\mathscr{D}(\mathbb{R}^n)$ and $\mathscr{D}'(\mathbb{R}^n)$. This is in line with the way the tempered distributions are introduced in the book by Rudin [10], which we will again be following to a large extend in this section.

3.1 The Space of rapidly decreasing functions

Definition 3.1 *(The space of rapidly decreasing functions)*
Let

$$\mathscr{S}(\mathbb{R}^n) := \left\{ f \in C^\infty(\Omega) \,:\, \sup_{x\in\Omega} |x^n \cdot D^\alpha f(x)| < \infty \;\forall\, n \in \mathbb{N},\, \forall \alpha \right\}$$

i.e. the space of all smooth functions such that the function and all it's derivatives decrease faster than any power of $\frac{1}{|x|}$.
$\mathscr{S}(\mathbb{R}^n)$ is called the space of rapidly decreasing functions.

It is obvious that every test function φ is also a rapidly decreasing function. We know that φ and all it's derivatives are compactly supported so obviously have a global maximum. Looking at the definition we can easily construct a family of norms on $\mathscr{S}(\mathbb{R}^n)$.

3.1.1 The topology of $\mathscr{S}(\mathbb{R}^n)$

Definition 3.2
We can define a countable family of norms on the space of rapidly decreasing functions by

$$\sigma_{m,n}(f) := \max_{|\alpha|\le m} \sup_{x\in\Omega} \left(1 + |x|^2\right)^n |D^\alpha f(x)|$$

This family of norms turns $\mathscr{S}(\mathbb{R}^n)$ into a locally convex space. We will denote the induced topology by σ.

Remark 3.3
Since $\left(1+|x|^2\right)^n$ is bounded on K for every fixed n, $(\sigma_{m,n})_{m\in\mathbb{N}}$ gives a family of norms which is equivalent to $(\|\cdot\|_m)$ on $\mathscr{D}_K$, so the subspace topology of σ on $\mathscr{D}_K$ is the same as τ_K.

As the topologies induced on $\mathscr{D}_K$ coincide, the embedding $i_{\mathscr{D}_K}$ of $\mathscr{D}_K$ into $\mathscr{S}(\mathbb{R}^n)$ is τ_K-to-$\sigma|_{\mathscr{D}_K}$-continuous. Hence $i : \mathscr{D}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is τ-to-σ-continuous.

Theorem 3.4

The test functions are dense in the rapidly decreasing functions, with respect to the topology σ.

$$\mathscr{D}(\mathbb{R}^n) \overset{d}{\hookrightarrow} \mathscr{S}(\mathbb{R}^n)$$

Proof:

Let $f \in \mathscr{S}(\mathbb{R}^n)$. For every $\varepsilon > 0$ and every $m, n \in \mathbb{N}$ we need to find $\varphi_{m,n,\varepsilon} \in \mathscr{D}(\mathbb{R}^n)$ such that $\sigma_{m,n}(f - \varphi_{m,n,\varepsilon}) < \varepsilon$. Fix m, n and ε and choose $\varphi \in \mathscr{D}(\mathbb{R}^n)$ such that $\varphi|_{B(0,1)} \equiv 1$. Now let

$$\varphi_t(x) := f(x) \cdot \varphi(t \cdot x)$$

For the difference $f - \varphi_t$ we find

$$\begin{aligned}(f - \varphi_t)(x) &= f(x) - f(x) \cdot \varphi(t \cdot x) \\ &= f(x) \cdot (\mathbb{1}(x) - \varphi(t \cdot x)) \\ &= f(x) \cdot (\mathbb{1} - \varphi)(t \cdot x) \end{aligned} \tag{16}$$

Differentiating (16) using the Leibniz formula and chain rule yields

$$D^\alpha (f - \varphi_t)(x) = \sum_{\beta \leq \alpha} \binom{|\alpha|}{|\beta|} (D^{\alpha-\beta} f)(x) \cdot t^{|\beta|} \cdot (D^\beta (\mathbb{1} - \varphi))(t \cdot x)$$

But $(\mathbb{1} - \varphi)(t \cdot x) = 0$ for every $t \in B\left(0, \frac{1}{t}\right)$ since φ was chosen to be identically 1 on the unit ball. Thus also $D^\alpha(f - \varphi_t)(x) = 0$ on $B\left(0, \frac{1}{t}\right)$, hence we can restrict attention to x outside this ball.

$$\sigma_{m,n}(f - \varphi_t) = \max_{|\alpha| \leq m} \sup_{x \notin B\left(0, \frac{1}{t}\right)} \left(1 + |x|^2\right)^n |D^\alpha (f - \varphi_t)(x)| \tag{17}$$

Now note that f is a rapidly decreasing function, so f and all it's derivatives decrease faster than any power of $\frac{1}{|x|}$ as x gets large. Therefore $(D^{\alpha-\beta} f)(x)$ will cancel out the growth of $\left(1 + |x|^2\right)^n$ for sufficiently large x. We can achieve this as, by making t small in (17), we can restrict the inner supremum to arbitrary large x.

As $t^{|\beta|}$ converges to zero when t goes to zero and $(\mathbb{1} - \varphi)$ is bounded, we see that $\sigma_{m,n}(f - \varphi_t)$ gets arbitrary small as $t \to 0$.

❏

Remark 3.5

One can also show the following inclusions

(a) $\mathscr{S}(\mathbb{R}^n) \overset{d}{\hookrightarrow} L_p(\mathbb{R}^n)$

(b) $\mathscr{S}(\mathbb{R}^n) \overset{d}{\hookrightarrow} C_0(\mathbb{R}^n)$

where $C_0(\mathbb{R}^n)$ denotes the space of continuous functions vanishing at infinity.

Again we would like our topology to be complete to get a well behaved dual space. This is indeed the case as stated in the following lemma.

Lemma 3.6
$\mathscr{S}(\mathbb{R}^n)$ is a Fréchet space.

Proof:
The topology is certainly metrisable as it is a countable family of norms. Hence we only need to check completeness.
Therefore let $(f_k) \subset \mathscr{S}(\mathbb{R}^n)$ be a σ-Cauchy sequence. We will show that this is also a Cauchy sequence in $C^\infty(\mathbb{R}^n)$ with respect to the topology τ_Ω which is complete. We will then show that the limit lies in $\mathscr{S}(\mathbb{R}^n)$ and that it also converges in $\mathscr{S}(\mathbb{R}^n)$ with respect to σ.

Part 1: *(σ-Cauchy $\Rightarrow$ τ_Ω-Cauchy)*
Note that (f_k) being a σ-Cauchy sequence means that for every $n \in \mathbb{N}$ and $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\sigma_{m,n}(f_{k_1} - f_{k_2}) < \varepsilon$ for all $k_1, k_2 \geq N$. So, as $\left(1+|x|^2\right)^n$ is bounded on any K, this means in particular that

$$\begin{aligned}
\varepsilon &> \max_{|\alpha|\leq m} \sup_{x\in\Omega} \left(1+|x|^2\right)^n |D^\alpha (f_{k_1} - f_{k_2})(x)| \\
&\geq \max_{|\alpha|\leq m} \sup_{x\in K} \left(1+|x|^2\right)^n |D^\alpha (f_{k_1} - f_{k_2})(x)| \\
&\geq \max_{|\alpha|\leq m} \sup_{x\in K} C \cdot |D^\alpha (f_{k_1} - f_{k_2})(x)|
\end{aligned}$$

By absorbing the constant C into ε we get that for every $n \in \mathbb{N}$ and $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\rho_m (f_{k_1} - f_{k_2}) < \varepsilon \text{ for all } k_1, k_2 \geq N$$

So (f_k) is also a Cauchy sequence in $C^\infty(\mathbb{R}^n)$ with the topology τ_Ω, which is complete. So $f_k \to f$ in $C^\infty(\mathbb{R}^n)$.

Part 2: *(f is a rapidly decreasing function)*
Every Cauchy sequence is bounded, so for every $m, n \in \mathbb{N}$ we have a constant $M_{m,n} \in \mathbb{R}$ such that

$$\sigma_{m,n}(f_k) = \max_{|\alpha|\leq m} \sup_{x\in\Omega} \left(1+|x|^2\right)^n |D^\alpha f_k(x)| < M_{m,n}$$

This means in particular that for every fixed $x \in \Omega$ the supremum over all α is bounded, which gives us

$$\begin{aligned}
&\sup_{|\alpha|\leq m} \left(1+|x|^2\right)^n |D^\alpha f_k(x)| < M_{m,n} \qquad \forall\, x \in \Omega \\
\Leftrightarrow \quad &\sup_{|\alpha|\leq m} |D^\alpha f_k(x)| < \frac{M_{m,n}}{(1+|x|^2)^n} \quad \forall\, x \in \Omega
\end{aligned} \tag{18}$$

But by the definition of τ_Ω all derivatives of (f_k) must converge pointwise to the respective derivative of f. Hence from (18) we can deduce that

$$\sup_{|\alpha|\leq m} |D^\alpha f(x)| < \frac{M_{m,n}}{(1+|x|^2)^n} \quad \forall\, x \in \Omega$$

But this means that $\sigma_{m,n}(f) < \infty$ for all m, n and hence that $f \in \mathscr{S}(\mathbb{R}^n)$.

Part 3: *(The sequence also converges in the space of rapidly decreasing functions)*
To show that also $f_k \underset{k\to\infty}{\longrightarrow} f$ in $\mathscr{S}(\mathbb{R}^n)$ we can without loss of generality assume that $f = 0$, as we can replace (f_k) by $(f_k - f)$ if necessary. This is obviously also a Cauchy sequence in $\mathscr{S}(\mathbb{R}^n)$ which converges if and only if (f_k) converges.
But $f_k \underset{k\to\infty}{\longrightarrow} 0$ in $C^\infty(\mathbb{R}^n)$ means that

$$\rho_m(f_k) = \max_{\alpha|\leq m} \sup_{x\in K_m} |D^\alpha f_k(x)| \underset{k\to\infty}{\longrightarrow} 0 \quad \forall\ m \in \mathbb{N}$$

which implies that

$$\begin{aligned} &\max_{|\alpha|\leq m} |D^\alpha f_k(x)| \underset{k\to\infty}{\longrightarrow} 0 \qquad \forall\, x \in \Omega \\ \Leftrightarrow \qquad & |D^\alpha f_k(x)| \underset{k\to\infty}{\longrightarrow} 0 \quad \forall x \in \Omega, \forall\, \alpha \end{aligned}$$

From this we can deduce that for any fixed $n \in \mathbb{N}$

$$\max_{|\alpha|\leq m} \left(1+|x|^2\right)^n |D^\alpha f_k(x)| \underset{k\to\infty}{\longrightarrow} 0 \quad \forall\, x \in \Omega$$

which implies that also

$$\sigma_{m,n}(f_k) = \max_{|\alpha|\leq m} \sup_{x\in\Omega} \left(1+|x|^2\right)^n |D^\alpha f_k(x)| \underset{k\to\infty}{\longrightarrow} 0$$

Hence (f_n) also converges in $\mathscr{S}(\mathbb{R}^n)$ to f, so $(\mathscr{S}(\mathbb{R}^n), \sigma)$ is complete.

❏

Remark 3.7
A metric on $\mathscr{S}(\mathbb{R}^n)$ is given by

$$d(f,g) := \sum_{m,n=0}^{\infty} 2^{-(m+n)} \frac{\sigma_{m,n}(f-g)}{1+\sigma_{m,n}(f-g)}$$

We have already seen in the proof of the fact that we can check continuity of a linear map on the space of distributions by checking sequential continuity, that it was crucial to have a Fréchet space as underlying structure. This will also occur again in this essay. Thus we see that the fact that $\mathscr{S}(\mathbb{R}^n)$ is a Fréchet space will simplify these proofs significantly as we don't need to restrict to the $\mathscr{D}_K$ spaces anymore with subsequently extending the result to the whole space.

3.2 The space of Tempered Distributions

3.2.1 Definition and basic properties

Definition 3.8 *(The space of Tempered Distributions)*

Let

$$\mathscr{S}'(\mathbb{R}^n) := \{T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{R} \,:\, T \text{ continuous, linear}\}$$

$\mathscr{S}'(\mathbb{R}^n)$ is called the space of tempered distributions.

Lemma 3.9

Let $T \in \mathscr{D}'(\mathbb{R}^n)$ then there is at most one $S \in \mathscr{S}'(\mathbb{R}^n)$ continuously extending T to $\mathscr{S}(\mathbb{R}^n)$.

Proof:

Assume there are two distinct continuous extensions of T, S_1 and S_2 say. For S_1 and S_2 to be distinct, they have to disagree on at least one function and therefore by continuity on a small open set $U \subset \mathscr{S}(\mathbb{R}^n)$. But by Theorem 4.3 the test functions are dense in $\mathscr{S}(\mathbb{R}^n)$, thus we can find $\varphi \in U \cap \mathscr{D}(\mathbb{R}^n)$ such that $S_1(\varphi) \neq S_2(\varphi)$. But this is a contradiction as S_1 and S_2 are both extensions of T.

❑

Remark 3.10

Certainly every $S \in \mathscr{S}'(\mathbb{R}^n)$, restricted to $\mathscr{D}(\mathbb{R}^n)$, is a distributions. Thus lemma 3.9 justifies to identify S with $S|_{\mathscr{D}(\mathbb{R}^n)} \in \mathscr{D}'(\mathbb{R}^n)$ and regard $\mathscr{S}'(\mathbb{R}^n)$ as a subspace of $\mathscr{D}'(\mathbb{R}^n)$.

3.2.2 The topology of $\mathscr{S}'(\mathbb{R}^n)$

Just as the space of distributions we equip $\mathscr{S}'(\mathbb{R}^n)$ with the weak* topology, which coincides with the subspace topology of the weak* topology on the space of distributions. Despite having a better underlying structure on $\mathscr{S}(\mathbb{R}^n)$, the weak* topology on $\mathscr{S}'(\mathbb{R}^n)$ still does not have a countable neighbourhood base.

Theorem 3.11

The weak* topology on $\mathscr{S}'(\mathbb{R}^n)$ is not first countable.

Proof:

Note that the proof of Theorem 2.31 did not use any specific properties of the space of test functions or distributions. In fact the result is true for the weak* topology on the dual space of any locally convex topological vector space of uncountable dimension.
But the space of test functions is contained in $\mathscr{S}(\mathbb{R}^n)$, so it has uncountable dimension and thus the weak* topology on $\mathscr{S}'(\mathbb{R}^n)$ is not first countable.

❑

3.3 Operations on $\mathscr{S}'(\mathbb{R}^n)$

3.3.1 Derivation and Multiplication

Similarly to the space of test functions, the space of rapidly decreasing functions is closed under derivation. Therefore the derivation works nicely in exactly the same way as for distributions. This is exactly as one hopes for when regarding the tempered distributions as a subspace.

For the multiplication of two tempered distributions everything we said about multiplication of distributions applies. In section 2.3.2 we proved that there cannot be an algebra extension of the continuous functions containing the Dirac delta distribution. But as the delta distribution is tempered, this immediately implies that a multiplication of tempered distributions cannot be defined in general.

3.3.2 Fourier Transformation

As we mentioned before a major reason to construct the space of tempered distributions is to extend the definition of the Fourier transformation. We will briefly state some of the basic properties and results in this section without any proofs. There is a large amount of literature available covering the Fourier transform on the space of tempered distributions in great detail, e.g. the books by Rudin [10] or Hörmander [9]. We will use the following notation:

Definition 3.12

Let $f \in L_1(\mathbb{R}^n)$ then

$$(\mathcal{F}f)(y) := \hat{f}(y) := (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x) \cdot e^{-y \cdot x} dx \quad y \in \mathbb{R}^n$$

is the Fourier transformation of f.

Examining the effect of the Fourier transform on $L_1(\mathbb{R}^n)$ one finds that there is no simple way to describe the image, other than that it will be a subset of the space of continuous functions vanishing at infinity.

Proposition 3.13

$\mathcal{F}$ is a bounded linear operator from $L_1(\mathbb{R}^n)$ into $C_0(\mathbb{R}^n)$.

$$\mathcal{F} \in \mathscr{L}\left(L_1(\mathbb{R}^n), C_0(\mathbb{R}^n)\right)$$

This in particular means that $\mathbb{1}$ cannot be the Forier transform of a function in $L_1(\mathbb{R}^n)$.

As we mentioned before the space of rapidly decreasing functions was chosen because it is closed under the Fourier transform. At first view one can also find several other good properties.

Proposition 3.14

The Fourier transform is a bounded linear operator from the space of rapidly decreasing functions into itself. It behaves well considering derivatives and has a fixed point in $\mathscr{S}(\mathbb{R}^n)$.
Let $f \in \mathscr{S}(\mathbb{R}^n), x \in$ then

(a) $\mathcal{F} \in \mathscr{L}\left(\mathscr{S}(\mathbb{R}^n), \mathscr{S}(\mathbb{R}^n)\right)$

(b) $\widehat{\xi^\alpha f}(x) = (-1)^{|\alpha|} D^\alpha \hat{f}(x)$

(c) $\widehat{D^\alpha f}(x) = x^\alpha \hat{f}(x)$

(d) $\mathcal{F}(e^{-|\cdot|^2}) = e^{-|\cdot|^2}$

In (b) $\xi^\alpha f$ denotes the map from $\mathbb{R}^n$ to $\mathbb{R}$ given by $x \mapsto \xi^\alpha \cdot f(\xi)$.

Closer examination reveals even more, the Forier transform actually is an automorphism on the space of rapidly decreasing functions. This has the consequence that it induces an automorphism on it's dual, the space of tempered distributions.

Proposition 3.15 *(Fourier inversion theorem for $\mathscr{S}(\mathbb{R}^n)$)*

$\mathcal{F} \in \mathscr{L}\left(\mathscr{S}(\mathbb{R}^n), \mathscr{S}(\mathbb{R}^n)\right)$ is bijective and continuous, so an automorphism.
For $f \in \mathscr{S}(\mathbb{R}^n)$ the inverse is given by

$$\left(\mathcal{F}^{-1} f\right)(x) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{y \cdot x} \cdot f(x) dy = \hat{f}(-x)$$

When we extended the derivative to the space of distributions, we did this by "pushing" the derivative onto the test function as the space of test functions is closed under derivation. As we have just seen, the space of rapidly decreasing functions is closed under Fourier transform, so following the same idea, we can now extend the Fourier transform to tempered distributions by "pushing" the Fourier transform onto the rapidly decreasing function.

Definition 3.16

For $S \in \mathscr{S}'(\mathbb{R}^n)$ define the Fourier transform by

$$\hat{S}(f) := S(\hat{f}) \quad f \in \mathscr{S}(\mathbb{R}^n)$$

$\hat{F}$ is again a tempered distribution.

Just as the derivation for distributions the Fourier transform on the tempered distributions inherits the good properties of the Fourier transform on the rapidly decreasing functions.

Proposition 3.17

$\mathcal{F} \in \mathscr{L}\left(\mathscr{S}'(\mathbb{R}^n), \mathscr{S}'(\mathbb{R}^n)\right)$ is bijective and continuous, so an automorphism.
Furthermore, for $S \in \mathscr{S}'(\mathbb{R}^n), x \in \mathbb{R}^n$

(a) $\widehat{\xi^\alpha S}(x) = (-1)^{|\alpha|} D^\alpha \hat{S}(x)$

(b) $\widehat{D^\alpha S}(x) = y^\alpha \hat{S}(x)$

In (a) $\xi^\alpha S$ denotes the map from $\mathscr{S}(\mathbb{R}^n)$ to $\mathbb{R}$ given by $f \mapsto S(\xi^\alpha f)$ using the notation from proposition 3.14.

4 Schwarz Kernel

In section 2.3.2 we saw that it is in general not possible to multiply two distributions. However there is a special case in which a general product can be defined. If we have two distributions which are defined on two independent sets there exists a well defined product, coherent with the corresponding product for functions. This case also sheds light on some other interesting properties.
For the rest of this chapter we will work with two independent open sets rather then one as before, i.e. we will consider $\Omega_1 \subset \mathbb{R}^{n_1}$ and $\Omega_2 \subset \mathbb{R}^{n_2}$. We will in general use the index $j = 1, 2$ for statements referring to both of these sets. In notation we will follow Blanchard and Brüning [2] as well as Hörmander [9], as these also provide most of the information presented throughout this chapter. Furthermore Trèves' book [15] has been an important source of information complementing the material presented in the former two books.

4.1 Tensor Products

To find a reasonable definition of a tensor product on the space of distributions we should start by examining the situation for functions. Especially we will take a close look at the tensor product of test functions, as it can be expected that this will shed light on the situation for distributions.

4.1.1 Tensor Product of functions

As a first step recall the tensor product of two functions.

Definition 4.1 *(Tensor Product of functions)*
Let $f_j : \Omega_j \to \mathbb{R}$. Then the tensor product of f_1 and f_2 is the function

$$(f_1 \otimes f_2)(x_1, x_2) := f_1(x_1) \cdot f_2(x_2)$$

for all $x_j \in \Omega_j$.
Clearly $f_1 \otimes f_2$ is well defined and lives on $\Omega_1 \times \Omega_2 \subset \mathbb{R}^{n_1+n_2}$.

Remark 4.2
It is obvious from above definition that the tensor product is supported on the product of the support of its factors and that it inherits regularity for differentiation (by the Leibniz rule) and integration (by Fubini) from its factors.

(a) $\text{supp}(f_1 \otimes f_2) = \text{supp}(f_1) \times \text{supp}(f_2)$.

(b) If $f_j \in C^{k_j}(\Omega_j)$ then $f_1 \otimes f_2 \in C^{\min(k_1,k_2)}(\Omega_1 \times \Omega_2)$

(c) If $f_j \in L_{1,loc}(\Omega_j)$ then $f_1 \otimes f_2 \in L_{1,loc}(\Omega_1 \times \Omega_2)$

This remark tells us that the tensor product of two test functions is again a test function on the product space. Taking all tensor products of elements of $\mathscr{D}(\Omega_j)$ we obtain a vector space which plays an important role in extending the tensor product to the space of distributions.

Notation 4.1:
We will write

$$\mathscr{D}(\Omega_1) \otimes \mathscr{D}(\Omega_2) := \operatorname{span}\{\varphi_1 \otimes \varphi_2 \,:\, \varphi_j \in \mathscr{D}(\Omega_j)\}$$

for the vector space spanned by all tensor products of two test functions on Ω_1 and Ω_2.

This space is obviously contained in $\mathscr{D}(\Omega_1 \times \Omega_2)$. The following proposition makes clear that it will be very useful for the definition of a tensor product of distributions.

Proposition 4.3
The tensor product of the test function spaces $\mathscr{D}(\Omega_1)$ and $\mathscr{D}(\Omega_2)$ is dense in the test function space on the product of Ω_1 and Ω_2.

$$\mathscr{D}(\Omega_1) \otimes \mathscr{D}(\Omega_2) \overset{d}{\hookrightarrow} \mathscr{D}(\Omega_1 \times \Omega_2)$$

4.1.2 Tensor product of distributions

Proposition 4.3 gives us a strong tool to extend the tensor product to the space of distributions, as we can reduce attention to the tensor product of test function spaces $\mathscr{D}(\Omega_1) \otimes \mathscr{D}(\Omega_2)$ to obtain results for the larger space of distributions on test functions on the product set $\mathscr{D}(\Omega_1 \times \Omega_2)$.
The functions in the tensor product of two test function spaces are still in a way separate. So as a step to understanding the situation for the test functions on the product space it is nearby to try to separate them, i.e. keep one of the variables fixed and see how a distribution acts on the other variable.

Proposition 4.4
Let $T_1 \in \mathscr{D}'(\Omega_1)$ and $\varphi \in \mathscr{D}(\Omega_1 \times \Omega_2)$. Fix $x_2 \in \Omega_2$ and regard φ as function of x_1 i.e.

$$\varphi_{x_2}(x_1) := \varphi(x_1, x_2) \in \mathscr{D}(\Omega_1)$$

Then $\langle \varphi_{x_2}, T_1 \rangle$ is well defined and identifying x_2 with $\langle \varphi_{x_2}, T_1 \rangle$

$$\begin{aligned} \psi_1 : \Omega_2 &\to \mathbb{R} \\ x_2 &\mapsto \langle \varphi_{x_2}, T_1 \rangle \end{aligned}$$

we get a test function on Ω_2, $\psi_1 \in \mathscr{D}(\Omega_2)$. The mapping

$$\begin{aligned} F_1 : \mathscr{D}(\Omega_1 \times \Omega_2) &\to \mathscr{D}(\Omega_2) \\ \varphi &\mapsto \psi_1 \end{aligned}$$

is a continuous linear mapping.
Clearly the same construction is possible for $T_2 \in \mathscr{D}'(\Omega_2)$ with Ω_1 and Ω_2 reversed, obtaining a continuous linear map

$$\begin{aligned} F_2 : \mathscr{D}(\Omega_1 \times \Omega_2) &\to \mathscr{D}(\Omega_1) \\ \varphi &\mapsto \psi_2 = (x_1 \mapsto \langle \varphi_{x_1}, T_2 \rangle) \end{aligned}$$

To simplify notation we will follow a suggestive notation widely used across literature, e.g. in [9] or [2], and write $\langle \varphi, T_1 \rangle$ for $F_1(\varphi) = \psi_1 = (x_2 \mapsto \langle \varphi_{x_2}, T_1 \rangle)$ and $\langle \varphi, T_2 \rangle$ for $F_2(\varphi)$.

The introduced notation already suggests the way to define the tensor product. We follow the construction in proposition 4.4 and, in a manner of speaking, apply one distribution after the other to the test function on the product space.

Theorem 4.5 *(Tensor Product of distributions)*
Let $T_j \in \mathscr{D}'(\Omega_j)$, then there exists a unique distribution $T_1 \otimes T_2 \in \mathscr{D}'(\Omega_1 \times \Omega_2)$ such that

$$\langle \varphi_1 \otimes \varphi_2, T_1 \otimes T_2 \rangle = \langle \varphi_1, T_1 \rangle \cdot \langle \varphi_2, T_2 \rangle \tag{19}$$

for all $\varphi_j \in \mathscr{D}(\Omega_j)$. This distribution is called the tensor product of T_1 and T_2.

Proof:
Following proposition 4.4 define a map T on $\mathscr{D}(\Omega_1 \times \Omega_2)$ by

$$\langle \varphi, T \rangle := \langle \langle \varphi, T_1 \rangle, T_2 \rangle$$

We have to show that $T = T_1 \otimes T_2$ is as stated in the theorem.
By proposition 4.4 T is well defined and linear. Furthermore, as the notation suggests we are, in a manner of speaking, applying two distributions one after the other, so it is clear that the map is also continuous. Thus is remains to show that T satisfies 19 and is uniquely determined by it.
Let $\varphi_j \in \mathscr{D}'(\Omega_j)$ and $\varphi = \varphi_1 \otimes \varphi_2$. By the construction in proposition 4.4 we see that

$$\langle \varphi, T_1 \rangle = \langle \varphi_1 \cdot \varphi_2, T_1 \rangle = \langle \varphi_1, T_1 \rangle \cdot \varphi_2$$

and thus as desired

$$\langle \varphi_1 \otimes \varphi_2, T \rangle = \langle \varphi_1, T_1 \rangle \cdot \langle \varphi_2, T_2 \rangle$$

This clearly defines T uniquely on $\mathscr{D}(\Omega_1) \otimes \mathscr{D}(\Omega_2)$. But this space is dense in $\mathscr{D}(\Omega_1 \times \Omega_2)$ by proposition 4.3, hence continuity of T implies that it is uniquely defined on the whole of $\mathscr{D}(\Omega_1 \times \Omega_2)$.

❑

The way we constructed $T_1 \otimes T_2$ in the proof indicates an immediate corollary, as the choice of order of applying T_1 and T_2 was arbitrary.

Corollary 4.6 *(Fubini's Theorem for Distributions)*
Let $T_j \in \mathscr{D}'(\Omega_j)$.

$$\langle (\varphi_1 \otimes \varphi_2, T_1 \otimes T_2 \rangle = \langle \langle \varphi, T_1 \rangle, T_2 \rangle = \langle \langle \varphi, T_2 \rangle, T_1 \rangle$$

for all $\varphi \in \mathscr{D}'(\Omega_1 \times \Omega_2)$. Particularly in view of regular distributions this can be regarded as a way of interchanging the order of integration, thus it is often called Fubini's Theorem for Distributions.

Proof:
When we constructed $T_1 \otimes T_2$ in the proof of theorem 4.5 we arbitrarily chose to apply first T_1 and then T_2. Obviously the same construction can be done with T_1 and T_2 interchanged. Clearly

$$\langle \varphi_1 \otimes \varphi_2, T \rangle = \langle \varphi_1, T_1 \rangle \cdot \langle \varphi_2, T_2 \rangle$$

implies that the two constructions coincide on $\mathscr{D}(\Omega_1) \otimes \mathscr{D}(\Omega_2)$ because of commutativity in $\mathbb{R}$. But then again by density of $\mathscr{D}(\Omega_1) \otimes \mathscr{D}(\Omega_2)$ in $\mathscr{D}(\Omega_1 \times \Omega_2)$ and continuity of $T_1 \otimes T_2$ they agree on the whole of $\mathscr{D}(\Omega_1 \times \Omega_2)$.

❑

Having two locally integrable functions we can now consider their tensor product or regard them as distributions and then form the distributional tensor product. Hence we need to check that these two products coincide.

Corollary 4.7
Let $f_j \in L_{1,loc}(\Omega_j)$, then

$$T_{f_1} \otimes T_{f_2} = T_{f_1 \otimes f_2}$$

Proof:
Fix an arbitrary $\varphi \in \mathscr{D}(\Omega_1 \times \Omega_2)$. By definition 4.1 we have $f_1 \otimes f_2 = f_1(x_1) \cdot f_2(x_2)$. Regarding this as a regular distribution yields

$$\begin{aligned} \langle \varphi, T_{f_1 \otimes f_2} \rangle &= \int_{\Omega_1 \times \Omega_2} f_1(x_1) \cdot f_2(x_2) \cdot \varphi(x_1, x_2) d(x_1, x_2) \\ &= \int_{\Omega_2} \int_{\Omega_1} f_1(x_1) \cdot f_2(x_2) \cdot \varphi(x_1, x_2) dx_1\, dx_2 \end{aligned}$$

Following the construction in proposition 4.4 we find that

$$\langle \varphi, T_{f_1} \rangle = \int_{\Omega_1} f_1(x_1) \cdot \varphi(x_1, x_2) dx_1$$

Using this and theorem 4.5 or corollary 4.6 we see that for $T_{f_1} \otimes T_{f_2}$ we have

$$\begin{aligned} \langle \varphi, T_{f_1} \otimes T_{f_2} \rangle &= \langle \langle \varphi, T_{f_1} \rangle, T_{f_2} \rangle \\ &= \langle \int_{\Omega_1} f_1(x_1) \cdot \varphi(x_1, x_2) dx_1, T_{f_2} \rangle \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f_1(x_1) \cdot \varphi(x_1, x_2) dx_1 \right) \cdot f_2(x_2) dx_2 \end{aligned}$$

So clearly, as φ was arbitrary, $T_{f_1} \otimes T_{f_2} = T_{f_1 \otimes f_2}$ for all $\varphi \in \mathscr{D}(\Omega_1 \times \Omega_2)$.

❑

4.2 The Schwartz Kernel Theorem

Having expanded the tensor product, obtaining distributions in two variables, we are now able to examine a connection to integral operator theory. Let us first recall the definition of integral operators.

Definition 4.8 *(Integral operator)*
Let F_j be function spaces on Ω_j respectively. A linear operator $T : F_2 \to F_1$ is called an integral operator or integral transform if it is of the form

$$(T(f))(x_1) = \int_{\Omega_2} K(x_1, x_2) \cdot f(x_2) dx_2 \quad f \in F_2$$

for some measurable function $K \in L_0(\Omega_1 \times \Omega_2)$. The function K is called the kernel function.

So every choice of a function of two variables K corresponds to an integral operator. Surely some caution is required in terms of the choice of K to make sure the integral always exists. The necessary regularity of K thus depends on the function space F_2, usual choices include continuous functions or square integrable functions.

Example 4.9
$K \in C(\Omega_1 \times \Omega_2)$ defines an integral operator from the space of continuous, compactly supported functions into the space of continuous functions.

$$T : C_c(\Omega_2) \to C(\Omega_1)$$

This can not only be extend to distributions, we can even write any bounded linear operator from the space of test functions in one variable into the space of distributions in a second variable as an integral operator. This is a very strong result, in general it is far from true that any bounded linear operator between two function spaces can be written as an integral transform.

Proposition 4.10 *(Schwartz kernel theorem)*
There is a one-to-one correspondence between $\mathscr{D}'(\Omega_1 \times \Omega_2)$ and $\mathscr{L}(\mathscr{D}(\Omega_2), \mathscr{D}'(\Omega_1))$. More precisely

- Every $K \in \mathscr{D}'(\Omega_1 \times \Omega_2)$ defines a $\mathcal{K} \in \mathscr{L}(\mathscr{D}(\Omega_2), \mathscr{D}'(\Omega_1))$ such that

$$\langle \varphi_1, \mathcal{K}(\varphi_2) \rangle = \langle \varphi_1 \otimes \varphi_2, K \rangle \tag{20}$$

- For every $\mathcal{K} \in \mathscr{L}(\mathscr{D}(\Omega_2), \mathscr{D}'(\Omega_1))$ there is exactly one distribution such that (20) holds

Thus the two spaces are isomorphic

$$\mathscr{D}'(\Omega_1 \times \Omega_2) \sim \mathscr{L}(\mathscr{D}(\Omega_2), \mathscr{D}'(\Omega_1))$$

As mentioned before this result is far from being obvious, in fact it lays quite deep. Let us make some comments on the proof. Trèves [15] gives a very structural proof that these spaces are isomorphic. First he establishes that the space of bounded linear operators is isomorphic to the completion of the tensor product of the two separate distribution spaces $\mathscr{D}'(\Omega_1)$ and $\mathscr{D}'(\Omega_2)$, denoted by $\mathscr{D}'(\Omega_1)\hat{\otimes}\mathscr{D}'(\Omega_2)$, by examining topological properties of the underlying space of test functions, its dual and its embedding into it's bidual. Subsequently he notes that the tensor product of the two distribution spaces is dense in both, it's completion and the space of distributions on $\mathscr{D}'(\Omega_1 \times \Omega_2)$. Consequently he shows that $\mathscr{D}'(\Omega_1)\hat{\otimes}\mathscr{D}'(\Omega_2)$ and $\mathscr{D}'(\Omega_1 \times \Omega_2)$ induce the same subspace topology on $\mathscr{D}'(\Omega_1) \otimes \mathscr{D}'(\Omega_2)$.
The approach in Hörmanders book [9] is more constructive, inspiring the first part of the proposition. The proof is rather lengthy, though. While it is straight forward that (20) defines a bounded linear map, the converse is not as easy, which was to be expected, as this is what makes the theorem so remarkable. That the condition uniquely defines a distribution is a fairly easy consequence of continuity and density, similar to what we have seen in the proof of theorem 4.5. It is however not as easy to proof the existence, the crack of the theorem. The proof given by Hörmander uses an approximating sequence of distributions. He assumes the distribution to exist and uses convolution of a distribution with a test function to construct the sequence. He than shows that the limit indeed exists and has the desired property.

Remark 4.11

The bounded linear maps in proposition 4.10 induce bilinear forms on the test functions. More precisely let $\varphi_j \in \mathscr{D}(\Omega_j)$ and $\mathcal{K} \in \mathscr{L}(\mathscr{D}(\Omega_2), \mathscr{D}'(\Omega_1))$ then

$$(\varphi_1, \varphi_2) \mapsto \langle \varphi_1, \mathcal{K}(\varphi_2) \rangle$$

is a bilinear map on $\mathscr{D}(\Omega_1) \times \mathscr{D}(\Omega_2)$.
In view of this, proposition 4.10 says that all these bilinear maps are already included in the general theory of distributions, using the structure we introduced in section 4.1.2, creating a two variable meaning for distributions.

5 Compactness in the space of Distributions

In this chapter we will closely examine the properties of compactness in the space of distributions. It will turn out that the situation is much better than one might expect at first.
As the unit ball is compact in the norm topology if and only if the space has finite dimension, the closed norm-balls are not compact in the space of distributions. Thus one doesn't have many norm compact sets. However, as we equipped the space with the weak* topology we get a large amount of w^*-compact sets by the Banach Alaoglu Bourbaki Theorem. As we are considering the weak* topology on $\mathscr{D}'(\Omega)$ speaking of compactness in this space will always refer to w^*-compactness.
Examining the situation closer we will find that we can in this case actually characterise the w^*-compactness by properties of sequences, which is not at all to be expected in the weak* topology.

5.1 The Banach Alaoglu Bourbaki Theorem

As $\mathscr{D}(\Omega)$ is not a normed spaces, we cannot apply the Banach Alaoglu theorem. Therefore we will state and prove the Banach Alaoglu Bourbaki theorem, which is a generalisation by Bourbaki which applies to locally convex spaces. To simplify notation, we will assume that we consider a vector space over $\mathbb{R}$, but the given proof will work for any scalar field. Before we can state the theorem we should recall the definition of a polar set.

Definition 5.1 *(Polar Set)*
Let X be a vector space, X' it's dual space. For any set $U \subset X$ the polar set in X' is the set

$$U^\circ := \left\{ x' \in X' \,:\, \sup_{x \in U} |\langle x, x' \rangle| \leq 1 \right\}$$

Having this definition we can state the Banach Alaoglu Bourbaki theorem. The proof given is a fairly standard proof which works like the proof for the Banach Alaoglu theorem to a large extend. This proof is most commonly to be found in literature (e.g. in [1]). The necessary adaptations of the proof of Bourbaki's generalisation given here can be found in this or similar form in Rudin's book [10].

Theorem 5.2 *(Banach Alaoglu Bourbaki Theorem)*
Let X be a vector space, X' it's dual space. Let U be a neighbourhood of 0 in X. Then $U^\circ \subset X'$ is w^*-compact.

Proof:
We will homeomorphically embed U° into a compact subset of $\mathbb{R}^X$ with the topology of pointwise convergence. This space is the space of X-indexed real sequences, which can also be regarded as the space of all $\mathbb{R}$-valued functions on X as they are canonically homeomorphic.
We will then show that the image of U° in $\mathbb{R}^X$ is closed. It is therefore also compact

as a closed subset of a compact set and, as the embedding was homeomorphic, we thus have that U° is compact.

Part 1: *(Embed U° homeomorphically into $\mathbb{R}^X$)*
Since U is a 0 neighbourhood we can find a $V \subset U$ such that V is absolutely convex. As every functional which is bounded by 1 on U is clearly also bounded by 1 on V we have that $U^\circ \subset V^\circ$. Hence compactness of V° implies compactness of U°. We can therefore assume U to be absolutely convex and absorbing.
Hence we can find a constant λ_x for every $x \in X$ such that $x \in \lambda_x \cdot U$. Since every x' in U° is bounded by 1 on U this tells us that it is bounded by λ_x on the whole of X.
Now define sets

$$F_x := \{\lambda \in \mathbb{R} \,:\, |\lambda| \leq \lambda_x\}$$

and form the product of these sets

$$F = \prod_{x \in X} F_x$$

Every F_x is a compact subset of $\mathbb{R}$ hence, by Tychonov's theorem, the product F is compact in the product topology.
Now U° embeds into F by the mapping

$$\theta : U^\circ \to F$$
$$\theta(x') = \big(x'(x)\big)_{x \in X}$$

It is obvious that θ is injective. The continuity of θ and θ^{-1} are immediate consequences of the universality property of weak topologies. Recall that for a space Y and a family of functions on Y, $\mathcal{E}$ say, the weak topology on Y generated by $\mathcal{E}$ is the smallest topology making every function in $\mathcal{E}$ continuous. The universality property then says that for any topological space Z a map $g : Z \to Y$ is continuous if and only if $f \circ g$ is continuous for every $f \in \mathcal{E}$
The product topology on F is generated by the coordinate projections

$$\pi_x : F \to F_x$$
$$\pi_x\left((\mu_x)_{x \in X}\right) = \mu_x$$

For continuity of θ it is thus enough to show that $\pi_x \circ \theta$ is continuous for every π_x. But by the definitions we have

$$\begin{aligned} \pi_x \circ \theta(x') &= \pi_x\big(\big(x'(x)\big)_{x \in X}\big) \\ &= x'(x) \\ &= \hat{x}|_{U^\circ} \end{aligned}$$

where $\hat{x}(x') = x'(x)$, the canonical embedding of x into X'', the bidual of X. This is clearly w^*-continuous.
On the other hand, the topology on X' and thus on U° is generated by $\{\hat{x} \in X'' \,:\, x \in X\}$. Therefore for continuity of θ^{-1} we need to check that $\hat{x} \circ \theta^{-1}$ is continuous for every $\hat{x}$. But using the definitions in this case we find

$$\begin{aligned} \hat{x} \circ \theta^{-1}\big(\big(x'(x)\big)_{x \in X}\big) &= \hat{x}(x') \\ &= x'(x) \\ &= \pi_x\big(\big(x'(x)\big)_{x \in X}\big) \end{aligned}$$

Thus θ^{-1} is continuous on the image of θ in F. This means that U° is homeomorphic to $\mathrm{Im}\theta \subset F$, which is compact in $\mathbb{R}^X$.

Part 2: *(The image of U° is closed in $\mathbb{R}^X$)*
As mentioned before, the space of real valued functions on X is homeomorphic to $\mathbb{R}^X$ by the canonical map

$$\Theta : (f : X \to \mathbb{R}) \to \mathbb{R}^X$$
$$(x \mapsto \nu_x) \mapsto (\nu_x)_{x \in X}$$

The map θ from part 1 is obviously the restriction of Θ to U°. So to show that the image of U° is closed, i.e. $\overline{\theta\left(U^\circ\right)} = \theta\left(U^\circ\right)$ we need to show that

$$\Theta^{-1}(\overline{\theta\left(U^\circ\right)}) = \Theta^{-1}(\theta\left(U^\circ\right)) = U^\circ$$

It is clear that $U^\circ \subset \Theta^{-1}(\overline{\theta\left(U^\circ\right)})$. Thus we fix $f_0 \in \Theta^{-1}(\overline{\theta\left(U^\circ\right)})$ and show that f_0 is in U°, i.e. linear and bounded by 1 on U.

We are considering $\mathbb{R}^X$ and hence the space of real valued functions on X with the topology of pointwise convergence so basic open neighbourhoods of f_0 are sets

$$V_{y,\varepsilon} = \{f : X \to \mathbb{R} \,:\, |f(y) - f_0(y)| < \varepsilon\}$$

for $y \in X$ and $\varepsilon > 0$.

- *Linearity of f_0:*
  For fixed $x, y \in X$ and $s, t \in \mathbb{R}$ the set

  $$V := V_{x,\varepsilon} \cap V_{y,\varepsilon} \cap V_{s\cdot x + t\cdot y,\varepsilon}$$

  is an open neighbourhood of f_0. As Θ^{-1} is continuous there exists $f \in \Theta^{-1}(\theta\left(U^\circ\right)) = U^\circ$ in V. So we have $f \in U_0$ such that $|f - f_0| < \varepsilon$ which we can use to estimate f_0 as follows

  $$\begin{aligned}
  &|f_0(s\cdot x + t\cdot y) - s\cdot f_0(x) - t\cdot f_0(y)| \\
  &= |f_0(s\cdot x + t\cdot y) - f(s\cdot x + t\cdot y) + s\cdot f(x) + t\cdot f(y) - s\cdot f_0(x) - t\cdot f_0(y)| \\
  &= |(f_0 - f)(s\cdot x + t\cdot y) + s\cdot (f - f_0)(x) + t\cdot (f - f_0)(y)| \\
  &< \varepsilon + |s|\cdot\varepsilon + |t|\cdot\varepsilon
  \end{aligned}$$

  But ε was arbitrary, so we find that f_0 is indeed linear.

- *f_0 bounded by 1 on U:*
  Fix an arbitrary $x \in U$ and consider the open neighbourhood $V_{x,\varepsilon}$. Then there exists $f \in \Theta^{-1}(\theta\left(U^\circ\right)) = U^\circ$ in $V_{x,\varepsilon}$, i.e. $|f(x) - f(x_0)| < \varepsilon$. But f is in U° so bounded by 1 on U thus in particular at x. As ε is arbitrary we must also have that f_0 is bounded by 1 at x. Since x was arbitrary f_0 is bounded by 1 on U and thus in U°.

We can conclude that $\Theta^{-1}(\overline{\theta\,(U^\circ)}) = U^\circ$. Hence $\theta\,(U^\circ)$ is closed and thus compact and so is U°.

❑

Remark 5.3
Note that

(a) $\mathscr{D}(\Omega)$ and $\mathscr{S}(\mathbb{R}^n)$ are locally convex spaces. Therefore theorem 5.2 applies to both, hence we get a large amount of compact sets in $\mathscr{D}'(\Omega)$ and $\mathscr{S}'(\mathbb{R}^n)$.

(b) For metric spaces the Banach Alaoglu Bourbaki theorem is equivalent to the Banach Alaoglu theorem. $\mathscr{S}(\mathbb{R}^n)$ is metrisable, thus the Banach Alaoglu theorem would suffice in this case.

This result tells us that we are provided with many compact sets in the spaces of distributions and tempered distributions, which is very desirable for applications. However it is often convenient to characterise compact sets by properties of sequences. This is a very common approach in metric spaces, as the different notions of compactness are all equivalent if the topology is metrisable. But for non-metrisable topologies the compactness definitions are not in general equivalent. Therefore we shall now examine how sequences in the space of distributions behave with respect to compact sets.

5.2 Sequential Compactness

We will start with an example that illustrates the necessity of having a closer look at the behaviour of the sequences in the compact sets we obtain from the Banach Alaoglu Bourbaki theorem. We consider the space of bounded sequences and its dual equipped with the weak* topology. In this case we can easily find a w^*-compact set which is not w^*-sequentially compact. The given example could be regarded as the standard example for a weak* compact but not weak* sequentially compact set, one can find it in e.g. [5].

Example 5.4

Let

$$l_\infty = \{x := (x_n) \text{ scalar sequence } : (x_n) \text{ bounded}\}$$

the space of bounded sequences equipped with the supremum norm

$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$$

and l'_∞ its dual space equipped with the weak* topology. As l_∞ is a normed space by Banach Alaoglu we know that the unit ball in l'_∞ is w^*-compact.
To show that the unit ball is not sequentially compact we need to find a sequence u_n in $B_{l'_\infty}$ without a converging subsequence. Recall that

$$u_n \overset{w^*}{\to} u \text{ in } l'_\infty \text{ if } \langle x, u_n\rangle \to \langle x, u\rangle \quad \forall\, x \in l_\infty$$

Now let

$$(u_n) \subset l'_\infty \quad \text{s.t.} \quad \langle x, u_n\rangle = x_n \quad \forall\, x \in l_\infty$$

Then

$$\|u_n\| = \sup_{x\in B_{l_\infty}} |\langle x, u_n\rangle| \overset{\text{def. } \|\cdot\|_\infty}{\leq} \sup_{x\in B_{l_\infty}} \|x\|_\infty \overset{x\in B_{l_\infty}}{\leq} 1$$

So actually $(u_n) \subset B_{l'_\infty}$ as desired. To show that (u_n) does not have a converging subsequence assume the contrary, that there exists a converging subsequence (u_{n_i}) say. To reach a contradiction it suffices to find one $x \in l_\infty$ such that $\lim_{i\to\infty} \langle x, u_{n_i}\rangle$ does not exist. Let

$$x_k := \begin{cases} (-1)^i & \text{if } k = n_i \\ 0 & \text{otherwise} \end{cases}$$

Then obviously $(x_k) \in l_\infty$ and

$$\lim_{i\to\infty} \langle (x_k), u_{n_i}\rangle = \lim_{i\to\infty} (-1)^i$$

which clearly doesn't exist. Hence there cannot exist a convergent subsequence of $(u_n) \subset B_{l'_\infty}$.
Therefore $B_{l'_\infty}$ is compact but not sequentially compact.

❑

This means that in general the result of Banach Alaoglu Bourbaki does not give us much control over sequences. However, in the space of distributions the situation turns out to be better than it might have been expected.

5.2.1 Sequential compactness of weak* bounded sets

Theorem 5.5

Let $A \subset \mathscr{D}'(\Omega)$ a set of distributions that is pointwise bounded, i.e. on any fixed test function $\varphi \in \mathscr{D}(\Omega)$. Then A is sequentially compact.

Proof:
We need to find a converging subsequence for any sequence of distributions in A. Therefore we can, without loss of generality, assume that A is already a countable sequence of distributions, if not we start by picking an arbitrary sequence inside A. Thus we have

$$A := \{T_n \,:\, n \in \mathbb{N}\} \qquad \sup_{n\in\mathbb{N}} |\langle \varphi, T_n\rangle| < \infty \;\forall\; \varphi \in \mathscr{D}(\Omega)$$

We have seen before that we could characterise continuity of linear maps on the space of test functions by restricting to the $\mathscr{D}_K$ spaces, because these spaces are Fréchet spaces, which gives us a lot more control. The fact that we are considering properties of sequences again suggests to do the same approach here. We will then use an inductive argument based on the constructions in lemma 2.3 to extend the result to the whole of $\mathscr{D}(\Omega)$.

Part 1: *(Establish the existence of a convergent subsequence for fixed $\mathscr{D}_K$)*
Fix $K_1 \subset \Omega$ compact. A is of course in particular pointwise bounded on every $\varphi \in \mathscr{D}_{K_1}$ and $\mathscr{D}_{K_1}$ is a Fréchet space by lemma 2.2. Therefore we can apply a generalised version of the Banach-Steinhaus theorem, which states that a set of continuous linear mappings from a Fréchet space into a normed space which is pointwise bounded is also equicontinuous.
Thus we know that A is pointwise bounded and equicontinuous as a set of continuous linear maps on $\mathscr{D}_{K_1}$. So if we now regard A as a subset of $C(\mathscr{D}_{K_1})$, the space of continuous functionals on $\mathscr{D}_{K_1}$ with the supremum norm, a general version of the Arzelà-Ascoli theorem applies and thus A is totally bounded with respect to the supremum norm. Therefore we know that there exists a subsequence (T_{n_i}) of A that converges in the supremum-norm, i.e. it converges uniformly, to some $f \in C(\mathscr{D}_{K_1})$. But this convergence is stronger then the weak* convergence, which is pointwise convergence on $\mathscr{D}_{K_j}$, whence it also converges w^*ly to f, i.e

$$T_{n_i} \xrightarrow[i\to\infty]{w^*} f$$

Part 2: *(Iteratively construct a nested sequence of convergent sequences)*
Now pick $K_2 \subset \Omega$ compact such that $K_1 \subset K_2^\circ$. By the argument in part 1 we can find a subsequence (T_{n_k}) of A that converges in the weak* topology. But by the choice of K_2 we have that $\mathscr{D}_{K_1} \subset \mathscr{D}_{K_2}$ and (T_{n_k}) converges pointwise on $\mathscr{D}_{K_2}$. Therefore it must in particular converge pointwise on $\mathscr{D}_{K_1}$. Thus we can without loss of generality assume, that (T_{n_k}) is a subsequence of the sequence (T_{n_i}) we found in the first step.
Iterating this argument we obtain a sequence of nested sequences in A which converge pointwise on some $\mathscr{D}_{K_j}$, i.e. in the respective weak* topology. Passing to the limit $K_j \nearrow K$ we obtain a sequence in A that converges pointwise, i.e. w^*ly on $\mathscr{D}(\Omega)$.

❑

Note that pointwise boundedness is exactly weak* boundedness as can be seen in proposition 2.30. As every weak* compact set is weak* bounded, theorem 5.5 tells us that every compact set we get from Banach Alaoglu Bourbaki is actually also sequentially compact. However the converse is unfortunately far from true, the space of distributions is not even sequential, as we mentioned before. Thus a sequentially compact, so in particular sequentially closed set might not even be weak* closed thus in particular not weak* compact.

5.2.2 The space of distributions is not sequential

To see that the space of distributions is indeed not sequential consider the following example. We will first state the example as a proposition in a general case which is also applicable to other non-sequential spaces and then present a concrete example for the space of distributions. To simplify things we will consider $\mathscr{D}'(\mathbb{R})$. This goes in line with the way Smolyanov stated the example in [14]. In this paper he states the example in its general case and mentions thereafter that the concrete example for the space of distributions fulfils the assumptions. Smolyanov doesn't give proves, the fact that the given example fulfils the assumptions and that it is in fact sequentially closed but not closed can be easily checked, as the conditions become fairly obvious once they are spelled out, which we will do as part of the example.

Proposition 5.6

Let E and F be locally convex spaces and

$$A := \{a_{k,n} \,:\, k, n \in \mathbb{N}\} \subset E$$

$$B := \{b_{n,k} \,:\, k, n \in \mathbb{N}\} \subset F$$

and let

$$C := \{c_{k,n} = a_{k,n} + b_{n,k} \,:\, k, n \in \mathbb{N}\} \subset E \oplus F$$

If

(i) $0 \notin A \cap B$

(ii) $a_{k,n} \underset{n\to\infty}{\longrightarrow} 0$ uniformly with respect to k

(iii) For fixed $n \in \mathbb{N}$ there exists no sequence k_i such that $a_{k_i,n}$ converges

(iv) For fixed $n \in \mathbb{N}$ $b_{n,k} \underset{k\to\infty}{\longrightarrow} 0$

(v) If b_{n_i,k_i} converges then $\sup_{i\in\mathbb{N}} n_i < \infty$

Then C is sequentially closed but not closed.

In the concrete example we will see that the crack of this example is that 0 is not in the sequential closure of the set as the set is sequentially closed and does not contain 0, but that 0 is contained in the closure. In fact condition (i) makes sure that $0 \notin C$ while conditions (ii) and (iv) make sure that 0 is "close" to C, i.e. contained in the closure. Conditions (iii) and (v) make sure that the set C is sequentially closed.
Let us now proceed to stating the example in $\mathscr{D}'(\mathbb{R})$ where these conditions can easily be checked.

Example 5.7

Let

$$A := \left\{ a_{k,n} = \delta_n^{(k)} \ : \ k, n \in \mathbb{N} \right\} \subset \mathscr{D}'(\mathbb{R})$$

$$B := \left\{ b_{n,k} = \frac{1}{k}\delta_0^{(n)} \ : \ k, n \in \mathbb{N} \right\} \subset \mathscr{D}'(\mathbb{R})$$

where $\delta^{(k)}$ and $\delta^{(n)}$ denote the k-th and n-th derivative of δ respectively. Now let

$$C := \{c_{k,n} = a_{k,n} + b_{n,k} \ : \ k, n \in \mathbb{N}\} \subset \mathscr{D}'(\mathbb{R})$$

Then C is weak* sequentially closed but not weak* closed. Indeed we have all properties from proposition 5.6:

(i) Clearly $a_{k,n} \neq 0$ and $b_{n,k} \neq 0$ for all $n, k \in \mathbb{N}$.

(ii)

$$a_{k,n} \xrightarrow[n\to\infty]{w^*} 0 \Leftrightarrow \langle \varphi, \delta_n^{(k)} \rangle \xrightarrow[n\to\infty]{} 0 \quad \forall\, \varphi \in \mathscr{D}(\mathbb{R})$$

This is indeed the case. If $\text{supp}(\varphi) \subset K$ then also $\text{supp}(D^k\varphi) \subset K$ for all k. Furthermore $\text{supp}(\delta_n^{(k)}) = \{n\}$ for all k, so there exists an $N \in \mathbb{N}$ such that $\text{supp}(\delta_n^{(k)}) \cap \text{supp}(D^k\varphi) = \emptyset$ for all $n \geq N$ and all $k \in \mathbb{N}$.
Thus $\langle \varphi, \delta_n^{(k)} \rangle = D^k\varphi(n) = 0$ for all $n > N$.

(iii)

$$a_{k,n} \xrightarrow[i\to\infty]{w^*} T \Leftrightarrow \langle \varphi, \delta_n^{(k_i)} \rangle \to \langle \varphi, T \rangle \quad \forall\, \varphi \in \mathscr{D}(\mathbb{R})$$

This clearly cannot happen for any unbounded sequence (k_i) as $\langle \varphi, \delta_n^{(k_i)} \rangle = (D^{k_i}\varphi)(n)$ and for fixed n we can clearly construct a test function such that the sequence of it's k_i-th derivatives evaluated at n does not converge.

(iv)

$$b_{n,k} \xrightarrow[k\to\infty]{w^*} 0 \Leftrightarrow \langle \varphi, \frac{1}{k}\delta_0^{(n)} \rangle \to 0 \quad \forall\, \varphi \in \mathscr{D}(\mathbb{R})$$

But this is true as $\langle \varphi, \frac{1}{k}\delta_0^{(n)} \rangle = \frac{1}{k}(D^n\varphi)(0)$ which clearly converges to 0 for any fixed n as k goes to infinity.

(v)

$$b_{n,k} \xrightarrow[i\to\infty]{w^*} T \Leftrightarrow \langle \varphi, \frac{1}{k_i}\delta_0^{(n_i)} \rangle \to \langle \varphi, T \rangle \quad \forall\, \varphi \in \mathscr{D}(\mathbb{R})$$

But $\langle \varphi, \frac{1}{k_i}\delta_0^{(n_i)} \rangle = \frac{1}{k_i}(D^{n_i}\varphi)(0)$, so if we allowed for the n_i to be unbounded we could construct a φ such that $(D^{n_i}\varphi)(0) > k_i^2 \cdot i$ for every i. Then clearly $\frac{1}{k_i}(D^{n_i}\varphi)(0)$ would not converge as i goes to infinity.

Putting these properties together we get the claim.
Assume that

$$(c_i) \subset C \qquad c_i = c_{k_i,n_i} = a_{k_i,n_i} + b_{n_i,k_i}$$

such that $c_i \to T$ for some $T \in \mathscr{D}'(\mathbb{R})$. This can clearly only happen if both summands converge. We know from (v) that if b_{n_i,k_i} converges, we must have $\sup_{i\in\mathbb{N}} n_i < \infty$. Thus, by passing on to a subsequence if necessary, we can assume that

the n_i are constant. But for fixed n_i (iii) tells us that there cannot be an increasing sequence k_i such that a_{k_i,n_i} converges. Therefore we can also deduce that the k_i must be constant. Thus c_i is constant, so trivially converging to $c_i \in C$. This shows that C is sequentially closed.
A basic open neighbourhood of $0 \in \mathscr{D}'(\mathbb{R})$ in the weak* topology is given by

$$V_{\varphi,\varepsilon} = \left\{T \in \mathscr{D}'(\mathbb{R}) \,:\, |\langle \varphi, T\rangle| < \varepsilon\right\}$$

By (ii) there exists $N \in \mathbb{N}$ big enough such that $N \notin \operatorname{supp}(\varphi)$ and then $\langle \varphi, a_{k,n}\rangle = 0$ for all k. But from (iv) we know that for fixed N the $b_{N,k}$ converge to 0. So there exists a $K \in \mathbb{N}$ such that $\langle \varphi, b_{N,k}\rangle < \varepsilon$ which means that

$$\langle \varphi, c_{K,N}\rangle = \langle \varphi, a_{K,N}\rangle + \langle \varphi, b_{N,K}\rangle < \varepsilon$$

and thus $c_{K,N} \in V_{\varphi,\varepsilon}$. So every neighbourhood of 0 contains an element of C, but from (i) we know that $0 \notin C$, so C is not weak* closed.

❏

A Notation

Topologies introduced within the essay:

- $\overline{\tau}$ - non-complete, metrisable topology on $\mathscr{D}(\Omega)$ - Notation 2.1
- τ_Ω - complete, metrisable topology on $C^\infty(\Omega)$ - Notation 2.2
- τ_K - complete, metrisable topology on $\mathscr{D}_K$ - Notation 2.2
- τ - complete, non-metrisable topology on $\mathscr{D}(\Omega)$ - Notation 2.3
- $\widetilde{\tau}$ - topology on $\mathscr{D}(\Omega)$, strictly finer than τ - Definition 2.15
- σ - complete, metrisable topology on $\mathscr{S}(\mathbb{R}^n)$ - Definition 3.2

Notation used potentially without explanation

- $\text{supp}(f) = \overline{\{f \neq 0\}}$ the support of a function
- $\|x\| = \|x\|_2 = \sum\limits_{i=1}^{n} |x_i|^2 \quad x \in \mathbb{R}^n$
- $L_{1,loc}(\Omega)$ the set of locally integrable functions on Ω
- $\mathscr{L}(A, B)$
  The space of bounded linear operators from A to B,where A, B function spaces
- $C_0(\mathbb{R}^n) = \left\{ f \in C(\mathbb{R}^n) \; : \; \lim\limits_{|x| \to \infty} f(x) = 0 \right\}$
  The space of continuous functions vanishing at infinity
- A° the interior of a set A
- $x \cdot y$ the scalar product of $x, y \in \mathbb{R}^n$
- L_0 the space of measurable functions
- $\mathcal{B}_n := B(0, n) := \{x \in \mathbb{R}^n \; : \; \|x\| \leq n\}$
  The closed ball in $\mathbb{R}^n$ centred at 0 with radius n

B References

[1] Graham R. Allan - Introduction to Banach Spaces and Algebras - Oxford University Press, 2011

[2] Philippe Blanchard, Erwin Brüning - Mathematical Methods in Physics, Distributions, Hilbert Space Operators, and Variational Methods - Birkhäuser, 2003

[3] Jean François Colombeau - Multiplication of Distributions, A tool in mathematics, numerical engineering and theoretical physics - Springer-Verlag, 1992

[4] John B. Conway - A Course in Functional Analysis - Springer-Verlag, 1990

[5] Carl L. DeVito - Functional Analysis - Academic Press, New York, 1978

[6] R. M. Dudley - On sequential convergence - Transactions of the American Mathematical Society Vol 112, 1964, pp. 483-507

[7] R. M. Dudley - Convergence of sequences of distributions - Proceedings of the American Mathematical Society Vol. 27, 1971, pp. 531-534

[8] Klaus Floret - Some aspects of the theory of locally convex inductive limits - North Holland Mathematical Studies Vol. 38, 1980, pp. 205-237

[9] Lars Hörmander - The Analysis of Linear Partial Differential Operators I, Second Edition - Springer-Verlag, 1990

[10] Walter Rudin - Functional Analysis, Second Edition - Mc Graw-Hill, 1991

[11] Laurent Schwartz - Sur l'impossibilité de la multiplications des distributions - Comptes Rendus de l'Académie de sciences Vol. 239, Paris, 1954, pp. 847-848

[12] Laurent Schwartz - Théorie Des Distributions - Hermann & C^{ie}, 1950

[13] Tameharu Shirai - Sur les Topologies des Espaces de L. Schwartz - Proceedings of the Japan Academy Vol. 35, 1959, pp. 31-36

[14] O. G. Smolyanov - Nonclosed sequentially closed subsets of locally convex spaces and applications - Note di Matematica Vol. XII, 1992, pp. 237-244

[15] François Trèves - Topological Vector Spaces, Distributions and Kernels - Academic Press, New York, 1967

[16] J. H. Webb - Sequential convergence in locally convex spaces - Mathematical Proceedings of the Cambridge Philosophical Society Vol. 64 Issue 02, 1968, pp. 341-364

[17] Albert Wilanski - Modern Methods in Topological Vector Spaces - Mc Graw-Hill, 1978