The course will focus on various aspects of linear parabolic and hyperbolic equations, with a particular emphasis on the construction and properties of parametrices and solution operators to such equations. Towards the end of the quarter, I’ll plan on giving an introduction to microlocal analysis, with a particular focus on applications to parabolic and hyperbolic problems.

Schedule

(may be subject to change depending on student interest)

- Weeks 1,2: Review of distribution theory and Fourier transform

Date: April 7, 2022.
• Weeks 3-5: Study of parabolic equations, with particular focus on parabolic regularity and the structure of the heat kernel
• Weeks 6,7: Study of hyperbolic equations
• Week 8: Introduction to microlocal analysis and parametrices for differential operators
• Week 9: Construction of parametrices for parabolic operators
• Week 10: Construction of parametrices for hyperbolic operators
1. Lecture 1 (03/29): Distribution Theory: Preliminaries

Distributions are generalizations of functions that work particularly well with differentiation, convolution, Fourier transforms, etc. We review the theory of distributions in this lecture.

The reference for this section is Hörmander’s *The Analysis of Partial Differential Operators I* [Hör90].

1.0. Conventions. We consider functions (either $\mathbb{R}$-valued or $\mathbb{C}$-valued; usually the difference is not significant) defined on an open subset $U$ of $\mathbb{R}^n$. A multi-index is an $n$-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers, with the corresponding differential operator $\partial^\alpha := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$ on $U$. The sum $|\alpha| = \alpha_1 + \cdots + \alpha_n$ is the order of the multi-index $\alpha$ / the differential operator $\partial^\alpha$. We denote $C^k(U)$ the space of functions $f$ on $U$ which has partial derivatives of order up to $k$, with $\partial^\alpha f$ continuous on $U$ for all $\alpha$ with $|\alpha| \leq k$. We denote $C^\infty_c(U) = \cap_k C^k(U)$. When needed, we’ll define $C^k_c(U)$ and $C^\infty_c(U)$ analogously, with the additional requirement that the partial derivatives be continuous up to the boundary.

For $U \subset \mathbb{R}^n$ open, we denote $C^\infty_c(U) := \{u \in C^\infty(\mathbb{R}^n) : \text{supp } u \text{ is a compact subset of } U\}$.

1.1. Definitions and properties.

**Definition 1.1.** Let $U \subset \mathbb{R}^n$ be open. A distribution on $U$ is a linear functional $u : C^\infty_c(U) \to \mathbb{C}$ which is continuous with respect to the topology on $C^\infty_c(U)$ (defined below). The space of all distributions on $U$ is denoted by $\mathcal{D}'(U)$.

For $u \in \mathcal{D}'(U)$ and $\phi \in C^\infty_c(U)$, we’ll write the application of $u$ against $\phi$ as $u(\phi)$ or $(u, \phi)_{U}$ (when emphasizing the domains on which the distributions live).

The topology on $C^\infty_c(U)$ is a bit involved to describe but the practical interpretation of the topology is as follows: a sequence $\{\phi_n\}$ in $C^\infty_c(U)$ converges to $\phi \in C^\infty_c(U)$ if:

- There exists a compact set $K \subset U$ such that $\text{supp } \phi_n \subset K$ for all $n$, and
- All derivatives of $\phi_n$ converge uniformly to the corresponding derivative of $\phi$ on $U$, i.e. for all multi-indices $\alpha$ we have

$$\sup_{x \in U} |\partial^\alpha \phi_n - \partial^\alpha \phi| \xrightarrow{n \to \infty} 0.$$
Thus, we say that a linear functional \( u : C^\infty_c(U) \to \mathbb{C} \) is\footnote{Strictly speaking, this should be called \textit{sequentially continuous}, as the topology on \( C^\infty_c(U) \) is not metrizable.} \textit{continuous} if, whenever \( \phi_n \to \phi \) with respect to the convergence defined above in \( C^\infty_c(U) \), we also have \( u(\phi_n) \to u(\phi) \) in \( \mathbb{C} \) as well.

**Lemma 1.2.** A linear functional \( u : C^\infty_c(U) \to \mathbb{C} \) is continuous if and only if, for every compact subset \( K \subset U \), there exists \( k \in \mathbb{N}_{\geq 0} \) and \( C > 0 \) such that

\[
|u(\phi)| \leq C \sum_{|\alpha| \leq k} \sup_K |\partial^\alpha \phi| \text{ for all } \phi \in C^\infty_c(U) \text{ with supp } \phi \subset K.
\]

**Proof.** The “if” part is straightforward to verify. For the “only if” part, suppose for some compact \( K \subset U \) that no such \( C \) and \( k \) were to exist to satisfy the above inequality. Then, for every \( k \in \mathbb{N}_{\geq 0} \), we could find \( \phi \in C^\infty_c(U) \) supported in \( K \) such that \( u(\phi) \) is arbitrarily large compared to \( \sum_{|\alpha| \leq k} \sup_K |\partial^\alpha \phi| \). In particular, for each \( k \) we can find \( \phi_k \) supported in \( K \) such that \( u(\phi_k) = 1 \), but \( \sum_{|\alpha| \leq k} \sup_K |\partial^\alpha \phi_k| \leq 1/k \). Note that this implies that \( \sup_K |\partial^\alpha \phi_k| \leq 1/k \) whenever \( k \geq |\alpha| \). In particular, we see that for each fixed \( \alpha \) we would have \( \sup_U |\partial^\alpha \phi_k| \to 0 \) as \( k \to \infty \), and hence \( \phi_k \to 0 \) in the topology of \( C^\infty_c(U) \). But then we should have \( u(\phi_k) \to u(0) = 0 \) as well due to the continuity of \( u \), which contradicts the assumption that \( u(\phi_k) = 1 \) for each \( k \). \( \square \)

Thus, we could have alternatively defined a distribution \( u \) as a linear functional satisfying

\[
|u(\phi)| \leq C \sum_{|\alpha| \leq k} \sup_K |\partial^\alpha \phi| \text{ for all } \phi \in C^\infty_c(U) \text{ with supp } \phi \subset K
\]

for some constants \( C \) and \( k \) for each compact \( K \subset U \) (note that \( C \) and \( k \) in general depend on \( K \)).

**Example 1.3.** The following are distributions (the continuity part is left as an exercise; the important aspect is viewing these as linear functionals):

- Any \( f \in L^1_{loc}(U) \) can be identified with a distribution
  \[
  T_f(\phi) := \int_U f(x)\phi(x) \, dx.
  \]
  (Note that the RHS makes sense for any \( \phi \in C^\infty_c(U) \), since \( f \in L^1(K) \) for any compact subset \( K \subset U \); in particular this is the case for \( K = \text{supp } \phi \).) In such cases, we’ll refer to the distribution as \( f \) as well, and we’ll say that a distribution \( u \) is in \( L^1_{loc}(U) \) (or \( L^p \), continuous, \( C^\infty \), etc.) if it can be identified with a function in \( L^1_{loc}(U) \) via the above identification.

**Remark 1.** This is slightly different than the complex inner product \( \langle f, \phi \rangle = \int_U f(x)\overline{\phi}(x) \, dx \)–note that \( \langle \cdot, \cdot \rangle \) is \( \mathbb{C} \)-anti-linear in the second variable, i.e. \( \langle f, \alpha \phi \rangle = \overline{\alpha}\langle f, \phi \rangle \) for \( \alpha \in \mathbb{C} \), whereas distributions are \( \mathbb{C} \)-linear.
• For any \( x_0 \in U \) and multi-index \( \alpha \), the functional
  \[ \phi \mapsto \partial^\alpha \phi(x_0) \]
  is a distribution. The case \( \alpha = 0 \) is called the Dirac delta at \( x_0 \), denoted \( \delta_{x_0} \)
  (the Dirac delta at the origin 0 is often just denoted \( \delta \)).

• In \( \mathbb{R} \), the distributions \( (x \pm i0)^{-1} \) are defined by
  \[ \phi \mapsto \lim_{\epsilon \to 0^+} \int_{\mathbb{R}} \frac{\phi(x)}{x \pm i\epsilon} \, dx. \]
  The limit on the RHS does indeed exist for any \( \phi \in C^\infty_c(\mathbb{R}) \), and the limits
  may be different depending on the sign \( \pm \). In fact, we have
  \[ (x - i0)^{-1} - (x + i0)^{-1} = 2\pi\delta_0. \]

The space of distributions \( \mathcal{D}'(U) \) also has a topology, given by the weak-* topology
viewing it as the dual space to \( C^\infty_c(U) \).

**Definition 1.4.** We say that a sequence \( u_n \) in \( \mathcal{D}'(U) \) converges to \( u \in \mathcal{D}'(U) \) as
distributions if, for every \( \phi \in C^\infty_c(U) \), we have
  \[ u_n(\phi) \to u(\phi) \quad (\text{in } \mathbb{C}). \]

In practice, convergence in \( \mathcal{D}'(U) \) is weaker than most kinds of convergence that can be considered. For example, if \( u_n \) has more structure, e.g. belongs to \( C^\infty_c(U) \), \( C^\infty(U) \), even \( L^1_{\text{loc}}(U) \), and it converges, e.g. uniformly or even in \( L^1 \) (or in \( L^1 \) on every compact set), then it also converges as distributions.

If \( V \subset U \) is open, then there is a continuous inclusion \( \iota : C^\infty_c(V) \hookrightarrow C^\infty_c(U) \).

**Definition 1.5.** If \( V \subset U \) is open, and \( u \in \mathcal{D}'(U) \), the restriction of \( u \) to \( V \) is the
distribution on \( V \) defined by
  \[ u|_V(\phi) := u(\iota \phi). \]

**Definition 1.6.** The support of a distribution \( u \in \mathcal{D}'(U) \) is the set
  \[ \text{supp } u := \{ x \in U : u|_V \text{ is not identically zero for any neighborhood } V \ni x \}. \]

**Example 1.7.** If \( f \) is a continuous function on \( U \), then the support of \( f \), viewing \( f \)
as a distribution, is the same as the support of \( f \), viewed as a function. That is,
  \[ \text{supp } f = \{ x \in U : f(x) \neq 0 \} \]
(here the closure is taken with respect to \( U \)).

---

5The only kind of convergence I can think of that is not stronger than convergence in \( \mathcal{D}'(U) \) is pointwise convergence. But even then, even a very weak bound on the sequence \( u_n \), combined with pointwise convergence of \( u_n \), is usually enough to give convergence in \( \mathcal{D}'(U) \), due to the extremely nice properties of \( C^\infty_c(U) \).
1.2. Operations on distributions: differentiation, multiplication. One essential operation that can be applied to distributions is differentiation. The idea is motivated by the integration by parts identity
\[ \int_U \partial^\alpha f(x) \phi(x) \, dx = (-1)^{|\alpha|} \int_U f(x) \partial^\alpha \phi(x) \, dx \]
which holds for all \( f \in C^\infty(U) \) and \( \phi \in C^\infty_c(U) \) (the compact support of \( \phi \) guarantees the lack of “boundary terms” that would normally arise from integration by parts). We thus define differentiation via this feature:

**Definition 1.8.** For \( 1 \leq j \leq n \) and \( u \in D'(U) \), the partial derivative \( \partial_j u \) is the distribution on \( U \) defined by
\[ \partial_j u(\phi) := -u(\partial_j \phi) \text{ for all } \phi \in C^\infty_c(U). \]

For a multi-index \( \alpha \), the derivative \( \partial^\alpha u \) is defined by iterating partial derivatives, in the same manner as iterating partial derivatives for smooth functions.

Note that \( \phi \in C^\infty_c(U) \Rightarrow \partial_j \phi \in C^\infty_c(U) \), so the right-hand side in the above definition does indeed make sense.

**Example 1.9.** For \( a > -1 \), define
\[ \chi^a_+(x) = \begin{cases} x^a & x > 0 \\ 0 & x \leq 0 \end{cases}. \]

Then \( \chi^a_+ \in L^1_{loc}(\mathbb{R}) \), and \( \frac{d}{dx} \chi^a_+ = \chi^{a-1}_+ \) for \( a > 0 \). Moreover, \( \chi^0_+(x) \) is the so-called “Heaviside function” \( H(x) \), and \( \frac{d}{dx} \chi^0_+ = \delta_0 \).

Similarly, if \( \rho \in C^\infty(U) \), then we have the identity
\[ \int_U (\rho(x)u(x)) \phi(x) \, dx = \int_U u(x)(\rho(x)\phi(x)) \, dx \]
for any function \( u \) just by rearranging terms; moreover note that \( \rho \phi \in C^\infty_c(U) \) if \( \phi \in C^\infty_c(U) \). Thus, we can define:

**Definition 1.10.** For \( \rho \in C^\infty(U) \) and \( u \in D'(U) \), the product \( \rho u \) is the distribution on \( U \) defined by
\[ (\rho u)(\phi) := u(\rho \phi). \]

**Remark 2.** Differentiation and multiplication by \( \rho \in C^\infty(U) \), as defined above, are in fact *continuous* linear operators \( D'(U) \to D'(U) \), such that the restriction of these operators to \( C^\infty_c(U) \subset D'(U) \) give the same results as the usual differentiation and multiplication on \( C^\infty_c(U) \). The reason why we insist on only multiplying by functions in \( C^\infty(U) \) is that it turns out there does not exist a *continuous* operator \( D'(\mathbb{R}^n) \times D'(\mathbb{R}^n) \to D'(\mathbb{R}^n) \) which extends the notion of pointwise multiplication defined initially on \( C^\infty_c(\mathbb{R}^n) \times C^\infty_c(\mathbb{R}^n) \).

\[ ^6 \text{At least } a \text{ priori it is not clear that multiplication by other functions will work. It turns out that we can multiply by slightly less regular functions under mild assumptions, but that this notion does not extend to all distributions.} \]
2. Lecture 2 (03/31): Composition, homogeneous distributions, and convolution

2.1. Composition with smooth maps and homogeneous distributions. In some cases, it is possible to define composition of distributions. We’ll focus on the case of composing a distribution in \( D' (\mathbb{R}^n) \) with a map \( f : U \to \mathbb{R}^m, U \subset \mathbb{R}^n, \) with \( m \leq n \).

We suppose first that there exist auxiliary functions \( y_{m+1}(x), y_{m+2}(x), \ldots, y_n(x) \) such that the function

\[
g : U \to \mathbb{R}^n, \quad g(x) = (f_1(x), \ldots, f_m(x), y_{m+1}(x), \ldots, y_n(x))
\]

has a \( C^\infty \) inverse \( h : g(U) \to U \). For \( y \in \mathbb{R}^n \), write \( y = (y', y'') \in \mathbb{R}^m \times \mathbb{R}^{n-m} \). Note that

\[
y = g(h(y)) \implies y' = f(h(y))
\]

since the first \( m \) components of \( g(x) \) are the \( m \) components of \( f(x) \). Then, if \( \phi \in C^\infty_c(U) \), we see that for \( u \in C^0(\mathbb{R}^m) \) we have\(^7\)

\[
\int_U u(f(x))\phi(x) \, dx = \int_{g(U)} u(f(h(y)))\phi(h(y)) \, |\det Dh(y)| \, dy
\]

\[
= \int_{\mathbb{R}^n} u(y')\phi(h(y)) \, |\det Dh(y)| \, dy
\]

\[
= \int_{\mathbb{R}^m} u(y')\tilde{\phi}(y') \, dy'
\]

where

\[
\tilde{\phi}(y') = \int_{\mathbb{R}^{n-m}} \phi(h(y', y'')) \, |\det Dh(y', y'')| \, dy''
\]

Thus, if we can construct auxiliary functions \( y_{m+1}, \ldots, y_n \) such that \( (f, y_{m+1}, \ldots, y_n) \) has a \( C^\infty \) inverse, then we can (uniquely) extend the notion of composition of a continuous function \( u : \mathbb{R}^m \to \mathbb{R} \) with \( f : U \to \mathbb{R}^m \) to distributions on \( U \) by defining

\[
(u \circ f, \phi)_U := (u, \tilde{\phi})_{\mathbb{R}^m}
\]

where \( \tilde{\phi} \) is defined as above.

In general, if the differential \( Df \) of \( f \) is surjective everywhere on \( U \), then this construction of auxiliary functions is possible locally, thanks to the Inverse Function Theorem. We can then take a locally finite partition of unity \( 1 = \sum \rho_k \), say with \( \text{supp} \rho_k \subset U_k \), with the property that for any compact subset \( K \subset U \), only finitely many intersections \( U_k \cap K \) are nonempty. Then, we can define

\[
(u \circ f, \phi)_U := \sum_k (u|_{U_k}, \rho_k \phi)_{U_k},
\]

\(^7\)In the second line, we made the substitution \( x = h(y) \), \( dx = |\det Dh(y)| \, dy \). In the third line, we can switch the region of integration to \( \mathbb{R}^n \), since \( \phi(h(y)) \) is nonzero only when \( h(y) \in \text{supp} \phi \subset U \implies y \in g(U) \).
where \((u|_{U_k}, \rho_k \phi)|_{U_k}\) is defined in the special case discussed in the previous paragraph (note that the assumption that the partition is locally finite guarantees the above sum is a finite sum for any \(\phi \in C^\infty_c(U)\)). We summarize this as follows:

**Theorem 2.1** (cf. Theorem 6.1.2 in [Hor90]). Suppose \(f : U \to \mathbb{R}^m\) is smooth, and \(Df(x)\) is surjective for all \(x \in U\). Then there is a unique continuous linear map \(f^* : \mathcal{D}'(\mathbb{R}^m) \to \mathcal{D}'(U)\) such that \(f^* u = u \circ f\) for all \(u \in C^0(\mathbb{R}^m)\). Moreover, for any partial derivative \(\partial_j\) we have

\[
\partial_j(f^* u) = \sum_{k=1}^{n} (\partial_j f_k) \cdot f^*(\partial_k u) \quad (i.e. \quad \partial_j(u \circ f) = \sum_{k=1}^{n} (\partial_k u \circ f) \partial_j f_k).
\]

**Remark 3.** If \(V \subset \mathbb{R}^m\) is an open set containing the image of \(f : U \to \mathbb{R}^m\), then the above theorem also holds replacing \(\mathbb{R}^m\) by \(V\).

**Example 2.2.** Consider \(f : \mathbb{R}^{2n} \to \mathbb{R}^n, f(x, x') = x - x'\) for \((x, x') \in \mathbb{R}^n \times \mathbb{R}^n\). Then \(g(x, x') = (x - x', x')\) admits a \(C^\infty\) inverse on \(\mathbb{R}^{2n}\), namely \(h(y', y'') = (y' + y'', y'')\) for \((y', y'') \in \mathbb{R}^n \times \mathbb{R}^n\) (then \(|\det Dh| = 1\) everywhere). Then

\[(u \circ f, \phi)_{\mathbb{R}^{2n}} = (u, \tilde{\phi})_{\mathbb{R}^n}
\]

where

\[
\tilde{\phi}(y) = \int_{\mathbb{R}^n} \phi(y' + y'', y'') \, dy''.
\]

In particular, the distribution \(\delta(x - x')\) is the distribution satisfying

\[(\delta(x - x'), \phi)_{\mathbb{R}^{2n}} = (\delta, \tilde{\phi})_{\mathbb{R}^n} = \tilde{\phi}(0) = \int_{\mathbb{R}^n} \phi(0 + y'', y'') \, dy'' = \int_{\mathbb{R}^n} \phi(x, x) \, dx
\]

for any \(\phi \in C^\infty_c(\mathbb{R}^{2n})\).

**Example 2.3.** Suppose \(U \subset \mathbb{R}^n\) is **conic**, meaning that \(x \in U \implies tx \in U\) for all \(t > 0\) (for example, \(U = \mathbb{R}^n\) or \(U = \mathbb{R}^n\setminus\{0\}\)). For \(t > 0\), the composition \(u(tx)\) for \(u \in \mathcal{D}'(U)\) is well-defined, and it satisfies

\[(u(tx), \phi) = t^{-n}(u, \phi(x/t)).
\]

(This is the same result if \(u\) is a continuous function and \(u(tx)\) is understood as a composition.)

**Definition 2.4.** We say that \(u \in \mathcal{D}'(U)\) is **homogeneous** of degree \(a \in \mathbb{R}\) if \(u(tx) = t^a u(x)\) for all \(t > 0\), where the composition \(u(tx)\) is defined above. Equivalently,

\[(u, \phi(x/t)) = t^{a+1}(u, \phi) \quad \text{or} \quad (u, \phi(tx)) = t^{-n-a}(u, \phi).
\]

**Example 2.5.** Consider the functions \(\chi_+^a\) defined in Example 1.9 for \(a > -1\). They define distributions which are homogeneous of degree \(a\).

**Example 2.6.** The Dirac delta \(\delta\) on \(\mathbb{R}^n\) is homogeneous of degree \(-n\). Indeed, \((\delta, \phi(tx)) = (\delta, \phi)\) for any \(t > 0\) since both sides evaluate to \(\phi(0)\), so \((\delta, \phi(tx)) = t^{-n-a}(\delta, \phi)\) for \(-n - a = 0\), i.e. \(a = -n\).
Example 2.7. Consider \( f : \mathbb{R}^{n+1}_t \to \mathbb{R} \) be given by \( f(t,x) = t^2 - |x|^2 \) for \((t,x) \in \mathbb{R} \times \mathbb{R}^n\). Note that \( Df \) is non-vanishing on \( \mathbb{R}^{n+1}_t \setminus \{0\} \), so \( u(t^2 - |x|^2) \) is well-defined as a distribution on \( \mathbb{R}^{n+1}_t \setminus \{0\} \). Moreover, if \( u \) is homogeneous of degree \( a \), then \( u(t^2 - |x|^2) \) is homogeneous of degree \( 2a \).

Note that given a distribution \( u \) on \( \mathbb{R}^n \setminus \{0\} \), it is not always possible to extend it to a distribution \( \tilde{u} \) on \( \mathbb{R}^n \) (i.e. we cannot always find \( \tilde{u} \in \mathcal{D}'(\mathbb{R}^n) \) such that \( \tilde{u}|_{\mathbb{R}^n \setminus \{0\}} = u \)). However, for homogeneous distributions this is possible:

**Theorem 2.8** (cf. Theorems 3.2.3 and 3.2.4 in [Hör90]). Suppose \( u \in \mathcal{D}'(\mathbb{R}^n \setminus \{0\}) \) is homogeneous of degree \( a \). Then there exists \( \tilde{u} \in \mathcal{D}'(\mathbb{R}^n) \) such that \( \tilde{u}|_{\mathbb{R}^n \setminus \{0\}} = u \). Moreover, if either \( a > -n \) or \( a \not\in \mathbb{Z} \), then such an extension is unique and also homogeneous of degree \( a \).

2.2. Convolution. Recall that the convolution of two functions \( f, g : \mathbb{R}^n \to \mathbb{R} \) is defined by

\[
(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y) \, dy = (f; g(x - \cdot))
\]

where \( g(x - \cdot) \) is the function \( y \mapsto g(x-y) \). This allows us to easily define convolution when \( f \) is a distribution and \( g \) is smooth with compact support:

**Definition 2.9.** Let \( u \in \mathcal{D}'(\mathbb{R}^n) \) and \( \phi \in C^\infty_c(\mathbb{R}^n) \). The convolution of \( u \) and \( \phi \) is the function \( u * \phi : \mathbb{R}^n \to \mathbb{R} \) defined by

\[
(u * \phi)(x) := (u, \phi(x - \cdot)).
\]

(Note that \( \phi(x - \cdot) \in C^\infty_c(\mathbb{R}^n) \) for any \( x \in \mathbb{R}^n \) if \( \phi \in C^\infty_c(\mathbb{R}^n) \).)

Some properties of convolution as defined above are as follows (proofs are in Section 4.1 of [Hör90]):

**Theorem 2.10.** Let \( u \in \mathcal{D}'(\mathbb{R}^n) \) and \( \phi, \psi \in C^\infty_c(\mathbb{R}^n) \). Then:

- \( u * \phi \in C^\infty(\mathbb{R}^n) \), with \( \text{supp} (u * \phi) \subset \text{supp} u + \text{supp} \phi \).
- For any multi-index \( \alpha \) we have
  \[
  \partial^\alpha (u * \phi) = (\partial^\alpha u) * \phi = u * (\partial^\alpha \phi).
  \]
- We have \( u * (\phi * \psi) = (u * \phi) * \psi \),
- \( \phi * \psi = \psi * \phi \) (viewing \( C^\infty_c(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n) \)).

Next time: We’ll define the composition of distributions (in limited contexts) by having it satisfy the associativity condition

\[
(u_1 * u_2) * \phi = u_1 * (u_2 * \phi).
\]

We’ll also discuss Schwartz kernels.

---

8For two subsets \( A, B \subset \mathbb{R}^n \), we define \( A + B = \{a + b : a \in A, b \in B\} \).

9Note in this equation that \( u * \phi \in C^\infty(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n) \), so we can again convolve it on the right with another function in \( C^\infty_c(\mathbb{R}^n) \), while \( \phi * \psi \in C^\infty_c(\mathbb{R}^n) \), so we can again convolve it on the left with a distribution.
3. Lecture 3 (04/05): Duality, Convolutions, and Schwartz Kernel

3.1. A comment on defining operations by duality. For all of the operations on distributions defined in the first two lectures, we can summarize the definitions in a common way. Suppose we have a continuous linear map \( L : C^\infty_c(X) \to D'(Y) \).

**Definition 3.1.** The adjoint of \( L : C^\infty_c(X) \to D'(Y) \) is the operator \( {}^tL : C^\infty_c(Y) \to D'(X) \) defined by

\[
({}^tL \psi, \phi)_X = (L \phi, \psi)_Y \quad \text{for} \quad \phi \in C^\infty_c(X), \psi \in C^\infty_c(Y).
\]

Note then that \( {}^t( {}^tL ) = L \).

**Example 3.2.** Some examples:

- For \( L = \partial_j \), we have \( {}^tL = -\partial_j \).
- For \( L = \rho \) (i.e. multiplication by \( \rho \in C^\infty \)), we have \( {}^tL = \rho \) (i.e. the same operator).
- If \( f : X \to Y \) is invertible with smooth inverse and \( L = f^* \) (i.e. \( L \phi = \phi \circ f \)), we have \( {}^tL = | \det f^{-1}|(f^{-1})^* \) (i.e. \( {}^tL \psi = (\psi \circ f^{-1})| \det f^{-1}| \)).

Suppose now that \( {}^tL \) now maps \( C^\infty_c(Y) \) not just into \( D'(X) \), but rather into \( C^\infty_c(X) \). Then, we can extend \( L \) to an operator \( \tilde{L} : D'(X) \to D'(Y) \), defined by

\[
(\tilde{L} u, \psi)_Y := (u, {}^tL \psi)_X.
\]

Then, \( \tilde{L} \) is continuous, and moreover it agrees with \( L \) on \( C^\infty_c(X) \). This is indeed how all of the operations defined so far (aside from convolution with \( C^\infty_c \)) have been defined.

If \( {}^tL \) does not map \( C^\infty_c(Y) \) into \( C^\infty_c(X) \), but rather just into \( C^\infty(X) \), we can still extend \( L \), albeit to a slightly different space.

**Definition 3.3.** The space \( E'(X) \) is the space of distributions \( u \in D'(X) \) such that \( \text{supp } u \) is compact.

**Theorem 3.4.** \( E'(X) \) is isomorphic, as topological vector spaces, to the dual space of \( C^\infty(X) \), where the topology of \( C^\infty(X) \) is that given by the seminorms

\[
\phi \mapsto \sum_{|\alpha| \leq k} \sup_K |\partial^\alpha \phi|, \quad k \in \mathbb{N}_0, K \subset X \text{ compact}.
\]

Thus, if \( {}^tL \) maps \( C^\infty_c(Y) \) into \( C^\infty(X) \), then \( L \) can be extended as a map \( E'(X) \to D'(Y) \). Similarly, if \( {}^tL \) maps into \( E'(X) \), then \( L \) can be extended as a map \( C^\infty(X) \to D'(Y) \).

3.2. Convolutions, continued. To define the convolution of two distributions \( u_1 \) and \( u_2 \), we could try to have it satisfy the “associativity” property above, i.e. for \( \phi \in C^\infty_c(\mathbb{R}^n) \) we would want \( u_1 * u_2 \) to satisfy

\[
(u_1 * u_2) * \phi = u_1 * (u_2 * \phi).
\]

There are two issues with doing so:
• First, the above condition gives a condition we’d like to be satisfied involving the convolution of our mystery distribution $u_1 * u_2$ against $\phi$; a priori it’s not clear how that defines the pairing $(u_1 * u_2, \phi)$.

• For the right-hand side to make sense, we want to somehow arrange for $u_2 * \phi$ to have compact support.

The second issue can be addressed by taking $u_2$ to have compact support. For the first issue, it turns out that knowledge of how a distribution convolves (i.e. knowledge of the operator $C_\infty^c(\mathbb{R}^n) \to C_\infty(\mathbb{R}^n)$, $\phi \mapsto u * \phi$) is enough information to determine the distribution itself. Indeed, just note that for any $\phi \in C_\infty^c(\mathbb{R}^n)$, if $\hat{\phi}(x) = \phi(-x)$, then

$$(u * \hat{\phi})(0) = (u, \phi).$$

However, there is another necessary condition that a convolution operator must satisfy. Namely, if for $h \in \mathbb{R}^n$ we let $\tau_h : C_\infty^c(\mathbb{R}^n) \to C_\infty^c(\mathbb{R}^n)$, $\tau_h \phi(x) = \phi(x-h)$, then $\tau_h(u * \phi) = u * \tau_h \phi$ (this follows basically from the definition). It turns out that this is essentially sufficient as well:

**Theorem 3.5** (cf. Theorem 4.2.1 of [Hör90]). If $U$ is a continuous linear map from $C_\infty^c(\mathbb{R}^n) \to C_\infty^c(\mathbb{R}^n)$, and $U \circ \tau_h = \tau_h \circ U$ for all $h \in \mathbb{R}^n$, then there exists a unique $u \in \mathcal{D}'(\mathbb{R}^n)$ such that $U\phi = u * \phi$ for all $\phi \in C_\infty^c(\mathbb{R}^n)$.

**Proof.** If such a $u$ were to exist, it must satisfy $(u, \phi) = (u * \hat{\phi})(0) = U(\hat{\phi})(0)$; hence define $u \in \mathcal{D}'(\mathbb{R}^n)$ by $(u, \phi) := U(\hat{\phi})(0)$. This is a distribution, i.e. is continuous, due to the continuity assumptions in the hypothesis. It remains to verify that $U\phi = u * \phi$ for all $\phi \in C_\infty^c(\mathbb{R}^n)$. This is where the assumption of commuting with translations comes in: just note that

$$(U\phi)(-h) = \tau_h(U\phi)(0) = U(\tau_h \phi)(0) = (u, \tau_h \phi) = (u * \tau_h \phi)(0) = \tau_h(u * \phi)(0) = (u * \phi)(-h)$$

for each $h \in \mathbb{R}^n$. Thus, $U\phi = u * \phi$ as functions in $C_\infty^c(\mathbb{R}^n)$.

As such, if $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$ with supp $u_2$ compact, we see that

$$U\phi := u_1 * (u_2 * \phi)$$

satisfies the assumptions in the theorem. Hence, we can make a definition:

**Definition 3.6.** Suppose $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$ with supp $u_2$ compact. The convolution $u_1 * u_2$ is the unique distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ satisfying

$$u * \phi = u_1 * (u_2 * \phi)$$

for all $\phi \in C_\infty^c(\mathbb{R}^n)$.

**Example 3.7.** The Dirac delta $\delta_0$ satisfies $\delta_0 * \phi = \phi$ for all $\phi \in C_\infty^c(\mathbb{R}^n)$. As a consequence, for all distributions $u \in \mathcal{D}'(\mathbb{R}^n)$ we also have $u * \delta_0 = u$.

By leveraging facts about convolution of functions, we can state:

**Theorem 3.8.** Let $u_1, u_2 \in \mathcal{D}'(\mathbb{R}^n)$ with supp $u_2$ compact. Then:

---

The topology on $C_\infty^c(\mathbb{R}^n)$ is the seminorm topology induced by the seminorms $\phi \mapsto \sum_{|\alpha| \leq k} \sup_K |\partial^\alpha \phi|$ over all $k \in \mathbb{N}_{\geq 0}$ and $K$ compact, i.e. a sequence of smooth functions converges if and only if it converges with respect to each of the preceding seminorms.
• If $\text{supp } u_1$ is also compact, then $u_1 * u_2 = u_2 * u_1$.
• We have $\text{supp } (u_1 * u_2) \subset \text{supp } u_1 + \text{supp } u_2$.
• If $u_3 \in \mathcal{D}'(\mathbb{R}^n)$ has compact support, then $u_1 * (u_2 * u_3) = (u_1 * u_2) * u_3$.
• We have $\partial^\alpha(u_1 * u_2) = (\partial^\alpha u_1) * u_2 = u_1 * (\partial^\alpha u_2)$. In particular, if $P = \sum a_\alpha \partial^\alpha$ is a constant-coefficient differential operator, then $P(u_1 * u_2) = (Pu_1) * u_2 = u_1 * (Pu_2)$.

An application of the last fact is the following: suppose $P$ is a constant-coefficient differential operator, and $u_2$ is compactly supported and satisfies $Pu_2 = \delta$ in the sense of distributions. Then, for any $u_1 \in \mathcal{D}'(\mathbb{R}^n)$, for $u = u_1 * u_2$ we have

$$Pu = P(u_1 * u_2) = u_1 * (Pu_2) = u_1 * \delta = u_1.$$ 

Thus, for any $u_1 \in \mathcal{D}'(\mathbb{R}^n)$, there exists a solution to $Pu = u_1$ in the sense of distributions, namely $u = u_1 * u_2$. This idea will be used more heavily next week.

There are many more situations in which the convolution of two distributions can be well-defined. One such situation is the following: suppose the map

$$\text{supp } u_1 \times \text{supp } u_2 \to \mathbb{R}^n, \quad (x, y) \mapsto x + y$$

is proper, meaning that the pre-image of compact sets is compact\textsuperscript{11} then the convolution $u_1 * u_2$ can be defined as follows: for a fixed $\phi \in C^\infty_c(\mathbb{R}^n)$, let $K = \text{supp } \phi$, and let $K_1$ and $K_2$ be the projections of the preimage of $K$ under the map $\text{supp } u_1 \times \text{supp } u_2 \to \mathbb{R}^n, (x, y) \mapsto x + y$. (Thus, if $(x, y) \in \text{supp } u_1 \times \text{supp } u_2$ and $x + y \in K$, then $x \in K_1$ and $y \in K_2$). Note that $K$, $K_1$, and $K_2$ are all compact, by the properness assumption. We then define

$$(u_1 * u_2, \phi) := ((\psi_1 u_1) * (\psi_2 u_2), \phi)$$

where $\psi_1, \psi_2 \in C^\infty_c(\mathbb{R}^n)$ are identically 1 in neighborhoods of $K_1$ and $K_2$, respectively. Note then that $\psi_i u_i$ are compactly supported distributions, so their convolution is well-defined. The idea behind the definition is to cut off the distributions $u_1$ and $u_2$ “only where they matter” when trying to evaluate the pairing of $u_1 * u_2$ against $\phi$.

To make this a well-defined definition, we need to check that this result is independent of the cutoffs $\psi_i$ chosen. For example, to check the result is independent of the choice of $\tilde{\psi}_2$, suppose $\psi_2$ and $\tilde{\psi}_2$ are both identically 1 in a neighborhood of $K_2$. We need to check that $((\psi_1 u_1) * (\psi_2 u_2), \phi) = ((\psi_1 u_1) * (\tilde{\psi}_2 u_2), \phi)$. Unraveling the definition of compositions, this amounts to checking that

$$(((\psi_1 u_1) * ((\tilde{\psi}_2 - \psi_2) u_2) * \tilde{\phi}))(0) = 0.$$ 

We note that $\tilde{\psi}_2 - \psi_2$ equals zero on a neighborhood of $K_2$, so $y \in \text{supp } (\tilde{\psi}_2 - \psi_2) \implies y \notin K_2$. In particular, if $x \in \text{supp } u_1$ and $y \in \text{supp } (\tilde{\psi}_2 - \psi_2)$, then $x + y \notin \text{supp } \phi$.

\textsuperscript{11}This can be phrased in other ways; two examples of which are the following: for any compact set $K \subset \mathbb{R}^n$:

• The set $(K - \text{supp } u_1) \cap \text{supp } u_2$ is compact.
• There exists $C > 0$ such that if $x \in \text{supp } u_1$ and $y \in \text{supp } u_2$, then $x + y \in K \implies |x| \leq C, |y| \leq C$. 

But we also know that
\[
supp \left( (\psi_1 u_1) \ast \left( (\tilde{\psi}_2 - \psi_2) u_2 \right) \ast \tilde{\phi} \right) \subset supp (\psi_1 u_1) + supp ((\tilde{\psi}_2 - \psi_2) u_2) + supp \tilde{\phi} \\
\subset supp u_1 + supp (\tilde{\psi}_2 - \psi_2) - supp \phi,
\]
and the latter set does not contain 0 by the discussion above. This shows that
\[
((\psi_1 u_1) \ast (\psi_2 u_2), \phi) \text{ does not depend on the choice of } \psi_2, \text{ so long as } \psi_2 \text{ is identically 1 in a neighborhood of } K_2.
\]
Similar methods show this is independent of the choice of \( \psi_1 \) as well.

**Example 3.9.** For \( \mathbb{R}_+ = [0, \infty) \), we have that \( \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}, \ (x, y) \rightarrow x + y \) is proper. Hence, if \( supp \ u_1, u_2 \subset \mathbb{R}_+ \), then the convolution \( u_1 \ast u_2 \) is well-defined.

### 3.3. Products and Schwartz Kernel.

One reason to study distributions, even if one is only interested in smooth solutions to a differential equation, is that they are intimately related to operators. First we consider products. In this subsection, let \( X \) and \( Y \) be open subsets of some Euclidean spaces.

**Definition 3.10.** For \( \phi \in C_\infty^c(X) \) and \( \psi \in C_\infty^c(Y) \), the tensor product is the function \( \phi \otimes \psi \in C_\infty^c(X \times Y) \) defined by
\[
(\phi \otimes \psi)(x, y) = \phi(x)\psi(y).
\]
From this, we can define the tensor product of distributions:

**Theorem 3.11.** Let \( u_1 \in \mathcal{D}'(X) \) and \( u_2 \in \mathcal{D}'(Y) \). Then there exists a unique distribution \( u \in \mathcal{D}'(X \times Y) \) satisfying
\[
u(\phi \otimes \psi) = u_1(\phi)u_2(\psi).
\]
This is called the tensor product of the distributions \( u_1 \) and \( u_2 \) and is denoted \( u_1 \otimes u_2 \).

We now consider the following situation: suppose we’re given a distribution \( K \) on the product space \( X \times Y \). Then the map
\[
\psi \mapsto (\phi \mapsto (K, \phi \otimes \psi)_{X \times Y})
\]
defines a linear operator \( T : C_\infty^c(Y) \rightarrow \mathcal{D}'(X) \). That is, given \( \psi \in C_\infty^c(Y) \), \( T\psi \) is a distribution on \( X \) satisfying \( (T\psi, \phi)_X = (K, \phi \otimes \psi)_{X \times Y} \). (In more informal terms, if we were to view \( K \) as a function \( K(x, y) \), then the operator in question is
\[
T\psi(x) = \int_{X \times Y} K(x, y)\psi(y) \, dy.
\]
Moreover, this operator \( T \) is continuous with respect to the respective topologies.

**Definition 3.12.** Let \( T : C_\infty^c(Y) \rightarrow \mathcal{D}'(X) \) be linear, and suppose \( K \in \mathcal{D}'(X \times Y) \) satisfies \( (T\psi, \phi)_X = (K, \phi \otimes \psi)_{X \times Y} \) for all \( \phi \in C_\infty^c(X) \) and \( \psi \in C_\infty^c(Y) \). Then we say that \( K \) is a Schwartz kernel of \( T \).

A remarkable fact is:

**Theorem 3.13** (Schwartz Kernel Theorem). Every continuous linear operator \( C_\infty^c(Y) \rightarrow \mathcal{D}'(X) \) has a unique Schwartz kernel associated to it.
Example 3.14. Some examples of Schwartz kernels (here the coordinates on $X$ and $Y$ are denoted $x$ and $y$):

- The identity operator $\text{Id} : C_c^\infty(X) \rightarrow C_c^\infty(X) \subset \mathcal{D}'(X)$ has Schwartz kernel $\delta(x-y)$.
- The differential operator $P = \sum a_\alpha(x) \partial^\alpha$ has Schwartz kernel $\sum a_\alpha(x) \partial^\alpha \delta(x-y)$.
- If $u \in \mathcal{D}'(\mathbb{R}^n)$, the convolution operator $C_c^\infty(\mathbb{R}^n) \rightarrow C_c^\infty(\mathbb{R}^n)$, $\phi \mapsto u \ast \phi$ has Schwartz kernel $u(x-y)$.
- If $T : C_c^\infty(Y) \rightarrow \mathcal{D}'(X)$ has Schwartz kernel $K \in \mathcal{D}'(X \times Y)$, then $\iota T : C_c^\infty(X) \rightarrow \mathcal{D}'(Y)$ has Schwartz kernel $\iota K \in \mathcal{D}'(Y \times X)$, where $\iota K(y,x) = K(x,y)^\prime$. Formally, if $\iota \psi(x,y) = \psi(y,x)$ for $\phi \in C_c^\infty(Y \times X)$, then $\iota \phi \in C_c^\infty(X \times Y)$, and
  $$(\iota K, \psi)_Y = (K, \iota \psi)_X.$$
- Suppose the Schwartz kernel, which a priori is a distribution on $X \times Y$, is actually in $L^2(X \times Y)$. Then the corresponding operator is a Hilbert-Schmidt operator. Such operators have some nice properties (e.g., they are compact operators).

The following was not covered during lecture but may be of interest for some students:

One application of studying the Schwartz kernel is the following: we can often give an upper bound on the set of singularities of $Tu$, if we know the singularities of the Schwartz kernel of $T$ and of $u$. We formalize the notions as follows:

Definition 3.15. Let $u \in \mathcal{D}'(X)$. The singular support of $u$, denoted $\text{sing supp} u$, is the set of $x \in X$ such that, for any neighborhood $V \ni x$, the restriction $u|_V$ does not agree with the restriction of any smooth function on $V$. (Equivalently, $x \in X$ is not in the singular support if there exists a neighborhood $V$ of $x$ such that $u|_V$ is smooth, i.e., agrees with the restriction of some smooth function).

Definition 3.16. Suppose $A \subset X \times Y$ and $B \subset Y$. The composition $A \circ B$ of sets is the set

$$A \circ B = \{ x \in X : \text{ there exists } y \in B \text{ such that } (x,y) \in A \}.$$ 

Equivalently,

$$A \circ B = \pi_X(A \cap \pi_Y^{-1}(B))$$

where $\pi_X : X \times Y \rightarrow X$ and $\pi_Y : X \times Y \rightarrow Y$ are the projections onto $X$ and $Y$.

Theorem 3.17. Suppose $T : C_c^\infty(Y) \rightarrow \mathcal{D}'(X)$ satisfies that $T$ maps into $C^\infty(X)$, and $\iota T$ maps continuously from $C_c^\infty(X)$ to $C^\infty(Y)$, so that $T$ can be extended to a map $T : \mathcal{E}'(Y) \rightarrow \mathcal{D}'(X)$. Let $K \in \mathcal{D}'(X \times Y)$ be the Schwartz kernel of $T$. Then,

$$\text{sing supp } Tu \subset \text{sing supp } K \circ \text{sing supp } u \quad \text{for all } u \in \mathcal{E}'(Y).$$

Proof Sketch. The proof boils down to the following statement, which will not be proven in this sketch:

if $T : \mathcal{E}'(Y) \rightarrow \mathcal{D}'(X)$ has Schwartz kernel $K \in C^\infty(Y \times X)$, then $Tu \in C^\infty(X)$ for all $u \in \mathcal{E}'(Y)$.
Assuming the statement, we can prove as follows. Suppose \((x, y) \not\in \text{sing supp } K \circ \text{sing supp } u\); we’d then like to show that \(x \not\in \text{sing supp } Tu\). Note that the assumptions then give that \(\{x\} \times \text{sing supp } u\) is disjoint from \(\text{sing supp } K\). Since the former set is compact and the latter set is closed, it follows that there exists open neighborhoods \(U\) and \(V\) of \(x\) and \(\text{sing supp } u\), respectively, such that \(U \times V\) is still disjoint from \(\text{sing supp } u\). Let \(U'\) and \(V'\) be neighborhoods of \(x\) and \(\text{sing supp } u\) compactly contained in \(U\) and \(V\), respectively, and let \(\phi \in C_c^\infty(U)\) and \(\psi \in C_c^\infty(V)\) be identically 1 on \(U'\) and \(V'\). Then the operator \(\phi T \psi\) has Schwartz kernel \((\phi \otimes \psi) \cdot K\) (or more colloquially \(\phi(x)K(x, y)\psi(y)\)), which is in \(C^\infty(X \times Y)\) since the support of \(\phi \otimes \psi\) is disjoint from the singular support of \(K\), and hence \(\phi T \psi u \in C^\infty(X)\). On the other hand, \(u - \psi u \in C_c^\infty(Y)\) since \(1 - \psi\) is supported away from \(\text{sing supp } u\), and hence \(\phi T (1 - \psi) u \in C^\infty(X)\) as well. Thus we have

\[
\phi Tu = \phi T \psi u + \phi T (1 - \psi) u \in C^\infty(X).
\]

Since \(\phi\) is identically 1 on \(U'\), it follows that the restriction of \(Tu\) to \(U'\) is equal to that of a \(C^\infty\) function on \(U'\), and hence \(x \not\in \text{sing supp } Tu\), as desired. \(\square\)

**Next Lecture:** Tempered distributions and Fourier transform.
4. Lecture 4 (04/07): Fourier Transform

4.1. Fourier Transform on functions and Schwartz space. Recall the Fourier transform on functions:

**Definition 4.1.** Let \( f \in L^1(\mathbb{R}^n) \). The *Fourier transform* of \( f \) is the function \( \hat{f} : \mathbb{R}^n \to \mathbb{C} \) defined by

\[
\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) \; dx.
\]

Some properties:

**Theorem 4.2.** Let \( f \in L^1(\mathbb{R}^n) \). Then:

1. We have \( \hat{f} \in C^0(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \).
2. If in addition \( f \in C^1(\mathbb{R}^n) \) and \( \partial_j f \in L^1(\mathbb{R}^n) \), then \( \hat{\partial_j f}(\xi) = i\xi_j \hat{f}(\xi) \).
3. If in addition \( x_j f \in L^1(\mathbb{R}^n) \), then \( \hat{f} \in C^1(\mathbb{R}^n) \) with \( \partial_j \hat{f}(\xi) = -ix_j \hat{f}(\xi) \).

It follows that the Fourier transform intertwines differentiation and multiplication by monomials. Hence, we are interested in a space of functions which behaves well under both operations:

**Definition 4.3.** The *Schwartz space*, denoted \( S \) or \( S(\mathbb{R}^n) \), is the set of all smooth functions \( \phi \in C^\infty(\mathbb{R}^n) \) satisfying the property that

\[
\sup_{x \in \mathbb{R}^n} |x^\beta \partial^\alpha \phi(x)| < \infty
\]

for all multi-indices \( \alpha, \beta \). It is a topological vector space, when equipped with the topology induced by the seminorms appearing in the left-hand side of the above inequality.

That is, the Schwartz space consists of functions that are not only infinitely differentiable, but in addition decay faster than any inverse polynomial rate, with their derivatives decaying that fast as well. Note that \( \phi \in S(\mathbb{R}^n) \implies x^\beta \partial^\alpha \phi \in S(\mathbb{R}^n) \) for any multi-indices \( \alpha \) and \( \beta \), i.e. \( S(\mathbb{R}^n) \) is closed under differentiation and multiplication by polynomials.

**Example 4.4.** We have the inclusion\(^{13}\) \( C^\infty_c(\mathbb{R}^n) \subset S(\mathbb{R}^n) \), i.e. any compactly supported smooth function will satisfy the above estimates.

---

\(^{12}\)There are multiple commonly used conventions regarding the definition/normalization of the Fourier transform. This is the convention we’ll use, since it works well with differentiation.

\(^{13}\)More accurately, we should write that there is an inclusion \( C^\infty_c(\mathbb{R}^n) \to S(\mathbb{R}^n) \) which is continuous with respect to the respective topologies on \( C^\infty_c(\mathbb{R}^n) \) and \( S(\mathbb{R}^n) \). Note that this inclusion, while continuous, is not an *embedding* of topological vector spaces, i.e. the topology on \( C^\infty_c(\mathbb{R}^n) \) is not the subspace topology obtained by the inclusion into \( S(\mathbb{R}^n) \). In particular, a sequence \( \{\phi_k\} \) in \( C^\infty_c(\mathbb{R}^n) \) may converge in \( S(\mathbb{R}^n) \) without converging in \( C^\infty_c(\mathbb{R}^n) \). To see this, fix a nonzero \( \phi \in C^\infty_c(\mathbb{R}^n) \), let \( \{a_k\} \) and \( \{b_k\} \) be decreasing sequences of positive numbers, and let \( \phi_k(x) = a_k \phi(b_k x) \). One can check that a sufficient condition for \( \phi_k \) to converge to 0 in \( S(\mathbb{R}^n) \) is for \( \lim_{k \to \infty} a_k b_k^m = 0 \) for all \( m \in \mathbb{Z} \); this can e.g. be arranged by taking \( a_k = e^{-k} \) and \( b_k = 1/k \). However, \( \phi_k \) does not converge to 0 in \( C^\infty_c(\mathbb{R}^n) \) since \( \text{supp} \phi_k = b_k^{-1} \text{supp} \phi \), so that in particular the supports of \( \phi_k \) are not all contained in some fixed compact set, thus violating a necessary condition for sequences to converge in \( C^\infty_c(\mathbb{R}^n) \).
Example 4.5. If $A$ is a symmetric positive definite $n \times n$ matrix, then $\phi(x) = e^{-(Ax,x)/2}$ is in $\mathcal{S}(\mathbb{R}^n)$.

The decay requirement on Schwartz functions gives that any Schwartz function is integrable, and hence we can consider the Fourier transform of Schwartz functions. We then have:

Lemma 4.6. For any $\phi \in \mathcal{S}(\mathbb{R}^n)$, we have $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$ as well.

Proof. It suffices to note that the intertwining of multiplication and differentiation allows us to conclude that

$$\xi^\beta \partial^{\alpha}_\xi \hat{\phi}(\xi) = \hat{\psi}(\xi), \quad \text{where } \psi = (-i\partial)^\beta_x ((-ix)^\alpha \phi).$$

Then $\psi$ is also a Schwartz function, and hence the Fourier transform is bounded. □

4.2. Tempered distributions and extending the Fourier transform. From Lemma 4.6, we see that the Schwartz space is a nice space of “test functions” which behaves well with respect to the Fourier transform. This motivates considering a class of distributions dual to this nice test space:

Definition 4.7. The space of tempered distributions, denoted $\mathcal{S}'(\mathbb{R}^n)$, is the dual space (i.e. space of continuous linear functionals into $\mathbb{C}$) of $\mathcal{S}(\mathbb{R}^n)$, where $\mathcal{S}(\mathbb{R}^n)$ is equipped with the seminorm topology. The space of tempered distributions is also a topological vector space, equipped with the weak-* topology.

Remark 4. Since $C^\infty_c(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset C^\infty(\mathbb{R}^n)$, with all inclusions continuous with respect to the respective topologies, it follows that we have inclusions $\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$, since $\mathcal{D}'(\mathbb{R}^n)$ and $\mathcal{E}'(\mathbb{R}^n)$ are the dual spaces of $C^\infty_c(\mathbb{R}^n)$ and $C^\infty(\mathbb{R}^n)$, respectively.

Remark 5. Note as well that $\mathcal{S}'(\mathbb{R}^n)$ is closed under differentiation, as well as multiplication by either functions in $\mathcal{S}(\mathbb{R}^n)$ or by polynomials, though not necessarily by arbitrary smooth functions.

Remark 6. It can be shown that $\mathcal{S}(\mathbb{R}^n)$ is in fact dense in $\mathcal{S}'(\mathbb{R}^n)$ (with respect to the weak-* topology on $\mathcal{S}'(\mathbb{R}^n)$). Thus, if we want to extend operators initially defined on $\mathcal{S}$ to continuous operators defined on $\mathcal{S}'$, such an extension would necessarily be unique due to the density of $\mathcal{S}$ in $\mathcal{S}'$.

We now ask how to define the Fourier transform for tempered distributions. We thus aim to find the adjoint of the Fourier transform, i.e. for $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$, see if we can rewrite the pairing $(\hat{\phi}, \psi)$ in terms of $\phi$ applied to an operator of $\psi$. Indeed, we see that

$$(\hat{\phi}, \psi) = \int_{\mathbb{R}^n} \hat{\phi}(\xi) \psi(\xi) d\xi = \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} e^{-i\xi \cdot x} \phi(x) dx \right) \psi(\xi) d\xi$$

$$= \int_{\mathbb{R}^n} \phi(x) \left( \int_{\mathbb{R}^n} e^{-i\xi \cdot x} \psi(\xi) d\xi \right) dx$$

$$= \int_{\mathbb{R}^n} \phi(x) \hat{\psi}(x) dx = (\phi, \hat{\psi})$$
by Fubini’s Theorem. Note as well that \( \hat{\psi} \in S(\mathbb{R}^n) \) if \( \psi \in S(\mathbb{R}^n) \). Thus, we define:

**Definition 4.8.** Given \( u \in S'(\mathbb{R}^n) \), the *Fourier transform* of \( u \) is the distribution \( \hat{u} \in S'(\mathbb{R}^n) \) defined by

\[
(\hat{u}, \phi) := (u, \hat{\phi}).
\]

We will sometimes denote the Fourier transform as an operator \( F : S' \to S' \) (i.e. \( F(u) = \hat{u} \)). Note that \( F \) is continuous both as a map \( S \to S \) and \( S' \to S' \).

Some important properties (often proven by proving the analogous properties for Schwartz functions):

**Theorem 4.9.** Let \( u, v \in \mathcal{D}'(\mathbb{R}^n) \). Then:

- If \( u \in L^1(\mathbb{R}^n) \), then the distribution \( \hat{u} \) defined in Definition 4.8 agrees with the continuous bounded function \( \hat{u} \) defined in Definition 4.1.
- If \( u \in L^2(\mathbb{R}^n) \), then the distribution \( \hat{u} \) is in fact in \( L^2(\mathbb{R}^n) \). Moreover, for \( v \in L^2(\mathbb{R}^n) \), we have the **Plancherel formula**

  \[
  (\hat{u}, \hat{v}) = (2\pi)^n (u, \overline{v}) \quad \text{(in particular } \|\hat{u}\|_{L^2} = (2\pi)^{n/2} \|u\|_{L^2}).
  \]

- If \( u \) is compactly supported, then \( \hat{u} \) is in fact a \( C^\infty \) function, and moreover it satisfies

  \[
  \hat{u}(\xi) = (u, e^{-i\xi \cdot x})
  \]

  (the RHS means \( (u, \chi(x)e^{-i\xi \cdot x}) \) for any \( \chi \in C^\infty_c(\mathbb{R}^n) \) which is identically 1 on \( \text{supp } u \).)

- If \( v \in \mathcal{E}'(\mathbb{R}^n) \), then

  \[
  \hat{u} \ast v = \hat{\hat{u}v}.
  \]

  (The formula continues to hold in many other situations as well.)

- If \( u \) and \( v \) are sufficiently nice (e.g. in \( S \)), then

  \[
  \hat{uv} = (2\pi)^{-n} \hat{u} \ast \hat{v}.
  \]

- In the sense of distributions, we have

  \[
  \partial_{x_j} \hat{u} = i\xi_j \hat{u}, \quad \hat{x_j} \hat{u} = i\partial_{\xi_j} \hat{u}.
  \]

- We have

  \[
  \hat{\hat{u}}(-x) = (2\pi)^n u,
  \]

  and hence

  \[
  F^{-1}u(x) = (2\pi)^{-n} F u(-x),
  \]

  or more colloquially the inverse Fourier transform is given by

  \[
  u(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \hat{u}(\xi) \, d\xi.
  \]
4.3. Examples.

Example 4.10. Let $\phi(x) = e^{-ax^2/2}$ on $\mathbb{R}$, with $a > 0$. To calculate $\hat{\phi}$, we first calculate $\hat{\phi}(0) = \int_{\mathbb{R}} e^{-ax^2/2} dx$. We recall that 

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi} \implies \hat{\phi}(0) = \int_{\mathbb{R}} e^{-ax^2/2} dx = \sqrt{\frac{2\pi}{a}}.$$ 

Moreover, noting that $x\phi(x) = xe^{-ax^2/2} = -\frac{1}{a} \phi'(x)$, we have 

$$(\hat{\phi})'(\xi) = -i\hat{x}\phi(\xi) = \frac{i}{a} \hat{\phi}'(\xi) = \frac{i}{a} (i\xi \hat{\phi})(\xi) = -\frac{\xi}{a} \hat{\phi}(\xi).$$

Recalling that $y'(t) = -cty(t) \implies y(t) = y(0)e^{-ct^2/2},$ it follows that 

$$\hat{\phi}(\xi) = \hat{\phi}(0)e^{-\xi^2/(2a)} = \sqrt{\frac{2\pi}{a}} e^{-\xi^2/(2a)}.$$ 

In particular, for $a = 1$ we see that $e^{-x^2/2}$ is an eigenfunction of the Fourier transform. This computes the Fourier transform of Gaussians in one dimension.

Suppose now that $\phi : \mathbb{R}^n \to \mathbb{R}$ is a multivariable Gaussian given by $\phi(x) = e^{-(Ax,x)/2}$, where $A$ is a symmetric positive definite $n \times n$ matrix. We now want to compute 

$$\hat{\phi}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi \cdot x} e^{-(Ax,x)/2} dx, \quad \xi \in \mathbb{R}^n.$$ 

We diagonalize $A = Q^{-1}DQ$, where $D$ is diagonal and $Q$ is orthogonal; note then that $\langle Ax, x \rangle = \langle DQx, Qx \rangle$. If we now let $y = Qx$ (then $dy = dx$ since $Q$ has determinant $\pm 1$), the above integral becomes 

$$\int_{\mathbb{R}^n} e^{-i\xi \cdot Q^{-1}Qx} e^{-(DQx,Qx)/2} dx = \int_{\mathbb{R}^n} e^{-i\xi \cdot Q^{-1}Q} e^{-(Dy,y)/2} dy.$$ 

Note that we can write $\xi \cdot Q^{-1}Q = \eta \cdot y$. If we let $\eta = Q\xi$, with $\eta = (\eta_1, \ldots, \eta_n)$, and let the diagonal values of $D$ be $a_1, \ldots, a_n$ (note these are all positive), then the above integral becomes 

$$\int_{\mathbb{R}^n} e^{-i\eta \cdot y} e^{-(\sum_{j=1}^n a_j y_j^2)/2} dy = \prod_{j=1}^n \left( \int_{\mathbb{R}} e^{-i\eta_j y_j} e^{-a_j y_j^2/2} dy_j \right)$$

$$= \prod_{j=1}^n \left( \sqrt{\frac{2\pi}{a_j}} e^{-\eta_j^2/(2a_j)} \right)$$

$$= \frac{(2\pi)^{n/2} e^{-(\sum_{j=1}^n \frac{\eta_j^2}{a_j})/2}}{\left( \prod_{j=1}^n a_j \right)^{1/2}}.$$

\[14\] Under different conventions, the exact choice of Gaussian that ends up being an eigenfunction may differ, but it will always be the case that some Gaussian is an eigenfunction.
Note that $\prod_{j=1}^n a_j = \det A$, and $\sum_{j=1}^n \eta_j^2 = \langle D^{-1} \eta, \eta \rangle = \langle D^{-1} Q \xi, Q \xi \rangle = \langle A^{-1} \xi, \xi \rangle$. It follows that we can write
\[
\hat{\phi}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi \cdot x} e^{-\langle Ax, x \rangle/2} dx = \frac{(2\pi)^{n/2}}{(\det A)^{n/2}} e^{-\langle A^{-1} \xi, \xi \rangle/2}.
\]

**Example 4.11.** Consider the distribution given by the constant function $1 \in C^\infty(\mathbb{R}^n)$. This does not belong to $L^1$ or $L^2$, so we need to compute its Fourier transform in the sense of distributions. Thus, we consider the distribution $(\hat{1}, \phi) = (1, \hat{\phi}) = \int_{\mathbb{R}^n} \hat{\phi}(\xi) d\xi$.

By the Fourier inversion formula, the right-hand side equals $(2\pi)^n \phi(0)$ (since $1 = e^{i(0, \xi)}$). It follows that $(\hat{1}, \phi) = (2\pi)^n \phi(0) \implies \hat{1} = (2\pi)^n \delta_0$. Similar logic yields $\hat{e^{i\xi_0 \cdot x}} = (2\pi)^n \delta_{\xi_0}$ for any $\xi_0 \in \mathbb{R}^n$.

Another way to compute the Fourier transform is by approximating the distribution by Schwartz functions and then take the limit (in the sense of distributions): this works because the Fourier transform is continuous as a map $S' \to S'$. As such, note that for $\epsilon > 0$, the Gaussians $e^{-\epsilon |x|^2/2}$ converge to $1$ in the space of distributions, meaning that $\lim_{\epsilon \to 0^+} (e^{-\epsilon |x|^2/2}, \phi) = (1, \phi)$ for all $\phi \in S(\mathbb{R}^n)$. Hence, the Fourier transforms of $e^{-\epsilon |x|^2/2}$ should also converge to the Fourier transform of $1$. From the previous example, we have
\[
e^{-\epsilon |x|^2/2}(\xi) = \left(\frac{2\pi}{\epsilon}\right)^{n/2} e^{-|\xi|^2/(2\epsilon)}.
\]

We note the following about the family of functions on the RHS:

- The integral of the RHS equals $(2\pi)^n$ for all $\epsilon$.
- As $\epsilon \to 0$, the RHS converges, uniformly outside any neighborhood of the origin, to zero.

These are enough to guarantee that $e^{-\epsilon |x|^2/2} \to (2\pi)^n \delta_0$ in $S'(\mathbb{R}^n)$ (cf. Problem 4 on HW 1).

**The following was not covered during lecture but may be of interest for some students:**

One technique often used in computing Fourier transforms of distributions is to consider analytic families of distributions:

**Definition 4.12.** Let $U \subset \mathbb{C}$ be open, and let $\{u_z\}_{z \in U}$ be a collection of tempered distributions in $S'(\mathbb{R}^n)$ indexed by $U$. We say that $\{u_z\}_{z \in U}$ is an analytic family of distributions on $U$ if, for any $\phi \in S(\mathbb{R}^n)$, the function
\[
U \ni z \mapsto (u_z, \phi) \in \mathbb{C}
\]
is a complex analytic function on $U$.

Most operations we’ve defined so far preserve the property of a family of distributions being analytic; in particular the Fourier transform of an analytic family of
distributions is also an analytic family of distributions, since for any $\phi \in \mathcal{S}(\mathbb{R}^n)$ the function $z \mapsto (\hat{\alpha}_z, \phi)$ is, by definition, the function $z \mapsto (u_z, \hat{\phi})$, which is analytic since $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$.

**Example 4.13.** Consider the function $u(x) = e^{iax^2/2}$ on $\mathbb{R}$, with $a \in \mathbb{R}\setminus\{0\}$. This defines a tempered distribution since $|u| = 1$ on $\mathbb{R}$. How do we compute its Fourier transform?

The trick is to use our computations for Gaussians $e^{-ax^2/2}$, $a > 0$ from before, and use analyticity to extend our results when “$a$ is complex”. Concretely, we note that for $U = \{\text{Re } z > 0\}$, the family $\{e^{-z^2/2}\}_{z \in U}$ is analytic (note that requiring $\text{Re } z > 0$ guarantees that $e^{-z^2/2}$ is bounded and thus defines a tempered distribution). Thus, its Fourier transform is also analytic in $U$. Moreover, the family of distributions

$$\left\{ \left( \frac{2\pi}{z} \right)^{1/2} e^{-\xi^2/(2z)} \right\}_{z \in U}$$

is also an analytic family of distributions (here the square root is well-defined on $U$ and sends $\mathbb{R}_+$ to $\mathbb{R}_+$; concretely $(re^{i\theta}) = r^{1/2}e^{i\theta/2}$ for $r > 0$, $-\pi/2 < \theta < \pi/2$, and $(2\pi)^{1/2}e^{-\xi^2/(2z)} = \mathcal{F}(e^{-z^2/2})$ when $z \in \mathbb{R}_+$. Thus by analytic continuation the two families must agree for all $z \in U$, i.e.

$$\mathcal{F}(e^{-z^2/2}) = \left( \frac{2\pi}{z} \right)^{1/2} e^{-\xi^2/(2z)}$$

for all $z$ with $\text{Re } z > 0$.

This does not quite give us our result, since we’d like to plug in $z = -ia$, which is not in this open set. Nonetheless, we note that $-ia + \epsilon \in U$ for $\epsilon > 0$, with $-ia + \epsilon \to -ia$ as $\epsilon \to 0^+$. Hence the Fourier transform of $e^{-(-ia+\epsilon)x^2/2}$ approaches that of $e^{iax^2/2}$, and hence

$$\mathcal{F}(e^{iax^2/2}) = \lim_{\epsilon \to 0^+} \left( \frac{2\pi}{-ia + \epsilon} \right)^{1/2} e^{-\xi^2/(2(-ia+\epsilon))}.$$

The only subtlety in evaluating the limit on the RHS is the square root:

- If $a > 0$, then $\frac{2\pi}{-ia + \epsilon} \to \frac{2\pi}{|a|} = \frac{2\pi}{|a|} e^{i\pi/2}$. Hence

$$\lim_{\epsilon \to 0^+} \left( \frac{2\pi}{-ia + \epsilon} \right)^{1/2} = \left( \frac{2\pi}{|a|} \right)^{1/2} e^{i\pi/4}.$$

- If $a < 0$, then $\frac{2\pi}{-ia + \epsilon} \to -\frac{2\pi}{|a|} = \frac{2\pi}{|a|} e^{-i\pi/2}$. Hence

$$\lim_{\epsilon \to 0^+} \left( \frac{2\pi}{-ia + \epsilon} \right)^{1/2} = \left( \frac{2\pi}{|a|} \right)^{1/2} e^{-i\pi/4}.$$

Putting it altogether, we obtain

$$\mathcal{F}(e^{iax^2/2}) = \left( \frac{2\pi}{|a|} \right)^{1/2} e^{i\frac{\pi}{2} \text{sgn } a} e^{-i\xi^2/(2a)}.$$
Similarly, if $A$ is a real symmetric non-singular $n \times n$ matrix, then similar arguments in Example 4.10 gives

$$
\mathcal{F}(e^{i\langle Ax, x \rangle / 2}) = \frac{(2\pi)^{n/2}}{|\det A|^{1/2}} e^{i \frac{x}{2} \text{sgn } A} e^{-i \langle A^{-1} \xi, \xi \rangle / 2},
$$

where $\text{sgn } A$ is the sum of the sign of its eigenvalues, i.e. the number of positive eigenvalues minus the number of negative eigenvalues.

**Next Lecture:** Introduction to parabolic equations.

**References**