Math 222A Lecture 11 Notes

Daniel Raban

September 30, 2021

1 Introduction to Distribution Theory

1.1 Weak solutions to PDEs

For the next month or so, our goal will be to study linear, constant coefficient PDEs

$$P(\partial)u = f, \qquad P(\partial) = \sum_{|\alpha| \leq m} c_\alpha \partial^\alpha.$$

We will first take a detour to study the theory of distributions. First, some motivation:

Example 1.1. Recall the transport equation

$$\begin{cases} (\partial_t + A_j \partial_j)u = 0 \\ u(0) = u_0 \end{cases}$$

with constant coefficients A_j. The characteristics are given by $\dot{x} = A$, which gives $x(t) = x(0) + tA$. This means that $\dot{u} = 0$ along these characteristics, so $u(x(t), t) = u(x(0), 0)$. In other words,

$$u(x, t) = u_0(x - tA).$$

Classically, if $u_0 \in C^1$, then $u \in C^1$. What if $u_0 \in C$? It doesn't make sense to say that the solution u is continuous because we need to take derivatives. If we interpret the equation as a directional derivative, $u_0 \in C$ gives a solution. This interpretation relies strongly on the specific problem. Can we treat this problem in general?

Suppose we have a smooth function $\varphi \in C_0^\infty$. We can write the equation as the condition

$$\int_{\mathbb{R}^{n+1}} (\partial_t + A_j \partial_j)u\varphi \, dx = 0,$$

where a function is 0 if it integrates to be 0 against all $\varphi \in C_0^\infty$. Now integrate by parts to get

$$-\int_{\mathbb{R}^{n+1}} u(\partial_t + A_j \partial_j)\varphi \, dx = 0, \qquad \forall \varphi \in C_0^\infty,$$

which applies to all $u \in C$. Our continuous solution will be a solution to this integral equation.

Definition 1.1. u is a **weak solution** to a PDE if the corresponding integral equation holds for all $\varphi \in C_0^\infty$.

Example 1.2. Recall the Burgers equation

$$u_t + uu_x = 0, \qquad u(0) = u_0.$$

The characteristics are given by $\dot{x} = u$ and $\dot{u} = 0$. The characteristics will intersect, and at the point of intersection of characteristics, the solution will start to develop a jump discontinuity, known as a **shock**.
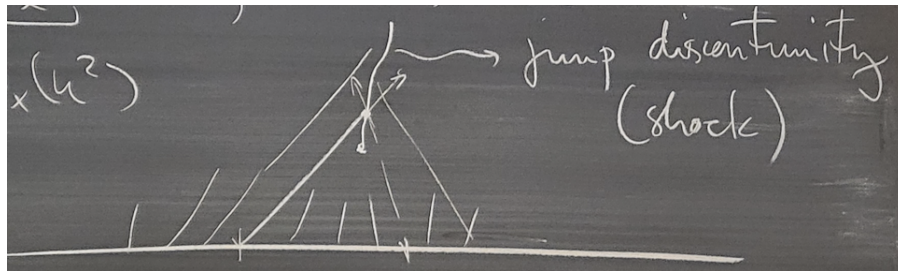
In this problem, if we think of the equation as a directional derivative, the derivative along the characteristics are different when they intersect, so we cannot get a solution. However, we can similarly look for a weak solution by integrating by parts as before. When we do this, we want to think of uu_x as $\frac{1}{2}\partial_x(u^2)$.

1.2 Topologies on vector spaces

The key idea in the theory of distribution is that we can think of a function $u : \mathbb{R}^n \to \mathbb{R}$ as a linear map on all $\varphi \in C_0(\mathbb{R}^n)$ via

$$u(\varphi) := \int_{\mathbb{R}^n} u \cdot \varphi \, dx.$$

Observe that if $u(\varphi) = 0$ for all φ, then $u = 0$.

We will use the notation $\mathcal{D} = C_0^\infty$ to refer to the smooth functions with compact support. Obesrve that $\mathcal{D}$ is a linear space. What is the topology of $\mathcal{D}$? Recall that C is a normed space, with

$$\|u\|_C = \sup_{x \in \mathbb{R}^n} |u(x)|.$$

Recall:

Definition 1.2. A **normed space** is a vector space V, with a norm map $\|\cdot\| : V \to \mathbb{R}$ (or $\mathbb{C}$) satisfying

(a) $\|u\| \geq 0$, with equality iff $u = 0$.

(b) $\|\lambda u\| = |\lambda| \|u\|$ for all $\lambda \in \mathbb{R}$ (or $\mathbb{C}$).

(c) $\|u + v\| \leq \|u\| + \|v\|$.

We obtain a metric space structure, given by $d(u, v) = \|u - v\|$. Recall that **completeness** of a metric space means that every Cauchy sequence is convergent.

Definition 1.3. A **Banach space** is a complete normed space.

Here is a special class:

Definition 1.4. A **Hilbert space** is a vector space with a complete inner product $\langle u, v \rangle = u \cdot v$.

In a Hilbert space, we get a norm by

$$u \cdot u = \|u\|^2 \geq 0.$$

Example 1.3. The L^2 **space** is given by

$$L^2(\mathbb{R}^n) = \left\{ u : \mathbb{R}^n \to \mathbb{R} \mid \int |u|^2 \, dx < \infty \right\}.$$

This space is a Hilbert space, given the inner product

$$u \cdot v = \int_{\mathbb{R}^n} uv \, dx$$

(with v replaced by $\bar{v}$ in the complex case).

Hilbert spaces are a special case of Banach spaces, but a single space can have different norm structures on it.

Example 1.4. We can equip $\mathbb{R}^n$ with the norm $\|v\|^2 = \sum_j v_j^2$ which comes from the usual dot product (a Hilbert space structure). We can also equip $\mathbb{R}^n$ with the L^p norm $\|v\|^p = \sum_j |v_j|^p$ with $1 \leq p < \infty$, which gives a Banach space structure.

Example 1.5. C^k is a Banach space with the norm

$$\|u\| = \sup_{|\alpha| \leq K} \sup_{x \in \mathbb{R}^n} |\partial^\alpha u(x)|.$$

Returning to our objective, what norm can we give $C^\infty(\mathbb{R}^n)$? We can define

$$\|u\|_\alpha = p_\alpha(u) := \sup_{x \in \mathbb{R}^n} |\partial^\alpha u(x)|.$$

The problem is that we have infinitely many of these. What would $u_n \to u$ mean in $C^\infty(\mathbb{R}^n)$? We want to say that $\partial^\alpha u_n \to \partial^\alpha u$ uniformly for all α.

The solution is to use all the $\|\cdot\|_\alpha$ as **seminorms**, which satisfy all the norm conditions except for $\|u\| = 0 \implies u = 0$.

Definition 1.5. Locally convex spaces are vector spaces equipped with a family of seminorms. A complete, locally convex space is called a **Fréchet space**.

In a locally convex space,

$$p_\alpha(u) = 0 \quad \forall \alpha \implies u = 0.$$

Why is this called a "locally convex space"? The idea is that each seminorm gives you neighborhoods of points, which may not be nested in each other for different seminorms. But these are all convex neighborhoods, and we can intersect these neighborhoods to get more convex neighborhoods around every point.

The picture of our function spaces looks like

$$\text{Hilbert spaces} \subsetneq \text{Banach spaces} \subsetneq \text{Fréchet spaces.}$$

Example 1.6. We will use the notation $\mathcal{E} = \{u \in \mathbb{R}^n \to \mathbb{R} \mid u \text{ is smooth}\}$. Here, we ask for nothing at ∞. What does $u_n \to u$ mean in $\mathcal{E}$? We can define this as $\partial^\alpha u_n \to \partial^\alpha u$ uniformly on compact sets. For this space, we need to use the collection of seminorms

$$p_{\alpha,K}(u) = \sup_{x \in K} |\partial^\alpha u(x)|, \qquad \alpha \in \mathbb{N}^d, K \text{ compact.}$$

We don't need to check *all* compact sets; it suffices to take nested balls with radius going to ∞. With this topology, $\mathcal{E}$ is a locally convex space.

For $\mathcal{D}$, we have an issue: if we have a sequence of functions of compact support, the support may grow to not be compact in the limit. To solve this, there is a notion called the **inductive limit** of locally convex spaces, essentially cooked up only to describe $\mathcal{D}$. To make a long story short, we describe convergence in $\mathcal{D}$ as $u_n \to u$ in $\mathcal{D}$ if

(a) $\partial^\alpha u_n \to \partial^\alpha u$ uniformly.

(b) There is a compact set K such that $\operatorname{supp} u_n \subseteq K$.

Remark 1.1. If $u \in C(\mathbb{R}^n)$ and $\varphi \in \mathcal{D}$, the map $\varphi \mapsto u(\varphi) = \int u\varphi \, dx$ is continuous.

Definition 1.6. The space of **distributions**, denoted $\mathcal{D}'$ or $\mathcal{D}^*$ is the space of linear, continuous $f : \mathcal{D} \to \mathbb{R}$.

This seems to separate us from our original goal. If we have a function, we can get a distribution, but if we have a distribution, we can't always get a function back; instead, we get *generalized functions*.[1]

---

[1]The term "distribution" comes from the French school, whereas the term "generalized functions" comes from the Russian school.

1.3 Examples of distributions

Here are some examples of distributions.

Example 1.7. The **Dirac mass** at 0 is

$$\delta_0(\varphi) = \varphi(0).$$

Example 1.8. Another distribution is

$$\delta_0'(\varphi) = -\varphi'(0).$$

The reason for the minus sign will become apparent later on. In general, we can define
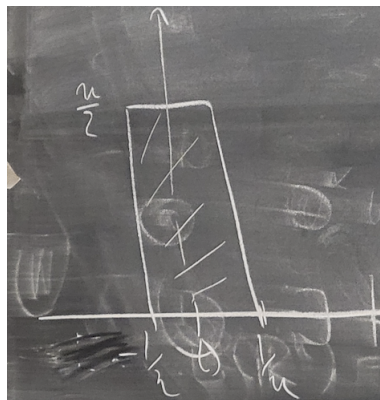
$$\delta_x^{(\alpha)}(\varphi) = (-1)^{|\alpha|}\partial^\alpha\varphi(x).$$

The space $\mathcal{D}'$ of distributions is a linear space.[2] It has the topology of **weak convergence**: $f_n \to f$ in $\mathcal{D}'$ if

$$f_n(\varphi) \to f(\varphi) \qquad \forall \varphi \in \mathcal{D}.$$

Example 1.9. Can we approximate δ_0 with functions? This may shed some light on what generalized functions look like. Let

$$u_n(x) = \begin{cases} n/2 & x \in [-1/n, 1/n] \\ 0 & \text{otherwise.} \end{cases}$$

Here, $\int u_n = 1$ for all n. If we try to take the limit in the sense of distributions, we get (in 1 dimension):

$$u_n(\varphi) = \int u_n \cdot \varphi \, dx$$

---

[2]You should think of the prime as a notion of duality of vector spaces.

$$= \frac{n}{2} \int_{-1/n}^{1/n} \varphi(x)\, dx$$

$$= \frac{1}{2} \int_{-1}^{1} \varphi(y/n)\, dx$$

$$\xrightarrow{\ n\to\infty\ } \varphi(0),$$

so $u_n(\varphi) \to \delta_0(\varphi)$. That is, $u_n \to \delta_0$.

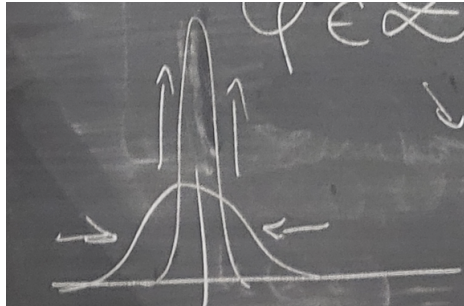**Remark 1.2.** In $\mathbb{R}^n$, we could use

$$u_\varepsilon = \frac{1}{\varepsilon^n c_n} \mathbb{1}_{B(0,\varepsilon)}, \qquad c_n = |B(0,1)|.$$

In n dimensions, this has size $\sim 1/\varepsilon^n$.

Remark 1.3. We could also use $\mathcal{D}$ functions. If $\varphi \in \mathcal{D}$ with $\int \varphi = 1$, then we can define the rescaled function (at scale ε)

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon} \varphi(x/\varepsilon).$$

Here is the picture:

By the same argument, $\varphi_\varepsilon \to \delta_0$ in $\mathcal{D}'$.

Next time, we will see how we can think of distributions as solutions to PDEs. This will require knowing things like how to differentiate distributions.