18.152 Introduction to PDEs, Fall 2011 Professor: Jared Speck

Class Meeting # 1: Introduction to PDEs

1. What is a PDE?

We will be studying functions $u = u(x^1, x^2, \cdots, x^n)$ and their partial derivatives. Here $x^1, x^2, \cdots, x^n$ are standard Cartesian coordinates on $\mathbb{R}^n$. We sometimes use the alternate notation $u(x,y)$, $u(x,y,z)$, etc. We also write e.g. $u(r,\theta,\phi)$ for spherical coordinates on $\mathbb{R}^3$, etc. We sometimes also have a "time" coordinate t, in which case $t, x^1, \cdots, x^n$ denotes standard Cartesian coordinates on $\mathbb{R}^{1+n}$. We also use the alternate notation $x^0 \stackrel{\text{def}}{=} t$.

We use lots of different notation for partial derivatives:

$$\frac{\partial}{\partial x^i} u = u_{x^i} = \partial_i u, \qquad 1 \leq i \leq n, \tag{1.0.1a}$$

$$\frac{\partial^2 u}{\partial x^i \partial x^j} = \frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j} u = u_{x^i x^j} = \partial_i \partial_j u, \qquad 1 \leq i, j \leq n. \tag{1.0.1b}$$

If $i = j$, then we sometimes abbreviate $\partial_i \partial_j u \stackrel{\text{def}}{=} \partial_i^2 u$. If u is a function of (x,y), then we also write $u_x = \frac{\partial}{\partial x} u$, etc.

Definition 1.0.1. A PDE in a single unknown u is an equation involving u and its partial derivatives. All such equations can be written as

$$F(u, u_{x^1}, \cdots, u_{x^n}, u_{x^1 x^1}, \cdots, u_{x^{i_1} \cdots x^{i_N}}, x^1, x^2, \cdots, x^n) = 0, \qquad i_1, \cdots, i_N \in \{1, 2, \cdots, n\} \tag{1.0.2}$$

for some function F.

Here N is called the *order* of the PDE. N is the maximum number of derivatives appearing in the equation.

Example 1.0.1. $u = u(t,x)$

$$-\partial_t^2 u + (1 + \cos u)\partial_x^3 u = 0 \tag{1.0.3}$$

is a third-order nonlinear PDE.

Example 1.0.2. $u = u(t,x)$

$$-\partial_t^2 u + 2\partial_x^2 u + u = t \tag{1.0.4}$$

is a second-order linear PDE.

We say that (1.0.4) is a *constant coefficient linear PDE* because u and its derivatives appear linearly (i.e. first power only) and are multiplied only by constants.

Example 1.0.3. $u = u(t,x)$

$$\partial_t u + 2(1+x^2)\partial_x^3 u + u = t \tag{1.0.5}$$

is a third-order linear PDE.

We say that (1.0.5) is a *variable coefficient linear PDE* because u and its derivatives appear linearly (i.e. first power only) and are multiplied only by functions of the coordinates (t,x).

Example 1.0.4. $u = u(t,x), v = v(t,x)$

$$\partial_t u + 2x\partial_x v = \sin(x^2), \tag{1.0.6a}$$

$$\partial_t v - x^2\partial_x u - 0 \tag{1.0.6b}$$

is a *system of PDEs* in the unknowns u, v.

2. The Goals of PDE (and of this course)

Suppose that we are interested in some physical system. A very fundamental question is:

- Which PDEs are good models for the system?

A major goal of *modeling* is to answer this question. There is no general recipe for answering it! In practice, good models are often the end result of confrontations between experimental data and theory. In this course, we will discuss some important physical systems and the PDEs that are commonly used to model them.

Now let's assume that we have a PDE that we believe is a good model for our system of interest. Then most of the time, the primary goals of PDE are to answer questions such as the following:

(1) Does the PDE have any solutions? (Some PDEs have NO SOLUTIONS whatsoever!!)
(2) What kind of "data" do we need to specify in order to solve the PDE?
(3) Are the solutions corresponding to the given data unique?
(4) What are the basic qualitative properties of the solution?
(5) Does the solution contain singularities? If so, what is their nature?
(6) What happens if we slightly vary the data? Does the solution then also vary only slightly?
(7) What kinds of quantitative estimates can be derived for the solutions?
(8) How can we define the size (i.e., "the norm") of a solution in way that is *useful* for the problem at hand?

3. Physical Examples

It is difficult to exaggerate how prevalent PDEs are. We will discuss some important physically motivated examples throughout this course. Here is a first look.

- $-\partial_t^2 u + \partial_x^2 u = 0$ *wave equation*, second-order, linear, homogeneous
- $-\partial_t u + \partial_x^2 u = 0$ *heat equation*, second-order, linear, homogeneous
- $\partial_x^2 u + \partial_y^2 u + \partial_z^2 u = 0$ *Laplace's equation*, second-order, linear, homogeneous
- $\partial_x^2 u + \partial_y^2 u + \partial_z^2 u = f(x,y,z)$ *Poisson's equation* with source function f, second-order, linear, inhomogeneous (unless $f = 0$)
- $\imath\partial_t u + \partial_x^2 u = 0$ *Schrödinger's equation*, second-order, linear, homogeneous
- $u_t + u_x = 0$, *transport equation*, first-order, linear, homogeneous
- $u_t + uu_x = 0$, *Burger's equation*, first-order, nonlinear, homogeneous

$\mathbf{E} = \big(E_1(x,y,z), E_2(x,y,z), E_3(x,y,z)\big)$, $\mathbf{B} = \big(B_1(x,y,z), B_2(x,y,z), B_3(x,y,z)\big)$ are vectors in $\mathbb{R}^3$

$$\partial_t \mathbf{E} - \nabla \times \mathbf{B} = 0, \qquad \nabla \cdot \mathbf{E} = 0, \tag{3.0.7a}$$

$$\partial_t \mathbf{B} + \nabla \times \mathbf{E} = 0, \qquad \nabla \cdot \mathbf{B} = 0 \tag{3.0.7b}$$

"Maxwell's equations" in a vacuum (i.e., matter-free spacetime), first-order, linear, homogeneous.

4. Linear PDEs

Before we dive into a specific model, let's discuss a distinguished class of PDEs that are relatively easy to study. The PDEs of interest are called *linear PDEs*. Most of this course will concern linear PDEs.

Definition 4.0.2. A *linear differential operator* $\mathcal{L}$ is a differential operator such that

$$\mathcal{L}(au + bv) = a\mathcal{L}u + b\mathcal{L}v \tag{4.0.8}$$

for all constants $a, b \in \mathbb{R}$ and all functions u, v.

Remark 4.0.1. The notation was introduced out of convenience and laziness. The definition is closely connected to the superposition principle.

Example 4.0.5. $\mathcal{L} \overset{\text{def}}{=} -\partial_t^2 + (t^2 - x^2)\partial_x^2$ is a linear operator: $\mathcal{L}u = -\partial_t^2 u + (t^2 - x^2)\partial_x^2 u$

Example 4.0.6. $u = u(x,y)$, $\mathcal{L}u = \partial_x^2 u + u^2 \partial_y^2 u$ does **NOT** define a linear operator:
$\mathcal{L}(u+v) = \partial_x^2(u+v) + (u+v)^2\partial_y^2(u+v) \neq \partial_x^2 u + u^2\partial_y^2 u + \partial_x^2 v + v^2 \partial_y^2 v = \mathcal{L}u + \mathcal{L}v$

Definition 4.0.3. A PDE is *linear* if it can be written as

$$\mathcal{L}u = f(x^1, \cdots, x^n) \tag{4.0.9}$$

for some linear operator $\mathcal{L}$ and some function f of the coordinates.

Definition 4.0.4. If $f = 0$, then we say that the PDE is *homogeneous*. Otherwise, we say that it is *inhomogeneous*.

Example 4.0.7. $u = u(t,x)$

$$\partial_t u - (1 + \cos t)\partial_x^2 u = tx \tag{4.0.10}$$

is a linear PDE.

Here is an incredibly useful property of linear PDEs.

Proposition 4.0.1 (Superposition principle). *If $u_1, \cdots, u_M$ are solutions to the linear PDE*

$$\mathcal{L}u = 0, \tag{4.0.11}$$

and $c_1, \cdots, c_M \in \mathbb{R}$, then $\sum_{i=1}^M c_i u_i$ is also a solution.

Proof.

$$\mathcal{L}\sum_{i=1}^{M} c_i u_i = \sum_{i=1}^{M} c_i \overbrace{\mathcal{L}u_i}^{=0} = 0. \tag{4.0.12}$$

□

Remark 4.0.2. This shows that the set of all solutions to $\mathcal{L}u = 0$ is a *vector space* when $\mathcal{L}$ is linear.

As we will see in the next proposition, inhomogeneous and homogeneous linear PDEs are closely related.

Proposition 4.0.2 (Relationship between the inhomogeneous and homogeneous linear PDE solutions). *Let S_h be the set of all solutions to the homogeneous linear PDE*

$$\mathcal{L}u = 0, \tag{4.0.13}$$

and let u_I be a "fixed" solution to the inhomogeneous linear PDE

$$\mathcal{L}u = f(x^1, \cdots, x^n). \tag{4.0.14}$$

Then the set S_I of all solutions to (4.0.14) *is the translation of S_H by u_I : $S_I = \{u_I + u_H \mid u_H \in S_H\}$.*

Proof. Assume that $\mathcal{L}u_I = f$, and let w be any other solution to (4.0.14), i.e., $\mathcal{L}w = f$. Then $\mathcal{L}(w - u_I) = f - f = 0$, so that $w - u_I \in S_H$. Thus, $w = u_I + \underbrace{(w - u_I)}_{\text{belongs to } S_H}$, and so $w \in S_I$ by definition. On the other hand, if $w \in S_I$, then $w = u_I + u_H$ for some $u_H \in S_H$. Therefore, $\mathcal{L}w = \mathcal{L}(u_I + u_H) = \mathcal{L}u_I + \mathcal{L}u_H = f + 0 = f$. Thus, w is a solution to (4.0.14). □

5. How to solve PDEs

- **There is no general recipe that works for all PDEs!** We will develop some tools that will enable us to analyze some important classes of PDEs.
- **Usually, we don't have explicit formulas for the solutions to the PDEs we are interested in!** Instead, we are forced to understand and *estimate* the solutions *without having explicit formulas.*

The two things that you typically need to study a PDE:

- You need to know the PDE.
- You need some "data."

6. Some simple PDEs that we can easily solve

6.1. **Constant coefficient transport equations.** Consider the first-order linear transport equation

$$a\partial_x u(x, y) + b\partial_y u(x, y) = 0, \tag{6.1.1}$$

where $a, b \in \mathbb{R}$. Let's try to solve this PDE by reasoning geometrically. Geometrically, this equation says that $\nabla u \cdot v = 0$, where $\nabla u \overset{\text{def}}{=} (\partial_x u, \partial_y u)$ and v is the vector $(a, b) \in \mathbb{R}^2$. Thus, the derivative of

u in the direction (a,b) is 0, which implies that u is constant along lines pointing in the direction of (a,b). The slope of such a line is $\frac{b}{a}$. Therefore, every such line can be described as the set of solutions to $bx - ay = c$, where $c \in \mathbb{R}$. Since u is constant along these lines, we know that u is a "function that depends only on the line c." Therefore $u(x,y) = f(c) = f(bx - ay)$ for some function f.

In order to provide more details about u, we would need to prescribe some "data." For example, if it is known that $u(x,0) = x^2$, then $x^2 = f(bx)$. Thus, $f(c) = b^{-2}c^2$, and $u(x,y) = (x - b^{-1}ay)^2$. In the future, we will discuss the kinds of data that can be specified in more detail. As we will see, the type of data will depend on the type of PDE.

6.2. **Solving a variable coefficient transport equations.** With only a bit of additional effort, the procedure from Section 6.1 can be extended to cover the case where the coefficients are pre-specified functions of x, y. Let's consider the following example:

$$y\partial_x u + x\partial_y u = 0. \tag{6.2.1}$$

Let P denote a point $P = (x,y)$, and let V denote the vector $V = (y,x)$. Using vector calculus notation, (6.2.1) can be written as $\nabla u(P) \cdot V = 0$, i.e., the derivative of u at P in the direction of V is 0. Thus, equation (6.2.1) implies that u is constant along the curve $\mathcal{C}$ passing through P that points in the same direction as V. This vector can be viewed as a line segment with slope $\frac{x}{y}$. Therefore, if the curve $\mathcal{C}$ is parameterized by $x \to (x, y(x))$ (where we are viewing y as a function of x along $\mathcal{C}$) then $\mathcal{C}$ has slope $\frac{dy}{dx}$, and y is therefore a solution to the following ODE:

$$\frac{dy}{dx} = \frac{x}{y}. \tag{6.2.2}$$

We can use the following steps to integrate (6.2.2), which you might have learned in an ODE class:

$$(6.2.2) \implies y\frac{dy}{dx} = x \implies \frac{1}{2}\frac{d}{dx}(y^2) = x \tag{6.2.3}$$

$$\implies \frac{y^2}{2} = \frac{x^2}{2} + c, \ c = \text{constant}. \tag{6.2.4}$$

Thus, the curve $\mathcal{C}$ is a hyperbola of the form $\{y^2 - x^2 = c\}$. These curves are called *characteristics*. We conclude that u is constant along the hyperbolas $\{y^2 - x^2 = c\}$, which implies that $u(x,y) = f(x^2 - y^2)$ for some function $f(c)$.

We can carry out the same procedure for a PDE of the form

$$a(x,y)\partial_x u + b(x,y)\partial_y u = 0, \tag{6.2.5}$$

as long as we can figure out how to integrate the ODE

$$\frac{dy}{dx} = \frac{b(x,y)}{a(x,y)}. \tag{6.2.6}$$

7. Some basic analytical notions and tools

We now discuss a few ideas from analysis that will appear repeatedly throughout the course.

7.1. **Norms.** In PDE, there are many different ways to measure the "size" of a function f. These measures are called *norms*. Here is a simple, but useful norm that will appear throughout this course.

Definition 7.1.1 (C^k **norms**). Let f be a function defined on a domain $\Omega \subset \mathbb{R}$. Then for any integer $k \geq 0$, we define the C^k norm of f on Ω by

$$\|f\|_{C^k(\Omega)} \overset{\text{def}}{=} \sum_{a=0}^{k} \sup_{x\in\Omega} |f^{(a)}(x)|, \tag{7.1.1}$$

where $f^{(a)}(x)$ is the a^{th} order derivative of $f(x)$. We often omit the symbol Ω when $\Omega = \mathbb{R}$.

Example 7.1.1.

$$\|\sin(x)\|_{C^7(\mathbb{R})} = 8. \tag{7.1.2}$$

The same notation is used in the case that $\Omega \subset \mathbb{R}^n$, but in this case, we now sum over all partial derivatives of order $\leq k$. For example, if $\Omega \subset \mathbb{R}^2$, then $\|f\|_{C^2(\Omega)} \overset{\text{def}}{=} \sup_{(x,y)\in\Omega} |f(x,y)| + \sup_{(x,y)\in\Omega} |\partial_x f(x,y)| + \sup_{(x,y)\in\Omega} |\partial_y f(x,y)| + \sup_{(x,y)\in\Omega} |\partial_x^2 f(x,y)| + \sup_{(x,y)\in\Omega} |\partial_x\partial_y f(x,y)| + \sup_{(x,y)\in\Omega} |\partial_y^2 f(x,y)|$.

If f is a function of more than one variable, then we sometimes want to extract different information about f in one variable compared to another. For example, if $f = f(t,x)$, then we use notation such as

$$\|f\|_{C^{1,2}} \overset{\text{def}}{=} \sum_{a=0}^{1} \sup_{(t,x)\in\mathbb{R}^2} |\partial_t^a f(t,x)| + \sum_{a=1}^{2} \sup_{(t,x)\in\mathbb{R}^2} |\partial_x^a f(t,x)|. \tag{7.1.3}$$

Above, the "1" in $C^{1,2}$ refers to the t coordinate, while the "2" refers to the x coordinate.

The next definition provides a very important example of another class of norms that are prevalent in PDE theory.

Definition 7.1.2 (L^p **norms**). Let $1 \leq p < \infty$ be a number, and let f be a function defined on a domain $\Omega \subset \mathbb{R}^n$. We define the L^p norm of f by

$$\|f\|_{L^p(\Omega)} \overset{\text{def}}{=} \left(\int_\Omega |f(x)|^p \, d^n x \right)^{1/p}. \tag{7.1.4}$$

We often write just L^p instead of $L^p(\mathbb{R}^n)$.

$\|\cdot\|_{L^p(\Omega)}$ has all the properties of a *norm*:

- Non-negativity: $\|f\|_{L^p(\Omega)} \geq 0$, $\|f\|_{L^p(\Omega)} = 0 \iff f(x) = 0$ almost everywhere[1]
- Scaling: $\|\lambda f\|_{L^p(\Omega)} = |\lambda| \|f\|_{L^p(\Omega)}$
- Triangle inequality: $\|f+g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}$

Similarly, $\|\cdot\|_{C^k(\Omega)}$ also has all the properties of a norm. All of these properties are very easy to show except for the last one in the case of $\|\cdot\|_{L^p(\Omega)}$. You will study the very important case $p = 2$ in detail in your homework.

[1] "Almost everywhere" is a term that would be precisely defined in a course on measure theory.

7.2. **The divergence theorem.** A lot of PDE results are derived using integration by parts (sometimes very fancy versions of it), which provides us with *integral identities*. This will become more apparent as the course progresses. Let's recall a very important version of integration by parts from vector calculus: the divergence theorem. We first need to recall the notion of a vectorfield on $\mathbb{R}^n$.

Definition 7.2.1 (Vectorfield). Recall that a vectorfield $\mathbf{F}$ on $\Omega \subset \mathbb{R}^n$ is an $\mathbb{R}^n-$valued (i.e. vector-valued) function defined on Ω. That is,

$$\begin{gathered}\mathbf{F} : \Omega \to \mathbb{R}^n, \\ \mathbf{F}(x^1, \cdots, x^n) = \big(F^1(x^1, \cdots, x^n), \cdots, F^n(x^1, \cdots, x^n)\big),\end{gathered} \tag{7.2.1}$$

where each of the F^i are scalar-valued functions on $\mathbb{R}^n$.

We also need to recall the definition of the divergence operator, which is a differential operator that acts on vectorfields.

Definition 7.2.2 (Divergence). Recall that $\nabla \cdot \mathbf{F}$, the *divergence of* $\mathbf{F}$, is the scalar-valued function on $\mathbb{R}^n$ defined by

$$\nabla \cdot \mathbf{F} \overset{\text{def}}{=} \sum_{i=1}^n \partial_i F^i. \tag{7.2.2}$$

We are now ready to recall the divergence theorem.

Theorem 7.1 (Divergence Theorem). *Let $\Omega \subset \mathbb{R}^3$ be a domain*[2] *with a boundary that we denote by $\partial\Omega$. Then the following formula holds:*

$$\int_\Omega \nabla \cdot \mathbf{F}(x,y,z)\, dxdydz = \int_{\partial\Omega} \mathbf{F}(\sigma) \cdot \hat{\mathbf{N}}(\sigma)\, d\sigma. \tag{7.2.3}$$

Above, $\hat{\mathbf{N}}(\sigma)$ is the ***unit outward*** *normal vector to $\partial\Omega$, and $d\sigma$ is the surface measure induced on $\partial\Omega$. Recall that if $\partial\Omega \subset \mathbb{R}^3$ can locally be described as the graph of a function $\phi(x,y)$ (e.g., $\partial\Omega = \{(x,y,z) \mid z = \phi(x,y)\}$), then*

$$d\sigma = \sqrt{1 + |\nabla\phi(x,y)|^2}dxdy, \tag{7.2.4}$$

where $\nabla\phi = (\partial_x\phi, \partial_y\phi)$ is the gradient of ϕ, and $|\nabla\phi| \overset{def}{=} \sqrt{(\partial_x\phi)^2 + (\partial_y\phi)^2}$ is the Euclidean length of $\nabla\phi$.

Remark 7.2.1. The divergence theorem holds in all dimensions, not just 3. In dimension 1, the divergence theorem is

$$\int_{[a,b]} \frac{d}{dx} F(x)\, dx = F(b) - F(a), \tag{7.2.5}$$

which is just the Fundamental Theorem of Calculus.

[2] Throughout this course, a *domain* is defined to be an open, connected subset of $\mathbb{R}^n$.