MATH 18.152 COURSE NOTES - CLASS MEETING # 15

18.152 Introduction to PDEs, Fall 2011 — Professor: Jared Speck

Class Meeting # 15: Classification of second order equations

1. Review of Three Important Examples of PDEs

Let's review some basic facts concerning the three PDEs we've examined in detail thus far.

Equation	Type	Well-posed problems	Features
$\Delta u(x) = f(x)$	Elliptic	Boundary value problems: All of $\mathbb{R}^n$ (with boundary conditions at ∞); finite boundaries under Dirichlet, Neumann, Robin, or Mixed boundary conditions	mean value properties; maximum principle; Harnack inequality
$\partial_t u(t,x) - \Delta u(t,x) = f(t,x)$	Diffusive (parabolic):	Initial value (Cauchy) problems: all of $\mathbb{R}^n$ at $t = 0$; Initial + boundary value problems: data at $t = 0$ + Dirichlet, Neumann, Robin, or Mixed boundary conditions	Infinite speeds of propagation; smoothing properties; maximum principle, $t^{-n/2}$ decay as $t \to \infty$ for the global Cauchy problem
$-\partial_t^2 u(t,x) + \Delta u(t,x) = f(t,x)$	Hyperbolic	Initial value (Cauchy) problems: all of $\mathbb{R}^n$ at $t = 0$; Initial + boundary value problems: data at $t = 0$ + Dirichlet, Neumann, Robin, or Mixed boundary conditions	Finite speed of propagation; domain of dependence and influence; energy identities; order $t^{(1-n)/2}$ decay as $t \to \infty$ for the global Cauchy problem

2. Motivating example

Let's consider the following second-order linear PDE on $\mathbb{R}^{1+n}$:

$$\mathcal{L}u \overset{\text{def}}{=} A^{\alpha\beta}\partial_\alpha\partial_\beta u + B^\alpha \partial_\alpha u + Cu = 0. \tag{2.0.1}$$

In (2.0.1), A, B, C are allowed to be functions of the coordinates $(x^0, \cdots, x^n)$. We will also use the standard notation $x^0 = t$. By the symmetry of the mixed partial derivatives, we can also assume that A is symmetric:

$$A^{\mu\nu} = A^{\nu\mu}. \tag{2.0.2}$$

The question we would like to address at the moment is the following: *what are the basic properties of solutions to* (2.0.1)? Is this equation most like a Laplace, heat, or wave equation? That is, is (2.0.1) elliptic, diffusive, or hyperbolic? As we will see, the most important part of equation (2.0.1) in this context is the *principal part* $A^{\alpha\beta}\partial_\alpha\partial_\beta u$, which involves the top-order derivatives.

To begin answering this question, let's start with a simple example on $\mathbb{R}^2$. Let's try to classify the following equation:

$$\mathcal{L}u \overset{\text{def}}{=} \partial_t^2 u - 4\partial_t\partial_x u + 2\partial_x^2 u = 0. \tag{2.0.3}$$

Note that it would be easy to answer our question if we were able to make a linear change of variables that eliminates the cross term $-4\partial_t\partial_x u$; the PDE would then look just like one of the other ones we have already studied. More precisely, let's try to eliminate the cross terms by making good choices for the constants a, b, c, d in the following linear change variables:

$$\widetilde{t} = at + bx, \tag{2.0.4a}$$

$$\widetilde{x} = ct + dx. \tag{2.0.4b}$$

In order to have a viable change of variables, we also need to achieve the following non-degeneracy condition from linear algebra:

$$ad - bc \neq 0. \tag{2.0.5}$$

(2.0.5) states the determinant of the above linear transformation is non-zero, and that the transformation is non-degenerate.

Then using the chain rule, we have that

$$\partial_t = \frac{\partial\widetilde{t}}{\partial t}\partial_{\widetilde{t}} + \frac{\partial\widetilde{x}}{\partial t}\partial_{\widetilde{x}} = a\partial_{\widetilde{t}} + c\partial_{\widetilde{x}}, \tag{2.0.6a}$$

$$\partial_x = \frac{\partial\widetilde{t}}{\partial x}\partial_{\widetilde{t}} + \frac{\partial\widetilde{x}}{\partial x}\partial_{\widetilde{x}} = b\partial_{\widetilde{t}} + d\partial_{\widetilde{x}}. \tag{2.0.6b}$$

Inserting (2.0.6a) - (2.0.6b) into (2.0.3), we compute that

$$\mathcal{L}u = (a^2 - 4ab + 2b^2)\partial_{\widetilde{t}}^2 u + (2ac + 4bd - 4ad - 4bc)\partial_{\widetilde{t}}\partial_{\widetilde{x}} u + (c^2 - 4cd + 2d^2)\partial_{\widetilde{x}}^2 u. \tag{2.0.7}$$

To make the cross term in (2.0.7) vanish, we now choose

$$a = 1, \qquad b = 0, \qquad c = 2, \qquad d = 1. \tag{2.0.8}$$

Note that (2.0.8) also verifies the non-degeneracy condition (2.0.5). We remark that other choices would also have worked. In the new coordinates, we have that

$$\mathcal{L}u = \partial_{\widetilde{t}}^2 u - 2\partial_{\widetilde{x}}^2 u. \tag{2.0.9}$$

Dividing by -2, we see that the PDE (2.0.3) was actually a "standard" linear wave equation in disguise:

$$-\frac{1}{2}\partial_{\widetilde{t}}^2 u + \partial_{\widetilde{x}}^2 = 0. \tag{2.0.10}$$

Relative to the coordinates $(\widetilde{t}, \widetilde{x})$, the "speed" associated to the wave equation (2.0.10) is $\sqrt{2}$.

Let's do another example. Consider the PDE

$$\mathcal{L}u \overset{\text{def}}{=} -2\partial_t^2 u - 2\partial_t\partial_x u - \partial_x^2 u + \partial_x u = 0. \tag{2.0.11}$$

Using (2.0.6a) - (2.0.6b) again, we compute that

$$\begin{aligned}\mathcal{L}u = {} & (-2a^2 - 2ab - b^2)\partial_{\widetilde{t}}^2 u + (-2ac - bd - 2ad - 2bc)\partial_{\widetilde{t}}\partial_{\widetilde{x}} u + (-2c^2 - 4cd - d^2)\partial_{\widetilde{x}}^2 u \\ & + b\partial_{\widetilde{t}} u + d\partial_{\widetilde{x}} u.\end{aligned} \tag{2.0.12}$$

Choosing

$$a = \frac{1}{\sqrt{2}}, \qquad b = 0, \qquad c = -1, \qquad d = 1, \tag{2.0.13}$$

we see that

$$\mathcal{L}u = -\partial_{\widetilde{t}}^2 u - \partial_{\widetilde{x}}^2 u + \partial_{\widetilde{x}} u. \tag{2.0.14}$$

Thus, multiplying by -1, we see that (2.0.11) is really just a Laplace-like equation in disguise:

$$\partial_{\widetilde{t}}^2 u + \partial_{\widetilde{x}}^2 u - \partial_{\widetilde{x}} u = 0. \tag{2.0.15}$$

Equation (2.0.11) is therefore elliptic. We remark that the first-order term in (2.0.15) does not affect the elliptic nature of the system.

Let's do one final example. Consider the PDE

$$\mathcal{L}u \overset{\text{def}}{=} \partial_t^2 u - 2\partial_t\partial_x u + \partial_x^2 u + \partial_x u = 0. \tag{2.0.16}$$

Using (2.0.6a) - (2.0.6b) again, we compute that

$$\begin{aligned}\mathcal{L}u = {} & (a^2 - 2ab + b^2)\partial_{\widetilde{t}}^2 u + (2ac + 2bd - 2ad - 2bc)\partial_{\widetilde{t}}\partial_{\widetilde{x}} u + (c^2 - 2cd + d^2)\partial_{\widetilde{x}}^2 u \\ & + b\partial_{\widetilde{t}} u + d\partial_{\widetilde{x}} u.\end{aligned} \tag{2.0.17}$$

Choosing

$$a = 1, \qquad b = 0, \qquad c = -1, \qquad d = -1, \tag{2.0.18}$$

we see that

$$\mathcal{L}u = \partial_{\widetilde{t}}^2 u - \partial_{\widetilde{x}} u. \tag{2.0.19}$$

Thus, (2.0.16) is equivalent to

$$-\partial_{\widetilde{x}} u + \partial_{\widetilde{t}}^2 u = 0. \tag{2.0.20}$$

Now observe that (2.0.20) is just the standard heat equation, with the variable $\widetilde{x}$ playing the role of "time" and $\widetilde{t}$ playing the role of "space." Equation (2.0.20) is therefore diffusive (parabolic).

3. A general framework

In this section, we will establish a general framework for classifying second order constant coefficient scalar PDEs. The framework will cover the three examples from the previous section as special cases. The proof will reveal that the classification is intimately connected to the theory of quadratic forms from linear algebra. Throughout this section, we will use the notation

$$x = (x^0, x^1, \cdots, x^n). \tag{3.0.21}$$

As above, we will investigate PDEs of the form

$$\mathcal{L}u \overset{\text{def}}{=} A^{\alpha\beta}\partial_\alpha\partial_\beta u + B^\alpha \partial_\alpha u + Cu = 0, \tag{3.0.22}$$

where $A^{\mu\nu} = A^{\nu\mu}$.

We begin by providing a simple version of Hadamard's classic definitions.

Definition 3.0.1 (Hadamard's classification of second order scalar PDEs). Equation (3.0.22) is respectively said to be *elliptic*, *hyperbolic*, or *parabolic* according to the following conditions on the $(1+n)\times(1+n)$ symmetric matrix A :

- All of the eigenvalues of A have the same sign - **elliptic**
- n of the eigenvalues of A have the same (non-zero) sign, and the remaining one has the opposite (non-zero) sign - **hyperbolic**
- n of the eigenvalues of A have the same (non-zero) sign, and the remaining one is 0 - **parabolic**

Remark 3.0.1. Many of the ideas in this section, including the definition above, can be generalized to include the case where A depends on (x), or even on the solution u itself; PDEs of the latter type are said to be *quasilinear.*

We now state and prove the main classification theorem.

Theorem 3.1 (Classification of second order constant-coefficient PDEs). *Consider the following second order **constant coefficient** PDE*

$$\mathcal{L}u(x) \overset{def}{=} A^{\alpha\beta}\partial_\alpha\partial_\beta u(x) + B^\alpha \partial_\alpha u(x) + Cu(x) = 0, \tag{3.0.23}$$

where $\partial_\alpha \overset{def}{=} \frac{\partial}{\partial x^\alpha}$. Then there exists a linear change of variables $y^\mu = M_\alpha{}^\mu x^\alpha$ such that

- *If all of the eigenvalues of $A^{\mu\nu}$ have the same (non-zero) sign, then (3.0.23) can be written as $\pm\mathcal{L}u = \Delta_y u(y) + \widetilde{B}^\alpha \frac{\partial}{\partial y^\alpha} u(y) + Cu(y) = 0$, where $\Delta_y \overset{def}{=} \sum_{\mu=0}^n \frac{\partial^2}{(\partial y^\alpha)^2}$.*
- *If n of the eigenvalues of A have the same (non-zero) sign, and the remaining one has the opposite (non-zero) sign, then (3.0.23) can be written as $\pm\mathcal{L}u = \square_y u(y) + \widetilde{B}^\alpha \frac{\partial}{\partial y^\alpha} u(y) + Cu(y) = 0$, where $\square_y \overset{def}{=} (m^{-1})^{\alpha\beta}\frac{\partial}{\partial y^\alpha}\frac{\partial}{\partial y^\beta}$ is the standard linear wave operator, and $(m)^{-1} = diag(-1, 1, 1, \cdots, 1)$ is the standard Minkowskian matrix.*

- *If n eigenvalues $\lambda^{(1)}, \cdots, \lambda^{(n)}$ of A have the same (non-zero) sign, and the remaining one is $\lambda^{(0)} = 0$, then (3.0.23) can be written as $\pm\mathcal{L}u = \widetilde{B}^0 \frac{\partial}{\partial y^0} u(y^0, y^1, \cdots, y^n) + \sum_{i=1}^n \frac{\partial^2}{(\partial y^i)^2} u(y^0, y^1, \cdots, y^n) + \sum_{i=1}^n \widetilde{B}^i \frac{\partial}{\partial y^i} u(y^0, y^1, \cdots, y^n) + Cy = 0$. Furthermore, let $v^{(0)}, v^{(1)}, \cdots, v^{(n)}$ be a corresponding diagonalizing unit-length co-vector basis. More precisely, this means that $\sum_{\alpha=0}^n |v_\alpha^{(\mu)}|^2 = 1$ for $0 \leq \mu \leq n$, that $A^{\alpha\beta} v_\alpha^{(\mu)} v_\beta^{(\nu)} = \lambda^{(\mu)}$ if $\mu = \nu$, and that $A^{\alpha\beta} v_\alpha^{(\mu)} v_\beta^{(\nu)} = 0$ if $\mu \neq \nu$ (standard linear algebraic theory guarantees the existence of such a basis). Then if the non-zero vector B satisfies $B^\alpha v_\alpha^{(0)} \neq 0$, we also have that $\widetilde{B}^0 \neq 0$.*

Remark 3.0.2. The "$\pm$" sign above distinguishes whether or not most of the eigenvalue of $A^{\mu\nu}$ are positive or negative. For example, if all of the eigenvalues of $A^{\mu\nu}$ are positive, then $\mathcal{L}u = \Delta_y u(y) + \cdots$, while if they are all negative, then $\mathcal{L}u = -\Delta_y u(y) + \cdots$ (and similarly for the other two cases).

Proof. Let's consider the first case, in which all of the eigenvalues have the same (non-zero) sign. Then by standard linear algebra, since $A^{\mu\nu}$ is symmetric and positive definite (perhaps after multiplying it by -1), there exists an invertible "change-of-basis" matrix $M_\mu{}^\nu$ such that

$$M_\alpha{}^\mu A^{\alpha\beta} M_\beta{}^\nu = I^{\mu\nu}, \tag{3.0.24}$$

where $I^{\mu\nu} \overset{\text{def}}{=} \text{diag}(1, 1, \cdots, 1)$ is the $(n+1) \times (n+1)$ identity matrix. In fact, we can choose

$$M_\alpha{}^\mu = \frac{1}{\sqrt{|\lambda^{(\mu)}|}} v_\alpha^{(\mu)} \text{ (no summation in } \mu), \tag{3.0.25}$$

where $\lambda^{(\mu)}$ is the "eigenvalue" of A corresponding to the unit-length covector $v_\alpha^{(\mu)}$ (i.e., $\sum_{\alpha=0}^n |v_\alpha^{(\mu)}|^2 = 1$) appearing in the statement of the theorem.

We now make the linear change of variables $y^\mu = M_\alpha{}^\mu x^\alpha$. Then by the chain rule, $\frac{\partial}{\partial x^\alpha} = \frac{\partial y^\mu}{\partial x^\alpha} \frac{\partial}{\partial y^\mu} = M_\alpha{}^\mu \frac{\partial}{\partial y^\mu}$. Therefore,

$$A^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \frac{\partial}{\partial x^\beta} u = A^{\alpha\beta} M_\alpha{}^\mu M_\beta{}^\nu \frac{\partial}{\partial y^\mu} \frac{\partial}{\partial y^\nu} u = I^{\mu\nu} \frac{\partial}{\partial y^\mu} \frac{\partial}{\partial y^\nu} u = \Delta_y u. \tag{3.0.26}$$

This completes the proof in the first case.

In the second case, in which n of the eigenvalues of A have the same (non-zero) sign, and the remaining one has the opposite (non-zero) sign, the proof is similar. The key difference is that because of the eigenvalue of opposite sign, (3.0.24) is replaced with

$$M_\alpha{}^\mu A^{\alpha\beta} M_\beta{}^\nu = (m^{-1})^{\mu\nu}, \tag{3.0.27}$$

where $(m^{-1})^{\mu\nu} \overset{\text{def}}{=} \text{diag}(-1, 1, 1, \cdots, 1)$ is the standard $(1+n) \times (1+n)$ Minkowski matrix. Therefore,

$$A^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \frac{\partial}{\partial x^\beta} u = A^{\alpha\beta} M_\alpha{}^\mu M_\beta{}^\nu \frac{\partial}{\partial y^\mu} \frac{\partial}{\partial y^\nu} u = (m^{-1})^{\mu\nu} \frac{\partial}{\partial y^\mu} \frac{\partial}{\partial y^\nu} u = \square_y u. \tag{3.0.28}$$

This completes the proof in the second case.

In the third case, in which n of the eigenvalues of A have the same (non-zero) sign, and the remaining one is 0, the proof is similar. The key difference is that because of the zero eigenvalue, (3.0.24) is replaced with

$$M_{\alpha}{}^{\mu} A^{\alpha\beta} M_{\beta}{}^{\nu} = D^{\mu\nu}, \tag{3.0.29}$$

where $D^{\mu\nu} \stackrel{\text{def}}{=} \text{diag}(0, 1, 1, \cdots, 1)$.

Therefore,

$$A^{\alpha\beta} \frac{\partial}{\partial x^{\alpha}} \frac{\partial}{\partial x^{\beta}} u = A^{\alpha\beta} M_{\alpha}{}^{\mu} M_{\beta}{}^{\nu} \frac{\partial}{\partial y^{\mu}} \frac{\partial}{\partial y^{\nu}} u = D^{\mu\nu} \frac{\partial}{\partial y^{\mu}} \frac{\partial}{\partial y^{\nu}} u = \sum_{i=1}^{n} \frac{\partial^2}{(\partial y^i)^2} u. \tag{3.0.30}$$

Furthermore, we have that

$$B^{\alpha} \frac{\partial}{\partial x^{\alpha}} u = M_{\alpha}{}^{\mu} B^{\alpha} \frac{\partial}{\partial y^{\mu}} u. \tag{3.0.31}$$

Thus, using using (3.0.25), we have that

$$\widetilde{B}^0 \stackrel{\text{def}}{=} M_{\alpha}{}^{0} B^{\alpha} = v_{\alpha}^{(0)} B^{\alpha} \neq 0. \tag{3.0.32}$$

$\square$

Example 3.0.1. In the first example from above,

$$A^{\mu\nu} = \begin{bmatrix} 1 & -2 \\ -2 & 2 \end{bmatrix}. \tag{3.0.33}$$

To calculate the eigenvalues of A, we first set

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & -2 \\ -2 & 2-\lambda \end{bmatrix} = \lambda^2 - 3\lambda - 2 = 0. \tag{3.0.34}$$

The solutions are

$$\lambda = \frac{3 \pm \sqrt{17}}{2}. \tag{3.0.35}$$

Since the eigenvalues are of opposite sign, the corresponding PDE is hyperbolic.

Example 3.0.2. In the second example from above,

$$A^{\mu\nu} = \begin{bmatrix} -2 & -1 \\ -1 & -1 \end{bmatrix}. \tag{3.0.36}$$

To calculate the eigenvalues of A, we first set

$$\det(A - \lambda I) = \det \begin{bmatrix} -2-\lambda & -1 \\ -1 & -1-\lambda \end{bmatrix} = \lambda^2 + 3\lambda + 1 = 0. \tag{3.0.37}$$

The solutions are

$$\lambda = \frac{-3 \pm \sqrt{5}}{2}. \tag{3.0.38}$$

Both of these eigenvalues are negative, and thus the corresponding PDE is elliptic.

Example 3.0.3. In the final example from above,

$$A^{\mu\nu} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \tag{3.0.39}$$

To calculate the eigenvalues of A, we first set

$$\det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{bmatrix} = \lambda^2 + 2\lambda = 0. \tag{3.0.40}$$

The solutions are

$$\lambda = 0, -2, \tag{3.0.41}$$

and so the corresponding PDE is parabolic.

18.152 Introduction to Partial Differential Equations.
Fall 2011