SOME REMARKS ON CHARACTERISTICS FOR LINEAR PDE

PLAMEN STEFANOV

1. Symbol and principal symbol

Let

$$P = \sum_{|\alpha|\le k} b_\alpha(x)\partial^\alpha \tag{1}$$

be a linear partial differential operator. The associated linear PDE is

$$Pu = f, \tag{2}$$

where f is given. The *symbol* $P(x,\xi)$ of P is a polynomial of ξ with coefficients depending on x. The modern way to define the symbol is through the correspondence

$$\xi \to -i\,\partial,$$

i.e., replace ∂ by $i\xi$. Then the symbol of P would be $P = \sum_{|\alpha|\le k} b_\alpha(x)(i\xi)^\alpha$. Very often, one writes

$$P = \sum_{|\alpha|\le k} a_\alpha(x)D^\alpha, \quad D := -i\,\partial, \tag{3}$$

and defines the symbol

$$P(x,\xi) = \sum_{|\alpha|\le k} a_\alpha(x)\xi^\alpha.$$

Clearly, $a_\alpha = i^{|\alpha|}b_\alpha$.

The *principal symbol* $P_0(x,\xi)$ of P is defined by

$$P_0(x,\xi) = \sum_{|\alpha|= k} a_\alpha(x)\xi^\alpha.$$

One can also use the correspondence

$$\xi \to \partial,$$

to define the *characteristic form* L by

$$L(x,\xi) = \sum_{|\alpha|= k} b_\alpha(x)\xi^\alpha.$$

Clearly, $P_0 = i^k L$. For the purpose of finding the characteristics, etc., we can work with L instead of P_0. The need for $i = \sqrt{-1}$ in the definition of the symbol becomes clear when we study Fourier Transform methods.

Example 1. $P = a(x)\Delta_x + b(x)\partial/\partial x_1 + c(x)$ has symbol $P(x,\xi) = -|\xi|^2 + b(x)i\xi_1 + c(x)$, the principal symbol is $P_0(x,\xi) = -|\xi|^2$, and $L = |\xi|^2$.

Definition 1. *P is called elliptic (in U), if $P_0(x,\xi)\neq 0$ for $\xi\neq 0$ (and any $x\in U$).*

Date: September 26, 2008.

Note that P does not need to be of order 2, and that the coefficients do not need to be real. If P is a matrix-valued operator (then (3) is a system), the condition is that $\det P_0(x,\xi) \neq 0$ for $\xi \neq 0$. Examples of elliptic operators are Δ, the operator in Example 1 above if $a \neq 0$, the d-bar operator $\bar\partial := \partial_x + i\partial_y$, the elasticity (matrix-valued) operator $\mu\Delta + (\lambda+\mu)\nabla\nabla\cdot$, where $\mu > 0$, $\lambda + 2\mu > 0$. The wave operator $\partial_t^2 - \Delta$ and the heat operator $\partial_t - \Delta$ are not elliptic.

2. Characteristic (co)vectors

Definition 2. *The characteristic variety Σ of P is the set of the zeros of its principal symbol, excluding $\xi = 0$, i.e.,*

$$\Sigma = \{(x,\xi);\ \xi \neq 0,\ P_0(x,\xi) = 0\}.$$

If P is a matrix-valued operator, then we replace P_0 above by $\det P_0(x,\xi)$. Sometimes one does not exclude $\xi = 0$ from Σ. Note that $(x,0)$ is always a zero of P_0. So P is elliptic, if its characteristic variety is empty.

One can see that (x,ξ) can be considered as a covector but we do not want to emphasize on this now.

Definition 3. (x,ξ) *is called a characteristic (co)vector if $(x,\xi) \in \Sigma$.*

We think of ξ as a (co)vector with base point at x.

3. Characteristic surfaces

By surface in $\mathbf{R}^n$ (sometimes called hypersurface), we mean a a set of points that locally can be given by an equation $F(x) = 0$, with $\nabla F \neq 0$. Then its dimension is $n-1$. A more modern understanding of what a surface is can be given using the language of differential geometry (manifolds, etc.) but we will avoid this at the moment. Note that in $\mathbf{R}^2$, surfaces are curves (and in $\mathbf{R}$ are points)!

Definition 4.

(a) The surface Γ is called characteristic for P at the point x of its normal vector ν at x is characteristic.

(b) Γ is called non-characteristic, if none of its points is characteristic.

In other words, the non-characteristic condition at $x \in \Gamma$ is

$$P_0(x,\nu) \neq 0.$$

It can also be written in the form

$$\sum_{|\alpha|=k} a_\alpha(x)\nu^\alpha \neq 0,$$

see also Evans' book, p. 225.

The Cauchy problem for a non-characteristic surfaces is locally solvable at least when "everything is analytic" (the Cauchy-Kovalevskaya theorem). Under the non-characteristic condition, we can always find all derivatives of the solution on Γ, if a solution exists. If (3) is linear, the solution is unique (locally) by the Holmgren's theorem. The Cauchy problem for elliptic equation is ill-posed (Hadamard's example).

If the analyticity condition in the CK theorem is violated, then there might be no solution. For example, the Laplace equation $\Delta u = 0$ with Cauchy data $u|_{x_n=0}$, $u_{x_n}|_{x_n=0} = g(x')$ has no solution if g is not analytic! See John's book, p. 98.

Suppose that we want to find $\phi(x)$ with non-zero gradient so that locally $\Gamma = \{\phi = \text{const.}\}$. Then $\nabla\phi$ is normal to Γ (basic calculus), so we get

$$P_0(x,\nabla\phi) = 0. \tag{4}$$

This is a version of the Hamilton-Jacobi equation. If P is elliptic, it has no solutions (with non-zero gradients).

Example 2. The wave operator is given by $P = \partial_t^2 - c^2\Delta$, $x \in \mathbf{R}^n$. The principal symbol, also equal to the symbol, is $P_0(\xi,\tau) = -\tau^2 + c^2|\xi|^2$. Here we think of t as x_{n+1} and we replace ξ_{n+1} by τ. The characteristic form is $L = \tau^2 - c^2|\xi|^2$. The characteristic vectors look like this

$$(\omega, c), \quad (\omega, -c), \quad \text{for any unit vector } \omega.$$

There are many characteristic surfaces with different shapes if $n \geq 2$, for example the planes $x\cdot\omega \pm ct =$ const., the cones $|x - x_0| = c(t - t_0)$, etc. If $n = 1$, then the only characteristic curves (they are curves now, we have two variables only: x and t) are the families of lines $x \pm ct =$ const. Note that they are no parallel to the characteristic vectors (unless $c = 1$), they are orthogonal to them! In that particular case, those lines are called characteristic lines.

4. Characteristic curves

O.K., what are characteristics (characteristic curves) and how to find them?

4.1. **Characteristics in two dimensions.** If $n = 2$, then the characteristic surfaces above are actually curves, called characteristics (characteristic curves). How to find them? Well, just solve (4). The solution(s) that you get give you the curves in the form $\phi(x,y) = 0$ (actually, you will get $\phi(x,y) =$ const.).

Let us see how this works for a 2nd order linear PDE of the form

$$au_{xx} + bu_{xy} + cu_{yy} + \cdots = 0, \tag{5}$$

where the dots represent linear terms of lower order. Note that a, b, c may depend on x, y. Then (4) reduces to

$$a\phi_x^2 + b\phi_x\phi_y + \phi_y^2 = 0. \tag{6}$$

Note that the dots in (5) do not matter for the equation above. If $b^2 - 4ac > 0$ we say that (5) is hyperbolic (if < 0, it is elliptic). Assume from now on, that it is hyperbolic. There are two (equivalent, of course) ways to solve it. First, treat it as a quadratic equation for ϕ_x/ϕ_y to get the following two 1st order PDEs

$$\phi_x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\phi_y. \tag{7}$$

Then find the general solutions, and you are done.

Another way is to do the following. On the curve $\phi = 0$, we have

$$\phi_x \mathrm{d}x + \phi_y \mathrm{d}y = 0.$$

Therefore, $(\mathrm{d}x, \mathrm{d}y)$ is normal to $\nabla\phi = (\phi_x, \phi_y)$, thus $\alpha(\mathrm{d}x, \mathrm{d}y) = (-\phi_y, \phi_x)$ for some smooth non-vanishing function α. Plug $\phi_x = \alpha\mathrm{d}y$, $\phi_y = -\alpha\mathrm{d}x$ in (7), cancel α^2 to get

$$a\,\mathrm{d}y^2 - b\,\mathrm{d}x\,\mathrm{d}y + c\,\mathrm{d}x^2 = 0$$

along the curve. Assume that at least locally, $\phi = 0$ is a graph of a function $y = y(x)$. Then divide formally by $\mathrm{d}x^2$ to get

$$a\Big(\frac{\mathrm{d}y}{\mathrm{d}x}\Big)^2 - b\frac{\mathrm{d}y}{\mathrm{d}x} + c = 0.$$

Note the negative sign in front of b! This is a quadratic equation for $\mathrm{d}y/\mathrm{d}x$, solve it to get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Solve those two equations to get

$$y_{1,2}(x) = \int \frac{b \pm \sqrt{b^2 - 4ac}}{2a}\,\mathrm{d}x + C_{1,2}. \tag{8}$$

So we get two families of characteristic curves.

4.2. **Normal form in two dimensions.** How to find a change of variables that puts (5) into its normal form $u_{\xi\eta} + \cdots = 0$? We still assume that (5) is hyperbolic. This is explained in Evans' book, see 7.2.5 there. The bottom line is the following. We are looking for a change

$$\xi = \phi_1(x, y), \quad \eta = \phi_2(x, y). \tag{9}$$

Then $\phi_{1,2}$ solve characteristic (OK, this term is overused) equation (6)! So we just need to solve (7). The geometric meaning of this is the following: the characteristics are the new coordinate axes $\xi =$ const. or $\eta =$ const. Here is why: the characteristics are given by $\phi = 0$, and also by $\phi =$ const., because adding a constant to ϕ does not change $\nabla\phi$, therefore, it is still as solution to (6). Here $\phi = \phi_1$ or $\phi = \phi_2$. By (9) those are the curves $\xi =$ const. or $\eta =$ const. Now, can we use the second approach that leads to (8) to find that change? Yes, but after you get $y_{1,2}$ you formally replace C_1 by ϕ_1 and C_2 by ϕ_2.

Example 3 (taken from McOwen's book). Consider the PDE

$$xu_{xx} + 2x^2u_{xy} - u_x + 1 = 0.$$

Clearly, it is hyperbolic for $x \neq 0$. One way to find a change of variables that reduces it to its normal form is to solve the characteristic equation (6) by solving (7). So we get

$$\phi_x = \frac{-2x^2 \pm \sqrt{4x^4}}{2x}\phi_y$$

that reduces to the following two equations (it is not a system)

$$\phi_x + 2x\phi_y = 0, \quad \phi_x = 0.$$

Find the general solutions (there are no boundary conditions), and that will give you (9). The problem here is that there are too many solutions, and we only need a one-parameter family for each one, so that the Jacobian of (9) to be non-zero (the latter follows from the hyperbolicity assumption, actually). For example, the general solution of the second equation is $\phi = h(y)$ for any function h. We will see in a moment that $h(y) = y$ only is enough. Why is this happening? Because there are many changes of variables that reduce the equation to the form $\alpha(\xi, \eta)u_{\xi,\eta} + \cdots = 0$ with a non-vanishing α so that we can divide by it. We just need one of them but the approach above gives us all of them.

It is simpler to use the second approach that leads to (8). We get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2x^2 \pm \sqrt{4x^4}}{2x} = \begin{cases} 2x \\ 0 \end{cases}$$

So,

$$y = x^2 + C_1 \quad \text{or} \quad y = C_2.$$

We replace the constants by ξ and η, respectively (actually, we replace C_1 by $-C_1$ first to get the same formula as in McOwen's book), to get

$$\xi = x^2 - y, \quad \eta = y. \tag{10}$$

This is the change of variable that we need. Now, make that change to get

$$u_{\xi\eta} = -\frac{1}{4}(\xi + \eta)^{-3/2}.$$

This is the normal form we have been looking for. By the way, one can solve this equation, fix ξ first, integrate, etc. This gives us the solution $u = (\xi + \eta)^{1/2} + F(\xi) + G(\eta)$, or, in the original variables,

$$u(x, y) = x + F(x^2 - y) + G(y).$$

Actually, we were a bit lucky, that change is global, and the characteristics are graphs of functions, indeed. On the other hand, locally, near any fixed point, either $y = y(x)$ works, or $x = x(y)$ does.

4.3. **Characteristics in dimensions 3 and higher.** If $n \geq 3$, then the characteristic surfaces have dimension at least 2, so they are not curves. Does the notion of characteristics make sense then? Is it useful for something?

The answer is affirmative. Characteristics in all dimensions $n \geq 2$ (including the case $n = 2$ that we studied above) are defined as follows. Consider the Hamiltonian $H(x, \xi) = P_0(x, \xi)$ (just a new fancy name for the principal symbol that actually has deep connections to classical mechanics and other fields). Solve the Hamiltonian system

$$\dot{x} = \partial H/\partial \xi, \quad \dot{\xi} = -\partial H/\partial x.$$

It is easy to see that H is constant along the solutions. We restrict ourselves to the solutions living on the "energy level" $H = 0$, assuming that P is not elliptic. The solutions $(x(t), \xi(t))$ are called *(zero) bicharacteristics*. One of the fundamental results in the theory of linear PDEs is Hörmander's theorem: the singularities of the solution of $Pu = f$ propagate along bicharacteristics. The precise formulation of this theorem requires a definition of singularities as points and directions (x, ξ) that leads to the notion of a wave fron set. The curves $x = x(t)$ are called characteristics.

It is easy to see that at any x, the (co)vector ξ is always normal to any characteristic. Indeed, $\xi \cdot \dot{x} = \xi \cdot \partial H/\partial \xi = kH = 0$ because H is homogeneous in ξ (Euler's identity). If $n = 2$, this means that the newly defined characteristics are the old ones (up to a change of the parametrization).

Example 4. Back to the wave operator $P = \partial_t^2 - c^2\Delta$, studied in Example 2. The lines $t \mapsto (x_0 + ct\omega, t)$, where $x_0 \in R^n$ and $|\omega| = 1$ are fixed, are all the characteristic curves (prove it!). If $n = 2$, they reduce to the lines $x \pm ct = \text{const}$. The projection to the x-space (in all dimensions) are the lines $x = x_0 + ct\omega$ and are often called light rays. They represent intuitively "photons" traveling along straight lines with the light speed c.