Chapter 8

Second Order PDEs: Reduction to Canonical Form

8.1 Classification of second order PDEs

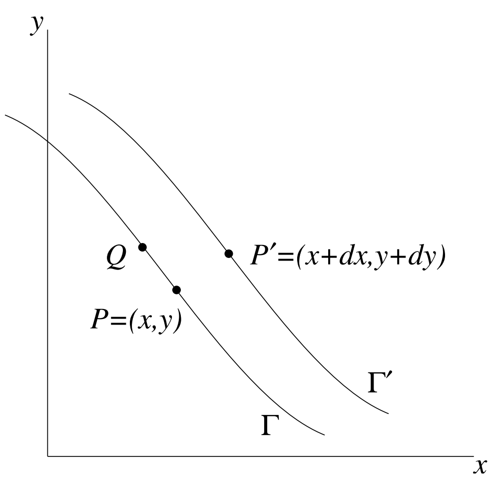
We give an intuitive description of how characteristics can be used to construct solutions of the general second order PDE in the form

$$au_{xx} + 2bu_{xy} + cu_{yy} = F, \tag{8.1}$$

where now a, b, c and F can all be functions of x, y, u, u_x and u_y. The equation is linear in the highest partial derivatives, even though it may be nonlinear in u, u_x and u_y.

We begin by assuming that we know, as initial conditions, u, u_x and u_y on some curve Γ in the xy-plane. The aim is to construct u, u_x and u_y on a neighbouring curve Γ', as illustrated in Figure 8.1. If we can do this then we can treat the values of u, u_x and u_y as new initial conditions and seek the solution on a new neighbouring curve Γ''. Continuing in this way a solution to (8.1) could be build up over a region of the xy-plane. Can we do this?

Let $P = (x, y)$ be a point on the curve Γ and suppose we want to know everything at the point $P' = (x + dx, y + dy)$ on the curve Γ'. If we know $u(x, y)$ at P, then we can calculate $u(x + dx, y + dy)$



at P' by Taylor expanding (and assuming that dx and dy are small). Thus,

$$u(x+dx, y+dy) \approx u(x,y) + u_x dx + u_y dy.$$

Note that u_x and u_y are evaluated at (x,y), which is on Γ and so they are known. So, for a given dx and dy, we can obtain u at P'.

If we next try to calculate u_x and u_y at P', using the same idea, we obtain

$$u_x(x+dx, y+dy) \approx u_x(x,y) + u_{xx}dx + u_{xy}dy \tag{8.2}$$

$$u_y(x+dx, y+dy) \approx u_y(x,y) + u_{xy}dx + u_{yy}dy, \tag{8.3}$$

where now the u_{xx} etc are again evaluated at P. However, at this point we do not know the second derivatives on Γ and so we cannot yet determine u_x and u_y at P'.

The solution is to note that we can obtain relations for u_{xx}, u_{xy} and u_{yy} that can be used to determine these quantities from the values of u, u_x and u_y on Γ. First, we know that together they satisfy the original equation (8.1). Further, they also satisfy (8.2) and (8.3) for any increment (dx, dy). In particular we can choose (dx, dy) such that the left hand side of (8.2) and (8.3) is evaluated on Γ, that is we take $(x+dx, y+dy)$ to be the point Q in the figure. Since Q lies on Γ, we know the u, u_x, and u_y there, and so we have two further relations involving the second derivatives of u. Thus, we have a total of three equations for the three unknowns u_{xx}, u_{xy} and u_{yy}, and can solve to obtain u_{xx}, u_{xy} and u_{yy} on the curve Γ. Then, with (dx, dy) chosen to go from curve Γ to curve Γ', we can calculate u_x and u_y on Γ' using (8.2) and (8.3) and the values of u_{xx}, u_{xy} and u_{yy} on Γ just obtained.

The three relations just described take the form

$$\begin{pmatrix} a & 2b & c \\ dx & dy & 0 \\ 0 & dx & dy \end{pmatrix} \begin{pmatrix} u_{xx} \\ u_{xy} \\ u_{yy} \end{pmatrix} = \begin{pmatrix} F \\ du_x \\ du_y \end{pmatrix}, \tag{8.4}$$

where $du_x = u_x(x+dx, y+dy) - u_x(x,y)$ and $du_y = u_y(x+dx, y+dy) - u_y(x,y)$. The top row of the matrix equation is (8.1), while the middle and bottom are (8.2) and (8.3), respectively. In these equations, dx and dy are chosen as the increment between points P and Q both lying on Γ, and so the left-hand side is known. Provided the equations can be solved, we obtain u_{xx}, u_{xy} and u_{yy} at the point (x,y). Once these are known, we choose dx and dy to be the increment between point P on Γ and point P' on Γ' to obtain u_x and u_y on Γ' using (8.2) and (8.3). We can then continue to build up the solution in a region of the plane.

This approach will work provided the rows on the left-hand side of (8.4) are linearly independent. The approach fails if they are *linearly dependent* and so it fails if

$$\det \begin{pmatrix} a & 2b & c \\ dx & dy & 0 \\ 0 & dx & dy \end{pmatrix} = 0. \tag{8.5}$$

Expanding the determinant, we see that the methods fails if

$$a dy^2 - 2b dy dx + c dx^2 = 0.$$

Dividing by dx^2, this becomes

$$a\left(\frac{dy}{dx}\right)^2 - 2b\frac{dy}{dx} + c = 0,$$

a quadratic equation with solutions

$$\frac{dy}{dx} = \frac{b \pm \sqrt{b^2 - ac}}{a}. \tag{8.6}$$

For use later, we can write these equations as

$$\frac{dy}{dx} = \frac{b + \sqrt{b^2 - ac}}{a} = \mu_+(x, y), \tag{8.7}$$

$$\frac{dy}{dx} = \frac{b - \sqrt{b^2 - ac}}{a} = \mu_-(x, y). \tag{8.8}$$

These differential equations are called the characteristic equations for the second order differential equation. We can think of each characteristic equation as giving rise to a family of characteristic curves. Our method fails if the curve Γ satisfies one of these equations at any point (x, y), i.e. if our initial conditions are given on a line that is somewhere tangent to a characteristic.

As with all solutions to quadratic equations, there are three cases.

1. $b^2 > ac$. In this case the characteristics are both real and (8.1) is called *hyperbolic* at the point $P = (x, y)$.
2. $b^2 = ac$. Here there is only one real characteristic and (8.1) is called *parabolic* at P.
3. $b^2 < ac$. There are no real characteristics and (8.1) is called *elliptic* at P.

These are local properties but:

- If a, b, c are constants, these properties hold everywhere in (x, y).
- If a, b, c depend on x, y, then these properties depend on (x, y). Hence, an equation can change from, say, hyperbolic to elliptic as x and y change.
- If a, b, c also depend on u, then these conditions depend on initial conditions and how the solution varies with x and y. This goes beyond the scope of this course.

8.2 Reduction to canonical form

Knowing whether the equation is of hyperbolic, parabolic or elliptic type is useful in that we can normally reduce the equation to a standard or *canonical* form. Further, to solve these equations numerically, we typically need to know the canonical form in order to choose the appropriate numerical method.

1. Hyperbolic equations are typified by the *wave equation*,

   $$u_{xx} - u_{yy} = 0, \quad a = 1, b = 0, c = -1,$$

   or

   $$u_{xy} = 0, \quad a = 0, b = 1, c = 0.$$

   The second form is usually regarded as the simpler form and gives rise to d'Alembert's solution of the wave equation.

2. Parabolic equations are typified by the *diffusion equation*, which is of the form

$$u_{xx} - u_y = 0, \quad a = 1, b = 0, c = 0.$$

3. Elliptic equations are typified by *Laplace's equation*

$$u_{xx} + u_{yy} = 0, \quad a = 1, b = 0, c = 1.$$

4. A mixed form of equation is *Tricomi's equation*,

$$yu_{xx} + u_{yy} = 0, \; a = y, b = 0, c = 1.$$

   Note that $b^2 - ac = -y$. Hence, the equation is elliptic for $y > 0$ but hyperbolic for $y < 0$.

We now consider a general second order partial differential equation of the form

$$au_{xx} + 2bu_{xy} + cu_{yy} = F. \tag{8.9}$$

We restrict attention to the case where a, b and c are functions of x and y only but allow

$$F = F\left(x, y, u, u_x, u_y\right).$$

The idea is to make a change of coordinates from (x, y) to (ξ, η) that reduces (8.9) to one of the canonical forms described above. We insist that the transformation is always locally invertible so that the Jacobian,

$$J = \frac{\partial(\xi, \eta)}{\partial(x, y)} \neq 0, \infty,$$

where J is defined by

$$J = \det \begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix}$$

To make the change of coordinates, we need to determine u_{xx} etc in terms of derivatives with respect to ξ and η. We do this using the chain rule for partial differentiation. Thus,

$$u_x = \xi_x u_\xi + \eta_x u_\eta.$$

Repeating this we have

$$u_{xx} = \xi_x \left(\xi_x u_\xi + \eta_x u_\eta\right)_\xi + \eta_x \left(\xi_x u_\xi + \eta_x u_\eta\right)_\eta.$$

Expanding the brackets and being careful with differentiating the products, we have

$$u_{xx} = \left(\xi_x\right)^2 u_{\xi\xi} + 2\xi_x\eta_x u_{\xi\eta} + \left(\eta_x\right)^2 u_{\eta\eta} + \text{ terms involving first derivatives } u_\xi, u_\eta$$

Similarly, we also determine

$$u_{yy} = \left(\xi_y\right)^2 u_{\xi\xi} + 2\xi_y\eta_y u_{\xi\eta} + \left(\eta_y\right)^2 u_{\eta\eta} + \text{ terms involving first derivatives } u_\xi, u_\eta,$$

and

$$u_{xy} = \xi_x\xi_y u_{\xi\xi} + (\xi_x\eta_y + \xi_y\eta_x)u_{\xi\eta} + \eta_x\eta_y u_{\eta\eta} + \text{ terms involving first derivatives } u_\xi, u_\eta.$$

Now substituting into Equation (8.9) and collecting all like second derivative terms together gives

$$Au_{\xi\xi} + 2Bu_{\xi\eta} + Cu_{\eta\eta} = \tilde{F}, \tag{8.10}$$

where

$$\begin{aligned}
A &= a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2,\\
B &= a\xi_x\eta_x + b(\xi_x\eta_y + \xi_y\eta_x) + c\xi_y\eta_y,\\
C &= a\eta_x^2 + 2b\eta_x\eta_y + c\eta_y^2,
\end{aligned}$$

and where

$$\tilde{F} = F + \text{ terms involving first derivatives } u_\xi, u_\eta.$$

It is *possible* to show that

$$B^2 - AC = (b^2 - ac)\left(\xi_x\eta_y - \xi_y\eta_x\right)^2 = (b^2 - ac)J^2,$$

and since the Jacobian, J, is non-zero, $B^2 - AC$ has the same sign as $b^2 - ac$. Hence, the nature of the PDE does not change when we change coordinates.

Our aim is to choose $\xi = \xi(x, y)$ and $\eta = \eta(x, y)$ so that the original PDE in x and y reduces to one of the canonical forms:

1. Hyperbolic, $u_{\xi\eta} =$ lower order derivatives.
2. Parabolic, $u_{\xi\xi} =$ lower order derivatives.
3. Elliptic, $u_{\xi\xi} + u_{\eta\eta} =$lower order derivatives.

1. Hyperbolic Equations. We try to write Equation (8.10) as $u_{\xi\eta} =$ lower order derivatives, and so we want $A = C = 0$. Thus, we need to choose $\xi(x, y)$ such that

$$A = a\xi_x^2 + 2b\xi_x\xi_y + c\xi_y^2 = 0.$$

This can be written as a quadratic for ξ_x/ξ_y

$$a\left(\frac{\xi_x}{\xi_y}\right)^2 + 2b\left(\frac{\xi_x}{\xi_y}\right) + c = 0.$$

Solving the quadratic equation gives

$$\begin{aligned}
\frac{\xi_x}{\xi_y} &= -\frac{b}{a} \pm \frac{1}{a}\sqrt{b^2 - ac}\\
&= -\mu_\pm\\
\implies \xi_x &= -\mu_\pm\xi_y,
\end{aligned}$$

where μ_+ and μ_- were defined in (8.7) and (8.8). Suppose we take the root given by μ_+. Now consider the small change in ξ, $d\xi$, where

$$d\xi = \xi_x dx + \xi_y dy = \xi_y\left(-\mu_+ dx + dy\right).$$

Then along the characteristic curves defined by $dy/dx = \mu_+$, we have that $d\xi = 0$ and hence that $\xi =$ constant.

In exactly the same way (using $C = 0$) we can show that η is also constant along the characteristic curves given by $\frac{dy}{dx} = \mu_-$. This means that we have two families of curves with

$$\begin{aligned} \xi = \text{constant} \quad \text{on} \quad \frac{dy}{dx} &= \mu_+ = \frac{b}{a} + \frac{1}{a}\sqrt{b^2 - ac}, \\ \eta = \text{constant} \quad \text{on} \quad \frac{dy}{dx} &= \mu_- = \frac{b}{a} - \frac{1}{a}\sqrt{b^2 - ac}. \end{aligned}$$

Transforming to coordinates given by the characteristic curves puts (8.10) into the form

$$2Bu_{\xi\eta} = \tilde{F}.$$

Since $B^2 > AC$ implies that $B > 0$, we can divide by $2B$ and get

$$u_{\xi\eta} = \text{lower order derivative terms.}$$

2. Parabolic Equations. To form the canonical form for a parabolic equation, we want $A = 0$ in (8.10). Since $b^2 = ac$, this means the expression for A is a perfect square, namely

$$a\left(\frac{\xi_x}{\xi_y}\right)^2 + 2b\left(\frac{\xi_x}{\xi_y}\right) + c = 0 \implies \frac{\xi_x}{\xi_y} = -\frac{b}{a}.$$

Again, $d\xi = 0$ implies that

$$\xi_x dx + \xi_y dy = 0 \implies \xi_y\left(-\frac{b}{a}dx + dy\right) = 0,$$

and so

$$\frac{dy}{dx} = \frac{b}{a}.$$

Hence, ξ is constant along the family of characteristic curves defined by $dy/dx = b/a$. This only gives one coordinate; the second coordinate η may be chosen to be any convenient $\eta(x, y)$ such that the Jacobian, J is non-zero.

Note that $B^2 = AC$ and since $A = 0$ we also have $B = 0$. Therefore the reduction to canonical form gives (8.10) as

$$Cu_{\eta\eta} = \text{ lower order derivative terms.}$$

In the case where the equation is linear in u, we get

$$u_{\eta\eta} + a_1 u_\xi + a_2 u_\eta + a_3 u = f(x, y).$$

We must have $a_1 \neq 0$ or else the equation would reduce to an ordinary differential equation in η: no derivative with respect to ξ would mean that ξ only appears as a parameter in the equation. The freedom to choose the coordinate η often means we can arrange the coefficient $a_2 = 0$. If $a_2 = a_3 = 0$, then we have the diffusion equation.

3. Elliptic Equations. For elliptic equations the aim is to choose ξ and η so that $A = C$ and $B = 0$, for which the transformed equation resembles Laplace's equation. In this case, however, there are no real characteristics.

The condition $A = C$ gives $A - C = 0$, which is

$$a\left(\xi_x^2 - \eta_x^2\right) + 2b\left(\xi_x\xi_y - \eta_x\eta_y\right) + c\left(\xi_y^2 - \eta_y^2\right) = 0.$$

The condition $B = 0$ gives

$$a\xi_x\eta_x + b\left(\xi_x\eta_y + \xi_y\eta_x\right) + c\xi_y\eta_y = 0.$$

Defining $\psi = \xi + i\eta$ and combining these two equations gives

$$a\psi_x^2 + 2b\psi_x\psi_y + c\psi_y^2 = 0.$$

As before, we can solve the quadratic to get

$$\frac{\psi_x}{\psi_y} = -\frac{b}{a} \pm \frac{1}{a}\sqrt{b^2 - ac} = -\frac{b}{a} \pm i\frac{D}{a},$$

where $D = \sqrt{ac - b^2}$. Thus,

$$a\psi_x = \psi_y(-b \pm iD).$$

Since $\psi = \xi + i\eta$, this gives

$$a(\xi_x + i\eta_x) = (\xi_y + i\eta_y)(-b \pm iD).$$

The real and imaginary parts of the equation give

$$\begin{aligned} a\xi_x &= -b\xi_y - D\eta_y, \\ a\eta_x &= -b\eta_y + D\xi_y. \end{aligned}$$

This can be rearranged, and noting that $D^2 + b^2 = ac$, gives

$$\xi_x = -D^{-1}\left(b\eta_x + c\eta_y\right), \tag{8.11}$$

$$\xi_y = D^{-1}\left(a\eta_x + b\eta_y\right). \tag{8.12}$$

These equations are known as *Beltrami's equations* and are a pair of linear, coupled first order PDEs. We will only encounter very simple cases where it is possible to spot a trivial solution.

Example 8.1 d'Alembert's solution to the wave equation.

Example 8.2 Reduce $2u_{xx} + 6u_{xy} - 8u_{yy} = 0$ to canonical form and solve.

Example 8.3 Reduce $u_{xx} - 2u_{xy} + u_{yy} + 2u_y + u = 0$ to canonical form.

Example 8.4 Reduce $u_{xx} - x^2u_{yy} = 0$ to canonical form; (*) and solve.