## 9 The Method of Characteristics

When we studied Laplace's equation $\nabla^2\phi = 0$ within a compact domain $\Omega \subset \mathbb{R}^n$, we imposed that $\phi$ obeyed one of the boundary conditions

$$\begin{aligned} \phi|_{\partial\Omega} &= f(x) \qquad \text{(Dirichlet)} \\ \mathbf{n}\cdot\nabla\phi|_{\partial\Omega} &= g(x) \quad \text{(Neumann)} \end{aligned}$$

for some specified functions $f, g : \partial\Omega \to \mathbb{C}$. We showed that there was a unique solution obeying Dirichlet boundary conditions, whereas the solution obeying Neumann conditions was unique up to the addition of a constant. On the other hand, in the case of a function $\phi : \Omega \times [0,\infty) \to \mathbb{C}$ that obeys the heat equation $\partial_t\phi = K\nabla^2\phi$ we imposed both a condition

$$\phi|_{\partial\Omega} \times (0,\infty) = f(x,t)$$

that holds on $\partial\Omega$ for all times, and also a condition

$$\phi|_{\Omega\times\{0\}} = g(x)$$

on the initial value of $\phi$ throughout $\Omega$. Finally, for $\phi : \Omega \times [0,\infty) \to \mathbb{C}$ obeying the wave equation $\partial_t^2\phi = c^2\nabla^2\phi$ we imposed the boundary condition

$$\phi_{\partial\Omega} \times (0,\infty) = f(x,t)$$

and initial conditions

$$\phi|_{\Omega\times\{0\}} = g(x) \qquad\qquad \partial_t\phi|_{\Omega\times\{0\}} = h(x)$$

on both the value and time derivative of $\phi$ at $t = 0$.

In all cases, we prescribe the value of $\phi$ or its derivatives on a co-dimension 1 surface of the domain of $\phi$ — that is, a surface where one of the coordinates (time or space) is held fixed. The choice of exactly what $\phi$ should look like on this surface (*e.g.* the functions $f, g$ and $h$ above) are known as the *Cauchy data* for the pde, and solving the pde subject to these conditions is said to be a *Cauchy problem.* According to Hadamard, the Cauchy problem is *well–posed* if

- A solution to the Cauchy problem exists
- The solution is unique
- The solution depends continuously on the auxiliary data.

The first two conditions are clear enough. We could violate the first condition if we try to impose too many conditions on our solution, overconstraining the problem, whilst the second can be violated by not restricting the solution enough. To understand the final condition properly would require us to introduce a topology on the space of functions (for example, we could use one induced by the inner product ( , ) ), but intuitively it means that a small change in the Cauchy data should lead to only a small change in the solution itself. This requirement is reasonable from the point of view of studying equations that arise in mathematical physics — since we can neither set up our apparatus nor measure our results with infinite precision, equations that can usefully model the physics had better obey this condition. (Some systems' behaviour, especially non-linear systems such as the weather, are exquisitely sensitive to the precise initial conditions. This is the arena of chaos theory.)

To gain some intuition for this final condition, let's consider a couple of examples where it is violated. Recall that evolution of some initial function via heat flow tends to smooth it: all sharp features become spread out as time progresses, and even two heat profiles that initially look very different end up looking very similar. For example, radiators and underfloor heating are both good ways to heat your room. On the other hand, suppose we specify $\phi(x,t)$ at some *late* time $t = T$ and try to evolve $\phi$ backwards in time using the heat equation, to see where our late–time profile came from. This problem will violate the final condition above, since even small flucutations in $\phi(x,T)$ will grow exponentially as time runs backwards, so that our early–time solutions will look very different.

For a second example, let $\Omega$ be the upper half plane $\{(x,y) \in \mathbb{R}^2 : y \geq 0\}$ and suppose $\phi : \Omega \to \mathbb{C}$ solves Laplace's equation $\nabla^2\phi = 0$ with boundary conditions

$$\phi(x,0) = 0 \qquad \text{and} \qquad \partial_y\phi(x,0) = g(x) \tag{9.1}$$

for some prescribed $g(x)$. Let's first take the case $g(x) = 0$ identically. Then the unique solution is $\phi = 0$ throughout $\Omega$. Now instead suppose $g(x) = \frac{\sin(Ax)}{A}$ for some constant $A \in \mathbb{R}$. Separation of variables shows that the solution in this case is

$$\phi(x,y) = \frac{\sin(Ax)\,\sinh(Ay)}{A^2} \tag{9.2}$$

and again this solution is unique. So far, all appears well, but consider taking the limit $A \to \infty$. In this limit our second choice of Cauchy data $\sin(Ax)/A \to 0$ everywhere along the $x$-axis and so becomes equal to the first. However, at $x = \pi/2A$ we have $\phi(\pi/2A, y) = \sinh(Ay)/A^2$ which for any finite $y$ grows exponentially as $A \to \infty$. Thus at large $A$ our second solution is very different from the first even for initial data that is very close. Hence the problem is ill–posed.

We'd like to understand how and where to specify our Cauchy data so as to ensure such ill–posed problems do not arise. One technique for thinking about this is known as the *method of characteristics*, which you met first in 1A Differential Equations. We'll look in some more detail at this here, beginning with the case of 1[st] order pdes with two independent variables.

## 9.1 Characteristics for first order pdes

We'll begin with the case of a 1[st] order pde. Suppose $\phi : \mathbb{R}^2 \to \mathbb{C}$ solves

$$\alpha(x,y)\,\partial_x\phi + \beta(x,y)\,\partial_y\phi = f(x,y) \tag{9.3}$$


**Figure 13**. *A parametrised curve $C \subset \mathbb{R}^2$, with its tangent vector at a point $s \in \mathbb{C}$.*

for some prescribed function $f : \mathbb{R}^2 \to \mathbb{C}$. We let $\mathbf{u}(x, y)$ be the vector field whose components are the coefficient functions in our pde, so

$$\begin{pmatrix} u_x \\ u_y \end{pmatrix} = \begin{pmatrix} \alpha(x,y) \\ \beta(x,y) \end{pmatrix} , \tag{9.4}$$

so our pde becomes just

$$\mathbf{u} \cdot \nabla\phi = f \, . \tag{9.5}$$

Thus, given an arbitrary starting point $(x_0, y_0) \in \mathbb{R}^2$, our equation governs how $\phi$ changes as we move infinitesimally along the direction in which $\mathbf{u}(x_0, y_0)$ points. The idea of the method of characteristics is to reduce the pde to an ode by first finding the behaviour of $\phi$ along a curve defined by the flow of the vector field $\mathbf{u}$.

### 9.1.1 Integral curves

In general, a (piecewise smooth) parameterised curve $C \subset \mathbb{R}^2$ can be viewed as a map $X : \mathbb{R} \to \mathbb{R}^2$ given by $X : s \mapsto (x(s), y(s))$. In this description, $s \in \mathbb{R}$ is a parameter telling us where we are along $C$ and $(x(s), y(s))$ tell us where this point of the curve sits inside $\mathbb{R}^2$ (see figure 13). The tangent vector to $C$ at $s$ is

$$\mathbf{v} = \begin{pmatrix} \dfrac{dx(s)}{ds} \\ \dfrac{dy(s)}{ds} \end{pmatrix} \tag{9.6}$$

which are just the derivatives wrt $s$ of the components of the image. If we're given a function $\phi : \mathbb{R}^2 \to \mathbb{C}$ then its *restriction to* $C$ is a function $\phi|_C : C \to \mathbb{C}$ given by $\phi|_C : s \mapsto \phi(x(s), y(s))$. The *directional derivative* of $\phi|_C$ along the curve $C$ is then

$$\frac{d\phi|_C(x(s), y(s))}{ds} = \frac{dx(s)}{ds}\, \partial_x\phi|_C + \frac{dy(s)}{ds}\, \partial_y\phi|_C \tag{9.7}$$

as follows from the chain rule.

To apply this to our pde, we need to find the curves whose tangent vector $\mathbf{v}$ is the vector of coefficients $\mathbf{u}$ defined by the pde. That is, we need to find the curves such that

$$\frac{dx(s)}{ds} = \alpha(x(s), y(s)) \qquad \text{and} \qquad \frac{dy(s)}{ds} = \beta(x(s), y(s)) \,. \tag{9.8}$$

These curves are called the *integral curves* of the vector field $\mathbf{u}$. For example, if $\mathbf{u}$ represents the velocity field of a stream, then the integral curves are the flow lines of a fluid, whilst if $\mathbf{u} = \mathbf{B}$ represents the magnetic field vector, then the integral curves are the field lines. In our pde context, these integral curves are known as the *characteristic curves* of the pde; they are integral curves specified by the equation itself. It's important to note that the integral curves are determined by the system (9.8) of $1^{\text{st}}$ order odes (in the variable $s$) and hence always exist, at least locally.

The pde only tells us how $\phi$ changes as we move along a given characteristic curve, and doesn't say anything about how $\phi$ varies from one characteristic curve to the next. Thus, if we want our problem to be well–posed, we'll need to specify Cauchy data for each characteristic curve. Furthermore, if we want to specify this data freely, then we shouldn't try to specify the value of $\phi$ at more than one point per characteristic curve. So altogether, we are free to specify Cauchy data along some other curve $B \subset \mathbb{R}^2$ that is *transverse* to all the characteristic curves, meaning that the tangent vector to $B$ is nowhere parallel to the tangent vectors $\mathbf{u}$ at the same point. Then $B$ will intersect the characteristics as shown in figure 14, and we will have a unique solution to our pde (at least locally).

It's usually convenient to use this initial data curve $B$ also to fix our parametrization of the characteristics by saying that the intersection point defines the origin $s = 0$ of the parameter measuring where we are along $C$. This condition ensures we have a unique solution to (9.8). If $B$ is parametrised by $t \in \mathbb{R}$, so that $t$ labels how far we are along $B$, then we can also label each member of our family of characteristic curves by the value of $t$ at which they intersect $B$. More specifically, the $t^{\text{th}}$ characteristic curve of $\mathbf{u}$ will be given by

$$C_t = \{(x = x(s,t), y = y(s,t)) \in \mathbb{R}^2\} \,, \tag{9.9}$$

where

$$\left.\frac{\partial x(s,t)}{\partial s}\right|_t = \alpha|_{C_t} \qquad \left.\frac{\partial y(s,t)}{\partial s}\right|_t = \beta|_{C_t} \tag{9.10}$$

subject to the condition that $(x(s,0), y(s,0))$ lies on the curve $B$. Finally, provided the Jacobian

$$J := \frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial x}{\partial t}\frac{\partial x}{\partial t} \tag{9.11}$$

is non–zero, then we can solve for $(t, s)$ in terms of $(x, y)$. Knowing $t(x, y)$ and $s(x, y)$ means that if we're given a point $(x, y) \in \mathbb{R}^2$ then we can say which curve we're on ($t$) and how far we are along that curve ($s$). Thus, if $J \neq 0$ then the family of integral curves of $\mathbf{u}$ are space–filling and non–intersecting, at least in some neighbourhood of $B$ (see figure 14).

From equation (9.7) we see that this equation is equivalent to the statement that the directional derivative of $\phi$ vanishes, and thus $\phi$ is constant, along the integral curves of the


**Figure 14**. *The curve $B$ is transverse to the family of curves.*

vector $\mathbf{u} = \binom{\alpha}{\beta}$. In the context of differential equations, these integral curves are called the *characteristic curves* of the p.d.e.. At the intersection point of any given characteristic curve with $B$, the Cauchy data fixes $\phi$ to be $h$, and since $\mathbf{u} \cdot \nabla\phi = 0$ along each integral curve, $\phi$ takes the same value $h(t)$ all along the $t^{\text{th}}$ integral curve, and so knowing where the integral curves actually are tells us what $\phi(x, y)$ is throughout (some region of) the plane.

Let's illustrate this with some examples. We start with a trivial case. Let $\phi : \mathbb{R}^2 \to \mathbb{C}$ obey

$$\partial_x \phi = 0 \tag{9.12}$$

subject to $\phi(0, y) = f(y)$ for some function $f$. Of course, we don't need any fancy method to solve this; clearly (9.12) says that $\phi(x, y)$ is a function of $y$ only and then the Cauchy data along the $y$–axis fixes $\phi(x, y) = f(y)$ throughout $\mathbb{R}^2$. It's instructive to see how the method of characteristics reproduces this result. In this case, the vector field $\mathbf{u}$ has components $(u_x, u_y) = (1, 0)$, so the integral curves are given by

$$\frac{dx}{ds} = 1 \qquad \frac{dy}{ds} = 0\,, \tag{9.13}$$

which has general solution $x = s + c$ and $y = d$ for some constants $c, d$. In this case, our Cauchy data is specified is the $y$–axis itself, which plays the role of our transverse curve $B$. We can parameterize this as $x = 0$, $y = t$ and the condition that the integral curves of $\mathbf{u}$ intersect $B$ at $s = 0$ fixes $c = 0$ and $d = t$. Thus our family of curves is defined by

$$C_t = \{(x = s, y = t)\} \subset \mathbb{R}^2 \tag{9.14}$$

so the $t^{\text{th}}$ characteristic is just a horizontal line at height $y = t$. According to the general theory, we write the differential equation (9.12) as $\partial_x \phi = \mathbf{u} \cdot \nabla\phi = 0$, or $\frac{\partial\phi}{\partial s}|_t = 0$ so that

$\phi$ is constant along each integral curve. Finally, the Cauchy data fixes $\phi(s,t) = f(t)$ on the $t^{\rm th}$ curve. Since $t = y$ and $x = s$, this is just $\phi(x,y) = f(y)$, recovering our previous solution.

For a slightly more interesting example, suppose $\phi : \mathbb{R}^2 \to \mathbb{R}$ solves

$$\mathrm{e}^x \partial_x \phi + \partial_y \phi = 0\,, \tag{9.15}$$

subject to $\phi(x,0) = \cosh x$. Let's first find the characteristics. The integral curves of the vector $(u_x, u_y) = (\mathrm{e}^x, 1)$ defined by the coefficients obey $dx/ds = \mathrm{e}^x$ and $dy/ds = 1$ and so are given by

$$\mathrm{e}^{-x} = -s + c \qquad y = s + d \tag{9.16}$$

for some constants $c, d$. In this example, the Cauchy data was specified along the $x$–axis, which we treat as a parametrised curve $B$ by setting $x = t$, $y = 0$. The condition that the characteristic curves intesect $B$ at $s = 0$ fixes the constants $c, d$ so that

$$\mathrm{e}^{-x} = -s + \mathrm{e}^{-t} \qquad y = s\,. \tag{9.17}$$

The differential equation (9.15) is $\mathbf{u} \cdot \nabla\phi = 0$, so $\phi$ is again simply constant along these characteristics, and the Cauchy data fixes $\phi$ to be $\cosh t$ on the $t^{\rm th}$ curve. Inverting the relations (9.17) to find $(s,t)$ as functions of $(x,y)$ shows that

$$s = y \qquad t = -\ln(y + \mathrm{e}^{-x}) \tag{9.18}$$

and therefore our solution is

$$\phi(x,y) = \cosh[\ln(y + \mathrm{e}^{-x})] \tag{9.19}$$

throughout $\mathbb{R}^2$. You can check by direct substitution that this does indeed solve our pde with the given boundary condition.

We can also use the method of characteristics to attack inhomogeneous problems such as

$$\partial_x \phi + 2\partial_y \phi = y\,\mathrm{e}^x \tag{9.20}$$

with $\phi = \sin x$ along the diagonal $y = x$. The $t^{\rm th}$ integral curve of $(u_x, u_y) = (1,2)$ is given by

$$C_t = \{(x = s + t,\ y = 2s + t)\} \subset \mathbb{R}^2\,, \tag{9.21}$$

where we've used the fact that the Cauchy data here is fixed along the curve $B$ given by $x = t$, $y = t$ to fix the values of $x(s)$ and $y(s)$ at $s = 0$. In this inhomogeneous example $\phi$ is no longer constant along the characteristic curves, but instead obeys

$$\mathbf{u} \cdot \nabla\phi = \left.\frac{\partial\phi}{\partial s}\right|_t = y\,\mathrm{e}^x = (2s + t)\,\mathrm{e}^{s+t} \tag{9.22}$$

with initial data that $\phi(0,t) = \sin t$ at the point $s = 0$ on the $t^{\rm th}$ curve. Solving this differential equation for $\phi$ is straightforward since $t$ is just a fixed parameter: we have

$$\phi(x,y) = (2 - t)\,\mathrm{e}^t + \sin t + (t + 2s - 2)\,\mathrm{e}^{s+t} \tag{9.23}$$

in terms of $(s, t)$. Finally, inverting the relations in (9.21) gives $s = y - x$ and $t = 2x - y$, so that our solution is

$$\phi(x, y) = (2 - 2x + y)\,\mathrm{e}^{2x-y} + \sin(2x - y) + (y - 2)\,\mathrm{e}^{x}\,, \tag{9.24}$$

in terms of the original variables.

Note the following features of the above construction:

- If any characteristic curve intersects the initial curve $B$ more than once then the problem is over–determined. In this case the value of $h(t)$ must be constrained at all such multiple intersection points or no solution will exist. For example, in the case of a homogeneous equation, we must have $h(t_1) = h(t_2)$ for any points $t_1, t_2 \in B$ that intersect the same characteristic curve.

- If the initial curve $B$ is itself a characteristic curve then either the solution either does not exist or, if it does, it will not be unique. The solution will fail to exist if the Cauchy data $h(t)$ does not vary along $B$ in the same way as the differential equation says our solution $\phi$ itself should vary as we move along this characteristic curve. If it does, then our solution is not unique because it is not determined on any other characteristic.

- If the initial curve is transverse to all characteristics and intersects them once only, then the problem is well–posed for any $h(t)$ and has a unique solution $\phi(x, y)$ (at least in a neighbourhood of $B$). Note that the initial data cannot be propagated from one characteristic to another. In particular, we see that if $h(t)$ is discontinuous, then these discontinuities will propagate along the corresponding characteristic curve.

In summary, to solve the quasi–linear[39] equation $\alpha\phi_x + \beta\phi_y = f(u, x, y)$ with $\phi|_B = h(t)$ on an initial curve $B$, we first write down the equations (9.8) which are o.d.e.s determining the characteristic curves. These are solved subject to the condition that they intersect $B$ when $s = 0$. We then algebraically invert these relations to obtain $t = t(x, y)$ and $s = s(x, y)$. Along any given characteristic curve $C_t$ the p.d.e. for $\phi$ becomes

$$\left.\frac{\partial \phi}{\partial s}\right|_t = f(\phi, x, y)|_{C_t} \tag{9.25}$$

which is just an o.d.e. in the variable $s$. We solve this o.d.e. subject to the initial condition $\phi(s = 0, t) = h(t)$, which gives $\phi$ as a function of $s$ and $t$. Finally, substituting in the relations $t = t(x, y)$ and $s = s(x, y)$ we obtain $\phi(x, y)$ for any $x, y$ in a neighbourhood of $B$.

## 9.2 Characteristics for second order pdes

New features emerge when we try to generalize the idea of characteristics to higher order pdes. The 'type' of equation we're dealing with determines what kind of Cauchy data should be imposed where in order to have a unique solution, and whether these solutions may develop singularities even starting from smooth Cauchy data.

[39] In a *quasi–linear* equation the coefficients of the leading

### 9.2.1 Classification of pdes

In this section we'll give a rough classification of second order pdes in a way that helps identify equations with similar properties. To start, suppose $\phi : \mathbb{R}^n \to \mathbb{C}$ and consider the general second–order linear differential operator $\mathcal{L}$ with

$$\mathcal{L}\phi := a^{ij}(x)\frac{\partial^2 \phi}{\partial x^i\,\partial x^j} + b^i(x)\frac{\partial \phi}{\partial x^i} + c(x)\phi \tag{9.26}$$

where $(x^1, x^2, \ldots, x^n)$ are coordinates on $\mathbb{R}^n$ and where the coefficient functions $a^{ij}$, $b^i$ and $c$ are real–valued. Since partial derivatives commute, we may assume $a^{ij} = a^{ji}$ without loss of generality. Introducting an auxiliary variable $k \in \mathbb{R}^n$, we define the *symbol* of $\mathcal{L}$ to be the polynomial

$$\sigma(x,k) := \sum_{i,j=1}^{n} a^{ij}(x)k_i k_j + \sum_{i=1}^{n} b^i(x)k_i + c(x) \tag{9.27}$$

in $k$. Likewise, the *principal part* of the symbol is the leading term

$$\sigma^{\mathrm{p}}(x,k) = \sum_{i,j} a^{ij}(x)k_i k_j\,. \tag{9.28}$$

Thus, for any fixed $x \in \mathbb{R}^n$, $\sigma^{\mathrm{p}}(x,k)$ defines a quadratic form in the $k_i$ variables. Note also that since $a^{ij} = a^{ji}$ this quadratic form is real and symmetric. For example, the symbol of the Laplacian $\nabla^2$ is $\sum_{i=1}^n (k_i)^2$ while the symbol of the heat operator $\partial/\partial x^0 - \nabla^2$ is $k_0 - \sum_{i=1}^n (k_i)^2$ where we treat the coordinate $x^0$ as time. The principal part of the symbol of the Laplacian is the same as the symbol itself, whilst for the heat operator the principal part of the symbol is $-\sum_i (k_i)^2$, the $k_0$ term being dropped.

We can similarly define symbols of arbitrary $p^{\mathrm{th}}$–order differential operators. The principal part of such symbols is always just the leading term, so will be a symmetric polynomial

$$\sigma^{\mathrm{p}}(x,k) = a^{i_1 i_2 \ldots i_p}(x)k_{i_1}k_{i_2}\cdots k_{i_p} \tag{9.29}$$

involving $p$ products of the $k_i$s. In particular, the principal part of the symbol of a first–order differential operator just takes the form $a^i(x)k_i$, so involves single powers of the $k_i$. The idea behind this definition is that principal part of the symbol tells us how the differential operator behaves when acting on very rapidly varying functions: for such functions we expect the higher–order derivatives to dominate over lower–order ones, or over the value of the function itself. We can also see that, if the coefficient functions are in fact constant then the symbol of $\mathcal{L}$ is essentially its Fourier transform. More precisely, if $\tilde{\phi}(k) = \int_{\mathbb{R}^n} \mathrm{e}^{-\mathrm{i}k\cdot x}\,\phi(x)\,\mathrm{d}^n x$ is the Fourier transform of $\phi(x)$ then $\sigma(\mathrm{i}k)\tilde{\phi}(k)$ is the Fourier transform of $\mathcal{L}\phi(x)$ if $\mathcal{L}$ has constant coefficients.

In the case of second–order operators, we treat $\sigma^{\mathrm{p}}(x,k)$ as a symmetric, real–valued quadratic form $\sigma^{\mathrm{p}}(x,k) = \mathbf{k}^{\mathrm{T}}\mathbf{A}\mathbf{k}$ where $\mathbf{A}$ is the matrix with entries $a^{ij}(x)$. We now classify these operators according to the eigenvalues of $\mathbf{A}$. The eigenvalues of a real, symmetric matrix are always real, and a second order differential operator of the form (9.26) is said to be

– *elliptic* if the eigenvalues of the principal part of the symbol all have the same sign,

– *hyperbolic* if all but one of the eigenvalues of the principal part of the symbol have the same sign,

– *ultrahyperbolic* if there is more than one eigenvalue with each sign, and

– *parabolic* if the quadratic form is degenerate (there is at least one zero eigenvalues).

This classification will be significant for the behaviour of solutions, especially in relation to their Cauchy data, via characteristics. Note that in general, the coefficient functions $a^{ij}(x)$ depend on the location $x \in \mathbb{R}^n$, so a single differential operator can be hyperbolic, ultrahyperbolic, parabolic or elliptic in different regions inside $\mathbb{R}^n$.

In this course, the main case of interest is that of a general, second–order linear differential operator on the plane:

$$\mathcal{L} = a(x,y)\frac{\partial^2}{\partial x^2} + 2b(x,y)\frac{\partial^2}{\partial x\,\partial y} + c(x,y)\frac{\partial^2}{\partial y^2} + d(x,y)\frac{\partial}{\partial x} + e(x,y)\frac{\partial}{\partial y} + f(x,y) \quad (9.30)$$

The principal part of the symbol of this differential operator is $\sigma^{\mathrm{p}}(x,k) = \mathbf{k}^{\mathrm{T}}\mathbf{A}\,\mathbf{k}$, where

$$\mathbf{A} = \begin{pmatrix} a(x,y) & b(x,y) \\ b(x,y) & c(x,y) \end{pmatrix} \quad (9.31)$$

is the matrix of coefficients of the second–order derivatives. The determinant of $\mathbf{A}$ is the product of its eigenvalues, so we see that the differential operator $\mathcal{L}$ is elliptic if $ac - b^2 > 0$, hyperbolic if $ac - b^2 < 0$ and parabolic if $ac - b^2 = 0$. (Ultrahyperbolic operators cannot arise in dimension $< 4$.) Thus the wave operator ($c = 1$, $b = 0$, $a = -(\text{wave speed})^2$) is hyperbolic, the heat operator ($a = 0$, $b = 0$, $c = -(\text{diffusion constant})$) is parabolic, while the Laplace operator ($a = c = 1$, $b = 0$) is elliptic.

### 9.2.2 Characteristic surfaces

We now introduce the notion of a characteristics for second–order pdes. To motivate the definition, recall that the curve defined by $f(x,y) = \text{const.}$ would be a characteristic of the first–order differential operator $\mathcal{L} = \alpha(x,y)\partial_x + \beta(x,y)\partial_y + \gamma(x,y)$ iff $f$ obeys $\alpha\partial_x f + \beta\partial_y f = 0$; this equation just says that $f$ doesn't change along the characteristic curves of $\mathcal{L}$, or in other words curves of constant $f$ are indeed characteristics. Similarly, for a first–order differential operator $\mathcal{L} = a^i(x)\partial_i$ in $n$ variables, the $(n-1)$-dimensional surface $f(x^1, x^2, \ldots, x^n) = \text{const.}$ is characteristic iff $a^i(x)\,\partial_i f = 0$ for all $x \in \mathbb{R}^n$. Note that the coefficient vector $a^i(x)$ is also what appears in the principal part of the symbol $\sigma^{\mathrm{p}}(x,k) = a^i(x)k_i$ here. Moving to the second–order case, similarly we say that the surface $C \in \mathbb{R}^n$ defined by $f(x^1, x^2, \ldots, x^n) = \text{const}$ is a *characteristic surface* of the operator $\mathcal{L}$ in (9.26) *at a point* $x \in \mathbb{R}^n$ if

$$a^{ij}(x)\,\frac{\partial f}{\partial x^i}\,\frac{\partial f}{\partial x^j} = 0\,, \quad (9.32)$$

or equivalently if

$$(\nabla f)^{\mathrm{T}}\,\mathbf{A}\,(\nabla f) = 0\,. \quad (9.33)$$

$C$ is a *characteristic surface* for $\mathcal{L}$ if it is characteristic everywhere.

Let's look for characteristics of our different types of differential operator. Firstly, if $\mathcal{L}$ is elliptic, the matrix $\mathbf{A}$ is definite, so the only solutions of (9.33) are when $\partial_i f = 0$ so that $f$ is identically constant. Since $f$ is independent of all coordinates *identically*, it does not define a surface. Consequently, an elliptic operator has no (real) characteristic surfaces and the method of characteristics is not applicable to elliptic pdes such as Laplace's equation.

Next, consider a parabolic operator. For simplicity, we'll just consider the case where it has only one zero eigenvalue, with all the remaining eigenvalues of the same sign. Let $\mathbf{n}$ be the normalised eigenvector with $\mathbf{An} = \mathbf{n}^{\mathrm{T}}\mathbf{A} = 0$. We decompose $\nabla f$ as

$$\begin{aligned} \nabla f &= (\nabla f - \mathbf{n}\,(\mathbf{n}\cdot\nabla f)) + \mathbf{n}\,(\mathbf{n}\cdot\nabla f) \\ &=: \nabla_{\perp} f + \mathbf{n}\,(\mathbf{n}\cdot\nabla f) \end{aligned} \tag{9.34}$$

where $\nabla_{\perp} f$ is orthogonal to $\mathbf{n}$ wrt the quadratic form $\mathbf{A}$. In terms of this decomposition,

$$\begin{aligned} (\nabla f)^{\mathrm{T}}\,\mathbf{A}\,(\nabla f) &= [\nabla_{\perp} f + \mathbf{n}\,(\mathbf{n}\cdot\nabla f)]^{\mathrm{T}}\ \mathbf{A}\ [\nabla_{\perp} f + \mathbf{n}\,(\mathbf{n}\cdot\nabla f)] \\ &= (\nabla_{\perp} f)^{\mathrm{T}}\ \mathbf{A}\ (\nabla_{\perp} f)\ , \end{aligned} \tag{9.35}$$

using the fact that $\mathbf{n}$ is a left– and right–eigenvector of $\mathbf{A}$, with eigenvalue 0. The remaining term involves only $\nabla_{\perp} f$ which lives in the (positive or negative) definite eigenspace of $\mathbf{A}$. Thus, just as in the elliptic case, the only solutions to the characteristic equation $(\nabla_{\perp} f)^{\mathrm{T}}\ \mathbf{A}\ (\nabla_{\perp} f) = 0$ is $\nabla_{\perp} f = 0$, so $f(x^1, x^2, \ldots, x^n)$ must in fact be independent of all the coordinates in directions orthogonal to $\mathbf{n}$. However, the value of $\mathbf{n}\cdot\nabla f$ is unconstrained by the characteristic equation, so surfaces whose normal vector is given by $\mathbf{n}$ are characteristic surfaces. Thus there is a unique characteristic surface through any point $x \in \mathbb{R}^n$ at which $\mathcal{L}$ is parabolic (with a single zero eigenvalue).

As an example, consider the heat operator $\mathcal{L} = \partial_t - \kappa\nabla^2$ where $\kappa$ is a diffusion constant and $\nabla^2$ the Laplacian on $\mathbb{R}^n$. For this operator

$$\mathbf{A} = \mathrm{diag}(0, -\kappa, -\kappa, \cdots, -\kappa) \tag{9.36}$$

because the time derivative only appears to first order. The zero eigenvector $\mathbf{n}$ points in the time direction, so surfaces of constant time are characteristic surfaces. Note that there is exactly one characteristic surface through any point $(x, t) \in \mathbb{R}^n \times [0, \infty)$.

Finally, we consider a hyperbolic operator, where all eigenvalues of $\mathbf{A}$ but one have the same sign. Suppose for definiteness that only one eigenvalue is negative and let $-\lambda$ be this negative eigenvalue, with $\mathbf{m}$ the corresponding unit eigenvector. Decomposing $\nabla f$ into its parts along and perpendicular to $\mathbf{m}$ as before, we now have

$$\begin{aligned} (\nabla f)^{\mathrm{T}}\,\mathbf{A}\,(\nabla f) &= [\nabla_{\perp} f + \mathbf{m}\,(\mathbf{m}\cdot\nabla f)]^{\mathrm{T}}\ \mathbf{A}\ [\nabla_{\perp} f + \mathbf{m}\,(\mathbf{m}\cdot\nabla f)] \\ &= (\mathbf{m}\cdot\nabla f)^2\left(\mathbf{m}^{\mathrm{T}}\mathbf{A}\mathbf{m}\right)\ +\ (\nabla_{\perp} f)^{\mathrm{T}}\ \mathbf{A}\ (\nabla_{\perp} f) \\ &= -\lambda(\mathbf{m}\cdot\nabla f)^2 + (\nabla_{\perp} f)^{\mathrm{T}}\ \mathbf{A}\ (\nabla_{\perp} f)\ , \end{aligned} \tag{9.37}$$

The characteristic condition $(\nabla f)^{\mathrm{T}}\,\mathbf{A}\,(\nabla f)=0$ thus determines $(\mathbf{m}\cdot\nabla f)$ in terms of $\nabla_\perp f$ via

$$\mathbf{m}\cdot\nabla f=\pm\sqrt{\frac{(\nabla_\perp f)^{\mathrm{T}}\ \mathbf{A}\ (\nabla_\perp f)}{\lambda}}\,. \tag{9.38}$$

Now, given any function $f(x^1,x^2,\ldots,x^n)$ defining a candidate for a characteristic surface, we calculate the value of the *rhs* of (9.38). Equation (9.38) can thus be regarded as an ode determining how $f$ must depend on the variable pointing along the direction of $\mathbf{m}$. Such first–order odes will always have solutions, and since we have two possible choices of sign, we can find two possible characteristic surfaces through any point $x\in\mathbb{R}^n$. In summary, there are *two* separate characteristic surfaces through any point $x\in\mathbb{R}^n$ at which a differential operator $\mathcal{L}$ is hyperbolic.

Let's again fix our ideas with a key example. Suppose $\mathcal{L}=-\partial_t^2+c^2\partial_x^2$ is the wave operator in 1+1 dimensions. Then

$$\mathbf{A}=\mathrm{diag}(-1,c^2) \tag{9.39}$$

and $\mathbf{m}$ points in the time direction, with $\mathbf{Am}=\mathbf{m}$. Thus, if $f(x,t)=$ const. is to be a characteristic surface, we need $\partial_t f=\pm\sqrt{\partial_x f\,A_{xx}\,\partial_x f}$, or in other words

$$(\partial_t\pm c\partial_x)f=0\,. \tag{9.40}$$

This says that curves (lines) of constant $x\pm ct$ are characteristics. At the beginning of the course, we saw that the general solution to the wave equation $\partial_t^2\phi=c^2\partial_x^2\phi$ with initial data $\phi(x,0)=f(x)$ and $\partial_t\phi(x,0)=g(x)$ was given by

$$\phi(x,t)=\frac{f(x+ct)+f(x-ct)}{2}+\frac{1}{2c}\int_{x-ct}^{x+ct}g(y)\,\mathrm{d}y\,. \tag{9.41}$$

as found by d'Alembert. We now see the key role played by characteristics in this solution. The value of the solution any point $(x,t)\in\mathbb{R}^{1,1}$ is fully determined by the behaviour of the initial functions $f,g$ in the interval $[x-ct,x+ct]$ of the $x$-axis, whose endpoints are the intersections of our initial data surface $t=0$ with the characteristics through the point $(x,t)$. This interval is called the *domain of dependence* for the solution at $(x,t)$. The initial value data $f$ itself propagates exactly along characteristics, so in particular a discontinuity (or other sharp features) in $f$ at some point $x_0$ will lead to a discontinuity in the solution along the characteristics emanating from $(x_0,0)$. Similarly, the value of the initial data $g$ at a point $x_0$ at time $t=0$ influences $\phi(x,t)$ at all points $(x,t)$ within the wedge–shaped region bounded by the characteristics $x\pm ct=x_0$ through $(x_0,0)$ — that is, in the region $x_0-ct<x<x_0+ct$ (see figure 15). Thus disturbances or signals travel only with speed $c$, as is familiar in special relativity.

Let's do one further example. Consider the equation $xy\,\partial_x^2\phi-\partial_y^2\phi=0$. In this case $\mathbf{A}=\mathrm{diag}(xy,-1)$ whose eigenvalues vary as we move around the plane. Since $\det\mathbf{A}=-xy$, the equation is hyperbolic in the first $(x,y>0)$ and third $(x,y<0)$ quadrants, elliptic in the second and fourth quadrants and parabolic along the axes $x=0$ or $y=0$. Let's find

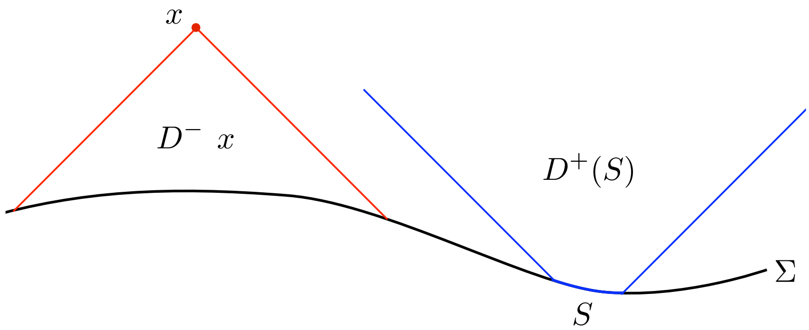


**Figure 15**. *Characteristics of the wave equation travel left and right with speed c. $D^-(p)$ is the (past) domain of dependence of a point p; the solution at $x$ is governed by the Cauchy data on $D^-(x) \cap \Sigma$. The range of influence $D^+(S)$ of a set $S \subset \Sigma$ is the set of points the Cauchy data on $S$ can influence.*

the characteristics in the first quadrant, where the negative eigenvector m points in the $y$-direction. The equation for a characteristic surface (curve) is thus

$$\partial_y f = \pm\sqrt{\partial_x f\, A_{xx}\, \partial_x f} = \pm\sqrt{xy}\partial_x f\,, \tag{9.42}$$

or equivalently

$$\frac{1}{\sqrt{y}}\frac{\partial f}{\partial y} \mp \sqrt{x}\frac{\partial f}{\partial x} = 0\,. \tag{9.43}$$

Letting $p = x^{1/2}$ and $q = y^{3/2}/3$, this is $(\partial_q \mp \partial_p)f = 0$, so the characteristic surfaces are curves of constant $q \pm p$. That is, the two families of characteristic curves are defined by

$$\begin{aligned} u &= \frac{1}{3}y^{3/2} + x^{1/2} \\ v &= \frac{1}{3}y^{3/2} - x^{1/2} \end{aligned} \tag{9.44}$$

for $u$ constant and $v$ constant.

#### 9.2.3 Black holes