24. The Method of Characteristics

Recall lecture 2, which began:

Most physical systems are governed by second order PDEs. In this course we discuss **Finite Difference Methods (FDMs)** for solving such equations. We want our algorithms to be able to reproduce the physics and so to begin with, we must understand the physical background. We consider the equation for $u(x,\, y)$

$$au_{xx} + bu_{xy} + cu_{yy} = f\ . \tag{24.1}$$

This equation is called **quasi-linear** provided the functions a, b, c and f do not depend on u_{xx}, u_{xy} or u_{yy}, namely that the highest order derivatives occur linearly. They may, however depend on x, y, u, u_x and u_y, so that (24.1) is not necessarily **linear** – for example it is perfectly allowable in the discussion that follows that

$$f = du_x + eu_y + hu + g, \tag{24.2}$$

provided d, e, h, g functions retain the *quasi-linear* requirement.

Suppose we know u, u_x and u_y along some curve Γ in (x, y)-space*. From a point P on Γ we move a small vector displacement $(dx,\, dy)$ to a new point Q not on Γ, as shown in Fig. 24.1. Under what circumstances can we determine uniquely the values of u, u_x and u_y at Q? We denote the change in these variables by du, $d(u_x)$ and $d(u_y)$. Then by the chain rule for partial derivatives $du = u_x dx + u_y dy$, which is known because u_x and u_y are known along Γ. Similarly,

$$\left.\begin{aligned} d(u_x) &= u_{xx}\, dx + u_{xy}\, dy \\ d(u_y) &= u_{xy}\, dx + u_{yy}\, dy \end{aligned}\right\}. \tag{24.3}$$

We combine (24.1) and (24.3) in matrix form:

$$\begin{pmatrix} a & b & c \\ dx & dy & 0 \\ 0 & dx & dy \end{pmatrix} \begin{pmatrix} u_{xx} \\ u_{xy} \\ u_{yy} \end{pmatrix} = \begin{pmatrix} f \\ d(u_x) \\ d(u_y) \end{pmatrix}, \tag{24.4}$$

a, b and c are known locally because u, u_x and u_y are known, and so the 3×3 matrix is known. Equation (24.4) will have a unique solution for u_{xx}, u_{xy} and u_{yy} unless the determinant of that matrix vanishes, that is unless

$$\begin{vmatrix} a & b & c \\ dx & dy & 0 \\ 0 & dx & dy \end{vmatrix} = a(dy)^2 - b\, dxdy + c(dx)^2 = 0. \tag{24.5}$$

If (24.5) holds, the equation (24.4) will have either no solution or infinitely many solutions. The condition for solutions to exist, which we will use later in the course, is that

$$a\,\frac{d(u_x)}{dx} + c\,\frac{d(u_y)}{dy} = f\ . \tag{24.6}$$

*We use the shorthand notation u_x, u_y to represent $\partial u/\partial x$ and $\partial u/\partial y$ respectively. Similarly u_{xx}, u_{xy}, u_{yy} to represent $\partial^2 u/\partial x^2$, $\partial^2 u/\partial x \partial y$ and $\partial^2 u/\partial y^2$ respectively

This is surprising. If we can choose a direction (dx, dy) which satisfies (24.5) we have the possibility that the second derivatives u_{xx} etc. may not be uniquely defined. In other words, the solution may have discontinuities across the line PQ, with u_{xx} taking different values on each side.

It is very important to know whether our solution can have this property. Equation (24.5) is called the **Characteristic equation** of (24.1). It is a quadratic in dy/dx with solution

$$\frac{dy}{dx} = \frac{b \pm \sqrt{b^2 - 4ac}}{2a} \,. \tag{24.7}$$

We can use equations (24.7) and (24.6) to find the solution. Equation (24.7) defines two families of characteristics, one with the +sign and one with the -sign. Starting from a given point (x, y) we can step (24.7) using some ODE-solver to find a new point $(x + dx, y + dy)$ which lies on the characteristic. Note that characteristics from the same family are approximately parallel – if ever they cross, then a **shock** develops and the solution becomes singular. However the two families have different gradients and so characteristics from different families meet all the time. So if we have a set of points P_n^j on some curve Γ^j at some "time", j, we can define points P_n^{j+1} to be the intersection of the +char from P_n^j and the -char from P_{n+1}^j as in the diagram, thus defining a new curve Γ^{j+1}.

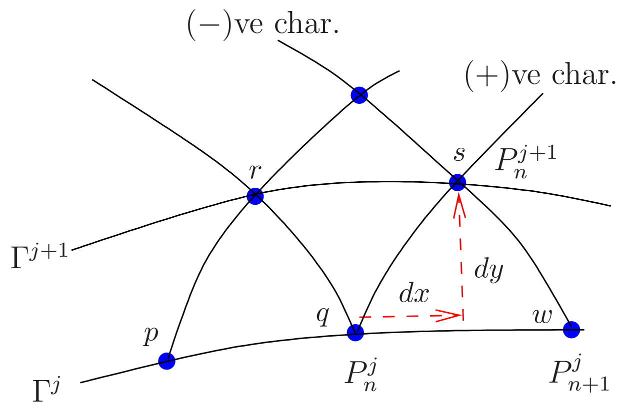

Figure 24.1: Problem description.

Now assume u, u_x and u_y are known on Γ^j . This means that u is known after an infinitesimal step (dx, dy) by $du = u_x dx + u_y dy$, but u_x and u_y need to be found. The key is to use the condition (24.7) along the characteristics, relating how much u_x and u_y change j. As we have two equations in two unknowns we can from the known values at P_n^j and P_{n+1}^j easily determine u_x and u_y at the new point P_n^{j+1} and hence we can advance the region where the solution is known to a larger region, bounded by Γ^{j+1}. This defines the method of characteristics.

As an example suppose the roots of (24.7) are real and distinct, and can be denoted by

$$\frac{dy}{dx} = F, \text{ and } \frac{dy}{dx} = G, \tag{24.8}$$

where initial values of u, u_x and u_y are known on Γ^j. As a first approximation, with reference to Figure 24.1, we may regard $p - r$ and $q - r$ as straight lines of slopes F_p and G_q. Then (24.8) can be approximated by

$$y_r - y_p = F_p(x_r - x_p), \text{ and } y_r - y_q = G_q(x_r - x_q), \tag{24.9}$$

giving two equations for the two unknowns (x_r, y_r). Next equation (24.1) can be used to show that the differentials $P = u_x$ and $Q = u_y$ are related by the equation

$$a\,\frac{dy}{dx}\,dP + c\,dQ - f\,dy = 0, \tag{24.10}$$

hence

$$a\,F\,dP + c\,dQ - f\,dy = 0, \text{ and } a\,G\,dP + c\,dQ - f\,dy = 0. \tag{24.11}$$

We approximate the first expression along $p - r$ by the equation

$$a_p\,F_p\,(P_r - P_p) + c_p\,(Q_r - Q_p) - f_p\,(y_r - y_p) = 0, \tag{24.12}$$

and the second expression along $q - r$ thus

$$a_q\,G_q\,(P_r - P_q) + c_q\,(Q_r - Q_q) - f_q\,(y_r - y_q) = 0. \tag{24.13}$$

Once (x_r, y_r) have been evaluated from (24.9), these are two equations for unknowns (P_r, Q_r). The value of u on r can then be obtained from

$$du = Pdx + Qdy, \tag{24.14}$$

on replacing values of (P, Q) along $p - r$ by

$$u_r - u_p = \frac{1}{2}(P_p + P_r)(x_r - x_p) + \frac{1}{2}(Q_p + Q_r)(y_r - y_p). \tag{24.15}$$

Solution of the above gives us a first approximation to u_r. We next improve upon this by replacing expressions (24.9) with new improved estimates of (x_r, y_r), using

$$y_r - y_p = \frac{1}{2}(F_p + F_r)(x_r - x_p), \text{ and } y_r - y_q = \frac{1}{2}(G_q + G_r)(x_r - x_q), \tag{24.16}$$

and equations (24.12-24.13) become

$$\left.\begin{aligned} \frac{1}{4}(a_p + a_r)\,(F_p + F_r)\,(P_r - P_p) + \frac{1}{2}(c_p + c_r)\,(Q_r - Q_p) - \frac{1}{2}(f_p + f_r)\,(y_r - y_p) = 0,\\ \frac{1}{4}(a_q + a_r)\,(G_q + G_r)\,(P_r - P_q) + \frac{1}{2}(c_q + c_r)\,(Q_r - Q_q) - \frac{1}{2}(f_q + f_r)\,(y_r - y_q) = 0.\end{aligned}\right\}. \tag{24.17}$$

An improved value for u_r then follows from (24.15). Repetition of the above steps, through iteration and correction thus provides u_r to sufficient accuracy. The number of iterations can be minimised by making q and p close to each other.

An advantage of the method is that by explicitly following characteristics we are modelling the physics well. If the solution does have discontinuities, we expect to follow them accurately. A disadvantage is that it does not generalise easily to 3D, when the characteristic curves become cones. Also, we can no longer have a regular grid – the equation decides where to place the points P_n^j , and they may bunch together. One could argue, that grid points are being carried to areas where they are needed, but there is the danger of thinning of points in some regions.

This leads us on to a more fundamental question for this course – we have always tried to use a regular grid, which affords us reliable 2-nd order accuracy. Should we relax this in some cases? We need points in order to resolve rapid variation. If we have a solution like $u = e^{-10x} + x^2$, we might want to have more points near $x = 0$ than near $x = 1$. To put it another way, if we use a small enough step-length to resolve the fast variation, we are being unnecessarily wasteful of computing power elsewhere. Suppose we introduce a new coordinate, $\xi = f(x)$. Then we can rewrite the PDE in terms of the new variable ξ and then use central differences in ξ, maintaining 2-nd order accuracy. We may clutter up the PDE with derivatives of f (which may be large in some regions), but this nevertheless gives us a methodical means of redistributing points to where we think they are needed – this is discussed in Lecture 23.