2. First-order linear equations

We want the solution $\theta(x,y)$ of

$$a(x,y)\frac{\partial\theta}{\partial x}+b(x,y)\frac{\partial\theta}{\partial y}=0 \tag{1}$$

where $a(x,y)$ and $b(x,y)$ are given functions. If a and b are constants, then $\theta{=}F(bx{-}ay)$ where F is an arbitrary function. If we regard y as time, then this is a wave traveling toward positive x at speed a/b.

If $a(x,y)$ and $b(x,y)$ are not constants, we may regard them as the components of a steady velocity field

$$\mathbf{v}=\big(a(x,y),b(x,y)\big)\equiv(u,v). \tag{2}$$

Our equation,

$$u(x,y)\frac{\partial\theta}{\partial x}+v(x,y)\frac{\partial\theta}{\partial y}=0 \tag{3}$$

tells us that θ is uniform along the streamlines of $\mathbf{v}$. To find these streamlines we solve the coupled ordinary differential equations,

$$\frac{dx}{dt}=u(x,y),\quad \frac{dy}{dt}=v(x,y) \tag{4}$$

for the "particle trajectory" $\big(x(t),y(t)\big)$ and then eliminate t to get $y(x)$ or $x(y)$. Not that the "time" t is an auxiliary variable, introduced for convenience. It plays no role in the final solution.

Example.

$$-y\frac{\partial\theta}{\partial x}+x\frac{\partial\theta}{\partial y}=0. \tag{5}$$

Then

$$\frac{dx}{dt}=-y,\quad \frac{dy}{dt}=x \quad\Rightarrow\quad dt=-\frac{dx}{y}=\frac{dy}{x} \quad\Rightarrow\quad x\,dx+y\,dy=0 \tag{6}$$

Thus

$$x^2+y^2=const \tag{7}$$

and the streamlines are concentric circles. The general solution of (5) is

$$\theta=F\left(x^2+y^2\right) \tag{8}$$

where F is an arbitrary function. To obtain a unique solution, we must specify θ at one (and only one) point on each circle.

Mathematicians call the "streamlines" *characteristics*. The generalization to n dimensions is obvious: To find the solution $\theta(x_1, x_2, \ldots, x_n)$ of

$$(9) \qquad u_1(x_1, x_2, \ldots, x_n)\frac{\partial\theta}{\partial x_1} + u_2(x_1, x_2, \ldots, x_n)\frac{\partial\theta}{\partial x_2} + \cdots + u_n(x_1, x_2, \ldots, x_n)\frac{\partial\theta}{\partial x_n} = 0$$

we solve the n equations

$$(10) \qquad \frac{dx_1}{ds} = u_1(x_1, x_2, \ldots, x_n), \quad \frac{dx_2}{ds} = u_2, \quad \ldots \quad \frac{dx_n}{ds} = u_n$$

where we have adopted the more neutral symbol s for the parameter along the trajectory. (We might want to use t later as one of the x_i's.) We may attempt to integrate the $n-1$ equations

$$(11) \qquad ds = \frac{dx_1}{u_1(\mathbf{x})} = \frac{dx_2}{u_2(\mathbf{x})} = \cdots = \frac{dx_n}{u_n(\mathbf{x})}$$

but this will only be possible if the $u_i(\mathbf{x})'s$ take a simple form. If we succeed in integrating the $n-1$ equations, then we will have $n-1$ relations of the form

$$(12) \qquad \phi_1(\mathbf{x}) = c_1, \quad \phi_2(\mathbf{x}) = c_2, \; \ldots \;, \phi_{n-1}(\mathbf{x}) = c_{n-1}$$

where the c_i are $n-1$ constants of integration. Each equation in (12) represents an $n-1$ dimensional surface in n-dimensional space. The intersection of these surfaces is a line. We have a different line for each set of c_i's. Thus the general solution of (9) is

$$(13) \qquad \theta = F(\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \ldots, \phi_{n-1}(\mathbf{x}))$$

where F is an *arbitrary* function of $n-1$ variables. However it is not *generally* possible to integrate (11). Thus (10) is the more important equation. It demonstrates that the solution of any first-order partial differential equation may be reduced to the solution of coupled ordinary differential equations. That is the fundamental point!

Example. To solve

$$(14) \qquad yx\frac{\partial\theta}{\partial x} + z\frac{\partial\theta}{\partial y} + \frac{\partial\theta}{\partial z} = 0$$

we write the trajectory equations

$$(15) \qquad \frac{dx}{ds} = yx, \quad \frac{dy}{ds} = z, \quad \frac{dz}{ds} = 1$$

and obtain 2 equations by eliminating ds between any pair in (15). For example, from the first 2 equations in (15), we obtain

$$(16) \qquad \frac{dx}{yx} = \frac{dy}{z}.$$

In attempting to integrate (16) students sometimes make the silly mistake of treating z as a constant:

$$(17) \qquad \frac{dx}{yx} = \frac{dy}{z} \quad \Rightarrow \quad \ln x = \frac{1}{2}\frac{y^2}{z} + C \qquad WRONG!$$

This is wrong because $x(s)$, $y(s)$ and $z(s)$ all vary along the trajectory. A correct approach would be to begin with an equation containing only 2 variables,

$$(18) \qquad \frac{dy}{z} = dz \quad \Rightarrow y = \frac{1}{2}z^2 + C$$

and use this to eliminate y from

$$(19) \qquad \frac{dx}{yx} = dz \quad \Rightarrow \quad \frac{dx}{\left(\frac{1}{2}z^2 + C\right)x} = dz \quad \Rightarrow \quad \ln x = \frac{1}{6}z^3 + Cz + C_1.$$

Thus, along the trajectories determined by (15), the quantities

$$(20) \qquad y - \frac{1}{2}z^2$$

and

$$(21) \qquad \ln x - \frac{1}{6}z^3 - \left(y - \frac{1}{2}z^2\right)z = \ln x + \frac{1}{3}z^3 - yz$$

are constant. Thus the general solution of (14) is

$$(22) \qquad \theta = F\left(y - \frac{1}{2}z^2, \ln x + \frac{1}{3}z^3 - yz\right)$$

where F is an arbitrary function of 2 variables. The reader should check this by direct substitution of (22) into (14).

Note once again that s is a parameter that helps us to find the solution, but it does not appear in the final result. We could use any other parameter $\tau(s)$ as long as $d\tau / ds \neq 0$. To make the solution (22) unique we must specify θ at one point along each characteristic.

Now suppose that (9) is further generalized to

(23) $$u_1(x_1,x_2,\ldots,x_n)\frac{\partial\theta}{\partial x_1}+u_2(x_1,x_2,\ldots,x_n)\frac{\partial\theta}{\partial x_2}+\cdots+u_n(x_1,x_2,\ldots,x_n)\frac{\partial\theta}{\partial x_n}=F(\theta,\mathbf{x})$$

where F is an arbitrary function. This includes the most general first-order, linear equation, for which $F = A(\mathbf{x})\theta + B(\mathbf{x})$. However, we can as easily discuss the case in which F depends nonlinearly on θ. In that case (23) is called *semi-linear*. When $F \neq 0$, (23) takes the form

(24) $$\frac{d\theta}{ds}=F(\theta,\mathbf{x})$$

where

(25) $$\frac{d}{ds}=u_1\frac{\partial}{\partial x_1}+u_2\frac{\partial}{\partial x_2}+\cdots+u_n\frac{\partial}{\partial x_n}.$$

Thus we solve (23) by solving the $n+1$ coupled ODEs,

(26) $$\frac{dx_1}{ds}=u_1(\mathbf{x}),\quad \frac{dx_2}{ds}=u_2(\mathbf{x}),\quad \ldots\quad \frac{dx_n}{ds}=u_n(\mathbf{x}),\quad \frac{d\theta}{ds}=F(\theta,\mathbf{x})$$

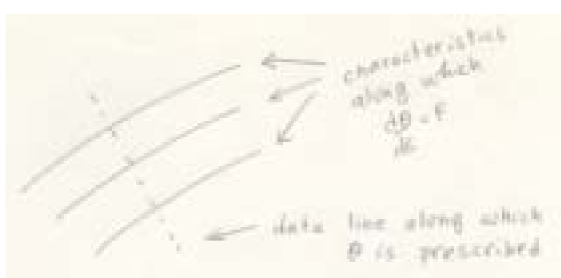
These equations trace out a trajectory in the (n+1)-dimensional space spanned by $(x_1,x_2,\ldots,x_n,\theta)$. As before, we may attempt to integrate (26) by writing

(27) $$(ds=)\ \frac{dx_1}{u_1(\mathbf{x})}=\frac{dx_2}{u_2(\mathbf{x})}=\cdots=\frac{dx_n}{u_n(\mathbf{x})}=\frac{d\theta}{F(\theta,\mathbf{x})}$$

but of course (27) may be too difficult to integrate.

Once again, to make the solution unique, we must specify θ at one point on each characteristic. However, if some of the characteristics are closed loops, then the case $F \neq 0$ could be problematic; the integration around the characteristic could lead to a different value of θ than we started with. This demonstrates that, although the method of characteristics has reduced our PDE to a system of ODEs, this system may have peculiarities.

Suppose that the characteristic are unclosed. Then, as "boundary conditions", we may specify θ along any line that cuts across the charcteristics as shown:



But consider what happens as the angle between the "data line" and the characteristics decreases:

When the data line coincides with a characteristic, the solution is *overdetermined* along that characteristic, and *undetermined* everywhere else. In that case we have two choices:

1.) Choose the boundary condition to be consistent with $d\theta / ds = F$ along the characteristic that coincides with the data line. Since θ remains arbitrary along every other characteristic, there are an infinite number of solutions for $\theta(\mathbf{x})$.

2.) Choose the boundary condition to be *in*consistent with $d\theta / ds = F$. Then there are *no* solutions for $\theta(\mathbf{x})$.

Mathematicians call these 2 choices the *Fredholm alternative*, after the first man who faced such a choice.

Although the geometrical picture given above is superior, we can rephrase all this more analytically as follows. In 2 dimensions (for example), we have

$$u(x,y)\frac{\partial \theta}{\partial x} + v(x,y)\frac{\partial \theta}{\partial y} = F. \tag{28}$$

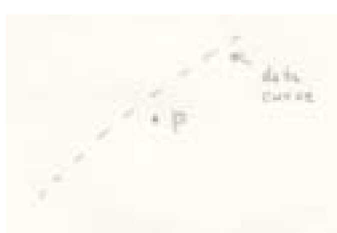
Our boundary condition is $\theta = \theta_0(r)$ along the data curve $x = x_0(r),\ y = y_0(r)$. That is,

$$\theta\big(x_0(r), y_0(r)\big) = \theta_0(r) \tag{29}$$

which implies

$$\frac{\partial \theta}{\partial x}\frac{dx_0}{dr} + \frac{\partial \theta}{\partial y}\frac{dy_0}{dr} = \frac{d\theta_0}{dr}. \tag{30}$$

Suppose we want to find θ at the point P, which is very near the data curve:



Obviously we need to know $\partial\theta / \partial n$, the derivative normal to the data line. We will know $\partial\theta / \partial n$ if we know $\partial\theta / \partial x$ and $\partial\theta / \partial y$. Equations (28) and (29) are 2 equations in these 2 quantities. These 2 equations have a unique solution if

$$(31) \qquad \det\begin{bmatrix} u & v \\ \dfrac{dx_0}{dr} & \dfrac{dy_0}{dr} \end{bmatrix} = (u,v) \wedge \left(\frac{dx_0}{dr},\frac{dy_0}{dr}\right) \neq 0$$

But this is just a fancy way of saying that data line cannot coincide with a characteristic. If the determinant (31) vanishes, there can still be a solution (many, in fact), but only if $d\theta / dr_0$ obeys a consistency condition that makes (28) equivalent to (29).

Thus, in general, characteristics may not be used as data lines. This property is sometimes taken to be the definition of characteristics. Another common definition is this: characteristics are lines along which discontinuites in the solution may propagate.

Finally, the problem can be further generalized to the *quasilinear* case,

$$(32) \qquad u_1(\mathbf{x},\theta)\frac{\partial\theta}{\partial x_1} + u_2(\mathbf{x},\theta)\frac{\partial\theta}{\partial x_2} + \cdots + u_n(\mathbf{x},\theta)\frac{\partial\theta}{\partial x_n} = F(\mathbf{x},\theta)$$

Equation (31) is nonlinear in the dependent varibale θ, but it remains linear in the derivatives of θ. Characteristic methods still provide a means of solution, but now there is a new twist: Because the direction of each characteristic depends not only on $\mathbf{x}$, but also on the value of θ swept along the characteristic, the characteristics can cross one another, leading to multi-valued solutions. We will consider this further in Lecture 4.

References. Ockendon pp 6-15, Zauderer pp 46-62.