

First order Partial Differential Equation

June 15, 2018

[1] A first-order PDE for an unknown function $u(x_1, x_2, \dots, x_n)$ has the following general form:

$$F(x_1, x_2, \dots, x_n, u, u_{x_1}, u_{x_2}, \dots, u_{x_n}) = 0 , \quad (1)$$

where F is a given function of $2n + 1$ variables. The solution method is based on the geometrical interpretation of u as a surface in an $(n + 1)$ -dimensional space.

Consider a surface in \mathbb{R}^3 whose graph is given by $u(x, y)$. The surface satisfies an equation of the form

$$F(x, y, u, u_x, u_y) = 0 , \quad (2)$$

The geometrical approach : Since $u(x, y)$ is a surface in \mathbb{R}^3 , and since the normal to the surface is given by the vector $(u_x, u_y, -1)$, the PDE (2) can be considered as an equation relating the surface to its normal (or alternatively its tangent plane). The main solution method will be a direct construction of the solution surface.

[2] Partial differential equations (PDEs) involve derivatives with respect to more than one independent variable; if the independent variables are x and y , a PDE in a dependent

variable $\varphi(x, y)$ will contain partial derivatives. Like ordinary derivatives, partial derivatives (of any order, including mixed derivatives) are linear operators, since they satisfy equations of the type

$$\frac{\partial[a\varphi(x, y) + b\varphi(x, y)]}{\partial x} = a\frac{\partial\varphi(x, y)}{\partial x} + b\frac{\partial\varphi(x, y)}{\partial x} . \quad (3)$$

Similar to the situation for ODEs, general differential operators, \mathcal{L} , which may contain partial derivatives of any order, pure or mixed, multiplied by arbitrary functions of the independent variables, are linear operators, and equations of the form

$$\mathcal{L}\varphi(x, y) = F(x, y) , \quad (4)$$

are linear PDEs. If the source term $F(x, y)$ vanishes, the PDE is termed **homogeneous**; if $F(x, y)$ is nonzero, it is **inhomogeneous**.

Homogeneous PDEs have the property that any linear combination of solutions will also be a solution to the PDE. This is the **superposition principle** which permits us to build specific solutions by the linear combination of suitable members of the set of functions constituting the general solution to the homogeneous PDE.

1 The method of characteristics

[1] Consider the general linear equation

$$a(x, y)u_x + b(x, y)u_y = c_0u(x, y) + c_1u(x, y) . \quad (5)$$

and the initial condition is parameterically:

$$\Gamma = \Gamma(s) = (x_0(s), y_0(s), u_0(s)), \quad s \in I = (\alpha, \beta) . \quad (6)$$

The curve Γ will be called the **initial curve**. It can be rewritten as

$$(a, b, c_0u + c_1) \cdot (u_x, u_y, -1) = 0 . \quad (7)$$

Since $(u_x, u_y, -1)$ is normal to the surface u , the vector $(a, b, c_0u + c_1)$ is in the tangent plane. The system of equations

$$\begin{aligned} \frac{dx}{dt}(t) &= a(x(t), y(t)) , \\ \frac{dy}{dt}(t) &= b(x(t), y(t)) , \\ \frac{du}{dt}(t) &= c_0(x(t), y(t))u(t) + c_1(x(t), y(t)) \end{aligned} \quad (8)$$

defines spatial curves lying on the solution surface (conditioned so that the curves start on the surface). It is a system of first-order ODEs, which are called the **system of characteristic equations** or, the **characteristic equations**. The solutions are called **characteristic curves** of the equation. The equations (8) are autonomous, i.e. there is **no explicit dependence upon the parameter t** . In order to determine a characteristic curve, an initial condition is needed. The initial point is required to lie on the initial curve Γ . Since each curve $(x(t), y(t), u(t))$ emanates from a different point $\Gamma(s)$, we shall explicitly write the curves in the form $(x(t, s), y(t, s), u(t, s))$. The initial conditions are written as:

$$x(0, s) = x_0(s), \quad y(0, s) = y_0(s), \quad u(0, s) = u_0(s) . \quad (9)$$

We selected the parameter t such that the characteristic curve is located on Γ when $t = 0$.

The method of characteristics applies to the quasilinear equation as well

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u) . \quad (10)$$

Each point on the initial curve Γ is a starting point for a characteristic curve. The

characteristic equations are

$$\begin{aligned}\frac{dx}{dt}(t) &= a(x, y, u) , \\ \frac{dy}{dt}(t) &= b(x, y, u) , \\ \frac{du}{dt}(t) &= c(x, y, u)\end{aligned}\tag{11}$$

and the initial condition

$$x(0, s) = x_0(s), \quad y(0, s) = y_0(s), \quad u(0, s) = u_0(s) .\tag{12}$$

The problem consisting of (10) and initial conditions (12) is called the **Cauchy problem** for quasilinear equations.

The main difference between the characteristic equations (8) derived for the linear equation, and the set (11) is that in the former case the first two equations of (8) are independent of the third equation and of the initial conditions. In the quasilinear case, this uncoupling of the characteristic equations is no longer possible, since the coefficients a and b depend upon u . In the linear case, the equation for u is always linear, and thus it is guaranteed to have a global solution (provided that the solutions $x(t)$ and $y(t)$ exist globally).

Each characteristic curve propagates independently of the other characteristic curves.

[2] Consider the following homogeneous linear first-order equation in two independent variables x and y , with constant coefficients a and b , and with dependent variable $\varphi(x, y)$:

$$\mathcal{L}\varphi = a\frac{\partial\varphi}{\partial x} + b\frac{\partial\varphi}{\partial y} = 0 .\tag{13}$$

This equation would be easier to solve if we could rearrange it so that it contained only one derivative; one way to do this is would be to rewrite PDE in terms of new coordinates

(s, t) such that one of them, say s , is such that $(\partial/\partial s)_t$ would expand into the linear combination of $\partial/\partial x$ and $\partial/\partial y$ in the original PDE, while the other new coordinate, t , is such that $(\partial/\partial t)_s$ does not occur in the PDE. Define $s = ax + by$ and $t = bx - ay$, and write $\varphi(x, y) = \varphi(x(s, t), y(s, t)) = \hat{\varphi}(s, t)$,

$$\begin{aligned}\left(\frac{\partial\varphi}{\partial x}\right)_y &= a\left(\frac{\partial\varphi}{\partial s}\right)_t + b\left(\frac{\partial\varphi}{\partial t}\right)_s, \\ \left(\frac{\partial\varphi}{\partial y}\right)_x &= b\left(\frac{\partial\varphi}{\partial s}\right)_t - a\left(\frac{\partial\varphi}{\partial t}\right)_s,\end{aligned}$$

so

$$a\frac{\partial\varphi}{\partial x} + b\frac{\partial\varphi}{\partial y} = (a^2 + b^2)\frac{\partial\hat{\varphi}}{\partial s}. \quad (14)$$

The PDE does not contain a derivative with respect to t . It has solution

$$\hat{\varphi}(s, t) = f(t), \quad \text{with } f(t) \text{ completely arbitrary.} \quad (15)$$

$$\varphi(x, y) = f(bx - ay), \quad (16)$$

where we stress that $f(t)$ is an arbitrary function of its argument. Since the satisfaction of this equation does not depend on the properties of the function f , we verify that $\varphi(x, y)$ as given in Eq. (16) is a solution of our PDE, irrespective of the choice of the function f . It is the general solution of the PDE.

Holding $t = bx - ay$ to a fixed value defines a line in the xy -plane on which solution φ is constant, with individual points on this line corresponding to different values of $s = ax + by$. The lines of constant s are orthogonal to those of constant t , and that s has the same coefficients as the derivatives in the PDE. The general solution to our PDE can be characterized as independent of s and with arbitrary dependence on t .

The curves of constant t are called characteristic curves, or more frequently just characteristics of PDE. The characteristic curves are the stream lines (flow lines) of s . They

are the lines that are traced out as the value of s is changed, keeping t constant. The characteristic can also be characterized by its slope,

$$\frac{dy}{dx} = \frac{b}{a}, \quad \text{for } \mathcal{L} \text{ in Eq. (30).} \quad (17)$$

The solution φ is constant along each characteristic. More general PDEs can be solved using ODE methods on characteristic lines, a feature that causes it to be said that PDE solutions propagate along the characteristics, giving further significance to the notion that in some sense these are lines of flow. If we know φ at any point on a characteristic, we know it on the entire characteristic line.

The characteristics have one additional (but related) property of importance. Ordinarily, if a PDE solution $\varphi(x, y)$ is specified on a curve segment (a boundary condition), one can deduce from it the values of the solution at nearby points that are not on the curve. If one introduces a Taylor expansion about some point (x_0, y_0) on the curve (thereby tacitly assuming that there are no singularities that invalidate the expansion), the value of φ at a nearby point (x, y) will be given by

$$\varphi(x, y) = \varphi(x_0, y_0) + \frac{\partial \varphi(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial \varphi(x_0, y_0)}{\partial y}(y - y_0) + \cdots. \quad (18)$$

To use Eq. (18), we need values of the derivatives of φ . To obtain these derivatives, note the following:

The specification of φ on a given curve, with the curve parametrically described by $x(l), y(l)$, means that the curve direction, i.e., dx/dl and dy/dl , is known, as is the derivative of φ along the curve, namely

$$\frac{d\varphi}{dl} = \frac{\partial \varphi}{\partial x} \frac{dx}{dl} + \frac{\partial \varphi}{\partial y} \frac{dy}{dl}, \quad (19)$$

which provides us with a linear equation satisfied by the two derivatives $\partial\varphi/\partial x$ and $\partial\varphi/\partial y$.

The PDE supplies a second linear equation, in this case

$$a\frac{\partial\varphi}{\partial x} + b\frac{\partial\varphi}{\partial y} = 0 . \quad (20)$$

Providing that the determinant of their coefficients is not zero, we can solve Eqs. (19) and (20) for $\partial\varphi/\partial x$ and $\partial\varphi/\partial y$ at (x_0, y_0) and therefore evaluate the leading terms of the Taylor series for $\varphi(x, y)$. The determinant of coefficients is

$$D = \begin{vmatrix} \frac{dx}{dl} & \frac{dy}{dl} \\ a & b \end{vmatrix} = b\frac{dx}{dl} - a\frac{dy}{dl} . \quad (21)$$

If φ was specified along a characteristic (for which $t = bx - ay = \text{constant}$),

$$b dx - a dy = 0 , \text{ or } b\frac{dx}{dl} - a\frac{dy}{dl} = 0 , \quad (22)$$

so that $D = 0$ and we cannot solve for the derivatives of φ . Our conclusions relative to characteristics, which can be extended to more general equations, are:

1. If the dependent variable φ of the PDE in Eq. (30) is specified along a curve (i.e., φ has a boundary condition specified on a boundary curve), this fixes the value of φ at a point of each characteristic that intersects the boundary curve, and hence at all points of each such characteristic;
2. If the boundary curve is along a characteristic, the boundary condition on it will ordinarily lead to inconsistency, and therefore, unless the boundary condition is redundant (i.e., coincidentally equal everywhere to the solution constructed from the value of φ at any one point on the characteristic), the PDE will not have a solution;
3. If the boundary curve has more than one intersection with the same characteristic, this will usually lead to an inconsistency, as the PDE may not have a solution that is simultaneously consistent with the values of φ at both intersections;

4. Only if the boundary curve is not a characteristic can a boundary condition fix the value of φ at points not on the curve. Values of φ specified only on a characteristic of the PDE provide no information as to the value of φ at points not on that characteristic.

Consider now a first-order PDE of a form more general

$$\mathcal{L}\varphi = a\frac{\partial\varphi}{\partial x} + b\frac{\partial\varphi}{\partial y} + q(x, y)\varphi = F(x, y) . \quad (23)$$

Make a transformation to new variables $s = ax + by$, $t = bx - ay$, and PDE becomes,

$$(a^2 + b^2) \left(\frac{\partial\varphi}{\partial s} \right) + \hat{q}(s, t)\hat{\varphi} = \hat{F}(s, t) . \quad (24)$$

Here $\hat{q}(s, t)$ is obtained by converting $q(x, y)$ to the new coordinates:

$$\hat{q}(s, t) = q \left(\frac{as + bt}{a^2 + b^2}, \frac{bs - at}{a^2 + b^2} \right) , \quad (25)$$

and \hat{F} is related in a similar fashion to F . Equation (24) is an ODE in s (containing what can be viewed as a parameter, t).

Given the three-dimensional (3-D) differential form

$$a\frac{\partial\varphi}{\partial x} + b\frac{\partial\varphi}{\partial y} + c\frac{\partial\varphi}{\partial z} , \quad (26)$$

apply a transformation to convert our PDE to the new variables $s = ax + by + cz$, $t = \alpha_1x + \alpha_2y + \alpha_3z$, $u = \beta_1x + \beta_2y + \beta_3z$, with α_i and β_i such that (s, t, u) form an orthogonal coordinate system. Then 3-D differential form is found equivalent to

$$(a^2 + b^2 + c^2) \frac{\partial\varphi}{\partial s} , \quad (27)$$

and the stream lines of s (those with t and u constant) are characteristics, along which we can propagate a solution φ by solving an ODE. Each characteristic can be identified

by its fixed values of t and u .

$$a \frac{\partial \varphi}{\partial x} + b \frac{\partial \varphi}{\partial y} + c \frac{\partial \varphi}{\partial z} = 0 , \quad (28)$$

$$(a^2 + b^2 + c^2) \frac{\partial \varphi}{\partial s} = 0 , \quad (29)$$

with solution $\varphi = f(t, u)$, with f a completely arbitrary function of its two arguments.

Consider next an attempt to solve 3-D PDE subject to a boundary condition fixing the values of the PDE solution φ on a surface. If the characteristic through a point on the surface lies in the surface, we have a potential inconsistency between the boundary condition and the solution propagated along the characteristic. We are also unable to extend φ away from the boundary surface because the data on the surface is insufficient to yield values of the derivatives that are needed for a Taylor expansion. To see this, note that the derivatives $\partial\varphi/\partial x$, $\partial\varphi/\partial y$, and $\partial\varphi/\partial z$ can only be determined if we can find two directions (parametrically designated l and l') such that we can solve Eq. (28) simultaneously and

$$\begin{aligned} \frac{\partial \varphi}{\partial l} &= \frac{\partial \varphi}{\partial x} \frac{dx}{dl} + \frac{\partial \varphi}{\partial y} \frac{dy}{dl} + \frac{\partial \varphi}{\partial z} \frac{dz}{dl} , \\ \frac{\partial \varphi}{\partial l'} &= \frac{\partial \varphi}{\partial x} \frac{dx}{dl'} + \frac{\partial \varphi}{\partial y} \frac{dy}{dl'} + \frac{\partial \varphi}{\partial z} \frac{dz}{dl'} , \end{aligned}$$

A solution can be obtained only if

$$D = \begin{vmatrix} \frac{dx}{dl} & \frac{dy}{dl} & \frac{dz}{dl} \\ \frac{dx}{dl'} & \frac{dy}{dl'} & \frac{dz}{dl'} \\ a & b & c \end{vmatrix} \neq 0 .$$

If a characteristic, with $dx/dl'' = a$, $dy/dl'' = b$, and $dz/dl'' = c$, lies in the two-dimensional (2-D) surface, there will only be one further linearly independent direction l ,

and D will necessarily be zero.

A boundary condition is effective in determining a unique solution to a first-order PDE only if the boundary does not include a characteristic, and inconsistencies may arise if a characteristic intersects a boundary more than once.

[3]

$$\frac{\partial w}{\partial t} + c \frac{\partial w}{\partial x} = 0 . \quad (30)$$

Consider the rate of change of $w(x(t), t)$ as measured by a moving observer, $x = x(t)$.

$$\frac{d}{dt}w(x(t), t) = \frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} \frac{dx}{dt} . \quad (31)$$

where $\frac{d}{dt}$ as measured by a moving observer is sometimes called the **substantial derivative**. $\frac{\partial w}{\partial t}$ represents the change in w at the fixed position, while $\frac{\partial w}{\partial x} \frac{dx}{dt}$ represents the change due to the fact that the observer moves into a region of possibly different w . Compared with the partial differential equation for w , if the observer moves with velocity c , that is, if

$$\frac{dx}{dt} = c , \quad (32)$$

then

$$\frac{d\omega}{dt} = 0 . \quad (33)$$

i.e. w is constant. An observer moving with this special speed c would measure no changes in w .

Integrating (32), yields

$$x = ct + x_0 , \quad (34)$$

the equation for the family of parallel characteristics¹ of (30). Note that at $t = 0$, $x = x_0$.

¹A characteristic is a curve along which a PDE reduces to an ODE.

$w(x, t)$ is constant along this line (not necessarily constant everywhere). w propagates as a wave with wave speed c .

If $w(x, t)$ is given initially at $t = 0$,

$$w(x, 0) = P(x) , \quad (35)$$

then determine w at the point (x, t) . Since w is constant along the characteristic,

$$w(x, t) = w(x_0, 0) = P(x_0) . \quad (36)$$

Given x and t , the parameter is known from the characteristic, $x_0 = x - ct$, and thus

$$w(x, t) = P(x - ct) , \quad (37)$$

which is called the general solution of (30). $P(x)$ can be an arbitrary function. The general solution of a first-order partial differential equation contains an arbitrary function, while the general solution to ordinary differential equations contains arbitrary constants.

At fixed t , the solution of the first-order wave equation is the same shape shifted a distance ct .

Consider

$$\frac{\partial w}{\partial t} + 3t^2 \frac{\partial w}{\partial x} = 2tw , \quad (38)$$

subject to the initial conditions $w(x, 0) = P(x)$. By the method of characteristics, if

$$\frac{dx}{dt} = 3t^2 , \quad (39)$$

then

$$\frac{dw}{dt} = 2tw . \quad (40)$$

The characteristics are not straight lines but satisfy

$$x = t^3 + x_0 , \quad (41)$$

where the characteristics start ($t = 0$) at $x = x_0$. Along the characteristics, by integrating the ODE,

$$w = ke^{t^2} . \quad (42)$$

To satisfy the initial condition at x_0 , $w(x_0, 0) = P(x_0)$, $P(x_0) = k$, the solution of the initial value problem by the method of characteristics is

$$w(x, t) = P(x_0)e^{t^2} = P(x - t^3)e^{t^2} . \quad (43)$$

Since $P(x - t^3)$ is an arbitrary function of $(x - t^3)$, it is the general solution of the partial differential equation.

2 Quasilinear equations

[1] **Quasilinear** : the derivatives of u appear in the equation linearly. The general form of a quasilinear equation is

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u) . \quad (44)$$

linear equations :

$$a(x, y)u_x + b(x, y)u_y = c_0(x, y)u + c_1(x, y) . \quad (45)$$

where a, b, c_0, c_1 are given functions of (x, y) .

[3] Most of this text describes methods for solving linear partial differential equations (separation of variables, eigenfunction expansions, Fourier and Laplace transforms,

Green's functions) that cannot be extended to nonlinear problems. The method of characteristics can be applied to partial differential equations of the form

$$\frac{\partial \rho}{\partial t} + c \frac{\partial \rho}{\partial x} = Q , \quad (46)$$

where c and Q may be functions of x, t , and ρ . When Q is not a linear function of ρ or, when the coefficient c depends on the unknown solution ρ , it is not linear. Superposition is not valid. Nonetheless it is called a **quasilinear partial differential equation**, since it is **linear in the first partial derivatives**, $\partial \rho / \partial t$ and $\partial \rho / \partial x$. Consider an observer moving in some prescribed way $x(t)$.

$$\frac{d\rho}{dt} = Q(\rho, x, t) , \quad (47)$$

if

$$\frac{dx}{dt} = c(\rho, x, t) . \quad (48)$$

The partial differential equation reduces to two coupled ordinary differential equations along the special trajectory or direction defined by (48), known as a **characteristic curve**, or simply a **characteristic**. The velocity defined by (48) is called the **characteristic velocity**, or **local wave velocity**. A characteristic starting from $x = x_0$ is determined from the coupled differential equations using the initial conditions $\rho(x, 0) = f(x)$. Along the characteristic, the solution ρ changes according to (47). Other initial positions yield other characteristics, generating a family of characteristics.

If the independent variables are x and y instead of x and t , then a quasilinear first-order partial differential equation is usually written as

$$a \frac{\partial \rho}{\partial x} + b \frac{\partial \rho}{\partial y} = c , \quad (49)$$

where a , b , and c may be functions of x , y , and ρ . The method of characteristics is

$$\frac{d\rho}{dx} = \frac{c}{a} , \quad (50)$$

if

$$\frac{dy}{dx} = \frac{b}{a} . \quad (51)$$

This is written in the following equivalent form:

$$\frac{dx}{a} = \frac{dy}{b} = \frac{d\rho}{dc} . \quad (52)$$

2.1 Traffic Flow

[3] Model a congested one-directional highway by a quasilinear partial differential equation. Introduce the traffic density $\rho(x, t)$, the number of cars per mile at time t located at position x . An easily observed and measured quantity is the traffic flow $q(x, t)$, the number of cars per hour passing a fixed place x (at time t).

consider an arbitrary section of roadway, between $x = a$ and $x = b$. If there are neither entrances nor exits on this segment of the road, then the number of cars between $x = a$ and $x = b$ [$N = \int_a^b \rho(x, t) dx$] might still change in time. The rate of change of the number of cars, dN/dt , equals the number per unit time entering at $x = a$ [the traffic flow $q(a, t)$ there] minus the number of cars per unit time leaving at $x = b$ [the traffic flow $q(b, t)$ there]:

$$\frac{d}{dt} \int_a^b \rho(x, t) dx = q(a, t) - q(b, t) . \quad (53)$$

which called the integral form of conservation of cars. Note that the boundary contribution may be expressed as an integral over the region:

$$q(a, t) - q(b, t) = - \int_a^b \frac{\partial}{\partial x} q(x, t) dx . \quad (54)$$

Thus, by taking the time derivative inside the integral (making it a partial derivative), it follows that

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 , \quad (55)$$

since a and b are arbitrary.

The number of cars per hour passing a place equals the density of cars times the velocity of cars. By introducing $u(x, t)$ as the car velocity,

$$q = \rho u . \quad (56)$$

The equations for the characteristics are

$$\frac{d\rho}{dt} = 0 , \quad (57)$$

along

$$\frac{dx}{dt} = c(\rho) . \quad (58)$$

The characteristic velocity c is not constant but depends on the density ρ . It is known as the density wave velocity. It follows that the density ρ remains constant along each as-yet undetermined characteristic. The velocity of each characteristic, $c(\rho)$, will be constant, since ρ is constant. Each characteristic is thus a straight line [as in the case in which $c(\rho)$ is a constant c_0]. However, different characteristics will move at different constant velocities because they may start with different densities. The characteristics, though each is straight, are not parallel to one another. Along the curve $dx/dt = c(\rho)$, $d\rho/dt = 0$ or ρ is constant. Initially ρ equals the value at $x = x_0$ (i.e., at $t = 0$). Thus, along this one characteristic,

$$\rho(x, t) = \rho(x_0, 0) = f(x_0) , \quad (59)$$

which is a known constant. The local wave velocity that determines the characteristic is a constant, $dx/dt = c(f(x_0))$. Consequently, this characteristic is a straight line,

$$x = c(f(x_0))t + x_0 , \quad (60)$$

since $x = x_0$ at $t = 0$. Different values of x_0 yield different straight-line characteristics. Along each characteristic, the traffic density ρ is a constant. To determine the density at some later time, the characteristic with parameter x_0 that goes through that space-time point must be obtained from (60).

2.2 Shock Waves

[3] For quasilinear partial differential equations, it is quite usual for characteristics to intersect. The resolution will require the introduction of moving discontinuities called shock waves. We restrict our attention to quasilinear partial differential equations with $Q = 0$, in which case

$$\frac{\partial \rho}{\partial t} + c(\rho) \frac{\partial \rho}{\partial x} = 0 . \quad (61)$$

In Fig. ??, one starting at $x = x_1$, with $\rho = f(x_1, 0) \equiv \rho_1$, and the other starting at $x = x_2$ with $\rho = f(x_2, 0) \equiv \rho_2$. These characteristics intersect if $c(\rho_1) > c(\rho_2)$, the faster catching up to the slower. The density is constant along characteristics. As time increases, the distance between the densities ρ_1 and ρ_2 decreases. Thus, this is called a compression wave. The density distribution becomes steeper as time increases. Eventually characteristics intersect; the theory predicts that the density is simultaneously ρ_1 and ρ_2 . If we continue to apply the method of characteristics, the faster-moving characteristic passes the slower. The method of characteristics predicts that the density becomes a

“multivalued” function of position; that is, at some later time, our mathematics predicts there will be three densities at some positions. The density wave breaks. However, in many physical problems (such as traffic flow) it makes no sense to have three values of density at one place. The density must be a single-valued function of position.

On the basis of the quasilinear partial differential equation, we predicted the physically impossible phenomenon that the density becomes multivalued. Since the method of characteristics is mathematically justified, it is the partial differential equation itself that must not be entirely valid. Some approximation or assumption that we used must at times be invalid. Assume that the density and velocity have a jump discontinuity, which we call a shock wave, or simply a shock. The shock occurs at some unknown position x_s and propagates in time, so that $x_s(t)$. We introduce the notation x_{s-} and x_{s+} for the position of the shock on the two sides of the discontinuity. The shock velocity, dx_s/dt , is as yet unknown.

On either side of the shock, the quasilinear partial differential equation applies, $\frac{\partial \rho}{\partial t} + c(\rho) \frac{\partial \rho}{\partial x} = 0$, where $c(\rho) = \frac{dq(\rho)}{d\rho}$. Determine how the discontinuity propagates. If ρ is conserved even at a discontinuity, then the flow relative to the moving shock on one side of the shock must equal the flow relative to the moving shock on the other side. This statement of relative inflow equaling relative outflow becomes

$$\rho(x_{s-}, t) \left[u(x_{s-}, t) - \frac{dx_s}{dt} \right] = \rho(x_{s+}, t) \left[u(x_{s+}, t) - \frac{dx_s}{dt} \right], \quad (62)$$

since flow equals density times velocity (here relative velocity). Solving for the shock velocity yields

$$\frac{dx_s}{dt} = \frac{q(x_{s+}, t) - q(x_{s-}, t)}{\rho(x_{s+}, t) - \rho(x_{s-}, t)} = \frac{[q]}{[\rho]}, \quad (63)$$

where we recall that $q = \rho u$ and where we introduce the notation $[q]$ and $[\rho]$ for the jumps in q and ρ , respectively. In gas dynamics, it is called the Rankine-Hugoniot condition. In summary, for the conservation law $\frac{\partial \rho}{\partial t} + c(\rho) \frac{\partial \rho}{\partial x} = 0$ (if the quantity $\int \rho dx$ is actually conserved), the shock velocity equals the jump in the flow divided by the jump in the density of the conserved quantity. At points of discontinuity, this shock condition replaces the use of the partial differential equation, which is valid elsewhere. However, we have not yet explained where shocks occur and how to determine $\rho(x_{s+}, t)$ and $\rho(x_{s-}, t)$.

The characteristics must flow into the shock on both sides. The characteristic velocity on the left ($2\rho = 8$) must be greater than the shock velocity ($\frac{dx_s}{dt} = 7$), and the characteristic velocity on the right ($2\rho = 6$) must be less than the shock velocity. This is a general principle, called the entropy condition,

$$c(\rho(x_{s-})) > \frac{dx_s}{dt} > c(\rho(x_{s+})) . \quad (64)$$

3 Nonlinear PDEs

[2]

References

- [1] Y. Pinchover and J. Rubinstein. *An Introduction to Partial Differential Equations*. Number v. 10 in An introduction to partial differential equations. Cambridge University Press, 2005.
- [2] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists, Seventh Edition: A Comprehensive Guide*. Academic Press, 7 edition, January 2012.

- [3] R. Haberman. *Applied Partial Differential Equations: With Fourier Series and Boundary Value Problems*. Featured Titles for Partial Differential Equations. PEARSON, 2013.