4.6 Characteristics and Translating Waves; Hamiltonian Systems

The propagation of 'information' that was illustrated in the previous Chapter can be understood (and generalized) by the concept of characteristics. These are easiest to describe if we write the governing equation as a system of first order equations. Figure 1 shows some cases that we have discussed already, and concludes with one that will be of key importance for understanding *Hamiltonian systems*.

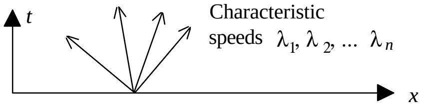
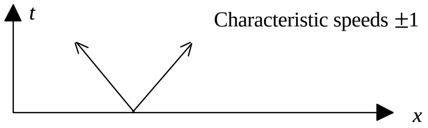
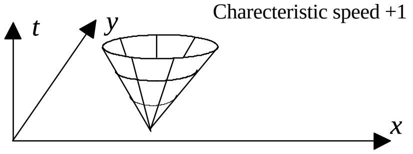
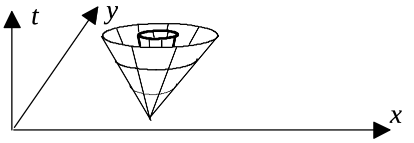
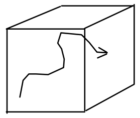
Case description	Written as first order system of equations	Illustration of characteristics
General system in 1-D	$\frac{\partial}{\partial t}\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \frac{\partial}{\partial x}\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$ Assume A has real and distinct eigenvalues $\lambda_1, ..., \lambda_n$	 Characteristic speeds $\lambda_1, \lambda_2, ... \lambda_n$
1-D wave equation	$\frac{\partial}{\partial t}\begin{bmatrix} u \\ v \end{bmatrix} = c\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\frac{\partial}{\partial x}\begin{bmatrix} u \\ v \end{bmatrix}$ To solve $\begin{cases} u_{tt} - c^2 u_{xx} = 0 \\ u(x,0) = f \\ u_t(x,0) = g \end{cases}$ initialize system $\begin{cases} u = f \\ v_x = g/c \end{cases}$	 Characteristic speeds ±1
2-D wave equation	$\frac{\partial}{\partial t}\begin{bmatrix} u \\ v \\ w \end{bmatrix} = c\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\frac{\partial}{\partial x}\begin{bmatrix} u \\ v \\ w \end{bmatrix} + c\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\frac{\partial}{\partial y}\begin{bmatrix} u \\ v \\ w \end{bmatrix}$ To solve $\begin{cases} u_{tt} - c^2(u_{xx} + u_{yy}) = 0 \\ u(x,y,0) = f \\ u_t(x,y,0) = g \end{cases}$ initialize system $\begin{cases} u = f \\ v_x + w_y = g/c \end{cases}$	Charecteristic speed +1 
2-D elastic wave equation	$\frac{\partial}{\partial t}\begin{bmatrix} u \\ v \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/\rho & 0 & 0 \\ 0 & 0 & 0 & 1/\rho & 0 \\ (\lambda+2\mu) & 0 & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 & 0 \\ \lambda & 0 & 0 & 0 & 0 \end{bmatrix}\frac{\partial}{\partial x}\begin{bmatrix} u \\ v \\ f \\ g \\ h \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 1/\rho & 0 \\ 0 & 0 & 0 & 0 & 1/\rho \\ 0 & \lambda & 0 & 0 & 0 \\ \mu & 0 & 0 & 0 & 0 \\ 0 & (\lambda+2\mu) & 0 & 0 & 0 \end{bmatrix}\frac{\partial}{\partial y}\begin{bmatrix} u \\ v \\ f \\ g \\ h \end{bmatrix}$ All 5 variables directly physically relevant	Charecteristic speeds $(\lambda+2\mu)^{1/2}$ and $\mu^{1/2}$ 
n-D scalar equation	$a_1(\underline{x})\frac{\partial u}{\partial x_1} + a_2(\underline{x})\frac{\partial u}{\partial x_2} + ... + a_n(\underline{x})\frac{\partial u}{\partial x_n} = 0$ n independent variables - none immediately associated with time	 Paths $x_i = x_i(s)$, $i = 1,...,n$ in n-space, parametrized by some auxiliary variable

Figure 1. Five different wave equations, all vritten as first order systems.

4.6.1 General hyperbolic system in 1-D

We can write the general system in Case 1, Figure 3, compactly as

$$\underline{u}_t - A\,\underline{u}_x = \underline{f} \qquad . \tag{1}$$

It is called	quasilinear	if	a_{ij} (the elements of A) are functions of x, t, $\underline{u}$,
	semilinear	if	a_{ij} does not depend on $\underline{u}$ (i.e. depends on x, t only), and
	linear	if	it is semilinear, and also f is a function of x, t only .

The arguments below work well if A depends smoothly on x and t. However, for simplicity, we will assume not only that A is constant but also that $\underline{f} = \underline{0}$.

The main idea behind the method of characteristics is to rearrange

$$\underline{u}_t - A\,\underline{u}_x = 0 \tag{2}$$

so that we

i. 'get rid of' the matrix A (while extracting from its eigenvalues an understanding of the general nature of the solution to (2)), and

ii. replace the system (2) of PDEs with ODEs along different paths (thereby gaining more precise insight in how solutions can evolve).

Multiplying (17) from the left by a row vector $\underline{v}_i^T$ amounts to linearly combining the different equations according to the elements of $\underline{v}_i^T$

$$\left[\quad \underline{v}_i^T \quad\right] \left\{ \frac{\partial}{\partial t}\left[\quad \underline{u} \quad\right] - \left[\quad A \quad\right] \frac{\partial}{\partial x}\left[\quad \underline{u} \quad\right] \right\} = 0 \qquad .$$

In case $\underline{v}_i^T$ is chosen as a *left eigenvector* to A, this combination simplifies the equation significantly:

$$\left[\quad \underline{v}_i^T \quad\right] \left\{ \frac{\partial}{\partial t}\left[\quad \underline{u} \quad\right] - \lambda_i \frac{\partial}{\partial x}\left[\quad \underline{u} \quad\right] \right\} = 0 \quad . \tag{3}$$

For each left eigenvector $\underline{v}_i^T$, $i = 1, 2, \ldots , n$, every one of the expressions inside the bracket of (3) takes the same form:

$$\frac{\partial}{\partial t} u_j - \lambda_i \frac{\partial}{\partial x} u_j \qquad , \qquad j = 1, 2, \ldots , n \ .$$

Instead of viewing this as a combination of two different derivatives (in the *t*- and *x*-directions respectively), we can view it as a single derivative in a single direction defined by

$$\frac{dx}{dt} = -\lambda_i \qquad . \tag{4}$$

This direction defines a straight line $x = x(t)$ in the x,t- plane. Along this line

$$\frac{d}{dt} u_j(x,t) = \frac{\partial u_j}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial u_j}{\partial t} = \frac{\partial u_j}{\partial t} - \lambda_i \frac{\partial u_j}{\partial x} \quad .$$

Therefore, assuming we follow any one of the different lines defined by (4) for $i = 1,2,\ldots,n$, we can write (3)

$$\left[\quad \underline{v}_i^T \quad \right] \frac{d}{dt} \left[\quad \underline{u} \quad \right] = 0 \quad ,$$

or in component form

$$\frac{d}{dt} \left[\sum_{j=1}^{n} v_{ij}\, u_j \right] = 0 \qquad , \qquad i = 1, 2, \ldots, n \;\; ,$$

i.e.

$$\sum_{j=1}^{n} v_{ij}\, u_j = \text{constant} \qquad (\text{along path } i \,,\; i = 1, 2, \ldots, n). \tag{5}$$

From this follows a theoretically important technique to advance $\underline{u}$ forward in time: Suppose we know everything at time $t = 0$, and want to find u_j , $j = 1, 2,\ldots,n$ at a certain point a time Δt later (cf. Figure 2).

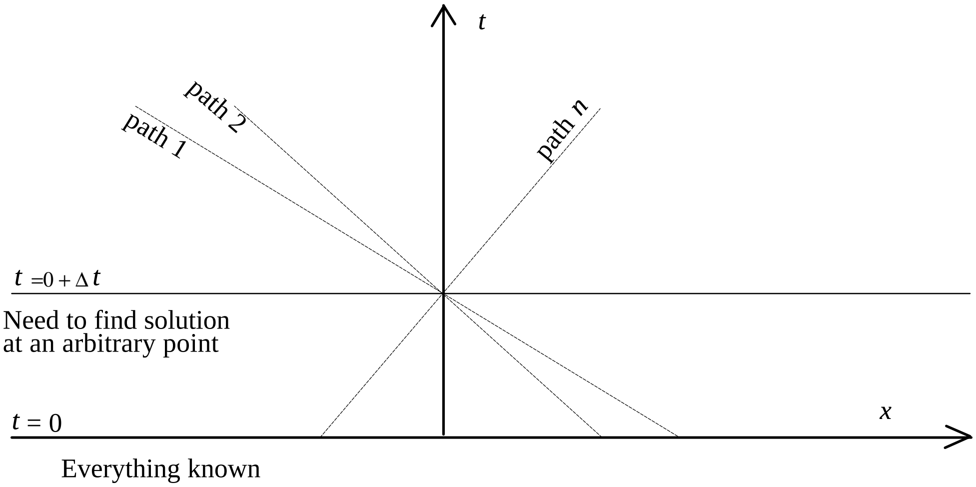


Figure 4. A possible way of using characteristics to advance the solution of a hyperbolic system forward in time.

Following the direction of path number 1 backwards, we find the spot from where we can read off the value of the quantity $\Sigma_{j=1}^{n} v_{1j} u_j$. Similarly, following each other path $i = 2,...,n$ backwards, we obtain at the new point the values for all of the n different linear combinations

$$\sum_{j=1}^{n} v_{ij}\, u_j(x,t) \qquad \text{for} \qquad i = 1, 2, \ldots, n\,.$$

This amounts to an $n \times n$ linear system for u_j, $j = 1, 2,...,n$. Hence, all the unknowns u_j can be determined at the new point.

> This procedure is exact, no matter how large time step Δt we take. If A had not been constant, or the RHS f had been present, this becomes instead an approximate (and quite ineffective) numerical procedure that requires many small time steps.

The main insight gained from this stepping along characteristics is that each eigenvalue of A corresponds to a distinct 'signal speed'. In particular, the extreme eigenvalues λ_1 and λ_n determine the fastest information can travel in forward and backward directions.

4.6.2 *n*-D Scalar Wave Equation

This case differs from our previous ones in that we

i. make no distinction of one independent variable as t (time) and the others as x_i (space); all are treated as spatial ones, and

ii. allow the coefficients to be functions of $\underline{x} = [x_1, x_2, \dots, x_n]$.

The scalar homogeneous equation in the n-D case becomes

$$a_1(\underline{x})\frac{\partial u}{\partial x_1} + a_2(\underline{x})\frac{\partial u}{\partial x_2} + \dots + a_n(\underline{x})\frac{\partial u}{\partial x_n} = 0 \tag{6}$$

Since u is a scalar, there is a single characteristic path $[x_1(s), x_2(s), \dots, x_n(s)]$, which we define through

$$\frac{dx_1}{ds} = a_1(\underline{x}),\ \frac{dx_2}{ds} = a_2(\underline{x}),\ \dots,\ \frac{dx_n}{ds} = a_n(\underline{x}) \qquad . \tag{7}$$

Having no 'time', we have here instead introduced an arbitrary variable s to parameterize the path. With $a_i(\underline{x})$ not constants, the path is no longer straight. Along it, the rate of change of a becomes

$$\frac{du}{ds} = \frac{\partial u}{\partial x_1}\cdot\frac{dx_1}{ds} + \frac{\partial u}{\partial x_2}\cdot\frac{dx_2}{ds} + \dots + \frac{\partial u}{\partial x_n}\cdot\frac{dx_n}{ds} = \qquad \text{by the chain rule}$$

$$= \frac{\partial u}{\partial x_1}\cdot a_1(x) + \frac{\partial u}{\partial x_2}\cdot a_2(x) + \dots + \frac{\partial u}{\partial x_n}\cdot a_n(x) = \qquad \text{by the path definition (7)}$$

$$= 0 \qquad , \qquad \text{by equation (6)}$$

i.e. the solution u to (6) remains unchanged along the path given through (7). If values for $u(\underline{x})$ are given along some boundary, following characteristic paths gives us the solution at other locations (Had (6) been non-homogeneous, say with RHS = $b(\underline{x})$, we would instead have obtained $\frac{du}{ds} = b(x)$, and we again find u by solving this ODE along the same characteristic path).

For the context of Chapter 2 (modeling of freak ocean waves), we need to generalize (6) to become an arbitrary function of u and of x_i , $\frac{\partial u}{\partial x_i}$, $i = 1,2,\dots,n$:

$$F(x_1, x_2, \dots, x_n, u, \tfrac{\partial u}{\partial x_1}, \tfrac{\partial u}{\partial x_2}, \dots, \tfrac{\partial u}{\partial x_n}) = 0 \tag{8}$$

or

$$F(x_1, x_2, \dots, x_n, u, p_1, p_2, \dots, p_n) = 0 \tag{9}$$

where we have used $p_1, p_2, ..., p_n$ to denote $\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, ..., \frac{\partial u}{\partial x_n}$ respectively. The characteristic paths (generalizing how (7) was obtained from (6)) become now

$$\frac{\partial x_i}{\partial s} = \frac{\partial F}{\partial p_i} \quad , \ i = 1,2,...,n . \tag{10}$$

Along these paths, we will next show that

$$\frac{\partial p_i}{\partial s} = -\left(\frac{\partial F}{\partial u} \cdot p_i + \frac{\partial F}{\partial x_i}\right) \quad , \ i = 1,2,...,n , \tag{11}$$

$$\frac{du}{ds} = \sum_{j=1}^{n} p_j \cdot \frac{\partial F}{\partial p_j} \quad . \tag{12}$$

These equations (10)-(12) forms a set of $2n+1$ coupled ODEs with s as the independent variable. They can be solved, giving paths through the $(x_1, x_2, ... , x_n)$-space and, along these paths, values for u and its derivatives p_i , $i = 1,2,...,n$.

Derivation of (11):

Differentiation of (9) with respect to x_i gives

$$0 = \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial u} \cdot \frac{\partial u}{\partial x_i} + \sum_{j=1}^{n} \frac{\partial F}{\partial p_j} \cdot \frac{\partial p_j}{\partial x_i} =$$

by chain rule on F

$$= \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial u} \cdot p_i + \sum_{j=1}^{n} \frac{\partial x_j}{\partial s} \cdot \frac{\partial^2 u}{\partial x_i \partial x_j} =$$

by the definitions of path (10) and of $p_j = \frac{\partial u}{\partial x_j}$

$$= \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial u} \cdot p_i + \sum_{j=1}^{n} \frac{\partial x_j}{\partial s} \cdot \frac{\partial p_i}{\partial x_j} =$$

swap order of differentiation in second-derivative terms, and use again $p_j = \frac{\partial u}{\partial x_j}$

$$= \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial u} \cdot p_i + \frac{\partial p_i}{\partial s}$$

by chain rule on p_i .

Derivation of (12):

$$\frac{\partial u}{\partial s} = \sum_{j=1}^{n} \frac{\partial u}{\partial x_j} \cdot \frac{\partial x_j}{\partial s} =$$

by chain rule on u

$$= \sum_{j=1}^{n} p_j \cdot \frac{\partial F}{\partial p_j}$$

by the definitions of path (10) and of $p_j = \frac{\partial u}{\partial x_j}$.

4.6.3 Hamiltonian equations

A particularly important special case of the pair of equations (11), (12) arises when (9) does not depend on *u*. Employing slightly different notation (to be more consistent with what is more commonly used in this case), we write (9)

$$H(\underline{x},\underline{p}) = 0 \tag{13}$$

where $\underline{p} = \nabla\phi\,(\underline{x})$ is the gradient of some scalar function $\phi\,(\underline{x})$. Calling the parameter t instead of s (since it in many applications will correspond to physical time), the equations (11) and (12) show that $\underline{x}(t)$ and $\underline{p}(t)$ (related through (9)) can be obtained through the 'Hamiltonian' system of ODEs

$$\left\{\begin{aligned} \frac{\partial x_i}{\partial t} &= \frac{\partial H}{\partial p_i} && (14)\\ \frac{\partial p_i}{\partial t} &= -\frac{\partial H}{\partial x_i} && (15) \end{aligned}\right. \qquad , \; i = 1, 2, \ldots$$

Relations of the form (13) are of fundamental importance, and arise very frequently in areas such as classical mechanics, nonlinear wave motions, and chaos. The function $H = 0$ often expresses a conservation law, such as conservation of energy. In the contect of this book, (13)-(15) provide the key tool in Chapter 2 for determining how water waves evolve as they encounter currents.