2. Method of Characteristics

In this section we explore the method of characteristics when applied to linear and nonlinear equations of order one and above.

2.1. Method of characteristics for first order quasilinear equations.

2.1.1. Introduction to the method.

A first order quasilinear equation in 2D is of the form

$$a(x,y,u)\,u_x + b(x,y,u)\,u_y = c(x,y,u); \tag{2.1}$$

in 3D is of the form

$$a(x,y,z,u)\,u_x + b(x,y,z,u)\,u_y + c(x,y,z,u)\,u_z = d(x,y,z,u). \tag{2.2}$$

One can easily generalize this to higher dimensions. In the following we will just illustrate the method in 2D. Generalizations to higher dimensions is straightforward.

The method works like this:

1. Write
$$\frac{\mathrm{d}x}{a(x,y,u)} = \frac{\mathrm{d}y}{b(x,y,u)} = \frac{\mathrm{d}u}{c(x,y,u)}. \tag{2.3}$$

2. Through manipulating the equalities, try to find two functions $\Phi(x,\ y,\ u)$, $\Psi(x,\ y,\ u)$ such that $\mathrm{d}\Phi = \mathrm{d}\Psi = 0$.

3. The general solution is then given by $F(\Phi, \Psi) = 0$ with F an arbitrary function.

Example 2.1. Find the general solution of the equation

$$a\,u_x + b\,u_y = 0; \qquad a, b \text{ are constants.} \tag{2.4}$$

Solution. The characteristic equations are

$$\frac{\mathrm{d}x}{a} = \frac{\mathrm{d}y}{b} = \frac{\mathrm{d}u}{0}. \tag{2.5}$$

What we need are two functions $\phi(x,y,u)$ and $\psi(x,y,u)$ such that $\mathrm{d}\phi = 0$, $\mathrm{d}\psi = 0$ along the characteristics.

Obviously we can take $\phi = u$. For ψ, notice that

$$\mathrm{d}(a\,y - b\,x) = a\,\mathrm{d}y - b\,\mathrm{d}x = 0, \tag{2.6}$$

thus we can take

$$\psi = a\,y - b\,x. \tag{2.7}$$

As a consequence, the solution satisfies

$$F(a\,y - b\,x, u) = 0 \tag{2.8}$$

for any function F. This means

$$u = f(a\,y - b\,x). \tag{2.9}$$

for an arbitrary function f.

Example 2.2. Find the general solution of

$$y\,u_y - x\,u_x = 2. \tag{2.10}$$

Solution. The characteristic equations are

$$\frac{\mathrm{d}x}{-x} = \frac{\mathrm{d}y}{y} = \frac{\mathrm{d}u}{2}. \tag{2.11}$$

Using

$$\frac{\mathrm{d}x}{-x} = \frac{\mathrm{d}y}{y} \tag{2.12}$$

we obtain

$$y\,\mathrm{d}x + x\,\mathrm{d}y = 0 \implies \mathrm{d}(x\,y) = 0. \tag{2.13}$$

Thus

$$\phi = x\,y; \tag{2.14}$$

On the other hand, from

$$\frac{\mathrm{d}y}{y} = \frac{\mathrm{d}u}{2} \tag{2.15}$$

we obtain

$$\mathrm{d}u = 2\,\mathrm{d}\log y \implies \mathrm{d}(u - 2\log y) = 0. \tag{2.16}$$

As a consequence we can take

$$\psi = u - 2\log y. \tag{2.17}$$

Putting these together we obtain

$$F(x\,y, u - 2\log y) = 0 \tag{2.18}$$

which gives

$$u = 2\log y + f(x\,y). \tag{2.19}$$

Example 2.3. Solve the initial value problem

$$x\,u_x + y\,u_y = u + 1 \qquad \text{with } u(x,y) = x^2 \text{ on } y = x^2. \tag{2.20}$$

Solution. The characteritic equations are

$$\frac{\mathrm{d}x}{x} = \frac{\mathrm{d}y}{y} = \frac{\mathrm{d}u}{u+1} \tag{2.21}$$

which easily lead to

$$\phi = \frac{y}{x}, \quad \psi = \frac{u+1}{x}. \tag{2.22}$$

Thus

$$u = x\,f\left(\frac{y}{x}\right) - 1. \tag{2.23}$$

Now the Cauchy data implies

$$x\,f(x) - 1 = u(x, x^2) = x^2 \tag{2.24}$$

thus

$$f(x) = x + x^{-1}. \tag{2.25}$$

As a consequence

$$u(x,y) = x\left(\frac{y}{x}\right) + x\left(\frac{y}{x}\right)^{-1} - 1 = y + \frac{x^2}{y} - 1. \tag{2.26}$$

Example 2.4. Solve the intial value problem

$$u_t + a(u)\,u_x = 0, \qquad u(x,0) = g(x). \tag{2.27}$$

Solution. We have

$$\frac{\mathrm{d}t}{1} = \frac{\mathrm{d}x}{a(u)} = \frac{\mathrm{d}u}{0} \tag{2.28}$$

which gives $\mathrm{d}u = 0$. Then it follows that $\mathrm{d}(x - a(u)\,t) = 0$. Thus the general solution is $F(u, x - a(u)\,t) = 0$ or $u = f(x - a(u)\,t)$. The solution is given implicitly.

Now setting $t = 0$ we conclude that $f = g$ so the final solution is

$$u = g(x - a(u)\,t). \tag{2.29}$$

We will discuss a lot more about this equation later.

But why dose the method work? In particular, what do we mean by

$$\frac{\mathrm{d}x}{a(x,y,u)} = \frac{\mathrm{d}y}{b(x,y,u)} = \frac{\mathrm{d}u}{c(x,y,u)}, \tag{2.30}$$

and why can we manipulate – add, subtract, multiply by functions, etc. – $\mathrm{d}x, \mathrm{d}y, \mathrm{d}u$ as we did above? To understand this, we need to understand the geometrical meaning of first order equations.

2.1.2. Geometry of first order equations.

Consider the general first-order quasi-linear equation

$$a(x,y,u)\,u_x + b(x,y,u)\,u_y = c(x,y,u). \tag{2.31}$$

Since we are thinking geometrically, we write the above as an orthogonality condition in $\mathbb{R}^3$ with coordinates x, y, u:

$$\begin{pmatrix} a(x,y,u) \\ b(x,y,u) \\ c(x,y,u) \end{pmatrix} \cdot \begin{pmatrix} u_x \\ u_y \\ -1 \end{pmatrix} = 0 \Longleftrightarrow \begin{pmatrix} a(x,y,u) \\ b(x,y,u) \\ c(x,y,u) \end{pmatrix} \perp \begin{pmatrix} u_x \\ u_y \\ -1 \end{pmatrix}. \tag{2.32}$$

Therefore all we need to do is to understand the relation between the vector $\begin{pmatrix} u_x \\ u_y \\ -1 \end{pmatrix}$ and the solution.

Consider the space $\mathbb{R}^3$ with coordinates x, y, u. Let $u(x,y)$ be the solution. Introduce a new function G: $\mathbb{R}^3 \mapsto \mathbb{R}$ through

$$G(x,y,u) = u(x,y) - u. \tag{2.33}$$

Note that in the right hand side of the above, the second u is a variable, the first u is a function. For example, suppose $u(x,y) = x^2 + y^2$, then the corresponding $G(x,y,u) = (x^2 + y^2) - u$.

Now we easily see that

$$\begin{pmatrix} u_x \\ u_y \\ -1 \end{pmatrix} = \begin{pmatrix} G_x \\ G_y \\ G_u \end{pmatrix} = \nabla G. \tag{2.34}$$

Recall that geometrically, ∇G is a normal vector of the surface $G = 0$ which is simply $u = u(x,y)$. As a consequence $\begin{pmatrix} u_x \\ u_y \\ -1 \end{pmatrix}$ is perpendicular to the solution surface $u = u(x,y)$.

On the other hand, from the equation we know that $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is perpendicular to the vector $\begin{pmatrix} u_x \\ u_y \\ -1 \end{pmatrix}$ which means $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ must be tangent to the surface $u = u(x,y)$.

Now we summarize. We have shown that the equation is equivalent to the geometrical requirement in the x-y-u space that the vector $\begin{pmatrix} a(x,y,u) \\ b(x,y,u) \\ c(x,y,u) \end{pmatrix}$ is tangent to the solution surface $u = u(x, y)$. As a consequence, any integral curve of $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, that is any $\begin{pmatrix} x(s) \\ y(s) \\ u(s) \end{pmatrix}$ satisfying

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(x,y,u) \tag{2.35}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = b(x,y,u) \tag{2.36}$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = c(x,y,u) \tag{2.37}$$

or equivalently

$$\frac{\dot{x}}{a} = \frac{\dot{y}}{b} = \frac{\dot{u}}{c}. \tag{2.38}$$

must be contained in one of the solution surfaces. Conversely, any surface "woven" by such integral curves is a solution surface.

The above understanding leads to the following "method of characteristics" due to Lagrange.

Theorem 2.5. *The general solution of a first-order, quasi-linear PDE*

$$a(x,y,u)\, u_x + b(x,y,u)\, u_y = c(x,y,u) \tag{2.39}$$

satisfies

$$F(\Phi, \Psi) = 0, \tag{2.40}$$

where F is an arbitrary function of $\Phi(x, y, u)$ and $\Psi(x, y, u)$, and any intersection of the level sets of Φ and Ψ is a solution of the characteristic equations

$$\frac{\mathrm{d}x}{a} = \frac{\mathrm{d}y}{b} = \frac{\mathrm{d}u}{c}. \tag{2.41}$$

The solutions to this equation are called "characteristics" or "characteristic curves".

Proof. We need to show the implicitly defined $u(x, y)$ satisfies the equation (2.39). To do this we differentiate $F(\Phi, \Psi) = 0$:

$$F_\Phi\, [\Phi_x + \Phi_z\, u_x] + F_\Psi\, [\Psi_x + \Psi_z\, u_x] = 0 \tag{2.42}$$

$$F_\Phi\, [\Phi_y + \Phi_z\, u_y] + F_\Psi\, [\Psi_y + \Psi_z\, u_y] = 0 \tag{2.43}$$

Since $(F_\Phi, F_\Psi) \neq (0, 0)$, we must have

$$\det \begin{pmatrix} \Phi_x + \Phi_z\, u_x & \Psi_x + \Psi_z\, u_x \\ \Phi_y + \Phi_z\, u_y & \Psi_y + \Psi_z\, u_y \end{pmatrix} = 0 \tag{2.44}$$

which simplifies to

$$(\Phi_z\, \Psi_y - \Phi_y\, \Psi_z)\, u_x + (\Phi_x\, \Psi_z - \Phi_z\, \Psi_x)\, u_y = \Phi_y\, \Psi_x - \Phi_x\, \Psi_y. \tag{2.45}$$

Now the fact that Φ, Ψ are solutions to (2.41) implies that

$$a\, \Phi_x + b\, \Phi_y + c\, \Phi_z = 0, \qquad a\, \Psi_x + b\, \Psi_y + c\, \Psi_z = 0. \tag{2.46}$$

which means

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} // \begin{pmatrix} \Psi_x \\ \Psi_y \\ \Psi_z \end{pmatrix} \times \begin{pmatrix} \Phi_x \\ \Phi_y \\ \Phi_z \end{pmatrix} = \begin{pmatrix} \Phi_z\, \Psi_y - \Phi_y\, \Psi_z \\ \Phi_x\, \Psi_z - \Phi_z\, \Psi_x \\ \Phi_y\, \Psi_x - \Phi_x\, \Psi_y \end{pmatrix}. \tag{2.47}$$

Combining (2.45) and (2.47), we reach

$$a\,u_x + b\,u_y = c \tag{2.48}$$

which ends the proof. □

2.1.3. A word on fully nonlinear equations.

The equation

$$a(x, y, u)\,u_x + b(x, y, u)\,u_y = c(x, y, u) \tag{2.49}$$

is called "quasi-linear" because it is linear with regard to the highest order derivatives u_x, u_y. In the most general case, the equation is "fully nonlinear":

$$F(x, y, u, u_x, u_y) = 0. \tag{2.50}$$

In this case it is still possible to obtain a system of ODEs which will lead to solution. There are two ways of getting there:

1. Realize that the essence of the method of characteristics is to study the equation along certain special curves along which the equation reduces to a system of ordinary differential equations and can be solved explicitly. Thus we try to figure out what kind of special curve $(x(s), y(s), u(s), p(s), q(s))$ can fulfill this.
2. From the geometrical point of view, we see that a fully nonlinear equation is not a simple orthogonality relation, but a nonlinear constraint for the normal vector $(u_x, u_y - 1)$ at every point (x, y, u). This constraint means at each point (x, y, u), the vectors that are penpendicular to a normal vector must belong to a cone (the quasilinear case this cone becomes a plane), called the *Monge cone*. To find a solution, is to find a surface which is tangent to the Monge cones at every point.

We will solve fully nonlinear equations using the first approach in Section 2.5, and leave the second approach to the "Problems" section.

References.

- G. Evans, J. Blackledge, P. Yardley "Analytic Methods for Partial Differential Equations", §3.2

Exercises.

Exercise 2.1. Solve

$$a\,u_x + b\,u_y + c\,u - d = 0 \tag{2.51}$$

with a, b, c, d constants.

Exercise 2.2. Find the solution of the following Cauchy problems.

a) $x\,u_x + y\,u_y = 2\,x\,y$, with $u = 2$ on $y = x^2$.

b) $u\,u_x - u\,u_y = u^2 + (x + y)^2$ with $u = 1$ on $y = 0$.

Exercise 2.3. Find the general solution of the following equations.

a) $y^2\,u_x - x\,y\,u_y = x\,(u - 2\,y)$.

b) $y\,z\,u_x - x\,z\,u_y + x\,y\,(x^2 + y^2)\,u_z = 0$.

Exercise 2.4. Consider a quasi-linear equation

$$a(x, y, u)\,u_x + b(x, y, u)\,u_y = c(x, y, u) \tag{2.52}$$

(without specifying any initial conditions). Let $u_1(x, y)$, $u_2(x, y)$ be two solutions. Assume that the surfaces $u - u_1(x, y) = 0$ and $u - u_2(x, y) = 0$ intersects along a curve Γ in the xyu space. Show that Γ must be a characteristic curve.

2.2. Initial Value Problem and Classical Solutions to Conservation Laws.

2.2.1. Initial value problem.

When discussing qualitative properties of the initial value problem,

$$a(x, y, u)\, u_x + b(x, y, u)\, u_y = c(x, y, u), \qquad u = u_0(\tau) \text{ along } x = x_0(\tau), y = y_0(\tau), \tag{2.53}$$

it is beneficial to take a more explicit approach.

1. Solve the ODE system

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(x, y, u) \tag{2.54}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = b(x, y, u) \tag{2.55}$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = c(x, y, u) \tag{2.56}$$

   with initial conditions

$$u(0, \tau) = u_0(\tau), x(0, \tau) = x_0(\tau), y(0, \tau) = y_0(\tau) \tag{2.57}$$

   to obtain

$$u = u(s, \tau), x = x(s, \tau), y = y(s, \tau). \tag{2.58}$$

2. Try to invert the relation:

$$x = x(s, \tau), y = y(s, \tau) \Longrightarrow s = S(x, y), \tau = \mathrm{T}(x, y). \tag{2.59}$$

   Substitute into u to obtain $u = u(S(x, y), \mathrm{T}(x, y))$.

Here what we are doing is, at every point $\tau = \tau_0$ on the curve $(x_0(\tau), y_0(\tau), u_0(\tau))$, we obtain another curve $(x(s, \tau_0), y(s, \tau_0), u(s, \tau_0))$ through solving the characteristics equation. Then all these curves would "weave" into a surface.

Remark 2.6. Consider the equation

$$a(x, y, u)\, u_x + b(x, y, u)\, u_y = c(x, y, u). \tag{2.60}$$

Consider the projection of the characteristics onto the x-y plane:

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(x, y, u); \qquad \frac{\mathrm{d}y}{\mathrm{d}s} = b(x, y, u). \tag{2.61}$$

From now on we will also call these curves in the x-y plane "characteristics". Note that along these characteristics, the equation reduces to (thanks to chain rule):

$$\frac{\mathrm{d}u}{\mathrm{d}s} = c(x, y, u). \tag{2.62}$$

So in conclusion the equation reduces to an ODE along characteristics. Note that when the equation is semilinear

$$a(x, y)\, u_x + b(x, y)\, u_y = c(x, y, u) \tag{2.63}$$

we can indeed first solve the characteristics first and then solve the ODE for u.

We illustrate the method with the following example.

Example 2.7. Solve

$$u\, u_x + u_y = 2, \qquad u(x, x) = x/2. \tag{2.64}$$

Solution. First to make things clearer, we write the problem as

$$u\,u_x + u_y = 2, \qquad u(\tau, \tau) = \frac{\tau}{2}. \tag{2.65}$$

The system for characteristics is

$$\frac{\mathrm{d}x}{\mathrm{d}s} = u \qquad x(0, \tau) = \tau \tag{2.66}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = 1 \qquad y(0, \tau) = \tau \tag{2.67}$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = 2 \qquad u(0, \tau) = \tau. \tag{2.68}$$

Studying the system we see that the u, y equations can be solved first:

$$u(s, \tau) = \tau + 2\,s; \qquad y(s, \tau) = \tau + s. \tag{2.69}$$

With $u(s, \tau)$ solved, we can obtain

$$x(s, \tau) = \tau + \tau\,s + s^2. \tag{2.70}$$

Now we need to write s, τ as functions of x, y. Substitute $s = y - \tau$ into the x formula:

$$x = \tau + \tau\,(y - \tau) + (y - \tau)^2 = y^2 + (1 - y)\,\tau \Longrightarrow \tau = \frac{x - y^2}{1 - y}. \tag{2.71}$$

Then

$$s = y - \tau = y - \frac{x - y^2}{1 - y} = \frac{y - x}{1 - y}. \tag{2.72}$$

Back to u, we have

$$u(s, \tau) = \tau + 2\,s = \frac{x - y^2}{1 - y} + 2\,\frac{y - x}{1 - y} = \frac{2\,y - y^2 - x}{1 - y}. \tag{2.73}$$

Let's discuss a bit about the existence and uniqueness of solutions. Our method consists of two steps, first solve the system

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(x, y, u) \tag{2.74}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = b(x, y, u) \tag{2.75}$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = c(x, y, u) \tag{2.76}$$

and then try to invert

$$x = x(s, \tau), y = y(s, \tau) \Longrightarrow s = S(x, y), \tau = \mathrm{T}(x, y). \tag{2.77}$$

If we assume that a, b, c are differentiable, then the first step can always be carried out. On the other hand, for the inversion we need certain conditions on the initial curve. One situation where the existence of S, T is guaranteed, is when Implicit Function Theorem applies, that is when

$$\det\begin{pmatrix} x_s & x_\tau \\ y_s & y_\tau \end{pmatrix} \neq 0 \tag{2.78}$$

along $s = 0$. It's important to realize that this determinant can be evaluated before solving the equation: Along $s = 0$, we have $x = x_0(\tau), y = y_0(\tau)$ which gives x_τ, y_τ. On the other hand using the equations for x, y we have $x_s = a, y_s = b$. As a consequence we have

Theorem 2.8. *Consider the initial value problem for quasilinear equation*

$$a(x, y, u)\,u_x + b(x, y, u)\,u_y = c(x, y, u), \qquad u(0, \tau) = u_0(\tau) \text{ along } (x_0(\tau), y_0(\tau)). \tag{2.79}$$

If

$$\det\begin{pmatrix} a(x_0(\tau), y_0(\tau), u_0(\tau)) & x_0'(\tau) \\ b(x_0(\tau), y_0(\tau), u_0(\tau)) & y_0'(\tau) \end{pmatrix} \neq 0 \tag{2.80}$$

for all τ, then there exists a unique solution in a small neighborhood of the initial curve.

Remark 2.9. It is fairly easy to convince yourself that if the initial condition is given along a characteristic curve, then there is either no solution or infinitely many solutions. On the other hand, if the initial curve $(x_0(\tau), y_0(\tau))$ only fails (2.80) at several isolated points, then a unique solution is still possible, although the solution may not be continuously differentiable everywhere.

2.2.2. Conservation laws.

A special class of initial value problems that is very important in science and engineering is the conservation law. Usually conservation laws are derived as follows. Consider the evolution of the density u of a certain substance:

- The total amount inside a set Ω at time t:

$$\int_\Omega u(t, x)\,\mathrm{d}x. \tag{2.81}$$

- Assumption: Change only occurs as this substance goes through the boundary. Quantified by a "flux" F. That is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u(t, x)\,\mathrm{d}x = -\int_{\partial\Omega} F \cdot n\,\mathrm{d}s \tag{2.82}$$

  where n is the outer normal.

- Gauss Theorem:

$$\int_{\partial\Omega} F \cdot n\,\mathrm{d}s = \int_\Omega \nabla \cdot F\,\mathrm{d}x \tag{2.83}$$

- We have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u(t, x)\,\mathrm{d}x + \int_\Omega \nabla \cdot F\,\mathrm{d}x = 0 \tag{2.84}$$

  as $\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega u(t, x) = \int_\Omega u_t\,\mathrm{d}x$, we reach

$$\int_\Omega [u_t + \nabla \cdot F]\,\mathrm{d}x = 0. \tag{2.85}$$

- As Ω is arbitrary, when $u_t + \nabla \cdot F$ is assumed to be regular enough, we have a differential equation

$$u_t + \nabla \cdot F = 0. \tag{2.86}$$

A *conservation law* is obtained when F is a function of u only:

$$u_t + \nabla \cdot F(u) = 0, \qquad u(0, x) = g(x) \tag{2.87}$$

Remark 2.10. Note that the above is in fact a very general framework. For example, if we think of ρ as temperature, then F according to Fourier's law should be $-\kappa\nabla\rho$, this leads to the heat equation.

2.2.3. Solving 1D conservation laws using method of characteristics.

Conservation laws in 2D and higher dimensions is still at the forefront of PDE research and far from well-understood. On the other hand, the theory for 1D conservation laws is much more complete. In particular, a 1D scalar conservation law

$$u_t + f(u)_x = 0, \qquad u(0, x) = g(x) \tag{2.88}$$

is almost fully understood. We will restrict ourselves to this case.

Denote $a(u) = f'(u)$, the equation can be written as

$$u_t + a(u)\, u_x = 0. \tag{2.89}$$

This is a very simple quasi-linear equation so we try to solve it using method of characteristics.

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1, \qquad t(0) = 0 \tag{2.90}$$

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(u), \qquad x(0) = x_0 \tag{2.91}$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = 0, \qquad u(0) = g(x_0) \tag{2.92}$$

The solution is given by

$$u(t, x) = g(x_0) \tag{2.93}$$

with

$$x = x_0 + a(g(x_0))\, t \tag{2.94}$$

Combine these we get an implicit formula for u:

$$u = g(x - a(u)\, t). \tag{2.95}$$

Although under certain conditions the implicit function theorem gives the existence of a uniquely determined u, we cannot really write it down explicitly.

However we can compute u_x explicitly:

$$u_x = g'\,(1 - a'\,t\,u_x) \Longrightarrow u_x = \frac{g'}{1 + g'\,a'\,t}. \tag{2.96}$$

We see that there is the possibility that u_x will become infinity – when $g'\,a' < 0$. This "blow-up" is very easy to understand when we look at the x-t plane characteristics:

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1, \qquad t(0) = 0 \tag{2.97}$$

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(u), \qquad x(0) = x_0 \tag{2.98}$$

Since u is unknown, this system seems cannot be solved. However for conservation law we have $\frac{\mathrm{d}u}{\mathrm{d}s} = 0$ which means along each characteristic curve $a(u)$ is a constant – the characteristics are straight lines!

Thus we see that a conservation law depicts the "carrying" of the initial values along straight lines whose slope are $1/a(u)$ (when discussing conservation laws we often use the "speed" $\frac{\mathrm{d}x}{\mathrm{d}t} = a(u)$ instead of slope, some books, such as G. B. Whitham "Linear and Nonlinear Waves", draw the coordinates such that t is the horizontal direction and x the vertical, to be able to keep talking about slopes). As $a(u)$ changes with u, characteristics corresponding to different u may cross each other, leading to the blow-up in finite time. In other words, the classical solution cannot be defined beyond a certain time.

Example 2.11. Consider the Burgers equation (the case $a(u) = u$)

$$u_t + u\, u_x = 0 \tag{2.99}$$

The solution is

$$u = g(x - u\, t). \tag{2.100}$$

Taking derivatives we get

$$u_x = \frac{g'(x_0)}{1 + t\, g'(x_0)}, \qquad u_t = -\frac{g'(x_0)\, g(x_0)}{1 + t\, g'(x_0)}. \tag{2.101}$$

We see that the maximum time for u to be differentiable is

$$T = \begin{cases} \left[\max_x \{-g'(x)\}\right]^{-1} & g'(x) < 0 \text{ for some } x \\ \infty & g'(x) \geqslant 0 \text{ for all } x \end{cases}. \tag{2.102}$$

Now consider

$$u(0, x) = g(x) = -x \tag{2.103}$$

We see that the maximum time for existence is 1, which can be verified by drawing characteristics. In fact we can solve the equation for this initial value to obtain

$$u(x, t) = \frac{x}{t-1} \tag{2.104}$$

References.

- Erich Zauderer, "Partial Differential Equations of Applied Mathematics", 2ed, §2.2, §2.3.

Exercises.

Exercise 2.5. (Zauderer) Solve the initial value problem

$$u_t + u_x = 0, \qquad u = g(t) \text{ along } x = t^2. \tag{2.105}$$

a) Does the solution exist for all (x, t)? If not, for what $g(t)$ does it exist for all (x, t)?

b) For those $g(t)$ guaranteeing the existence of $u(x,\ t)$ for all $(x,\ t)$, do u_x and u_t remain bounded everywhere? If not, specify the location where u_x or u_t (or both) becomes infinite, and explain.

Exercise 2.6. (Zauderer) Solve the following equations using method of characteristics. Discuss what caution should be taken when assigning initial conditions.

$$x\, u_x + t\, u_t = -u; \qquad t\, u_x - x\, u_t = u. \tag{2.106}$$

Exercise 2.7. (Zauderer) Show that the initial value problem

$$u_t + u_x = 0, \qquad u = x \text{ on } x^2 + t^2 = 1 \tag{2.107}$$

has no solution. However, if the initial data are given only over the semicircle that lies in the half-plane $x + t \leqslant 0$, the solution exists but is not differentiable along the characteristic base curves that issue from the two end points of the semicircle.

Exercise 2.8. (Zauderer) Solve the initial value problem

$$v_t + e^x\, v_x = 0, \qquad v(x, 0) = x. \tag{2.108}$$

Exercise 2.9. (Zauderer) Show that the initial value problem

$$(t - x)\, u_x - (t + x)\, u_t = 0, \qquad u(x, 0) = f(x), \qquad x > 0 \tag{2.109}$$

has no solution if $f(x)$ is an arbitrary function.

Exercise 2.10. (Zauderer) Show that the initial value problem

$$u_t + u_x = x, \qquad u(x, x) = 1 \tag{2.110}$$

has no solution. Observe that the initial curve $t = x$ is a characteristic base curve and explain why this is not a characteristic initial value problem.

Exercise 2.11. Consider the quasilinear initial value problem

$$a(x,y,u)\,u_x+b(x,y,u)\,u_y=c(x,y,u),\qquad u=u_0(\tau)\text{ along }(x_0(\tau),y_0(\tau)). \tag{2.111}$$

We know that in general the solution can only exist in a small neighborhood of the initial curve and may develop singularity a finite distance from it.

However, consider the method: We obtain the solution surface by "weaving" together solutions to

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(x,y,u) \tag{2.112}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = b(x,y,u) \tag{2.113}$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = c(x,y,u) \tag{2.114}$$

which, according to the existence/uniqueness theory of ODEs, should exist for all s if we assume a,b,c to be Lipschitz functions.

Does this contradict the fact that the solution surface only exists in a neighborhood of the initial curve? If not, why?

Exercise 2.12. (Zauderer) Consider the wave equation

$$u_{tt}-u_{xx}=0,\qquad u(x,0)=g(x),\quad u_t(x,0)=h(x). \tag{2.115}$$

Show that

a) If we set $v(x,t)=u_t-u_x$, then v satisfies

$$v_t+v_x=0,\qquad v(x,0)=h(x)-g'(x). \tag{2.116}$$

b) Use method of characteristics to solve the v equation and then the u equation. Show that the solution is given by the d'Alembert's formula

$$u(x,t)=\frac{1}{2}\,[g(x+t)+g(x-t)]+\frac{1}{2}\int_{x-t}^{x+t}h(s)\,\mathrm{d}s. \tag{2.117}$$

Exercise 2.13. (Zauderer) Solve the initial value problem

$$u_t+u\,u_x=x,\qquad u(x,0)=f(x) \tag{2.118}$$

for $f(x)=1$ and $f(x)=x$.

Exercise 2.14. Consider $u_t+f(u)_x=0$, with $u(x,\,0)=g(x)$. Find the maximum time for u to be differentiable.

Exercise 2.15. Analyze

$$u_t-u^2\,u_x+c\,u=0,\qquad u(x,0)=g(x). \tag{2.119}$$

Exercise 2.16. Solve the Burgers equation

$$u_t+u\,u_x=0,\qquad u(x,0)=1-x^2. \tag{2.120}$$

Exercise 2.17. Consider the first order quasi-linear equation in higher dimensions:

$$a_1(x_1,\dots,x_n)\,u_{x_1}+\cdots+a_n\,(x_1,\dots,x_n)\,u_{x_n}=a(x_1,\dots,x_n). \tag{2.121}$$

Develop a theory for its initial value problem and obtain a theorem similar to Theorem 2.8.

2.3. Weak Solutions to Conservation Laws.

2.3.1. Weak solutions.

We have seen that when solving conservation law

$$u_t + f(u)_x = 0, \qquad u(x,0) = g(x) \tag{2.122}$$

or equivalently[2.1]

$$u_t + a(u)\, u_x = 0, \qquad u(x,0) = g(x), \tag{2.123}$$

the solution in general may develop singularities in finite time. In other words, u_x and u_t will become infinite. In fact, one may continue the solution along the characteristics, thus allowing the solution to be multi-valued. However this clearly is not appropriate for most situations modelled by conservation laws – for example traffic flows can be modelled by conservation laws, where allowing multi-valued solutions would be allowing cars to pile up!

There are two possible ways to get out of this situation.

- Accept the fact as an indication that the model is not an accurate reflection of physical reality, and try to revise the model.
- Try to "generalize" the idea of solutions, define "weak" solutions which exists for all time.

Either one could be a good choice. It is indeed possible to avoid finite time singularities through more accurate models obtained from putting back small terms that have been neglected. However, those models are all much more complicated, and more importantly, a good understanding of the classical solutions to those models in fact requires a good understanding of discontinuous solutions to (2.122)!

Now we give the definition of "weak solutions" which is allowed to be discontinuous:

Definition 2.12. (Weak solution for conservation laws) *A locally integrable function u is called a weak solution of*

$$u_t + f(u)_x = 0, \qquad u(x,0) = u_0 \tag{2.124}$$

if

$$\int\int_{t>0} [u\, \phi_t + f(u)\, \phi_x]\, \mathrm{d}x\, \mathrm{d}t + \int_{\mathbb{R}} u_0\, \phi\, \mathrm{d}x = 0. \tag{2.125}$$

holds for every $\phi \in C_0^1(\mathbb{R} \times \mathbb{R}^+)$.

Here C_0^1 denotes functions that are continuously differentiable and compact supported.[2.2]

Remark 2.13. With some additional assumptions, one can show the equivalence of (2.125) to the following integral relation

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_a^b u\, \mathrm{d}x + f(u(b,t)) - f(u(a,t)) = 0 \tag{2.126}$$

which is exactly how we derive conservation laws in the first place! The reason why we use (2.125) instead of (2.126) is that on one hand the latter requires more regularity on u, on the other the seemingly more complicated (2.125) is in fact more efficient in complicated situations.

It can be shown that the "weak solutions" is indeed a more general concept than classical solutions. Here by "classical solution" we mean a solution $u(x,t) \in C^1$ whose derivatives u_t and u_x indeed satisfies the equation everywhere.

Lemma 2.14. *Any classical solution is a weak solution. On the other hand, let* $u \in C^1$ *be a weak solution, then it is also a classical solution.*

Proof. For the first part simply multiply the equation by $\phi \in C_0^1$ and integrate by parts; For the second, start from the definition of weak solutions, integrate by parts to put derivatives to u. The details are left as exercises. ☐

2.1. Equivalent only when the solution is continuously differentiable! That is, the two formulations are equivalent only when we are discussing classical solutions.

2.2. A function $f(x)$ is said to have compact support if its support – the set $\{x: f(x) \neq 0\}$ – is compact, that is if there is $R > 0$ such that $f(x) = 0$ for all $|x| > R$.

As soon as we allow discontinuous functions to be solutions, we have to allow such initial conditions too. As can be seen in Section 2.4, it may be very physical to consider such initial conditions. Let's look at two examples which indicate the key issue that our weak solution theory should deal with.

Example 2.15. Solve

$$u_t + u\,u_x = 0, \qquad u(x,0) = g(x) \tag{2.127}$$

with

$$g(x) = \begin{cases} 0 & x<0 \\ 1 & x>0 \end{cases}, \text{ and } g(x) = \begin{cases} 1 & x<0 \\ 0 & x>0 \end{cases}. \tag{2.128}$$

using characteristics.

We easily see that for the first initial condition, we only have

$$u(x,t) = \begin{cases} 0 & x<0 \\ ? & 0<x<t \\ 1 & x>t \end{cases} \tag{2.129}$$

while for the second initial condition we have multi-valuedness right from $t>0$.

Thus the theory of weak solution needs to solve two problems:

1. The method of characteristics may leave regions of solution undefined.
2. The method of characteristics may cause ambiguity in the value of u.

In the following we will fix these through introducing rarefaction waves and shock waves.

- Filling the void: Rarefaction waves.

  It turns out that to fill the void, it suffices to "fill in" a "patch" of solution of the form $u = U(x/t)$. Consider the example:

  $$u_t + u\,u_x = 0, \qquad u(x,0) = g(x) = \begin{cases} 0 & x<0 \\ 1 & x>0 \end{cases}. \tag{2.130}$$

  We look for solutions of the form $U(x/t)$. Substituting into the equation we get

  $$-t^{-2} x\, U' + U\,(t^{-1} U') = 0 \Longrightarrow U(x/t) = x/t. \tag{2.131}$$

  This gives

  $$u(t,x) = \begin{cases} 0 & x<0 \\ x/t & 0<x<t \\ 1 & x>t \end{cases} \tag{2.132}$$

  One can check that it is indeed a weak solution.

  Remark 2.16. In general, if a "void fan" starts at (x_0, t_0), the solution to fill in should look like $U\left(\frac{x - x_0}{t - t_0}\right)$.

  We will discuss more in Section 2.4.

- Resolving the ambiguity: Shock waves.
  - Instead of letting characteristics cross, we introduce a discontinuity curve which "cuts" the characteristics. That is, we try to get u which is piecewise C^1 with jump discontinuities along certain curves. Turns out that the requirement of u being a weak solution totally determines these curves.
  - Consider one such curve, denote it by Γ. Let $\phi \in C_0^1$ be supported in a small ball centering on Γ. The ball is so small that it does not intersect with the x-axis and u is C^1 everywhere in the ball except along Γ.

    Denote this ball by D, which is divided into two parts D_1, D_2 by Γ. As $\phi = 0$ along the x-axis, the definition of weak solutions becomes

    $$\int\int_D u\,\phi_t + f(u)\,\phi_x \,\mathrm{d}x\,\mathrm{d}t = 0. \tag{2.133}$$

We write the left hand side as $\int\int_{D_1} + \int\int_{D_2}$ and try to use integration by parts.

Since u is C^1 in D_1, D_2, we have

$$\int\int_{D_1} u\,\phi_t + f(u)\,\phi_x\,\mathrm{d}x\,\mathrm{d}t = -\int\int [u_t + f(u)_x]\,\phi\,\mathrm{d}x\,\mathrm{d}t + \oint_{\partial D_1} [u\,n_t + f(u)\,n_x]\,\phi\,\mathrm{d}S \tag{2.134}$$

Since u solves the equation in the classical sense in D_1 (see exercise) we have

$$\int\int_{D_1} u\,\phi_t + f(u)\,\phi_x\,\mathrm{d}x\,\mathrm{d}t = \oint_{\partial D_1} [u\,n_t + f(u)\,n_x]\,\phi\,\mathrm{d}S. \tag{2.135}$$

Similarly

$$\int\int_{D_2} u\,\phi_t + f(u)\,\phi_x\,\mathrm{d}x\,\mathrm{d}t = \oint_{\partial D_2} [u\,n_t + f(u)\,n_x]\,\phi\,\mathrm{d}S. \tag{2.136}$$

Since ϕ vanishes on ∂D_1 except along Γ, we finally obtain

$$\int_\Gamma [[u]\,n_t + [f(u)]\,n_x]\,\phi\,\mathrm{d}S = 0 \tag{2.137}$$

where $[u]$ is the "jump" of u across Γ.

Now let Γ be determined by $\frac{\mathrm{d}x}{\mathrm{d}t} = s(x,t)$. We have $\frac{n_t}{n_x} = -s$ which gives

$$\int_\Gamma [-s\,[u] + [f(u)]]\,\phi\,\mathrm{d}S = 0. \tag{2.138}$$

Due to the arbitrariness of ϕ, the weak solution must satisfy

$$[f(u)] = s\,[u]. \tag{2.139}$$

This is called the *jump condition*.

In the special case of gas dynamics, this condition is referred to as *Rankine-Hugoniot condition*.

Theorem 2.17. *Consider the scalar conservation law*

$$u_t + f(u)_x = 0, \qquad u(x,0) = g(x). \tag{2.140}$$

If a function $u(x,t)$ satisfies:

i. It is piecewise smooth;

ii. It satisfies the jump condition along curves of discontinuity;

iii. It satisfies the initial condition,

then $u(x,t)$ is a weak solution to the problem.

2.3.2. Entropy conditions.

It turns out that weak solutions are in general not unique. For example, consider the Burgers equation with initial data $u(x,0) = \begin{cases} 0 & x<0 \\ 1 & x>0 \end{cases}$, it turns out that both

$$u_1(x,t) = \begin{cases} 0 & x < t/2 \\ 1 & x > t/2 \end{cases}, \qquad u_2(x,t) = \begin{cases} 0 & x<0 \\ x/t & 0<x<1 \\ 1 & x>t \end{cases} \tag{2.141}$$

are weak solutions.

The fix to this situation is the introduction of the so-called "entropy" condition

$$\frac{u(x+a,t) - u(x,t)}{a} \leqslant \frac{E}{t}, \qquad \forall a>0, t>0 \tag{2.142}$$

where E is independent of x, t. A solution satisfying this entropy condition is called an "entropy solution". For an entropy solution, if it has a discontinuity, then necessarily $u_l > u_r$. Since we are considering the case $f'' > 0$, we always have

$$f'(u_l) > s > f'(u_r) \tag{2.143}$$

where s is the speed of the discontinuity (that is, the discontinuity is the curve $\frac{\mathrm{d}x}{\mathrm{d}t} = s(x,t)$).

Remark 2.18. If we draw the characteristics, the entropy condition requires characteristics to "meet" at the discontinuity instead of "emanating" from it. Physically speaking, each characteristic curve is a carrier of information, the requirement that they "meet" at any discontinuity is the same as saying information must decrease across any shocks. This is consistent with the Second Law of thermodynamics. This point of view helps in appreciating the following discussion on the irreversibility of entropy solutions.

We can also give the following mathematical justification of the entropy condition.

Remark 2.19. We consider the following situations regarding the relations between $f'(u_l)$, $f'(u_r)$ and s.

- $f'(u_l) < \dot{s}$. In this case there are characteristics starting from the left side of $x = s(t)$ and entering into Ω_L, Therefore we need one condition at the left side of the discontinuity;
- $f'(u_l) > \dot{s}$. No condition along the left side is needed;
- $f'(u_r) < \dot{s}$. No condition along the right side of the discontinuity is needed;
- $f'(u_r) > \dot{s}$. One condition along the right side is needed.

Now for the scalar conservation law, the jump condition

$$\dot{s} = \frac{f(u_l) - f(u_r)}{u_l - u_r} \tag{2.144}$$

is only one equation and therefore all it can do is to determine $\dot{s}$. As a consequence, we do not have any extra condition and the solution is determined only when $f'(u_l) > \dot{s}$ and at the same time $f'(u_r) < \dot{s}$.

Example 2.20. Consider

$$u_t + u\,u_x = 0, \qquad u(x,0) = \begin{cases} 1 & x < 0 \\ \dfrac{1-x}{\varepsilon} & 0 < x < \varepsilon \\ 0 & x > \varepsilon \end{cases}. \tag{2.145}$$

References.

- L. C. Evans, "Partial Differential Equations".

Exercises.

Exercise 2.18. Construct entropy solutions for the following initial value problems

a)

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \qquad u(0,x) = \begin{cases} 1 & x < 3 \\ 0 & x > 3 \end{cases} \tag{2.146}$$

b)

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \qquad u(0,x) = \begin{cases} 2 & x < 0 \\ 1 & 0 < x < 1 \\ 0 & x > 1 \end{cases} \tag{2.147}$$

c)

$$u_t + \left(\frac{u^4}{4}\right) x = 0, \qquad u(0,x) = \begin{cases} 1 & x < 0 \\ 0 & x > 0 \end{cases}. \tag{2.148}$$

Exercise 2.19. Compute explicitly the unique entropy solution of

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \qquad u(0,x) = g \tag{2.149}$$

for

$$g(x) = \begin{cases} 1 & x < -1 \\ 0 & -1 < x < 0 \\ 2 & 0 < x < 1 \\ 0 & x > 1 \end{cases}. \tag{2.150}$$

Draw a picture of your answer. Be sure to illustrate what happens for all times $t>0$.

Exercise 2.20. Prove that any classical solution is a weak solution. On the other hand, let $u\in C^1$ be a weak solution, then it is in fact a classical solution.

Exercise 2.21. Consider the following approach for the construction of weak solution for the scalar conservation law

$$u_t+f(u)_x=0, \qquad u(x,0)=g(x). \tag{2.151}$$

- Write the equation as $u_t+a(u)\,u_x=0$. Multiply the equation by $a'(u)$ and set $v=a(u)$, we reach

$$v_t+v\,v_x=0, \qquad v(x,0)=a(g(x)). \tag{2.152}$$

- Then apply the theory of Burgers equation.

Exercise 2.22. Prove that

$$u(t,x)=\begin{cases} 0 & x<0 \\ x/t & 0<x<t \\ 1 & x>t \end{cases} \tag{2.153}$$

is a weak solution to the problem

$$u_t+u\,u_x=0, \qquad u(0,x)=\begin{cases} 0 & x<0 \\ 1 & x>0 \end{cases}. \tag{2.154}$$

Exercise 2.23. Consider the scalar conservation law

$$u_t+f(u)_x=0, \qquad u(x,0)=\begin{cases} 0 & x<0 \\ 1 & x>0 \end{cases}. \tag{2.155}$$

Construct the entropy solution for the problem for the following two cases: $f''(u)>0$ everywhere, $f''(u)<0$ everywhere.

Exercise 2.24. (Evans) Assume $f(0)=0$, u is a continuous weak solution of the conservation law

$$u_t+f(u)_x=0, \qquad u(0,x)=g \tag{2.156}$$

and u has compact support in $\mathbb{R}\times[0,T]$ for each $T>0$ (meaning: for each $T>0$ there is R – may depend on T – such that $u=0$ outside $[-R,R]\times[0,T]$. Prove

$$\int_{-\infty}^{\infty} u(t,x)\,\mathrm{d}x=\int_{-\infty}^{\infty} g(x)\,\mathrm{d}x. \tag{2.157}$$

for all $t>0$.

Exercise 2.25. (Evans) Show that

$$u(t,x)=\begin{cases} -\frac{2}{3}\left(t+\sqrt{3\,x+t^2}\right) & 4\,x+t^2>0 \\ 0 & 4\,x+t^2<0 \end{cases} \tag{2.158}$$

is an (unbounded) entropy solution of $u_t+\left(\frac{u^2}{2}\right)_x=0$.

Exercise 2.26. Consider the following problem

$$u_t+\left(\frac{u^2}{2}\right)_x=0, \qquad u(0,x)=\begin{cases} 1 & x<-\varepsilon \\ \frac{\varepsilon-x}{2\,\varepsilon} & -\varepsilon<x<\varepsilon \\ 0 & x>\varepsilon \end{cases} \tag{2.159}$$

Construct the entropy solution and study what happens when $\varepsilon\searrow 0$.

2.4. Traffic flow.

One way to model the traffic on a road is as a fluid flow and describe it by means of density of cars ρ, average speed v and flux f. It is intuitive that we have

$$f = v\,\rho. \tag{2.160}$$

Now if we assume that the speed only depends on the density ρ, we would reach

$$f = f(\rho) = v(\rho)\,\rho. \tag{2.161}$$

To obtain a scalar conservation law, we need the following assumptions:

1. There is only one lane and overtaking is not allowed.
2. No car "sources" or "sinks" – no exit or entrance.

Under these assumptions we could obtain the following model for traffic:

$$\rho_t + f(\rho)_x = 0. \tag{2.162}$$

We may attach to it an initial condition

$$\rho(x, 0) = g(x). \tag{2.163}$$

Remark 2.21. It is clear that these assumptions are either very restrictive, or even unrealistic, in particular that $v = v(\rho)$. However as we will see, such a simple model can already reveal interesting phenomena consistent with everyday observation. On the other hand, one can try to make the model more realistic by relaxing or even dropping one or more assumptions.

Note. It is clear from common sense that

$$v'(\rho) \leqslant 0. \tag{2.164}$$

So in some sense the behavior of the equation will be "opposite" to that of Burgers, which is the paradigm example for the case $v'(\rho) > 0$. Those who have developed some intuitive understanding of the Burgers equation should keep this difference in mind.

In the following we will consider the simplest model for $v(\rho)$:

$$v(\rho) = v_m\left(1 - \frac{\rho}{\rho_m}\right). \tag{2.165}$$

In this case we have

$$f(\rho) = v_m\,\rho\left(1 - \frac{\rho}{\rho_m}\right). \tag{2.166}$$

2.4.1. Characteristics.

Using the above $f(\rho)$ the equation can be written as

$$\rho_t + v_m\left(1 - \frac{2\,\rho}{\rho_m}\right)\rho_x = 0 \tag{2.167}$$

which can be easily solved to get

$$\rho(x, t) = g(x - f'(g(x_0))\,t) \tag{2.168}$$

or, using the explicit formula for f',

$$\rho(x, t) = g\left(x - v_m\left(1 - \frac{2\,g(x_0)}{\rho_m}\right)t\right). \tag{2.169}$$

We see that this is wave propagation with velocity

$$v_m\left(1 - \frac{2\,g(x_0)}{\rho_m}\right). \tag{2.170}$$

It is important to understand that this represents the velocity of any disturbance and is different from the velocity of each car, which is given by

$$v(\rho) = v_m\left(1 - \frac{\rho}{\rho_m}\right) = v_m\left(1 - \frac{g(x_0)}{\rho_m}\right) \tag{2.171}$$

for the car starting at x_0. Comparing these two speeds we see clearly that the "wave velocity" is smaller than that of the individual car. This is consistent with everyday experience.

From theory of conservation law we know that classical solutions "blow-up" in finite time, which in this context would mean clash between cars. However this does not always happen in real life. One way to resolve this is to consider weak solutions. In fact, in the traffic problem consideration of discontinuous solutions is imperative, as can be seen from the following.

2.4.2. The Green light problem.

Suppose that bumper-to-bumper traffic is standing at a red light, placed at $x=0$, while the road ahead is empty. This gives the initial condition

$$g(x)=\begin{cases}\rho_m & x\leqslant 0\\ 0 & x>0\end{cases}. \tag{2.172}$$

At $t=0$ the traffic light turns green and we try to understand the evolution of car-density.

Recall that $a(\rho):=f'(\rho)=v_m\left(1-\frac{2\rho}{\rho_m}\right)$, we have

$$f'(g(x_0))=\begin{cases}-v_m & x_0\leqslant 0\\ v_m & x_0>0\end{cases} \tag{2.173}$$

so we have to fit in a rarefaction wave. The solution is then

$$\rho(x,t)=\begin{cases}0 & x\geqslant v_m t\\ r(x/t) & -v_m t<x<v_m t\\ \rho_m & x\leqslant -v_m t\end{cases} \tag{2.174}$$

where $r(\xi)=\frac{1}{2}\left(1-\frac{\rho_m}{v_m}\xi\right)$ is the inverse function of $a(\cdot)$.

2.4.3. Traffic jam problem.

Let's say the initial density is

$$g(x)=\begin{cases}\rho_m/8 & x<0\\ \rho_m & x>0\end{cases}. \tag{2.175}$$

In this case we have

$$f'(g(x_0))=\begin{cases}\dfrac{3}{4}v_m & x_0<0\\ -v_m & x_0>0\end{cases} \tag{2.176}$$

which leads to the characteristics

$$x=\frac{3}{4}v_m t+x_0, \qquad x=-v_m t+x_0 \tag{2.177}$$

for $x_0<0$ and $x_0>0$ respectively.

We see that the characteristics cross as soon as $t>0$, and the initial cross is at $x=0$. Consequently, we put in a shock emanating from $(0,0)$ to resolve this.

Applying the jump condition we see that

$$\dot{s}(t)=-\frac{1}{8}v_m t. \tag{2.178}$$

This is also qualitatively consistent with everyday experience.

Remark 2.22. Note that red light can be treated as special case of traffic jam.

References.

- Sandro Salsa "Partial Differential Equations in Action: From Modelling to Theory" §4.3.
- G. B. Whitham, "Linear and Nonlinear Waves" §3.1.

Exercises.

Exercise 2.27. Consider the general velocity profile $v(\rho)$ with $v'\leqslant 0$, $v''\leqslant 0$, $v(0)=v_m$ and $v(\rho_m)=0$. Repeat the arguments of this section.

2.5. Method of characteristics for first order nonlinear equations.

In this lecture we try to solve the first order equation

$$F(x, y, u, u_x, u_y) = 0, \qquad u = u_0(\tau) \quad \text{along } (x_0(\tau), y_0(\tau)). \tag{2.179}$$

Here $F(x, y, z, p, q)$ is a general nonlinear function. The only assumption we put on it is that it is a smooth function, so that we can take as many derivatives of it as necessary.

For such fully nonlinear equations the method of characteristics we have discussed so far clearly doesn't work anymore. However, the key idea of the method: reduce the PDE to a system of ODEs along particular curves, still works.

2.5.1. The method.

Consider a curve $((x(s), y(s)))$ to be determined. We try to find out whether it is possible to simplify the PDE along this particular curve. Let $z(s), p(s), q(s)$ be the restriction of u, u_x, u_y along this curve, that is

$$z(s) = u((x(s), y(s))), \qquad p(s) = u_x((x(s), y(s))), \qquad q(s) = u_y((x(s), y(s))). \tag{2.180}$$

We try to find enough relations between them so that they can be solved – more specifically, we try to obtain a system of five ODEs. We should notice that $p(s), q(s)$ cannot be determined by $z(s)$ alone.

- First notice

$$\frac{\mathrm{d}z}{\mathrm{d}s} = u_x \frac{\mathrm{d}x}{\mathrm{d}s} + u_y \frac{\mathrm{d}y}{\mathrm{d}s} = p(s) \frac{\mathrm{d}x}{\mathrm{d}s} + q(s) \frac{\mathrm{d}y}{\mathrm{d}s}. \tag{2.181}$$

- Now consider $p(s)$ and $q(s)$. We have

$$\frac{\mathrm{d}p}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}(u_x) = u_{xx} \frac{\mathrm{d}x}{\mathrm{d}s} + u_{xy} \frac{\mathrm{d}y}{\mathrm{d}s}; \tag{2.182}$$

Similarly

$$\frac{\mathrm{d}q}{\mathrm{d}s} = u_{xy} \frac{\mathrm{d}x}{\mathrm{d}s} + u_{yy} \frac{\mathrm{d}y}{\mathrm{d}s}. \tag{2.183}$$

We can write it into matrix form:

$$\begin{pmatrix} \frac{\mathrm{d}p}{\mathrm{d}s} \\ \frac{\mathrm{d}q}{\mathrm{d}s} \end{pmatrix} = \begin{pmatrix} u_{xx} & u_{xy} \\ u_{xy} & u_{yy} \end{pmatrix} \begin{pmatrix} \frac{\mathrm{d}x}{\mathrm{d}s} \\ \frac{\mathrm{d}y}{\mathrm{d}s} \end{pmatrix}. \tag{2.184}$$

To cancel the second order derivatives, we differentiate the equation $F(x, y, u, u_x, u_y) = 0$ to get

$$0 = \frac{\partial F}{\partial x} = F_x + F_z u_x + F_p u_{xx} + F_q u_{xy}; \tag{2.185}$$

$$0 = \frac{\partial F}{\partial y} = F_y + F_z u_y + F_p u_{xy} + F_q u_{yy}. \tag{2.186}$$

Now if we require (recall that our purpose is to find a special curve along which the equation gets simplified)

$$\frac{\mathrm{d}x}{\mathrm{d}s} = F_p, \qquad \frac{\mathrm{d}y}{\mathrm{d}s} = F_q \tag{2.187}$$

we would reach

$$\frac{\mathrm{d}p}{\mathrm{d}s} = -F_x - F_z p; \qquad \frac{\mathrm{d}q}{\mathrm{d}s} = -F_y - F_z q. \tag{2.188}$$

Thus we have obtained five equations for the five unknown functions:

$$\frac{\mathrm{d}x}{\mathrm{d}s} = F_p(x, y, z, p, q) \tag{2.189}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = F_q(x, y, z, p, q) \tag{2.190}$$

$$\frac{\mathrm{d}z}{\mathrm{d}s} = \frac{\mathrm{d}x}{\mathrm{d}s} p + \frac{\mathrm{d}y}{\mathrm{d}s} q = F_p(x, y, z, p, q)\, p + F_q(x, y, z, p, q)\, q \tag{2.191}$$

$$\frac{\mathrm{d}p}{\mathrm{d}s} = -F_x - F_z p \tag{2.192}$$

$$\frac{\mathrm{d}q}{\mathrm{d}s} = -F_y - F_z q. \tag{2.193}$$

Now we need initial conditions. It is clear that $x(0,\tau)=x_0(\tau)$, $y(0,\tau)=y_0(\tau)$, $z(0,\tau)=u_0(\tau)$. For p, q the situation is less explicit. Note that first we have

$$\frac{\mathrm{d}u_0(\tau)}{\mathrm{d}\tau}=p_0(\tau)\,\frac{\mathrm{d}x_0}{\mathrm{d}\tau}+q_0(\tau)\,\frac{\mathrm{d}y_0}{\mathrm{d}\tau}. \tag{2.194}$$

But we need one more equation to determine p_0, q_0. This other equation is simply

$$F(x_0,y_0,u_0,p_0,q_0)=0. \tag{2.195}$$

Therefore we obtain p_0, q_0 from solving the above two equations.

The method of characteristics for fully nonlinear first order equation then proceeds as follows:

1. Identify F, and then solve the system (2.189 – 2.193) with initial conditions as discussed above.
2. Find the inverse

$$s=S(x,y), \qquad \tau=\mathrm{T}(x,y) \tag{2.196}$$

substitute into $z(s,\tau)$ to obtain the solution.

Remark 2.23. We see that in the above process we did many differentiations and algebraic manipulations. It is now quite unclear why the solution we obtained indeed solved the original equation. We will discuss this issue in Section 2.8.1.

Remark 2.24. It is possible to obtain the system (2.189 – 2.193) more deductively (less guessing) using the geometric meaning of the equation. Recall that u is a solution to the quasi-linear equation

$$a(x,y,u)\,u_x+b(x,y,u)u_y=c(x,y,u) \tag{2.197}$$

if the normal vector to the surface $u-u(x,y)=0$ is perpendicular to the vector (a,b,c). This means the surface must be tangent to one single direction (a,b,c).

Correspondingly, u solves a fully nonlinear equation

$$F(x,y,u,u_x,u_y)=0 \tag{2.198}$$

if and only if the normal vector $(p,q,-1)$ satisfy the nonlinear algebraic equation

$$F(x,y,u,p,q)=0. \tag{2.199}$$

Instead of a plane, this would require the solution surface to be tangent to a certain cone, called Monge cone, at each point (x,y,u). Note that the single direction (a,b,c) can be seen as a (degenerate) cone.

More discussion about Monge cone can be found in Section 2.8.4.

Example 2.25. (Quasi-linear equation) A quasi-linear equation

$$a(x,y,u)\,u_x+b(x,y,u)\,u_y=c(x,y,u), \qquad u=u_0(\tau) \quad \text{along } (x_0(\tau),y_0(\tau)). \tag{2.200}$$

can be put into the fully nonlinear equation framework through setting

$$F(x,y,z,p,q)=a(x,y,z)\,p+b(x,y,z)\,q-c(x,y,z). \tag{2.201}$$

Now (2.189 – 2.193) becomes

$$\frac{\mathrm{d}x}{\mathrm{d}s} = a(x,y,z) \tag{2.202}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = b(x,y,z) \tag{2.203}$$

$$\frac{\mathrm{d}z}{\mathrm{d}s} = a\,p+b\,q=c(x,y,z) \tag{2.204}$$

The p,q equations are quite complicated, but we don't need them as we can already solve the system.

Furthermore, the initial values of p,q are determined through

$$\frac{\mathrm{d}x_0}{\mathrm{d}\tau}\,p_0+\frac{\mathrm{d}y_0}{\mathrm{d}\tau}\,q_0 = \frac{\mathrm{d}u_0}{\mathrm{d}\tau} \tag{2.205}$$

$$a\,p_0+b\,q_0 = c \tag{2.206}$$

The condition for the unique solvability of p_0, q_0 is

$$\det \begin{pmatrix} \frac{\mathrm{d}x_0}{\mathrm{d}\tau} & \frac{\mathrm{d}y_0}{\mathrm{d}\tau} \\ a & b \end{pmatrix} \neq 0, \tag{2.207}$$

which is exactly "the initial curve is not characteristics".

We see that the first two equations and the third equation are decoupled. Thus we can solve the first two equations and obtain the solution. There is no need to solve the 3rd equation.

Example 2.26. Solve

$$(u_x)^2 + (u_y)^2 = 1, \qquad u = 0 \text{ along } x^2 + y^2 = 1. \tag{2.208}$$

Solution. We have

$$F(x, y, z, p, q) = p^2 + q^2 - 1. \tag{2.209}$$

Thus the system for characteristics is

$$\frac{\mathrm{d}x}{\mathrm{d}s} = F_p = 2\,p \tag{2.210}$$

$$\frac{\mathrm{d}y}{\mathrm{d}s} = F_q = 2\,q \tag{2.211}$$

$$\frac{\mathrm{d}z}{\mathrm{d}s} = F_p\,p + F_q\,q = 2\,p^2 + 2\,q^2 = 2 \tag{2.212}$$

$$\frac{\mathrm{d}p}{\mathrm{d}s} = -F_x - F_z\,p = 0 \tag{2.213}$$

$$\frac{\mathrm{d}q}{\mathrm{d}s} = -F_x - F_z\,q = 0. \tag{2.214}$$

Solving this we see that

$$x(s,\tau) = 2\,p_0(\tau)\,s; \quad y(s,\tau) = 2\,q_0(\tau)\,s; \quad z(s,\tau) = 2\,s; \quad p(s,\tau) = p_0(\tau); \quad q(s,\tau) = q_0(\tau). \tag{2.215}$$

It is clear that all we need to do is to figure out p_0, q_0. The equations for them are

$$\frac{\mathrm{d}x_0(\tau)}{\mathrm{d}\tau}\,p_0 + \frac{\mathrm{d}y_0(\tau)}{\mathrm{d}\tau}\,q_0 = \frac{\mathrm{d}u_0(\tau)}{\mathrm{d}\tau} = 0 \tag{2.216}$$

$$p_0^2 + q_0^2 = 1. \tag{2.217}$$

At this stage we have to parametrize $x^2 + y^2 = 1$: $x_0(\tau) = \cos\tau$; $y_0(\tau) = \sin\tau$. With such parametrization we easily obtain

$$p_0 = \cos\tau,\ q_0 = \sin\tau; \qquad \text{or } p_0 = -\cos\tau,\ q_0 = -\sin\tau \tag{2.218}$$

Using the first (p_0, q_0) we get

$$x(s,\tau) = (1 + 2\,s)\cos\tau, \qquad y(s,\tau) = (1 + 2\,s)\sin\tau, \qquad z(s,\tau) = 2\,s. \tag{2.219}$$

We solve it with a trick: It is clear that $(1 + u)^2 = x^2 + y^2$ which leads to $u = (x^2 + y^2)^{1/2} - 1$.[2.3]

If we take the other pair of (p_0, q_0) we would get $u = 1 - (x^2 + y^2)^{1/2}$.

Remark 2.27. Note that the solution to the above example has a singular point at the origin.

Example 2.28. (Copson) Solve $u_x\,u_y = x\,y$, $u = x$ along $y = 0$.

Solution. The characteristic equations are

$$\dot{x} = q; \dot{y} = p; \dot{u} = 2\,p\,q; \dot{p} = y; \dot{q} = x \tag{2.220}$$

with initial condition $u(\tau) = x(\tau) = \tau, y(\tau) = 0$.

Solving this system we get

$$x = A\,e^t + B\,e^{-t}, y = C\,e^t + D\,e^{-t}, u = A\,C\,e^{2t} - B\,D\,e^{-2t} + E, p = C\,e^t - D\,e^{-t}, q = A\,e^t - B\,e^{-t}. \tag{2.221}$$

Note that $p\,q + x\,y = 0 \Longrightarrow A\,D + B\,C = 0$.

Initial condition gives $p_0 = 1, q_0 = 0$. Thus we have

$$A + B = s, C + D = 0, A\,C - B\,D + E = s, C - D = 1, A - B = 0, A\,D + B\,C = 0. \tag{2.222}$$

2.3. The other root, $u = -1 - (x^2 + y^2)^{1/2}$ does not satisfy the boundary condition.

This leads to

$$x = s\cosh t,\, y = \sinh t,\, u = s\cosh^2 t,\, p = \cosh t,\, q = s\sinh t. \tag{2.223}$$

Solving it we have $x^2/s^2 - y^2 = 1$ which gives finally $u^2 = x^2\,(y^2 + 1)$.

Example 2.29. (Hamilton-Jacobi equation) Consider the one dimensional Hamilton-Jacobi equation

$$u_t + H(x, u_x) = 0. \tag{2.224}$$

Let

$$F(x, t, z, p, q) = q + H(x, p). \tag{2.225}$$

Then we have

$$\frac{\mathrm{d}x}{\mathrm{d}s} = F_p = H_p; \tag{2.226}$$

$$\frac{\mathrm{d}t}{\mathrm{d}s} = F_q = 1; \tag{2.227}$$

$$\frac{\mathrm{d}z}{\mathrm{d}s} = F_p\, p + F_q\, q = H_p\, p + q = H_p\, p - H. \tag{2.228}$$

$$\frac{\mathrm{d}p}{\mathrm{d}s} = -F_z\, p - F_x = -H_x; \tag{2.229}$$

$$\frac{\mathrm{d}q}{\mathrm{d}s} = -F_z\, q - F_t = 0. \tag{2.230}$$

As $\dot{t} = 1$, we can replace s by t. Thus the characteristics equations become

$$\frac{\mathrm{d}x}{\mathrm{d}t} = H_p, \quad \frac{\mathrm{d}z}{\mathrm{d}t} = H_p\, p - H, \quad \frac{\mathrm{d}p}{\mathrm{d}t} = -H_x, \quad \frac{\mathrm{d}q}{\mathrm{d}t} = 0. \tag{2.231}$$

In particular we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H}{\partial p}; \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial x} \tag{2.232}$$

which are the so-called Hamiltonian equations which governs the evolution of particles (x is the location and p is the momentum of the particle). This understanding is important in deriving the solution formula for the H-J equation.

Exercises.

Exercise 2.28. (Zauderer) Solve the following problems:

a) $u_x^2\, u_y - 1 = 0$, $u(x, 0) = x$.

b) $u_x\, u_y = 4, u(x, 0) = 2$.

c) $u_t + u_x^2 = t$, $u(x, 0) = 0$.

d) $u_t + u_x^2 + u = 0$, $u(x, 0) = x$.

e) $u_t + u_x^2 = 0$, $u(x, 0) = -x^2$. Show that the solution breaks down when $t = 1/4$.

Exercise 2.29. (Zauderer) (Snell's law) Consider the eiconal equation

$$u_x^2 + u_y^2 = n(x, y)^2, \qquad n(x, y) = \begin{cases} n_1 & y < 0 \\ n_2 & y > 0 \end{cases}. \tag{2.233}$$

Here $n_2 > n_1$ are constants. Let the initial condition be $u(x, 0) = n_1\, x \cos\theta$.

a) Solve the equation.

b) By considering the directions ∇u, confirm Snell's law.

Exercise 2.30. Generalize the method in this section to fully nonlinear equations in n-dimensions:

$$F(x_1, \ldots, x_n, u, u_{x_1}, \ldots, u_{x_n}) = 0. \tag{2.234}$$

References.

- E. T. Copson, "Partial Differential Equations", §1.3, 1.4.
- Lawrence C. Evans, "Partial Differential Equations", 1ed or 2ed, §3.2
- Erich Zauderer, "Partial Differential Equations of Applied Mathematics", 2ed, §2.4.

2.6. Classification of 2nd Order equations.

In this section we review classification and reduction to canonical form for 2nd order equations.

2.6.1. Reduction to canonical forms.

Consider a general linear 2nd order equation:

$$A\,u_{xx} + B\,u_{xy} + C\,u_{yy} + D\,u_x + E\,u_y + F u = G. \tag{2.235}$$

where each coefficient $A, B, C, \ldots$ is a function of x, y.

It turns out that one can simplify the 2nd order terms to one of the following three so-called "canonical" forms

1. hyperbolic: u_{xy} or $u_{xx} - u_{yy}$;
2. parabolic: u_{xx} or u_{yy};
3. elliptic: $u_{xx} + u_{yy}$.

The properties of the equations/solutions are very different for each category.

General strategy.

The idea is to apply a change of variables

$$\xi = \xi(x, y), \qquad \eta = \eta(x, y). \tag{2.236}$$

This gives

$$\begin{aligned}
u_x &= u_\xi\,\xi_x + u_\eta\,\eta_x, & (2.237)\\
u_y &= u_\xi\,\xi_y + u_\eta\,\eta_y, & (2.238)\\
u_{xx} &= u_{\xi\xi}\,\xi_x^2 + 2\,u_{\xi\eta}\,\xi_x\,\eta_x + u_{\eta\eta}\,\eta_x^2 + u_\xi\,\xi_{xx} + u_\eta\,\eta_{xx}, & (2.239)\\
u_{xy} &= u_{\xi\xi}\,\xi_x\,\xi_y + u_{\xi\eta}\,(\xi_x\,\eta_y + \xi_y\,\eta_x) + u_{\eta\eta}\,\eta_x\,\eta_y + u_\xi\,\xi_{xy} + u_\eta\,\eta_{xy}, & (2.240)\\
u_{yy} &= u_{\xi\xi}\,\xi_y^2 + 2\,u_{\xi\eta}\,\xi_y\,\eta_y + u_{\eta\eta}\,\eta_y^2 + u_\xi\,\xi_{yy} + u_\eta\,\eta_{yy}. & (2.241)
\end{aligned}$$

Substituting these into the equation we obtain

$$A^* u_{\xi\xi} + B^* u_{\xi\eta} + C^*\,u_{\eta\eta} + D^* u_\xi + E^*\,u_\eta + F^*\,u = G^*. \tag{2.242}$$

with

$$\begin{aligned}
A^* &= A\,\xi_x^2 + B\,\xi_x\,\xi_y + C\,\xi_y^2, & (2.243)\\
B^* &= 2\,A\,\xi_x\,\eta_y + B\,(\xi_x\,\eta_y + \xi_y\,\eta_x) + 2\,C\,\xi_y\,\eta_y, & (2.244)\\
C^* &= A\,\eta_x^2 + B\,\eta_x\,\eta_y + C\,\eta_y^2, & (2.245)\\
D^* &= A\,\xi_{xx} + B\,\xi_{xy} + C\,\xi_{yy} + D\,\xi_x + E\,\xi_y, & (2.246)\\
E^* &= A\,\eta_{xx} + B\,\eta_{xy} + C\,\eta_{yy} + D\,\eta_x + E\,\eta_y, & (2.247)\\
F^* &= F, & (2.248)\\
G^* &= G. & (2.249)
\end{aligned}$$

Now recall that our purpose is to reduce the equation to canonical form. In other words, we would explore the possibility of choosing appropriate ξ, η such that

1. $A^* = C^* = 0$, $B^* \neq 0$, or
2. $B^* = 0$, exactly one of A^* and $C^* = 0$, or
3. $A^* = C^* \neq 0$, $B^* = 0$.

Remark 2.30. It may happen that $A^* = B^* = C^* = 0$. But that just means that the equation is first order.

We note that

$$(B^*)^2 - 4\,A^*\,C^* = (B^2 - 4\,A\,C)\,(\xi_x\,\eta_y - \xi_y\,\eta_x)^2. \tag{2.250}$$

Therefore the sign of $B^2 - 4\,A\,C$ is invariant under change of variables. We will classify using this quantity:

1. $B^2 - 4\,A\,C > 0$: Hyperbolic case;

2. $B^2 - 4AC = 0$: Parabolic case;
3. $B^2 - 4AC < 0$: Elliptic case.

Hyperbolic case.

In this case we are able to find ξ, η independent such that

$$A^* = A\xi_x^2 + B\xi_x\xi_y + C\xi_y^2 = 0 \implies A\left(\frac{\xi_x}{\xi_y}\right)^2 + B\left(\frac{\xi_x}{\xi_y}\right) + C = 0, \tag{2.251}$$

and

$$C^* = A\eta_x^2 + B\eta_x\eta_y + C\eta_y^2 = 0 \implies A\left(\frac{\eta_x}{\eta_y}\right)^2 + B\left(\frac{\eta_x}{\eta_y}\right) + C = 0. \tag{2.252}$$

This is possible because, thanks to $B^2 - 4AC > 0$, there are two different solutions $r_1 \neq r_2$ for the equation

$$A r^2 + B r + C = 0. \tag{2.253}$$

Then we require

$$\frac{\xi_x}{\xi_y} = r_1, \qquad \frac{\eta_x}{\eta_y} = r_2. \tag{2.254}$$

Under this change of variables we have $A^* = C^* = 0$, therefore the equation can be reduced to

$$u_{\xi\eta} = H. \tag{2.255}$$

If we let

$$\alpha = \xi + \eta, \quad \beta = \xi - \eta, \tag{2.256}$$

the equation becomes

$$u_{\alpha\alpha} - u_{\beta\beta} = H_1(\xi, \eta, u, u_\xi, u_\eta). \tag{2.257}$$

Parabolic case.

In this case there is only one double root for

$$A r^2 + B r + C = 0. \tag{2.258}$$

Denote it by r_0. Then $\xi(x, y)$ can be obtained through

$$\frac{\xi_x}{\xi_y} = r_0. \tag{2.259}$$

Once ξ is obtained, any $\eta(x, y)$ independent of ξ, that is $\xi_x\eta_y - \xi_y\eta_x \neq 0$, will reduce the equation to its canonical form

$$u_{\xi\xi} = H_2(\xi, \eta, u, u_\xi, u_\eta). \tag{2.260}$$

Elliptic case.

The remaining case is $B^2 - 4AC < 0$. In this case we can obtain two complex roots. In other words

$$\frac{\xi_x}{\xi_y} = \left(\frac{\eta_x}{\eta_y}\right)^*. \tag{2.261}$$

We see that this holds when ξ and η are complex conjugates.

In this case, we can introduce new variables

$$\alpha = \frac{1}{2}(\xi + \eta), \qquad \beta = \frac{1}{2i}(\xi - \eta) \tag{2.262}$$

or equivalently

$$\xi = \alpha + \beta i, \quad \eta = \alpha - \beta i. \tag{2.263}$$

Using the change of variables formula we have

$$u_{\alpha\alpha} + u_{\beta\beta} = 4u_{\xi\eta} \tag{2.264}$$

as a consequence, the canonical form in real variables α, β is

$$u_{\alpha\alpha} + u_{\beta\beta} = H_3(\alpha, \beta, u, u_\alpha, u_\beta). \tag{2.265}$$

Remark 2.31. In practice, it is more efficient to obtain $\alpha = \alpha(x, y)$, $\beta = \beta(x, y)$ and then transform the equation, without first reducing the equation into $u_{\xi\eta} = H_4$.

Remark 2.32. The level sets of ξ and η, that is the curves $\xi(x, y) = c$, $\eta(x, y) = c$, are called "characteristic curves" of the equations (in the parabolic case the only characteristic curves are $\xi(x, y) = c$).

2.6.2. Solving 2nd order equations.

To solve a 2nd order linear PDE, we follow the following steps.

1. Solve $A\,u_{xx} + B\,u_{xy} + C\,u_{yy} \Longrightarrow A\,(\mathrm{d}y)^2 - B\,(\mathrm{d}x)\,(\mathrm{d}y) + C\,(\mathrm{d}x)^2 = 0$. Note the sign change.[2.4] Obtain
   a) new variables ξ and η when $B^2 - 4\,A\,C > 0$ (hyperbolic);
   b) new variable ξ, choose any η with $\det\begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} \neq 0$, when $B^2 - 4\,A\,C = 0$ (parabolic);
   c) two complex functions ξ and η, set $\alpha = (\xi + \eta)/2$, $\beta = (\xi - \eta)/2\,i$ as new variables, when $B^2 - 4\,A\,C < 0$ (elliptic).
2. Perform change of variables and reduce the equation to canonical forms using the following formulas:

$$\begin{aligned} u_x &= u_\xi\,\xi_x + u_\eta\,\eta_x, && (2.270)\\ u_y &= u_\xi\,\xi_y + u_\eta\,\eta_y, && (2.271)\\ u_{xx} &= u_{\xi\xi}\,\xi_x^2 + 2\,u_{\xi\eta}\,\xi_x\,\eta_x + u_{\eta\eta}\,\eta_x^2 + u_\xi\,\xi_{xx} + u_\eta\,\eta_{xx}, && (2.272)\\ u_{xy} &= u_{\xi\xi}\,\xi_x\,\xi_y + u_{\xi\eta}\,(\xi_x\,\eta_y + \xi_y\,\eta_x) + u_{\eta\eta}\,\eta_x\,\eta_y + u_\xi\,\xi_{xy} + u_\eta\,\eta_{xy}, && (2.273)\\ u_{yy} &= u_{\xi\xi}\,\xi_y^2 + 2\,u_{\xi\eta}\,\xi_y\,\eta_y + u_{\eta\eta}\,\eta_y^2 + u_\xi\,\xi_{yy} + u_\eta\,\eta_{yy}. && (2.274)\end{aligned}$$

3. Try to obtain the general solution of the canonical form equation.
4. (For Cauchy problems) Substitute the Cauchy data into the general solution and determine the solution.

Remark 2.33. The approach

$$A\,(\mathrm{d}y)^2 - B\,(\mathrm{d}x)\,(\mathrm{d}y) + C\,(\mathrm{d}x)^2 = 0 \tag{2.275}$$

is more efficient when solving specific problems, while (denote $r_{1,2}$ be the solutions to $A\,r^2 + B\,r + C = 0$) considering

$$\xi_x - r_1\,\xi_y = 0; \qquad \eta_x - r_2\,\eta_y = 0 \tag{2.276}$$

is often better suited for theoretical study.

Example 2.34. Determine the region in which the given equation is hyperbolic, parabolic, or elliptic, and transform the equation in the respective region to canonical form.

1.

$$x\,u_{xx} + u_{yy} = x^2. \tag{2.277}$$

2.4. We explain a bit. Remember that $\frac{\xi_x}{\xi_y}$ and $\frac{\eta_x}{\eta_y}$ solve

$$A\,r^2 + B\,r + C = 0. \tag{2.266}$$

Let r_1, r_2 be the two roots, then the equation can be written as

$$A\,(r - r_1)\,(r - r_2) = 0 \Longrightarrow B = -A\,(r_1 + r_2), \quad C = A\,r_1\,r_2. \tag{2.267}$$

Now as $\frac{\mathrm{d}y}{\mathrm{d}x} = -r_1, -r_2$, they solve the equation

$$A\,(r + r_1)\,(r + r_2) = 0 \Longleftrightarrow A\,r^2 - B\,r + C = 0. \tag{2.268}$$

In other words, $\mathrm{d}y, \mathrm{d}x$ satisfy

$$A\,(\mathrm{d}y)^2 - B\,(\mathrm{d}x)\,(\mathrm{d}y) + C\,(\mathrm{d}x)^2 = 0. \tag{2.269}$$

Solution. We have $A = x, B = 0, C = 1$. Thus

$$B^2 - 4AC = -4x. \tag{2.278}$$

– $x < 0$: Hyperbolic.

We solve

$$x(\mathrm{d}y)^2 + (\mathrm{d}x)^2 = 0. \tag{2.279}$$

This gives

$$\mathrm{d}x \pm \sqrt{-x}\,\mathrm{d}y = 0 \tag{2.280}$$

which leads to

$$\mathrm{d}[y \pm 2\sqrt{-x}] = 0 \tag{2.281}$$

therefore

$$\xi = y + 2\sqrt{-x}, \qquad \eta = y - 2\sqrt{-x} \tag{2.282}$$

which gives

$$\xi_x = -\frac{1}{\sqrt{-x}}, \ \ \xi_y = 1, \ \ \xi_{xx} = -\frac{1}{2(\sqrt{-x})^3}, \ \ \xi_{xy} = 0, \ \ \xi_{yy} = 0; \tag{2.283}$$

$$\eta_x = \frac{1}{\sqrt{-x}}, \ \ \eta_y = 1, \ \ \eta_{xx} = \frac{1}{2(\sqrt{-x})^3}, \ \ \eta_{xy} = 0, \ \ \eta_{yy} = 0. \tag{2.284}$$

We compute

$$u_{xx} = u_{\xi\xi}\,\xi_x^2 + 2\,u_{\xi\eta}\,\xi_x\,\eta_x + u_{\eta\eta}\,\eta_x^2 + u_\xi\,\xi_{xx} + u_\eta\,\eta_{xx} = -\frac{u_{\xi\xi}}{x} + \frac{2\,u_{\xi\eta}}{x} - \frac{u_{\eta\eta}}{x} + \frac{u_\eta - u_\xi}{2(\sqrt{-x})^3}, \tag{2.285}$$

$$u_{xy} = u_{\xi\xi}\,\xi_x\,\xi_y + u_{\xi\eta}(\xi_x\,\eta_y + \xi_y\,\eta_x) + u_{\eta\eta}\,\eta_x\,\eta_y + u_\xi\,\xi_{xy} + u_\eta\,\eta_{xy} = \frac{u_{\eta\eta} - u_{\xi\xi}}{\sqrt{-x}}, \tag{2.286}$$

$$u_{yy} = u_{\xi\xi}\,\xi_y^2 + 2\,u_{\xi\eta}\,\xi_y\,\eta_y + u_{\eta\eta}\,\eta_y^2 + u_\xi\,\xi_{yy} + u_\eta\,\eta_{yy} = u_{\xi\xi} + 2\,u_{\xi\eta} + u_{\eta\eta}. \tag{2.287}$$

Thus the equation becomes

$$4\,u_{\xi\eta} + \frac{u_\xi - u_\eta}{2\sqrt{-x}} = x^2. \tag{2.288}$$

From the change of variables we obtain

$$\sqrt{-x} = \frac{\xi - \eta}{4} \tag{2.289}$$

and as a consequence

$$u_{\xi\eta} = \frac{1}{4}\left(\frac{\xi - \eta}{4}\right)^4 - \frac{1}{2}\left(\frac{1}{\xi - \eta}\right)(u_\xi - u_\eta). \tag{2.290}$$

– $x = 0$: parabolic. In this case the equation becomes

$$u_{yy} = x^2 \tag{2.291}$$

which is already in canonical form.

– $x > 0$: elliptic. In this case we still have

$$x(\mathrm{d}y)^2 + (\mathrm{d}x)^2 = 0. \tag{2.292}$$

which gives

$$\pm i\sqrt{x}\,\mathrm{d}y + \mathrm{d}x = 0 \implies \mathrm{d}[2\sqrt{x} \pm i\,y] = 0. \tag{2.293}$$

Thus

$$\xi = 2\sqrt{x} + i\,y, \qquad \eta = 2\sqrt{x} - i\,y. \tag{2.294}$$

We then have

$$\alpha = \frac{\xi + \eta}{2} = 2\sqrt{x}, \qquad \beta = \frac{\xi - \eta}{2\,i} = y. \tag{2.295}$$

This leads to

$$\alpha_x = \frac{1}{\sqrt{x}},\ \ \alpha_y = 0,\ \ \alpha_{xx} = -\frac{1}{2\,(\sqrt{x})^3},\ \ \alpha_{xy} = 0,\ \ \alpha_{yy} = 0, \tag{2.296}$$

$$\beta_x = 0,\ \ \beta_y = 1,\ \ \beta_{xx} = \beta_{xy} = \beta_{yy} = 0. \tag{2.297}$$

Consequently

$$u_{xx} = \frac{u_{\alpha\alpha}}{x} - \frac{u_\alpha}{2\,\sqrt{x}^3}, \qquad u_{yy} = u_{\beta\beta} \tag{2.298}$$

and the equation becomes

$$u_{\alpha\alpha} + u_{\beta\beta} = x^2 + \frac{u_\alpha}{2\,\sqrt{x}} = \frac{u_\alpha}{\alpha} + \left(\frac{\alpha}{2}\right)^4. \tag{2.299}$$

2.

$$x^2\,u_{xx} - 2\,x\,y\,u_{xy} + y^2\,u_{yy} = 2\,e^x. \tag{2.300}$$

Solution. We have

$$B^2 - 4\,A\,C \equiv 0 \tag{2.301}$$

So the equation is of parabolic type. We solve the characteristics equation

$$x^2\,(\mathrm{d}y)^2 + 2\,x\,y\,\mathrm{d}x\,\mathrm{d}y + y^2\,(\mathrm{d}x)^2 = 0 \tag{2.302}$$

which reduces to

$$(x\,\mathrm{d}y + y\,\mathrm{d}x)^2 = 0 \implies \xi = x\,y. \tag{2.303}$$

Thus the Jacobian is

$$\det\begin{pmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{pmatrix} = \det\begin{pmatrix} y & x \\ \eta_x & \eta_y \end{pmatrix}. \tag{2.304}$$

We can take for example $\eta = x$ to make the Jacobian nonzero. We have

$$\xi_x = y,\ \ \xi_y = x,\ \ \xi_{xx} = 0,\ \ \xi_{xy} = 1,\ \ \xi_{yy} = 0, \tag{2.305}$$

$$\eta_x = 1,\ \ \eta_y = \eta_{xx} = \eta_{xy} = \eta_{yy} = 0. \tag{2.306}$$

Thus we have

$$u_{xx} = y^2\,u_{\xi\xi} + 2\,y\,u_{\xi\eta} + x^2\,u_{\eta\eta},\ \ u_{xy} = x\,y\,u_{\xi\xi} + x\,u_{\xi\eta} + u_\xi,\ \ u_{yy} = x^2\,u_{\xi\xi} \tag{2.307}$$

which leads to

$$x^4\,u_{\eta\eta} - 2\,x\,y\,u_\xi = 2\,e^x. \tag{2.308}$$

So the canonical form is

$$u_{\eta\eta} = \frac{2\,\xi}{\eta^4}\,u_\xi + \frac{2}{\eta^4}\,e^\eta. \tag{2.309}$$

Example 2.35. Obtain the general solution of the following equation:

$$4\,u_{xx} + 12\,u_{xy} + 9\,u_{yy} - 9\,u = 9. \tag{2.310}$$

Solution. First we reduce it to canonical form. As $B^2 - 4\ A\ C = 0$, the equation is parabolic. The characteristics equation is

$$4\,(\mathrm{d}y)^2 - 12\,(\mathrm{d}x)\,(\mathrm{d}y) + 9\,(\mathrm{d}x)^2 = 0 \implies 2\,\mathrm{d}y - 3\,\mathrm{d}x = 0. \tag{2.311}$$

Thus we have

$$\xi = 2\,y - 3\,x. \tag{2.312}$$

We can simply take $\eta = y$.

Thus

$$\xi_x = -3,\ \ \xi_y = 2,\ \ \xi_{xx} = \xi_{xy} = \xi_{yy} = 0; \qquad \eta_y = 1,\ \ \eta_x = \eta_{xx} = \eta_{xy} = \eta_{yy} = 0. \tag{2.313}$$

Under this change of variables, we have

$$u_{xx} = u_{\xi\xi}\,\xi_x^2 + 2\,u_{\xi\eta}\,\xi_x\,\eta_x + u_{\eta\eta}\,\eta_x^2 + u_\xi\,\xi_{xx} + u_\eta\,\eta_{xx} = 9\,u_{\xi\xi}; \tag{2.314}$$

$$u_{xy} = u_{\xi\xi}\,\xi_x\,\xi_y + u_{\xi\eta}\,(\xi_x\,\eta_y + \xi_y\,\eta_x) + u_{\eta\eta}\,\eta_x\,\eta_y + u_\xi\,\xi_{xy} + u_\eta\,\eta_{xy} = -6\,u_{\xi\xi} - 3\,u_{\xi\eta}; \tag{2.315}$$

$$u_{yy} = u_{\xi\xi}\,\xi_y^2 + 2\,u_{\xi\eta}\,\xi_y\,\eta_y + u_{\eta\eta}\,\eta_y^2 + u_\xi\,\xi_{yy} + u_\eta\,\eta_{yy} = 4\,u_{\xi\xi} + 4\,u_{\xi\eta} + u_{\eta\eta}; \tag{2.316}$$

Thus the equation reduces to

$$9\,u_{\eta\eta} - 9\,u = 9 \iff u_{\eta\eta} - u = 1. \tag{2.317}$$

We see that the general solution is

$$u(\xi, \eta) = f(\xi)\,e^{\eta} + g(\xi)\,e^{-\eta} - 1. \tag{2.318}$$

Or in (x, y) variables

$$u(x, y) = f(2\,y - 3\,x)\,e^{y} + g(2\,y - 3\,x)\,e^{-y} - 1. \tag{2.319}$$

Example 2.36. Obtain the general solution of the following equation:

$$u_{xx} + u_{xy} - 2\,u_{yy} - 3\,u_x - 6\,u_y = 9\,(2\,x - y). \tag{2.320}$$

Solution. We compute

$$B^2 - 4\,A\,C = 1 - 4\,(-2) = 9 > 0 \tag{2.321}$$

thus the equation is hyperbolic. The characteristics equation is

$$(\mathrm{d}y)^2 - (\mathrm{d}x)\,(\mathrm{d}y) - 2\,(\mathrm{d}x)^2 = 0 \iff (\mathrm{d}y - 2\,\mathrm{d}x)\,(\mathrm{d}y + \mathrm{d}x) = 0 \tag{2.322}$$

which gives

$$\xi = y - 2\,x, \qquad \eta = y + x. \tag{2.323}$$

We have

$$\xi_x = -2,\ \ \xi_y = 1; \qquad \eta_x = \eta_y = 1 \tag{2.324}$$

and all second order derivatives are 0. As a consequence

$$\begin{aligned}
u_{xx} &= u_{\xi\xi}\,\xi_x^2 + 2\,u_{\xi\eta}\,\xi_x\,\eta_x + u_{\eta\eta}\,\eta_x^2 + u_\xi\,\xi_{xx} + u_\eta\,\eta_{xx} = 4\,u_{\xi\xi} - 4\,u_{\xi\eta} + u_{\eta\eta}, && (2.325)\\
u_{xy} &= u_{\xi\xi}\,\xi_x\,\xi_y + u_{\xi\eta}\,(\xi_x\,\eta_y + \xi_y\,\eta_x) + u_{\eta\eta}\,\eta_x\,\eta_y + u_\xi\,\xi_{xy} + u_\eta\,\eta_{xy} = -2\,u_{\xi\xi} - u_{\xi\eta} + u_{\eta\eta}, && (2.326)\\
u_{yy} &= u_{\xi\xi}\,\xi_y^2 + 2\,u_{\xi\eta}\,\xi_y\,\eta_y + u_{\eta\eta}\,\eta_y^2 + u_\xi\,\xi_{yy} + u_\eta\,\eta_{yy} = u_{\xi\xi} + 2\,u_{\xi\eta} + u_{\eta\eta}, && (2.327)\\
u_x &= u_\xi\,\xi_x + u_\eta\,\eta_x = -2\,u_\xi + u_\eta, && (2.328)\\
u_y &= u_\xi\,\xi_y + u_\eta\,\eta_y = u_\xi + u_\eta, && (2.329)\\
2\,x - y &= -\xi. && (2.330)
\end{aligned}$$

The equation reduces to

$$-9\,u_{\xi\eta} - 9\,u_\eta = -9\,\xi \iff (u_\xi + u)_\eta = u_{\xi\eta} - u_\eta = \xi. \tag{2.331}$$

The general solution can be obtained via

$$u_\xi + u = \xi\,\eta + h(\xi) \implies (e^\xi\,u)_\xi = e^\xi\,\xi\,\eta + e^\xi\,h(\xi) \implies e^\xi\,u = \eta\,e^\xi\,(\xi - 1) + f(\xi) + g(\eta). \tag{2.332}$$

Therefore

$$u(\xi, \eta) = \eta\,(\xi - 1) + f(\xi) + g(\eta)\,e^{-\xi} \tag{2.333}$$

and

$$u(x, y) = (y + x)\,(y - 2\,x - 1) + f(y - 2\,x) + g(y + x)\,e^{2x - y} \tag{2.334}$$

where f, g are arbitrary functions.

A limited number of non-constant coefficient equations can also be solved.

Example 2.37. Obtain the general solution.

$$x^2\,u_{x\,x}+2\,x\,y\,u_{x\,y}+y^2\,u_{y\,y}+x\,y\,u_x+y^2\,u_y=0. \tag{2.335}$$

Solution. We check $B^2-4\,A\,C=(2\,x\,y)^2-4\,x^2\,y^2=0$ so the equation is parabolic. The characteristics equation is

$$x^2\,(\mathrm{d}y)^2-2\,x\,y\,(\mathrm{d}x)\,(\mathrm{d}y)+y^2\,(\mathrm{d}x)^2=0 \implies x\,\mathrm{d}y-y\,\mathrm{d}x=0. \tag{2.336}$$

Thus

$$\xi=\frac{y}{x}. \tag{2.337}$$

We compute

$$J=\det\begin{pmatrix}\xi_x & \xi_y\\ \eta_x & \eta_y\end{pmatrix}=\det\begin{pmatrix}-\frac{y}{x^2} & \frac{1}{x}\\ \eta_x & \eta_y\end{pmatrix} \tag{2.338}$$

and we can take $\eta=x$ to guarantee $J\neq 0$. Now we have

$$\xi_x=-\frac{y}{x^2},\;\; \xi_y=\frac{1}{x},\;\; \xi_{x\,x}=\frac{2\,y}{x^3},\;\; \xi_{x\,y}=-\frac{1}{x^2},\;\; \xi_{y\,y}=0. \tag{2.339}$$

$$\eta_x=1,\;\; \eta_y=\eta_{x\,x}=\eta_{x\,y}=\eta_{y\,y}=0. \tag{2.340}$$

This gives

$$u_{x\,x} = u_{\xi\,\xi}\,\xi_x^2+2\,u_{\xi\,\eta}\,\xi_x\,\eta_x+u_{\eta\,\eta}\,\eta_x^2+u_\xi\,\xi_{x\,x}+u_\eta\,\eta_{x\,x}=\frac{y^2}{x^4}\,u_{\xi\,\xi}-\frac{2\,y}{x^2}\,u_{\xi\,\eta}+u_{\eta\,\eta}+\frac{2\,y}{x^3}\,u_\xi, \tag{2.341}$$

$$u_{x\,y} = u_{\xi\,\xi}\,\xi_x\,\xi_y+u_{\xi\,\eta}\,(\xi_x\,\eta_y+\xi_y\,\eta_x)+u_{\eta\,\eta}\,\eta_x\,\eta_y+u_\xi\,\xi_{x\,y}+u_\eta\,\eta_{x\,y}=-\frac{y}{x^3}\,u_{\xi\,\xi}+\frac{1}{x}\,u_{\xi\,\eta}-\frac{1}{x^2}\,u_\xi, \tag{2.342}$$

$$u_{y\,y} = u_{\xi\,\xi}\,\xi_y^2+2\,u_{\xi\,\eta}\,\xi_y\,\eta_y+u_{\eta\,\eta}\,\eta_y^2+u_\xi\,\xi_{y\,y}+u_\eta\,\eta_{y\,y}=\frac{1}{x^2}\,u_{\xi\,\xi}, \tag{2.343}$$

$$u_x = u_\xi\,\xi_x+u_\eta\,\eta_x=-\frac{y}{x^2}\,u_\xi+u_\eta, \tag{2.344}$$

$$u_y = u_\xi\,\xi_y+u_\eta\,\eta_y=\frac{1}{x}\,u_\xi. \tag{2.345}$$

The equation becomes

$$x^2\,u_{\eta\,\eta}+x\,y\,u_\eta=0 \implies u_{\eta\,\eta}+\xi\,u_\eta=0. \tag{2.346}$$

We solve the equation

$$u_{\eta\,\eta}+\xi\,u_\eta=0 \implies u_\eta+\xi\,u=h(\xi) \implies (e^{\xi\,\eta}\,u)_\eta=e^{\xi\,\eta}\,h(\xi) \tag{2.347}$$

which leads to

$$e^{\xi\,\eta}\,u=\xi^{-1}\,e^{\xi\,\eta}\,h(\xi)+g(\xi) \implies u(\xi,\eta)=\xi^{-1}\,h(\xi)+e^{-\xi\,\eta}\,g(\xi). \tag{2.348}$$

So finally

$$u(x,y)=f\left(\frac{y}{x}\right)+g\left(\frac{y}{x}\right)e^{-y} \tag{2.349}$$

with f,g arbitrary functions.

Example 2.38. Obtain the general solution.

$$r\,u_{t\,t}-c^2\,r\,u_{r\,r}-2\,c^2\,u_r=0 \tag{2.350}$$

where c is a constant.

Solution. We check $B^2-4\,A\,C=0+4\,c^2\,r^2>0$ so that equation is hyperbolic. The characteristics equation is

$$r\,(\mathrm{d}r)^2-c^2\,r\,(\mathrm{d}t)^2=0 \implies \mathrm{d}r\pm c\,\mathrm{d}t=0 \tag{2.351}$$

so we take

$$\xi=r+c\,t, \qquad \eta=r-c\,t. \tag{2.352}$$

From this we obtain

$$\xi_r=1,\;\; \xi_t=c; \qquad \eta_r=1,\;\; \eta_t=-c \tag{2.353}$$

and all second order derivatives are zero.

Now we compute

$$u_{tt} = u_{\xi\xi}\,\xi_t^2 + 2\,u_{\xi\eta}\,\xi_t\,\eta_t + u_{\eta\eta}\,\eta_t^2 + u_\xi\,\xi_{tt} + u_\eta\,\eta_{tt} = c^2\,u_{\xi\xi} - 2\,c^2\,u_{\xi\eta} + c^2\,u_{\eta\eta}, \tag{2.354}$$

$$u_{rr} = u_{\xi\xi}\,\xi_r^2 + 2\,u_{\xi\eta}\,\xi_r\,\eta_r + u_{\eta\eta}\,\eta_r^2 + u_\xi\,\xi_{rr} + u_\eta\,\eta_{rr} = u_{\xi\xi} + 2\,u_{\xi\eta} + u_{\eta\eta} \tag{2.355}$$

$$u_r = u_\xi\,\xi_r + u_\eta\,\eta_r = u_\xi + u_\eta. \tag{2.356}$$

The equation then reduces to

$$-4\,r\,c^2\,u_{\xi\eta} - 2\,c^2\,u_\xi - 2\,c^2\,u_\eta = 0 \Longrightarrow 2\,r\,u_{\xi\eta} + u_\xi + u_\eta = 0 \Longrightarrow (\xi + \eta)\,u_{\xi\eta} + u_\xi + u_\eta = 0. \tag{2.357}$$

It turns out that the equation can be rewritten to

$$[(\xi + \eta)\,u]_{\xi\eta} = 0. \tag{2.358}$$

Therefore the general solutions are

$$u(\xi, \eta) = (\xi + \eta)^{-1}\,[f(\xi) + g(\eta)] \tag{2.359}$$

or equivalently

$$u(r, t) = r^{-1}\,[f(r + c\,t) + g(r - c\,t)]. \tag{2.360}$$

Remark 2.39. In fact, if we let $v = r\,u$ from the very start, we can reduce the equation immediately to the wave equation.

2.6.3. Classification in higher dimensions.

Consider the general linear 2nd order equation in $\mathbb{R}^n$:

$$\sum_{i,j=1}^{n} (a_{ij}\,u_{x_i})_{x_j} + \sum_{i=1}^{n} b_i\,u_{x_i} + c\,u + d = 0 \tag{2.361}$$

with all coefficients functions of $x_1, \ldots, x_n$. We further assume that $a_{ij} = a_{ji}$. On first look it may seem that we are dealing with a special class of equations here. However this is not so. See exercise.

We hope that the equation can be reduced to some “canonical form” through a change of variables $\boldsymbol{\xi} = \boldsymbol{\xi}(\boldsymbol{x})$, that is $\xi_i = \xi_i(x_1, \ldots, x_n)$. The equation with variable ξ_i now reads

$$\sum_{i,j=1}^{n} \tilde{a}_{ij}\,u_{\xi_i\,\xi_j} + \text{lower order terms} = 0 \tag{2.362}$$

where

$$\tilde{a}_{ij} = \sum a_{kl}\,\frac{\partial \xi_i}{\partial x_k}\,\frac{\partial \xi_j}{\partial x_l}. \tag{2.363}$$

If we introduce the matrices

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \qquad R = \begin{pmatrix} \frac{\partial \xi_1}{\partial x_1} & \cdots & \frac{\partial \xi_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial \xi_n}{\partial x_1} & \cdots & \frac{\partial \xi_n}{\partial x_n} \end{pmatrix} = (\ \nabla \xi_1 \ \ \cdots \ \ \nabla \xi_n\ )^T, \tag{2.364}$$

then the matrix $\tilde{A} := (\tilde{a}_{ij})$ is

$$\tilde{A} = R\,A\,R^T. \tag{2.365}$$

From linear algebra we know that there is a nonsingular matrix R such that

$$R\,A\,R^T = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}. \tag{2.366}$$

If we let $R^T = (\boldsymbol{r}_1 \ \cdots \ \boldsymbol{r}_n)$, then the equation reduces to

$$\sum_{i=1}^{n} \lambda_i (\boldsymbol{r}_i \cdot \nabla)[(\boldsymbol{r}_i \cdot \nabla) u] + \text{lower order terms=0}. \tag{2.367}$$

where each λ_i is either 1 or -1 or 0. If we further denote

$$D_i := \boldsymbol{r}_i \cdot \nabla \tag{2.368}$$

the equation can be formally written into

$$\sum_{i=1}^{n} \lambda_i D_i D_i u + \sum_{i=1}^{n} \tilde{b_i} D_i u + c u + d = 0. \tag{2.369}$$

The situation seems simple enough: Find ξ_i such that $(\nabla\xi_1 \ \cdots \ \nabla\xi_n) = R^T$. However this is in general not possible. Intuitively the reason is the following.

We notice that a change of variable is n relations. On the other hand the matrix $a_{i j}$, being symmetric, has $n (n + 1)/2$ independent variables. Thus to reduce it to diagonal form, we need $n \geqslant n (n - 1)/2$ which means $n \leqslant 3$. Such counting also reveals why when $n = 2$ we can further reduce the equations to the canonical forms: In this case $n = 2 > 1 = n (n - 1)/2$.

However one can show that when the equation is constant-coefficient, such reduction is always possible. Therefore the above "counting" argument is not satisfactory.

One possibly better explanation is as follows. We consider the case where $\boldsymbol{r}_i \perp \boldsymbol{r}_j$ for each $i \neq j$ (from linear algebra we know that this is possible).

Clearly ξ_i should be defined as follows:

$$D_j \xi_i = \boldsymbol{r}_j \cdot \nabla \xi_i = 0 \qquad j \neq i \tag{2.370}$$

This way $\nabla \xi_i \parallel \boldsymbol{r}_i$. According to Frobenius Theorem in differential geometry, for ξ_i to exist, we need

$$D_k D_j - D_j D_k = \sum_{l \neq i} C_l^{k,j} D_l \tag{2.371}$$

for all $k, j \neq i$. When the equation is constant-coefficient, clearly (2.371) will be satisfied; Similarly, in 2D it is trivially satisfied.

Definition 2.40. *The PDE is called*

i. elliptic if all $\lambda_i > 0$ or all $\lambda_i < 0$.

ii. hyperbolic if all but one are of the same sign.

iii. parabolic if at least one $\lambda_i = 0$.

Remark 2.41. It is clear that the three categories do not cover all possible cases. Interestingly, until today nobody has ever found a significant second order PDE that doesn't belong to any of the three categories.

Remark 2.42. Note that Definition 2.40 requires the following: The number of positive, negative, and zero eigenvalues remains the same under change of variables. The is Sylvester's law of inertia in the quadratic form theory of linear algebra. Basically, what we need to show is

$$R^T \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_{n-p-q} \end{pmatrix} R = \begin{pmatrix} I_{p'} & & \\ & -I_{q'} & \\ & & 0_{n-p'-q'} \end{pmatrix} \tag{2.372}$$

for some nonsingular matrix R, then $p = p'$, $q = q'$.

We prove by contradiction. Assume that $p < p'$. Set $y = R x$. Then for all $x \in \mathbb{R}^n$,

$$x_1^2 + \cdots + x_{p'}^2 - x_{p'+1}^2 - \cdots - x_{p'+q'}^2 = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_{p+q}^2. \tag{2.373}$$

As $p<p'$, we can find $x_1, \ldots, x_{p'}$ not all zero such that

$$y = R\begin{pmatrix} x_1 \\ \vdots \\ x_{p'} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y_{p+1} \\ \vdots \\ y_n \end{pmatrix} \tag{2.374}$$

for some $y_{p+1}, \ldots, y_n$ which may not may not be zeroes.

Then for this particular pair of x, y, we have

$$x_1^2 + \cdots + x_{p'} = -y_{p+1}^2 - \cdots - y_{p+q} \Longrightarrow x_1 = \cdots = x_{p'} = 0 \tag{2.375}$$

Contradiction.

Exercises.

Exercise 2.31. Given a general linear 2nd order equation in $\mathbb{R}^n$:

$$\sum_{i,j}^{n} a_{ij}\, u_{x_i x_j} + \sum_{i=1}^{n} b_i\, u_{x_i} + c\, u + d = 0. \tag{2.376}$$

Show that there are $\tilde{a}_{ij}, \tilde{b}_i, \tilde{c}, \tilde{d}$ with $\tilde{a}_{ij} = \tilde{a}_{ji}$ such that the equation can be written as

$$\sum_{i,j=1}^{n} (\tilde{a}_{ij}\, u_{x_i})_{x_j} + \sum_{i=1}^{n} \tilde{b}_i\, u_{x_i} + \tilde{c}\, u + \tilde{d} = 0. \tag{2.377}$$

Exercise 2.32. Reduce the following equations to canonical form. Then use further transformation

$$u(\xi, \eta) = \exp(\alpha\, \xi + \beta\, \eta)\, v(\xi, \eta) \tag{2.378}$$

and choose the constants α, β to eliminate the first derivative terms.

a) $u_{xx} + 4\, u_{xy} + 3\, u_{yy} + 3\, u_x - u_y + u = 0$;

b) $u_{xx} + 2\, u_{xy} + u_{yy} + 5\, u_x + 3\, u_y + 2\, u = 0$;

c) $u_{xx} - 6\, u_{xy} + 12\, u_{yy} + 4\, u_x + 2\, u = x\, y$.

Do you think a generalization of (2.378) can eliminate first derivative terms for the general constant coefficient case in higher dimensions? Justify your answer.

Exercise 2.33. Consider the general linear 2nd order equation in $\mathbb{R}^n$:

$$\sum_{i,j}^{n} a_{ij}\, u_{x_i x_j} + \sum_{i=1}^{n} b_i\, u_{x_i} + c\, u + d = 0. \tag{2.379}$$

with constant coefficients. Prove that there is a change of variables which reduce the equation to canonical form.

Exercise 2.34. Consider the general linear 2nd order equation in $\mathbb{R}^n$:

$$\sum_{i,j}^{n} a_{ij}\, u_{x_i x_j} + \sum_{i=1}^{n} b_i\, u_{x_i} + c\, u + d = 0. \tag{2.380}$$

Let $\xi_i = \xi_i(x_1, \ldots, x_n)$ be a change of variables. Let the equation in the new coordinates be

$$\sum_{i,j}^{n} \tilde{a}_{ij}\, u_{\xi_i \xi_j} + \sum_{i=1}^{n} \tilde{b}_i\, u_{\xi_i} + \tilde{c}\, u + \tilde{d} = 0. \tag{2.381}$$

If we now classify the equation using eigenvalues of $(\,\tilde{a}_{ij}\,)$, do we get the same result? Justify your answer.

References.

2.7. Meaning of Characteristics.

We have seen that in the case of higher dimensional 2nd order equations, we can no longer reduce the equation to a simple "canonical form". Then what should we mean by talking "characteristics"? The key to extend this idea to more general cases, is to understand the role of characteristics in propagating singularities. We have seen in exercise that for quasilinear 1st order equations, if two solution surfaces intersect along a curve, then this curve must be characteristic. Now imagine a solution that is smooth but with discontinuous derivatives along a curve.

2.7.1. Singularities propogate along characteristics.

Let's return to the first order linear equations

$$a(x, y)\, u_x + b(x, y)\, u_y = c(x, y)\, u + d(x, y). \tag{2.382}$$

and try to look at its characteristics (or more precisely, the projection of its characteristics into the x-y plane).

We know that its characteristics are tangent to (a, b). Now if we represent the characteristics as level sets for a function $\varphi(x, y)$, we would have $\nabla\varphi$ perpendicular to (a, b) thus

$$a(x, y)\, \varphi_x + b(x, y)\, \varphi_y = 0. \tag{2.383}$$

Note that the lower order terms are not involved.

Now consider the following situation. u is smooth everywhere but the the derivatives have a "jump" along a curve give implicitly by $\Phi(x, y) = 0$. Also assume along this curve u itself is still conintuous. This means that the tangential derivative is always continuous across the curve, the "jump" is in the normal derivative, or $\nabla\Phi \cdot \nabla u$ has a jump. But we already know that $(a, b) \cdot \nabla u$ is continuous, so we must have $\nabla\Phi \perp (a, b)$, that is the curve is characteristic.

Alternatively one can argue as follows. Consider (locally) a change of variables $\xi = \Phi(x, y), \eta = \Psi(x, y)$ which is orthonormal. Now applying chain rule we reach

$$(a\, \Phi_x + b\, \Phi_y)\, u_\xi + \ldots = 0 \tag{2.384}$$

Now consider the "jump" across the discontinuity:

$$[(a\, \Phi_x + b\, \Phi_y)\, u_\xi + \ldots] = 0. \tag{2.385}$$

As all other terms are continuous across the surface, we have

$$(a\, \Phi_x + b\, \Phi_y)\, [u_\xi] = 0. \tag{2.386}$$

It follows that

$$a\, \Phi_x + b\, \Phi_y = 0. \tag{2.387}$$

This is exactly the same as (2.383).

Next let's consider the 1D equation in nD:

$$a_1\, u_{x_1} + \cdots + a_n\, u_{x_n} = a. \tag{2.388}$$

Arguing similarly as before, we conclude that a surface along which the derivative is discontinuous must satisfy

$$a_1\, \Phi_{x_1} + \cdots + a_n\, \Phi_{x_n} = 0. \tag{2.389}$$

Note that, if we assume u has discontinuity at the kth derivative along a curve, then we can also decide that the curve must be characteristic.

Now consider the 2nd order equation in 2D:

$$A\, u_{xx} + B\, u_{xy} + C\, u_{yy} + D\, u_x + E\, u_y + F u = G. \tag{2.390}$$

We assume that u is smooth everywhere except along $\Phi=0$, all its first order derivatives are continuous across the curve, while there is a "jump" in its second order derivatives.

Similar to the first order case, we choose Ψ so that the level sets of Ψ are orthogonal to those of Φ, and do the change of variables:

$$(x,y)\mapsto(\xi=\Phi,\eta=\Psi). \tag{2.391}$$

The discontinuity now is along $\xi=0$. As all first order derivatives are continuous across it, we see that the only "jump" we have is $u_{\xi\xi}$.

Changing the variables, we reach

$$(A\,\Phi_x^2+B\,\Phi_x\,\Phi_y+C\,\Phi_y^2)\,u_{\xi\xi}+[\text{terms continuous across } \xi=0]=0. \tag{2.392}$$

Taking the difference across $\xi=0$ we conclude the condition for Φ as

$$A\,\Phi_x^2+B\,\Phi_x\,\Phi_y+C\,\Phi_y^2=0. \tag{2.393}$$

Again we see that it means $\Phi=0$ is a characteristic curve.

Similarly, for higher dimensional case

$$\sum_{i,j}^n a_{ij}\,u_{x_ix_j}+\sum_{i=1}^n b_i\,u_{x_i}+c\,u+d=0. \tag{2.394}$$

we reach

$$\sum_{i,j}^n a_{ij}\,\Phi_{x_i}\,\Phi_{x_j}=0. \tag{2.395}$$

Interestingly, this seems not related to the eigenvectors for the matrix (a_{ij}).

Finally we consider the fully nonlinear case

$$F(x,y,u,u_x,u_y)=0 \tag{2.396}$$

with solution u smooth everywhere but with jump in first order derivative across $\Phi=0$. We still do the change of variables to reach

$$F(\cdots,u_\xi\,\varphi_x+u_\eta\,\psi_x,u_\xi\,\varphi_y+u_\eta\,\psi_y)=0. \tag{2.397}$$

Here .... stands for terms that are continuous across $\xi=0$. Now taking difference across $\xi=0$ we get

$$F(\ldots,[u_\xi]\,\varphi_x+\cdots,[u_\xi]\,\varphi_y+\cdots)=0. \tag{2.398}$$

Here we have to assume that $\varphi=0$ is such that it allows arbitrary sized jump $[u_\xi]$. In other words we assume that the above is true for all possible $[u_\xi]$. This means

$$F_p\,\varphi_x+F_q\,\varphi_y=0. \tag{2.399}$$

2.7.2. Linear first order systems.

We consider the system

$$A(x,y)\,\boldsymbol{u}_x+B(x,y)\,\boldsymbol{u}_y=C(x,y)\,\boldsymbol{u}+\boldsymbol{d}(x,y) \tag{2.400}$$

where $\boldsymbol{u},\boldsymbol{d}$ are vector functions and A,B,C matrices.

Apply the same analysis, we reach the condition for $\Phi=0$ to be a characteristic curve:

$$\det(A\,\varphi_x+B\,\varphi_y)=0. \tag{2.401}$$

Or if the curve is $y=h(x)$, then the condition is

$$\det(B-h'(x)\,A)=0. \tag{2.402}$$

If all the roots $h'(x)$ are real and distinct in some region, the system is said to be totally hyperbolic in that region. Denoting these roots by $\lambda_i(x,y)$ we obtain n characteristic curves:

$$y' = \lambda_i(x,y). \tag{2.403}$$

If all the roots are complex, then it's called elliptic.

When the system is hyperbolic, it can be transformed into canonical form as follows. Let $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$ be eigenvectors of $B^{-1} A$. Let the matrix $R = \left(\boldsymbol{r}_1 \; \ldots \; \boldsymbol{r}_n \right)$. Then let $\boldsymbol{u} = R\,\boldsymbol{v}$. We reach

$$\boldsymbol{v}_y + (R^{-1} B^{-1} A\, R)\, \boldsymbol{v}_x = \hat{C}\, \boldsymbol{v} + \boldsymbol{d} \tag{2.404}$$

but $R^{-1} B^{-1} A\, R = \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ is diagonal. So the system can be replaced by

$$\frac{\partial v_i}{\partial y} + \lambda_i \frac{\partial v_i}{\partial x} = (C\, R\, \boldsymbol{v} + \boldsymbol{d})_i. \tag{2.405}$$

If we further introduce the characteristics given by $\frac{\mathrm{d}x}{\mathrm{d}y} = \lambda_i(x,y)$, the above can be further simplified to

$$\frac{\mathrm{d}v_i}{\mathrm{d}s} = \sum c_{ij}\, v_j + d_i. \tag{2.406}$$

Thus in principle the equation can be solved.

Remark 2.43. (Higher dimensional problem) For higher dimensional problems, such as in 3D

$$A\, \boldsymbol{u}_x + B\, \boldsymbol{u}_y + C\, \boldsymbol{u}_z = \cdots \tag{2.407}$$

such reduction is not possible, since it is in general not possible to simultaneously diagonalize two matrices.

Remark 2.44. (Quasi-linear system) For quasi-linear system

$$A(\boldsymbol{u}, x, y)\, \boldsymbol{u}_y + B(\boldsymbol{u}, x, y)\, \boldsymbol{u}_x = \boldsymbol{c}(\boldsymbol{u}, x, y) \tag{2.408}$$

reduction to form (2.405) may not be possible. The reason is that the eigenvectors may be function of $\boldsymbol{u}$ and consequently when differentiating $\boldsymbol{u}_y = (R\,\boldsymbol{v})_y$, new terms with derivatives of y will appear. In such case, one idea is as follows.

We introduce $\boldsymbol{l}_i$ be left eigenvectors of $B^{-1} A$. Set $v_i = \boldsymbol{l}_i \cdot \boldsymbol{u}$. Now try to find r_i such that $\boldsymbol{l}_i = \mu_i \nabla_{\boldsymbol{u}} r_i$, if such $r_1, \ldots, r_n$ can be found, then the equation can be reduced to

$$\frac{\mathrm{d}r_i}{\mathrm{d}s} = f_i(r_1, \ldots, r_n, x, y) \tag{2.409}$$

along each characteristic curve. Such r_i's are called "Riemann invariants" and are crucial in the study of quasi-linear systems such as gas dynamics. It turns out that in 2D one can always find Riemann invariants.

2.7.3. Classification for general equations.

In the most general case, classification is based on existence of characteristics. For example, if there is no real characteristics, we say the equation is elliptic.

Exercises.

Exercise 2.35. Prove that if u solves

$$\sum_{i,j}^{n} a_{ij}\, u_{x_i x_j} + \sum_{i=1}^{n} b_i\, u_{x_i} + c\, u + d = 0 \tag{2.410}$$

and has a jump discontinuity in 2nd order derivatives across $\Phi = 0$, then

$$\sum_{i,j}^{n} a_{ij}\, \Phi_{x_i}\, \Phi_{x_j} = 0. \tag{2.411}$$

Exercise 2.36. Consider the biharmonic equation

$$u_{xxxx} + 2\, u_{xxyy} + u_{yyyy} = 0. \tag{2.412}$$

Derive the equation for its characteristics and show that the equation is elliptic.

Exercise 2.37. Consider the wave equation

$$u_{tt} = u_{xx} + u_{yy}. \tag{2.413}$$

Assume the characteristic surfaces are given by $t - \varphi(x, y) = c$. Find the equation for φ.

Exercise 2.38. Consider the fully nonlinear 2nd order equation

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0 \tag{2.414}$$

for smooth function $F(x, y, z, p, q, r, s, t)$. Derive the equation for its characteristics.

Exercise 2.39. (Zauderer) Consider the first order linear equation

$$a(x, y)\, u_x + b(x, y)\, u_y = c(x, y)\, u + d(x, y). \tag{2.415}$$

Assume that there is a jump in u_x along a characteristic curve, while u itself is continuous. Denote the differentiation along the characteristic by $\frac{\mathrm{d}}{\mathrm{d}s}$ and the jump in u_x by $[u_x] := u_{x,\mathrm{right}} - u_{x,\mathrm{left}}$. Prove that

$$\frac{\mathrm{d}}{\mathrm{d}s}[u_x] = \left(c - a_x + \left(\frac{a}{b}\right) b_x\right) [u_x] \tag{2.416}$$

and conclude that for linear equations, any jump in u_x can only be propagated, but not generated.

Exercise 2.40. (Zauderer) Let $\varphi(x, y) =$ constant be a family of characteristics for

$$a(x, y)\, u_x + b(x, y)\, u_y = c(x, y)\, u + d(x, y). \tag{2.417}$$

Let $\xi = \varphi(x, y)$ and $\eta = \psi(x, y)$ be perpendicular to it. Show that the equation reduces to

$$(a\, \psi_x + b\, \psi_y)\, u_\eta = c\, u + d. \tag{2.418}$$

Now assume u is continuous across $\xi = 0$ while u_ξ has a jump there, then the jump $[u_\xi]$ satisfies

$$(a\, \psi_x + b\, \psi_y)[u_\xi]_\eta = c\, [u_\xi]. \tag{2.419}$$

Thus the propagation of jumps are determined by a equation.

Exercise 2.41. Let u be a continuous solution of the heat equation $u_t = D\, u_{xx}$. Show that the characteristic curves are $t =$ constant. Then show that no derivative of u can "jump" along a smooth curve. In other words, if u is piecewise smooth, then u is smooth.

Exercise 2.42. Consider the first order quasi-linear equation:

$$a(x, y, u)\, u_x + b(x, y, u)\, u_y = c(x, y, u). \tag{2.420}$$

Assume that u is smooth everywhere except that along $\Phi = 0$ there is a "jump" in its 2nd order derivatives. Derive the equation for Φ. Then consider the case when the "jump" is in its kth derivative and all $(k-1)$th derivatives are continuous across the curve.

References.

2.8. Problems.

2.8.1. Proof of method of characteristics.

Notice that (2.189 – 2.193) for the fully nonlinear equation

$$F(x, y, u, u_x, u_y) = 0 \tag{2.421}$$

are only necessary conditions: If $u(x, y)$ is a solution, then along $\frac{\mathrm{d}x}{\mathrm{d}s} = F_p(x, y, u, u_x, u_y)$, $\frac{\mathrm{d}y}{\mathrm{d}s} = F_q(x, y, u, u_x, u_y)$, $z(s) := u(x(s), y(s))$, $p(s) := u_x(x(s), y(s))$, $q(s) := u_y(x(s), y(s))$ must satisfy (2.191 – 2.193). However, effort is needed to show that they are also sufficient: As long as we can invert $s = S(x, y), \tau = \mathrm{T}(x, y)$, then $u(x, y) := z(S(x, y), \mathrm{T}(x, y))$ indeed solves the original equation. We do this in the following steps.

1. Show that if $F(x, y, z, p, q) = 0$ when $s = 0$, then $F(x, y, z, p, q) = 0$ for all s.

2. Show that if the initial curve $(x_0(\tau), y_0(\tau))$ is not characteristic, then locally we can invert

$$x(s, \tau), y(s, \tau) \Longrightarrow s = S(x, y), \tau = \mathrm{T}(x, y). \tag{2.422}$$

   We need to make sure that s, τ are defined for all $(x,\ y)$ in neighborhood of the initial curve $(x_0(\tau), y_0(\tau))$.

3. Set

$$u(x, y) = z(S(x, y), \mathrm{T}(x, y)). \tag{2.423}$$

   We need to show that

$$u_x(x, y) = p(S(x, y), \mathrm{T}(x, y)), \qquad u_y(x, y) = q(S(x, y), \mathrm{T}(x, y)). \tag{2.424}$$

   Once this is done, the proof ends.

4. Show (2.424) is true. One way to do this is to show that[2.5]

$$U = z_s - p\, x_s - q\, y_s, \qquad V = z_\tau - p\, x_\tau - q\, y_\tau \tag{2.425}$$

   are identically zero. For U this follows directly from (2.189 – 2.193). For V we need to use the relation

$$V_s = U_\tau + p_\tau x_s - p_s\, x_\tau + q_\tau y_s - q_s\, y_\tau \tag{2.426}$$

   on top of (2.189 – 2.193) to get a manageable equation.

2.5. One may be attempted to show that $u_x(x, y) = z_s\, S_x + z_\tau\, \mathrm{T}_x$ satisfies the same equation along the characteristics as p. However the problem is we do not know the equation for z_τ.

2.8.2. Solving Burgers with general initial condition.

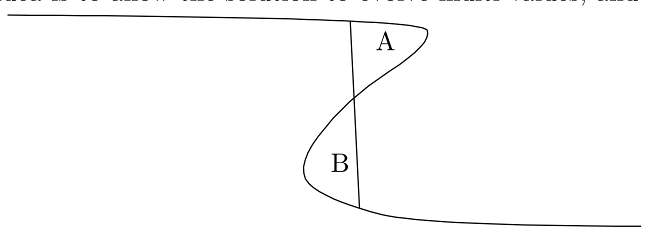
Consider the Burgers equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \qquad u(x,0) = g(x) \tag{2.427}$$

with general $g(x)$. We try to construct entropy solution. There are several issues to settle.

1. For a short time there should be no problem. There should be a time $T > 0$ such that no new discontinuity appears before $t < T$.
2. Any new discontinuity should be a shock, that is $u_L > u_R$.
3. The main problem is how to design a shock in the general case, where u is not constant on each side of the discontinuity. Thus we need to figure out u_L, u_R and the location of the shock $x = s(t)$ simultaneously.

   The idea is to allow the solution to evolve multi-values, and then introduce a "cut":

   The location of the "cut" is the make the areas A and B the same.

   The multivalued solution is given by $u(x,t) = g(\xi)$, $x = \xi + g(\xi)\,t$. We try to find ξ_1, ξ_2 such that

   $$\frac{g(\xi_1) + g(\xi_2)}{2}\,(\xi_2 - \xi_1) = \int_{\xi_1}^{\xi_2} g(\xi)\,\mathrm{d}\xi \tag{2.428}$$

   Together with the shock position:

   $$s(t) = \xi_i + g(\xi_i)\,t \qquad i = 1, 2. \tag{2.429}$$

   One can check that

   $$\dot{s}(t) = \frac{g(\xi_1) + g(\xi_2)}{2}. \tag{2.430}$$

4. Now we apply this idea to the case $g(\xi) = \begin{cases} \geqslant c, \text{“single hump”} & |\xi| \leqslant L \\ c & |\xi| > L \end{cases}$ and try to understand its long time behavior. When t is large, we can expect $\xi_2 > L$ so

   $$\frac{1}{2}\,(g(\xi_1) - c)\,(\xi_2 - \xi_1) = \int_{\xi_1}^{L} (g(\xi) - c)\,\mathrm{d}\xi. \tag{2.431}$$

   From the $s(t)$ equations we can obtain $t = (\xi_2 - \xi_1)/(g(\xi_1) - c)$ which gives

   $$\frac{1}{2}\,[g(\xi_1) - c]^2\,t = \int_{\xi_1}^{L} (g(\xi) - c)\,\mathrm{d}\xi. \tag{2.432}$$

   Now as $t \longrightarrow \infty$, we should expect $\xi_1 \searrow -L$ and obtain the asymptotic behavior of the shock as $s(t) \sim c\,t + \sqrt{2\,A\,t}$.

5. Similar argument can be applied to the case where $g(\xi)$ looks like $\begin{cases} \sin(\xi/L_1) & 0 < \xi \leqslant L_1 \\ \sin(\xi/L_2) & -L_2 \leqslant \xi \leqslant 0 \\ 0 & \text{elsewhere} \end{cases}$.

Reference.

- G. B. Whitham, "Linear and Nonlinear Waves", §2.8, §2.9.

2.8.3. Viscous Burgers equation.

Consider the viscous Burgers equation

$$u_t + u\, u_x = \nu\, u_{xx}, \qquad u(x,0) = g(x). \tag{2.433}$$

If we introduce a new unknown function v through the Cole-Hopf transformation

$$u = -2\,\nu\,\frac{v_x}{v}, \tag{2.434}$$

then the equation becomes

$$v_t = \nu\, v_{xx}. \tag{2.435}$$

Solving this heat equation leads to

$$u(x,t) = \frac{\displaystyle\int_{-\infty}^{\infty} \frac{x-y}{t}\, e^{-G(x,y,t)/2\nu}\, \mathrm{d}y}{\displaystyle\int_{-\infty}^{\infty} e^{-G(x,y,t)/2\nu}\, \mathrm{d}y} \tag{2.436}$$

where

$$G(x,y,t) = \int_0^y g(\xi)\, \mathrm{d}\xi + \frac{(x-y)^2}{2\,t}. \tag{2.437}$$

We study the behavior of $u(x,t)$ as $\nu \longrightarrow 0$.

Applying method of steepest descents[2.6] we obtain, when there is only one stationary point, $u(x,t) \sim \frac{x-y}{t}$ which is the same as $u = g(y), x = y + F(y)\, t$. This is exactly the solution for the inviscid Burgers equation $u_t + u\, u_x = 0$ using method of characteristics.

Since $\frac{\partial G}{\partial y} = g(y) - \frac{x-y}{t}$, when characteristics cross one another, the number of stationary points of $G(x,y,t)$ becomes more than one. In the case that there are two, y_1, y_2, one can actually derive that

$$\frac{1}{2}\,[g(y_1) + g(y_2)]\,(y_1 - y_2) = \int_{y_2}^{y_1} g(\xi)\, \mathrm{d}\xi \tag{2.438}$$

which leads to the jump condition.

Reference.

- G. B. Whitham, "Linear and Nonlinear Waves", Chap. 4.

2.6. Basic idea: the main contribution of the integral in (2.436) comes from around stationary points of $G(x,y,t)$.

2.8.4. Monge cone.

Consider the fully nonlinear first order equation

$$F(x, y, u, u_x, u_y) = 0. \tag{2.439}$$

An alternative approach (more deductive) to derive the characteristics equations (2.189 – 2.193) is the following. We still use the notation $F = F(x, y, z, p, q)$.

- First look at $F(x,\ y,\ z,\ p,\ q) = 0$ as a algebraic equation. Assuming $F_p^2 + F_q^2 \neq 0$ (why is this reasonable?), we can apply Implicit Function Theorem to represent, at a fixed point (x_0, y_0, u_0), q as a function of p: $q = q(p)$.

- Recall that $(p,\ q,\ -1)$ should be the normal direction to the (to be determined) solutions surface $u(x, y) - u = 0$. Thus the solution surface should has a tangent vector contained in the plane

$$u - u_0 - p\,(x - x_0) - q(p)\,(y - y_0) = 0. \tag{2.440}$$

- Consider the envelope of all such planes. It is a cone called Monge cone, and the solution surface should be tangent to it.

- Review the basic theory of envelope: Let $F(x, y, \ldots; \lambda) = 0$ be a class of surfaces parametrized by λ. Its envelope $G(x, y, \ldots) = 0$ is a single surface which is tangent to each of the original surfaces at some point(s). So basically $G = F(x, y, \ldots; \lambda(x, y, \ldots))$ for some specific choice of λ as a function of $x, y, \ldots$. Now for the two surfaces to be tangent, we need their normals to be the same:

$$\begin{pmatrix} F_x(\cdots;\lambda) \\ \vdots \end{pmatrix} = \begin{pmatrix} G_x \\ \vdots \end{pmatrix}. \tag{2.441}$$

  This can be fulfilled if we require $F_\lambda = 0$ at their common points. Consequently the envelope is obtained through cancelling λ using $F = 0$ and $F_\lambda = 0$.

- Now observe that $F(x, y, u, p, q(p)) = 0$ implies $\frac{\mathrm{d}F}{\mathrm{d}p} = 0$ which gives $F_p + q'(p)\,F_q = 0$. Substituting this value of $q'(p)$ into the envelope system we get

$$\frac{x - x_0}{F_p} = \frac{t - t_0}{F_q} = \frac{u - u_0}{p\,F_p + q\,F_q}. \tag{2.442}$$

  This gives

$$\frac{\mathrm{d}x}{\mathrm{d}s} = F_p; \qquad \frac{\mathrm{d}y}{\mathrm{d}s} = F_q; \qquad \frac{\mathrm{d}u}{\mathrm{d}s} = p\,F_p + q\,F_q. \tag{2.443}$$

- Finally determine the equations for p, q by applying ∂_x and ∂_y to $F(x, y, u(x, y), p(x, y), q(x, y)) = 0$ and use the fact $p_y = q_x$ and (2.443).

Remark 2.45. Note that, the equations for p, q are only necessary conditions for them to be u_x, u_y. See Section 2.8.1 for justification.

Reference.

- Erich Zauderer "Partial Differential Equations of Applied Mathematics" 2ed 1998. §2.4.

2.8.5. Complete Integral.

Consider the general nonlinear first order equation

$$F(x, y, u, u_x, u_y) = 0. \tag{2.444}$$

One way to find its general solution is through the following method:

- Find a two-parameter family of solutions:

$$u = \phi(x, y; a, b). \tag{2.445}$$

  This sometimes can be easily done, for example for the Clairaut's equation $u = x\, u_x + y\, u_y + f(u_x, u_y)$ we can easily see that $u = a\ x + b\ y + f(a,\ b)$ solves the equation. Such an expression is called a "complete integral" of the equation. Note that there are two arbitrary constants instead of one arbitrary function.

- The general solution can be obtained as follows. Relate b and a arbitrarily through $b = b(a)$. Then solve the envelope equations

$$\begin{aligned} u &= \phi(x, y; a, b(a)); & (2.446) \\ \phi_a(x, y; a, b(a)) + \phi_b(x, y; a, b(a))\, b'(a) &= 0. & (2.447) \end{aligned}$$

  The second equation gives $a = a(x, y)$ and then $b(a(x, y))$ becomes the arbitrary function in the formula for general solution. To prove that this indeed give a solution to (2.444) we only need to check $u_x = \phi_x$ and $u_y = \phi_y$.

- It is also OK to introduce a parameter t and set $a = a(t)$, $b = b(t)$ as arbitrary functions. Why this will not give a general solution with two (instead of one) arbitrary functions?

- One can also consider the envelope of the fulll two-parameter family (2.445) and solve

$$u = \phi(x, y; a, b), \quad \phi_a(x, y; a, b) = 0, \quad \phi_b(x, y; a, b) = 0. \tag{2.448}$$

  This will give us a "singular solution" which cannot be obtained through the general solution. This can be done when $\det \begin{pmatrix} \phi_{a\,a} & \phi_{a\,b} \\ \phi_{b\,a} & \phi_{b\,b} \end{pmatrix} \neq 0$. Note that this condition means that in general $\begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix}$ is nonsingular that is b is not a function of a.

- One can further show that for any fixed a, the solution to the system $(2.446 - 2.447)$ gives a characteristic curve. More specifically, for fixed a write (2.447) as $\phi_a = \lambda\ \sigma$, $\phi_b = \mu\ \sigma$ for one parameter σ, and write

$$x = x(\sigma), y = y(\sigma), u = u(\sigma), p = \phi_x(x, y; a, b) = p(\sigma), q = \phi_y(x, y; a, b) = q(\sigma), \tag{2.449}$$

  and further assume $F_u \neq 0$, $\det \begin{pmatrix} \phi_{a\,x} & \phi_{a\,y} \\ \phi_{b\,x} & \phi_{b\,y} \end{pmatrix} \neq 0$, differentiating ϕ_a, ϕ_b with respect to σ and then differentiating $F(x, y, \phi, \phi_x, \phi_y) = 0$ with respect to a, b, we can conclude (after introducting a new parameter s) that

$$\frac{\mathrm{d}x}{\mathrm{d}s} = F_p, \qquad \frac{\mathrm{d}y}{\mathrm{d}s} = F_q. \tag{2.450}$$

- The above analysis breaks down when $F_u = 0$, that is when u does not explicitly appear in the equation. In this case we need to have complete integral of the form $u = \phi(x, y; a) + b$.

- The above can be generalized to the higher dimensional case, where a complete integral looks like $u = \phi(x_1, \ldots, x_n; a_1, \ldots, a_n)$.

Reference.

- Paul R. Garabedian "Partial Differential Equations" §2.3.

2.8.6. 2nd Order Equations in 2D.

It turns out to be possible to reduce 2nd order equations in 2D to "canonical form".

- First consider the quasi-linear case:

$$a\,u_{xx}+b\,u_{xy}+c\,u_{yy}+d=0 \tag{2.451}$$

where a,b,c,d are functions of $x,y,u,p:=u_x,q:=u_y$.

  - Now consider such a curve in the plane along which the 2nd order derivatives cannot be uniquely determined. Taking (2.451) together with

$$\mathrm{d}p=u_{xx}\,\mathrm{d}x+u_{xy}\,\mathrm{d}y;\qquad \mathrm{d}q=u_{xy}\,\mathrm{d}x+u_{yy}\,\mathrm{d}y \tag{2.452}$$

  we discover that for u_{xx},u_{xy},u_{yy} to be not unique, we need

$$\det\begin{pmatrix} a & b & c\\ \mathrm{d}x & \mathrm{d}y & 0\\ 0 & \mathrm{d}x & \mathrm{d}y\end{pmatrix}=0\Longleftrightarrow a\,(\mathrm{d}y)^2-b\,(\mathrm{d}x)\,(\mathrm{d}y)+c\,(\mathrm{d}x)^2=0. \tag{2.453}$$

  Solving this quadratic equation – assuming that we are in the hyperbolic case – we obtain two families of curves $\xi=\text{constant}$, $\eta=\text{constant}$.

  - Writing the two roots as $r_{1,2}$, we conclude that

$$y_\xi-r_1\,x_\xi=0;\qquad y_\eta-r_2\,x_\eta=0. \tag{2.454}$$

  Note that r_1,r_2 are functions of x,y,u,p,q. Now we need three more equations regarding u, p, q to form a system of 5 equations and 5 unknowns.

  - Recall the equations satisfied by u_{xx}, u_{yy}, u_{xy}: (2.451),(2.452) as well as (2.453), we realize that for this system to have solution we need the rank of

$$\begin{pmatrix} a & b & c & d\\ \mathrm{d}x & \mathrm{d}y & 0 & -\mathrm{d}p\\ 0 & \mathrm{d}x & \mathrm{d}y & -\mathrm{d}q\end{pmatrix} \tag{2.455}$$

  to be at most 2. This means

$$\det\begin{pmatrix} a & c & d\\ \mathrm{d}x & 0 & -\mathrm{d}p\\ 0 & \mathrm{d}y & -\mathrm{d}q\end{pmatrix}=0\Longleftrightarrow a\,\mathrm{d}y\,\mathrm{d}p+c\,\mathrm{d}x\,\mathrm{d}q+d\,\mathrm{d}x\,\mathrm{d}y=0. \tag{2.456}$$

  This, together with (2.453), gives two equations for p,q:

$$p_\xi+r_2\,q_\xi+\frac{d}{a}\,x_\xi=0;\qquad p_\eta+r_1\,q_\eta+\frac{d}{a}\,x_\eta=0. \tag{2.457}$$

  - Finally, the equation for u is simply one of the following two[2.7]

$$u_\xi-p\,x_\xi-q\,y_\xi=0,\qquad u_\eta-p\,x_\eta-q\,y_\eta=0. \tag{2.458}$$

  - One can prove that the solution to this 5×5 system indeed gives the solution to the original equation.
  - One can further differentiate the 5 equations to obtain a 5×5 system for the mixed derivatives. It turns out that this system has unique solution:

$$x_{\xi\eta}+\cdots=0,\ y_{\xi\eta}+\cdots=0,\ u_{\xi\eta}+\cdots=0,\ p_{\xi\eta}+\cdots=0,\ q_{\xi\eta}+\cdots=0. \tag{2.459}$$

  Here $\cdots$ denotes lower order terms. Thus we see that this generalizes the reduction to canonical form from linear equations to quasi-linear equations.

- In the case $b^2-4\,a\,c<0$, similar to the linear case, one can reach the canonical form similar to (2.459) but with $x_{\xi\eta}$ etc. replaced by $x_{\alpha\alpha}+x_{\beta\beta}$ etc.

2.7. Because they are not independent.

- This approach also works for fully nonlinear equations

$$F(x, y, u, p, q, r, s, t) = 0 \tag{2.460}$$

resulting in a system of 8 first order equations.

- Examples.

Example 2.46. Consider the equation of minimal surfaces:

$$(1 + q^2)\, r - 2\, p\, q\, s + (1 + p^2)\, t = 0. \tag{2.461}$$

The canonical system reads

$$x_{\alpha\alpha} + x_{\beta\beta} = 0, \qquad y_{\alpha\alpha} + y_{\beta\beta} = 0, \qquad u_{\alpha\alpha} + u_{\beta\beta} = 0 \tag{2.462}$$

together with

$$x_\alpha^2 + y_\alpha^2 + u_\alpha^2 = x_\beta^2 + y_\beta^2 + u_\beta^2, \qquad x_\alpha\, x_\beta + y_\alpha\, y_\beta + u_\alpha\, u_\beta = 0. \tag{2.463}$$

Thus α, β are the conformal coordinates of the surface.

Reference.

- Paul R. Garabedian "Partial Differential Equations" §3.4.