

First-Order Partial Differential Equations and Characteristics

N-20260619

§1.1 What a first-order PDE asks for

A partial differential equation is an equation for an unknown function of several variables and its partial derivatives. A first-order scalar equation for $u = u(x, t)$ has the schematic form

$$F(x, t, u, u_x, u_t) = 0.$$

The simplest useful class is the linear transport equation

$$u_t + au_x = b(x, t),$$

where a is a constant. It is called a transport equation because it moves the initial profile without changing its shape, except for the forcing term b .

The geometric method for first-order equations is the method of characteristics. Instead of trying to understand u on the whole (x, t) -plane at once, one looks for curves along which the PDE becomes an ordinary differential equation.

Definition 1.1.1 (Characteristic curves). For the constant-coefficient transport equation

$$u_t + au_x = b(x, t),$$

a characteristic is a curve $x = x(t)$ satisfying

$$\frac{dx}{dt} = a.$$

Along such a curve,

$$\frac{d}{dt}u(x(t), t) = u_x(x(t), t)x'(t) + u_t(x(t), t) = au_x + u_t = b(x(t), t).$$

Thus a first-order PDE has been replaced by two ordinary differential equations:

$$\frac{dx}{dt} = a, \quad \frac{du}{dt} = b(x(t), t).$$

The first equation describes the characteristic line, and the second describes the value of the solution along that line.

§1.2 The unforced transport equation

Consider

$$u_t - cu_x = 0.$$

This is the sign convention used by the first-order wave example in the original notes. It has the form $u_t + au_x = 0$ with $a = -c$, so the characteristics satisfy

$$\frac{dx}{dt} = -c.$$

Hence

$$x(t) = x_0 - ct, \quad x_0 = x + ct.$$

Along each such line,

$$\frac{d}{dt}u(x(t), t) = u_t - cu_x = 0.$$

Therefore u is constant on the line $x + ct = x_0$.

Theorem 1.2.1 (Solution of the constant-speed transport equation)

Let f be a differentiable function. The solution of

$$u_t - cu_x = 0, \quad u(x, 0) = f(x)$$

is

$$u(x, t) = f(x + ct).$$

Proof. The initial point of the characteristic through (x, t) is $(x + ct, 0)$. Since u is constant along the characteristic,

$$u(x, t) = u(x + ct, 0) = f(x + ct).$$

A direct check gives

$$u_t = cf'(x + ct), \quad u_x = f'(x + ct),$$

so $u_t - cu_x = 0$. □

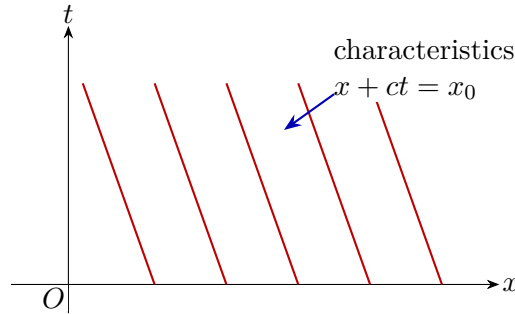


Figure 1.1: Characteristics for $u_t - cu_x = 0$. The solution is constant along each red line and the initial profile moves to the left as t increases.

§1.3 Example: translating a parabola

Let

$$u_t - cu_x = 0, \quad u(x, 0) = x^2 - 3.$$

The theorem gives

$$u(x, t) = (x + ct)^2 - 3.$$

For each fixed time t , the graph in the (x, u) -plane is the same parabola shifted left by ct .

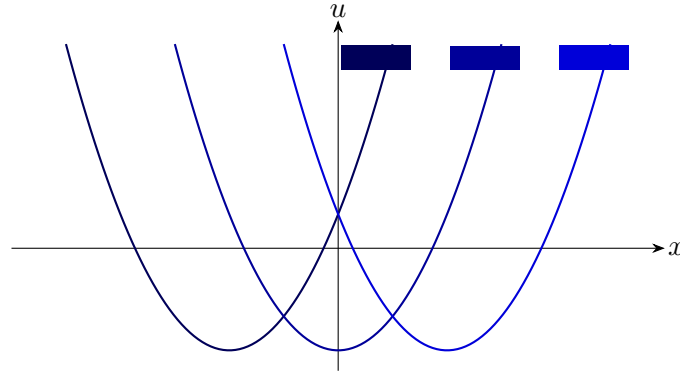


Figure 1.2: Time slices of $u(x, t) = (x + ct)^2 - 3$, shown as translated copies of the initial profile.

§1.4 Forced transport

Now consider

$$u_t + au_x = b(x, t), \quad u(x, 0) = f(x).$$

The characteristic through (x, t) has the form

$$x(s) = x - a(t - s), \quad 0 \leq s \leq t,$$

because $x(t) = x$. Along this path,

$$\frac{d}{ds}u(x(s), s) = b(x(s), s).$$

Integrating from 0 to t gives the formula

$$u(x, t) = f(x - at) + \int_0^t b(x - a(t - s), s) ds.$$

Example 1.4.1 (A forced first-order wave equation)

Solve

$$u_t + 5u_x = e^{-t}, \quad u(x, 0) = e^{-x^2}.$$

Here $a = 5$, so

$$x(s) = x - 5(t - s).$$

The forcing depends only on time, and therefore

$$\int_0^t e^{-s} ds = 1 - e^{-t}.$$

The initial point is $x - 5t$. Hence

$$u(x, t) = e^{-(x-5t)^2} + 1 - e^{-t}.$$

Equivalently, if one writes the characteristic label as $x_0 = x - 5t$, then $u = f(x_0) + 1 - e^{-t}$. The sign of the last term is fixed by differentiating the proposed solution back into the PDE.

§1.5 What can go wrong

For first-order equations, the initial curve must meet the characteristic curves transversely. If the data are prescribed on a characteristic, then the PDE may fail to determine a unique solution. For constant-speed transport, prescribing $u(x, 0) = f(x)$ is safe because the initial line $t = 0$ intersects each characteristic once. This is the first place where PDEs differ from ordinary differential equations: the geometry of the set on which data are given is part of the problem.

§1.6 General linear first-order equations

The constant-coefficient equation is only the first model. A more general linear first-order equation is

$$a(x, t)u_x + b(x, t)u_t = c(x, t)u + d(x, t).$$

The characteristic curves now satisfy

$$\frac{dx}{ds} = a(x, t), \quad \frac{dt}{ds} = b(x, t).$$

Along such a curve, the value of u satisfies

$$\frac{du}{ds} = c(x(s), t(s))u + d(x(s), t(s)).$$

Thus even when the coefficients vary, the method is still the same: solve a characteristic system in the independent variables, then solve an ordinary differential equation for the dependent variable along the characteristic.

Example 1.6.1 (Variable speed transport)

Consider

$$u_t + xu_x = 0, \quad u(x, 0) = f(x).$$

The characteristics solve

$$\frac{dx}{dt} = x,$$

so $x(t) = x_0 e^t$. The initial footpoint through (x, t) is therefore $x_0 = x e^{-t}$. Since u is constant along characteristics,

$$u(x, t) = f(x e^{-t}).$$

A direct check gives

$$u_t = -x e^{-t} f'(x e^{-t}), \quad x u_x = x e^{-t} f'(x e^{-t}),$$

and hence $u_t + x u_x = 0$.

§1.7 Quasilinear equations

A first-order equation is called quasilinear if it is linear in the derivatives but the coefficients may depend on u :

$$a(x, t, u)u_x + b(x, t, u)u_t = c(x, t, u).$$

The characteristic system is now

$$\frac{dx}{ds} = a(x, t, u), \quad \frac{dt}{ds} = b(x, t, u), \quad \frac{du}{ds} = c(x, t, u).$$

This is already a nonlinear system of ordinary differential equations. In this sense, first-order nonlinear PDEs are not an entirely new object: they are families of coupled ODEs whose curves fill the plane.

Example 1.7.1 (Inviscid Burgers equation)

The equation

$$u_t + uu_x = 0$$

has characteristic speed u itself. Along a characteristic,

$$\frac{du}{dt} = u_t + uu_x = 0,$$

so u is constant. If $u(x, 0) = f(x)$, then the characteristic starting from x_0 has

$$x = x_0 + f(x_0)t, \quad u = f(x_0).$$

Thus the implicit solution is obtained from

$$x_0 = x - ut, \quad u = f(x - ut).$$

This formula is useful only as long as the map $x_0 \mapsto x_0 + f(x_0)t$ remains one-to-one.

§1.8 Crossing characteristics and shocks

For linear transport with constant speed, characteristics never cross. For nonlinear equations such as Burgers' equation, faster characteristics can overtake slower ones. When this happens, the classical solution becomes multi-valued if one continues the characteristic construction naively.

Differentiate the characteristic map for Burgers' equation:

$$x = x_0 + f(x_0)t.$$

The map stops being locally invertible when

$$\frac{\partial x}{\partial x_0} = 1 + f'(x_0)t = 0.$$

Thus if $f'(x_0) < 0$ somewhere, characteristic crossing occurs at time

$$t_* = -\frac{1}{\min f'(x_0)}$$

provided the minimum is negative. This is the first appearance of shock formation.

In an introductory PDE course, one often stops here and says that classical differentiable solutions break down. A more advanced treatment replaces them by weak solutions and adds an entropy condition to select the physically relevant shock. The point for this note is more elementary: first-order PDEs are controlled by the geometry of characteristic curves, and loss of uniqueness in that geometry is visible before one writes any distribution theory.

§1.9 Initial curves and well-posedness

Let initial data be prescribed on a curve Γ in the (x, t) -plane. For the characteristic method to determine a solution near Γ , the curve must not be tangent to the characteristic direction. For

$$a(x, t)u_x + b(x, t)u_t = c(x, t, u),$$

the characteristic direction in the (x, t) -plane is (a, b) . If Γ has tangent vector τ , then the non-characteristic condition is

$$\tau \nparallel (a, b).$$

Equivalently, the characteristic curves must cross the data curve rather than slide along it.

Example 1.9.1 (Bad data curve)

For $u_t + u_x = 0$, the characteristics are lines $x - t = \text{const.}$. If data are prescribed on the line $x - t = 0$, then the data are given along a characteristic. The equation says the solution is constant on that line, but it does not determine what happens on neighboring characteristic lines. Therefore the problem is not locally well-posed from that data curve.

§1.10 Conservation-law viewpoint

Some first-order equations are written in conservation form:

$$u_t + \partial_x F(u) = 0.$$

For smooth solutions, this is the same as

$$u_t + F'(u)u_x = 0.$$

The characteristic speed is $F'(u)$. Burgers' equation corresponds to $F(u) = u^2/2$. Conservation form becomes important when shocks appear, because it still has meaning after differentiability is lost.

Integrating over an interval $[a, b]$, we obtain

$$\frac{d}{dt} \int_a^b u(x, t) dx = F(u(a, t)) - F(u(b, t)).$$

Thus the amount of u in the interval changes only by flux through the boundary. This is the conceptual reason that conservation laws are the natural nonlinear continuation of transport equations.