# 4 Introduction to First-Order Partial Differential Equations

You have already encountered examples of first-order PDEs in Section 3 when we discussed the pure advection mechanism, and mentioning the interesting model case of traffic flow theory. Here are a few more examples.

*Example 1*: From the glass industry, the drawing of an optical fiber of cross-sectional area $A(x,t)$, with velocity $v(x,t)$, can be modeled by what is called an "extensional flow". (See Figure 1.) The principle viscous resistance is a normal stress proportional to $\partial v/\partial x$, giving a force proportional to $A\frac{\partial v}{\partial x}$. From conservation of mass and momentum balance, then

$$\frac{\partial A}{\partial t} + \frac{\partial}{\partial x}(Av) = 0$$ [1]

(A two-dimensional plate version of this, with $2h(x,y,t)$ being plate thickness, becomes

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(uh) + \frac{\partial}{\partial y}(vh) = 0;$$

however, throughout these Notes the problems we study will be of one space variable and one time variable, or two space variables, for convenience of getting our solution methodology ideas across.)

*Example 2*: Flow of a thin layer of paint down a wall[2] (see Figure 2)

Assumptions:

1. The paint layer is thin; hence, velocity $u$ is approximately unidirectional;
2. Gravity is resisted by viscosity of paint; therefore, shear force is proportional to $\frac{\partial u}{\partial y}$;

[1] The author once developed a model for the swimming behavior of an Aplysia utilizing this same equation.

[2] The author got this example from *Applied Partial Differential Equations*, by J. Ockendon, et al, Oxford Univ. Press, 1999.

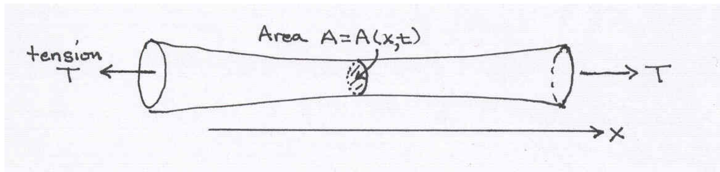


Figure 1: Forming a optical fiber

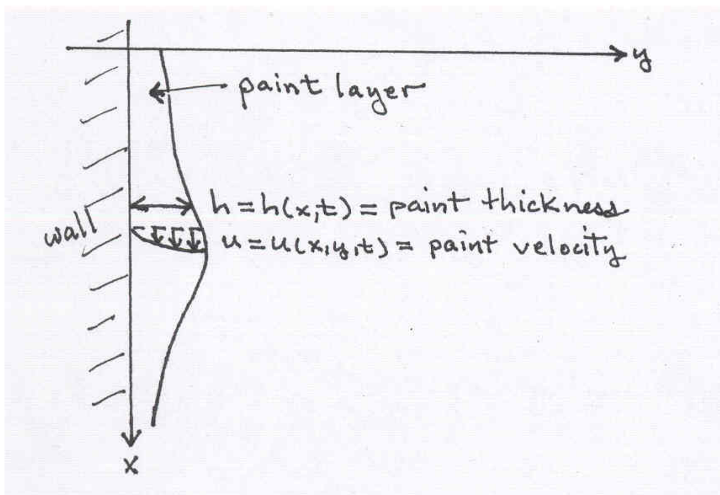


Figure 2: Paint running down wall model

3. A force balance on a small fluid element shows $\frac{\partial^2 u}{\partial y^2}$ = constant $= -c$, with $c$ proportional to acceleration of gravity.

4. Shear force equals zero at the paint surface: $\frac{\partial u}{\partial y}_{|y=h} = 0$;

5. Paint sticks to the wall: $u_{|y=0} = 0$, which implies $u = \frac{1}{2}cy(2h - y)$ from the above considerations;

6. The rate of change of paint thickness must be balanced by the x-variation of the paint flowing down the wall. This flux is $q = q(x,t) = \int_0^h u\, dy$.

By a mass conservation argument (avoided here), we obtain

$$\frac{\partial h}{\partial t} + \frac{\partial q}{\partial x} = 0 \quad \text{or} \quad \frac{\partial h}{\partial t} + \frac{\partial}{\partial x}\left(\int_0^h u\, dy\right) = 0 \quad \text{or}$$

$$\frac{\partial h}{\partial t} + ch^2 \frac{\partial h}{\partial x} = 0 \ .$$

*Example 3*: As an example of one connection between pdes and probability theory, one often considers a probability of something happening to $n$ "things" at time $t$, $p_n(t)$, and where these probabilities are related through an (infinite) sequence of differential equations. By employing a *generating function* approach, a first-order pde arises.

Suppose, for example, that $p_n(t)$ represents the probability that $n$ fish are caught by time $t$; for those who have had some probability, we are discussing a Poisson process that describes the fish catching probability. From linear birth-death theory, $p_n(t)$ satisfies equations ($\lambda$ is a positive constant)

$$\frac{dp_n}{dt} = -n\lambda p_n + (n-1)\lambda p_{n-1}, \quad n = 1, 2, \ldots$$

$$p_1(0) = 1, \quad p_j(0) = 0 \quad j > 1 \quad (p_0 \equiv 0)$$

Let $F(x,t)$ be the generating function for the sequence $\{p_n\}$; that is, $F(x,t) = \sum_{n=0}^{\infty} p_n(t) x^n$. Thus, multiplying the ode for $p_n$ by $x^n$ and summing, we obtain

$$\sum_{n\geq 0} x^n \frac{dp_n}{dt} = -\sum_{n\geq 0} n x^n \lambda p_n + \sum_{n\geq 0} (n-1) x^n \lambda p_{n-1} \quad \text{or}$$

$$\frac{\partial}{\partial t} \sum_{n\geq 0} x^n p_n = -\lambda x \frac{\partial}{\partial x} \sum_{n\geq 0} x^n p_n + \lambda x^2 \frac{\partial}{\partial x} \sum_{n\geq 1} x^{n-1} p_{n-1} \ , \quad \text{or}$$

$$\frac{\partial F}{\partial t} = -\lambda x \frac{\partial F}{\partial x} + \lambda x^2 \frac{\partial F}{\partial x} \ ,$$

or

$$\frac{\partial F}{\partial t} + \lambda x (1-x) \frac{\partial F}{\partial x} = 0 \ , \quad \text{with} \quad F(0,x) = x \ . \tag{1}$$

*Exercise*: Show that $F(t,x) = \frac{xe^{-\lambda t}}{1-x+xe^{-\lambda t}}$ is the solution to problem (1).

*Example 4*: Although we will not discuss systems of first-order equations, systems are important in practice. One example is a special case of one-dimensional unsteady gas dynamics[3], namely

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) = 0 \ , \qquad \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = -\frac{1}{\rho} \frac{\partial p}{\partial x} \ ,$$

[3]see for example, reference 2 on page 1.

$$\rho(\frac{\partial}{\partial t} + u\frac{\partial}{\partial x})(\frac{p}{(\gamma-1)\rho} + \frac{1}{2}u^2) + \frac{\partial}{\partial x}(pu) = 0$$

where $\gamma > 1$. Here $\rho, p, u$ are gas density, pressure, and velocity, respectively, and these equations represent mass, force, and energy balances, respectively.

The first and third examples provide linear pdes, the type we concentrate on in this section, and ones for which we only have to specify "initial conditions".

## 4.1 Classification of First-Order PDEs

From the previous discussion on the one-space dimension conservation of mass equation, equations involving just first derivatives can arise in a variety of circumstances. The general equation involving functions of two independent variables has the form

$$F(x, y, u(x, y), \frac{\partial u}{\partial x}(x, y), \frac{\partial u}{\partial y}(x, y)) = 0 \ ,$$

but this is too general of an equation to attack at this point. Consider instead the equation

$$a\frac{\partial u}{\partial x} + b\frac{\partial u}{\partial y} = c \quad . \tag{2}$$

Then

1. If coefficients $a, b$ depend at most on $x, y$, and $c$ has the form $c = c_1(x, y)u + c_2(x, y)$, then (2) is a **linear** pde;
2. If $a, b$ depend as most on $x, y$, and $c$ does not necessarily depend linearly on $u$, then (2) is a **semilinear** equation;
3. If $a, b, c$ depend on $x, y, u$, then (2) is still linear in the derivatives of $u$, so (2) is called a **quasilinear** equation.

The method of solution we discuss below holds for both linear and semilinear equations, but it has to be modified for quasilinear equations, like example 2 above.

## 4.2 Characteristic Equations for (2)

The characteristic equations for (2) are given by

$$\left.\begin{aligned} \tfrac{\partial x}{\partial \tau} &= a(x, y) \\ \tfrac{\partial y}{\partial \tau} &= b(x, y) \\ \tfrac{\partial u}{\partial \tau} &= c(x, y, u) \end{aligned}\right\} \tag{3}$$

The solution curves are called **characteristic curves** of the pde (2). The reason for the form of equation (2) is that it is a dot product:

$$(a, b, c) \cdot (u_x, u_y, -1) = 0 \ .$$

Over the $x, y$-domain in which the equation is suppose to hold, one can imagine a solution surface $z = u(x, y)$ in $x, y, z$ space. From multivariable calculus we know that $(u_x, u_y, -1)$ is the normal vector to this surface, so $(a, b, c)$ being orthogonal to it must be tangent to the surface; since $(x_\tau, y_\tau, u_\tau)$ is also tangent, we arrive at system (3).

## 4.3 The Cauchy Problem

We assume the coefficients $a, b, c$ are continuous functions in their arguments $x, y$ throughout a specified domain $\Omega \subset \mathbb{R}^2$, and let $\Gamma$ be a simple curve in $\Omega$; that is, it is non-self intersecting. We can then index $\Gamma$ by a parameter $s$: $\Gamma := \{(x, y) = (\varphi(s), \psi(s)) : s \in \mathbb{I} \subset \mathbb{R}\}$. Here $\varphi, \psi$ are considered differentiable functions of $s$. Then assume $u = w(s)$ is prescribed on curve $\Gamma$. We need a condition to guarantee we have a solution $u(x, y)$ in a neighborhood of $\Gamma$, but let us proceed formally first to give a recipe for finding a solution to (2) and work out a few examples. Given a specific equation (2) and specific initial data $w$ on specified $\Gamma$,

1. From the initial condition, write down a parameterization of $\Gamma$ in terms of $(\varphi(s), \psi(s))$.

2. Write down the characteristic equations for $x$ and $y$, and solve them, subject to $x = \varphi(s), y = \psi(s)$ at $\tau = 0$. The solution is formally given by $x = f(\tau, s), y = g(\tau, s)$.

3. Invert these functions to obtain $\tau = \hat{f}(x, y), s = \hat{g}(x, y)$.

4. Solve the equation for $u$ in terms of $(\tau, s)$,
$$\frac{\partial u}{\partial \tau} = c(f(\tau,s), g(\tau,s), u(\tau,s)) \ , \text{ subject to } \quad u|_{\tau=0} = w(s) \ .$$
Say the solution is $u = h(\tau, s)$.

5. Substituting in from step (3), we finally have $u(x,y) = h(\hat{f}(x,y), \hat{g}(x,y))$ as the solution to the Cauchy problem.

A point to emphasize here is that in solving the pde (2), the $x, y$ coordinates are not the convenient coordinates in which to solve the problem. We must transform over to the local $\tau, s$ coordinates. At any given point in the domain we are on a characteristic curve $\tau$ units away from the initial curve $\Gamma$. We can think of running back down the curve until we get to $\tau = 0$, at which point we are somewhere on $\Gamma$ at location $s$.

*Example 5*: Linear advection example

$$\begin{aligned} &\frac{\partial \rho}{\partial t} + c\frac{\partial \rho}{\partial x} = 0 \quad |x| < \infty, t > 0 \quad c > 0 \quad \text{is a constant} \\ &\rho(x,0) = \rho_0(x) \quad , \quad \rho_0 \quad \text{is continuously differentiable} \end{aligned}$$

Here $\Gamma$ is the $x$-axis, so $x = \varphi(s) = s, t = \psi(s) = 0$. Then

$$\frac{dt}{d\tau} = 1, \quad t_{|\tau=0} = 0 \to t = \tau, \quad \text{so} \quad \tau = t$$
$$\frac{dx}{d\tau} = c, \quad x_{|\tau=0} = s \to x = c\tau + s, \quad \text{so} \quad s = x - c\tau = x - ct$$
$$\frac{d\rho}{d\tau} = 0, \ (\text{so } \rho = \rho(s) \text{ only}) \quad , \rho_{|\tau=0} = \rho_0(s) \ .$$

Thus, $\rho(x,t) = \rho_0(x - ct)$. For example, if $\rho_0(x) = e^{-x^2}$, then $\rho(x,t) = e^{-(x-ct)^2}$. That is, the solution at time $t$ just translates the "bump" $e^{-x^2}$ with maximum at $x = 0$ at time 0 to a bump with maximum at $x = ct$ at time $t$.

*Example 6*: $u_x + u_y = 1$, with $u(x,0) = e^x$.
Again, $\Gamma = x$-axis, i.e. $\Gamma = \{(x,y) = (s,0) : s \in \mathbb{R}\}$. Then

$$\frac{dx}{d\tau} = 1 \ , \quad x_{|\tau=0} = s \quad \to x = \tau + s$$

$$\frac{dy}{d\tau} = 1 \ , \quad y_{|\tau=0} = 0 \qquad \to y = \tau$$
$$\text{so} \quad \tau = y, s = x - y \ .$$

Now
$$\frac{du}{d\tau} = 1 \ , \quad u_{|\tau=0} = e^s \quad u = \tau + e^s \ ,$$
so finally, $u(x, y) = y + e^{x-y}$. (Note: $u_{|\tau=0} = e^x|_{x=x(s,0)} = e^s$.)

*Example 7*: $xu_x - yu_y = -u + x$, $u(x, x) = x^2$, $x > 0$.
Now $\Gamma = \{(x, y) = (s, s) : s > 0\}$, so $u$ on $\Gamma$ is just $s^2$. The characteristic equations for $x$ and $y$ are

$$\frac{dx}{d\tau} = x; \ , \quad x_{|\tau=0} = s$$

$$\frac{dy}{d\tau} = -y; \ , \quad y_{|\tau=0} = s$$

Therefore, $x = se^\tau$, and $y = se^{-\tau}$. Hence, $e^\tau = x/s = s/y$, so $s^2 = xy$, or $s = \sqrt{xy}$. Thus, $e^\tau = \sqrt{xy}/y = \sqrt{x/y}$. Now

$$\frac{du}{d\tau} + u = x = se^\tau \ , \ \text{so} \quad \frac{d}{d\tau}(e^\tau u) = se^{2\tau} \ ,$$

and integrating gives $e^\tau u = \frac{s}{2}e^{2\tau} + u_0(s)$,

so $u_{|\tau=0} = s^2 = \frac{s}{2} + u_0(s)$, or $u(\tau, s) = \frac{s}{2}e^\tau + (s^2 - s/2)e^{-\tau}$. Finally,

$$u(x, y) = \frac{1}{2}\sqrt{xy}\sqrt{\frac{x}{y}} + (xy - \frac{\sqrt{xy}}{2})\sqrt{\frac{y}{x}} = \frac{x}{2} + x^{1/2}y^{3/2} - \frac{y}{2} \ .$$

*Example 8*: $(x + 2)u_x + 2yu_y = 2u$, $u(-1, y) = \sqrt{y}$, $y > 0$.

Write $\Gamma = \{(x, y) = (-1, s) : s > 0\}$.

$$\frac{dx}{d\tau} = x + 2 \ , \quad x_{|\tau=0} = -1 \to x = -2 + e^\tau \ ,$$

hence $\tau = \ln(x+2)$.

$$\frac{dy}{d\tau} = 2y \ , \quad y_{|\tau=0} = s \to y = se^{2\tau} \ .$$

Thus, $s = \frac{y}{e^{2\tau}} = \frac{y}{(x+2)^2}$. Also

$$\frac{du}{d\tau} = 2u \ , \quad u_{|\tau=0} = \sqrt{s}, \ \text{ so } \ u(\tau, s) = \sqrt{s}e^{2\tau} \ .$$

This gives $u(x, y) = \sqrt{y}(x+2)$.

*Example 9*: $xu_x + 2xu_y + u = 0$, $u(x, 0) = x$.

Write $\Gamma = \{(x, y) = (s, 0) : s \in \mathbb{R}\}$.

$$\frac{dx}{d\tau} = x \ , \quad x_{|\tau=0} = s \to x = se^{\tau} \ .$$

$$\frac{dy}{d\tau} = 2se^{\tau} \ , \quad y_{|\tau=0} = 0 \ .$$

Note that the characteristic equation, like all characteristic equations here, must be expressed in the local coordinates $\tau, s$. Now $y = 2se^{\tau} + C = 2s(e^{\tau} - 1)$. Also

$$\frac{du}{d\tau} = -u \ , \quad u_{|\tau=0} = s \to u = se^{-\tau} \ .$$

Thus, $y = 2s(x/s - 1) = 2x - 2s$, so $s = x - y/2$, which gives $u(x, y) = (x - y/2)^2/x$.

It is easy to check that each of these expressions for $u$ in the last five examples satisfy their respective pde and initial condition.

*Remark on local existence*: In our procedure we obtain $x = f(\tau, s), y = g(\tau, s)$, and then we invert these to obtain $\tau = \hat{f}(x, y), s = \hat{g}(x, y)$; then we write $u(x, y) = h(\hat{f}(x, y), \hat{g}(x, y))$. But, in order to do that inversion we need an *inverse-function theorem* from calculus, which states that a sufficient condition for solvability of $x = f, y = g$ in a neighborhood of $\Gamma$, since

$f, g$ are differentiable, to obtain $\tau = \hat{f}, s = \hat{g}$, is for the Jacobian of the transformation, namely

$$\frac{\partial(x,y)}{\partial(\tau,s)} = \det \begin{bmatrix} \frac{\partial f}{\partial \tau} & \frac{\partial f}{\partial s} \\ \\ \frac{\partial g}{\partial \tau} & \frac{\partial g}{\partial s} \end{bmatrix} = \frac{\partial f}{\partial \tau}\frac{\partial g}{\partial s} - \frac{\partial f}{\partial s}\frac{\partial g}{\partial \tau} \neq 0 \text{ along } \Gamma \ .$$

Then the Jacobian is nonvanishing in a neighborhood of $\Gamma$, and we have at least a *local* solution to our first order Cauchy problem. Note that in examples 6,7,8 above, the Jacobian (evaluated at $\tau = 0$) is $-1, 2s > 0$, and 1, respectively.

*Example 10*: $u_x + 2u_y = u^2$, with $u(x,0) = F(x)$.
Notice that in all the previous examples the equations were linear, but here the equation is semilinear. $\Gamma = \{(x,y) = (s,0) : s \in \mathbb{R}\}$. We assume here that $F$ is continuously differentiable. Now

$$\frac{dy}{d\tau} = 2 \ , y_{|\tau=0} = 0 \rightarrow y = 2\tau \quad \text{so } \tau = y/2$$

$$\frac{dx}{d\tau} = 1 \ , \ x_{|\tau=0} = s \rightarrow x = \tau + s \quad \text{so } s = x - y/2$$

and

$$\frac{du}{d\tau} = u^2 \ , \ \ u_{|\tau=0} = F(s) \ .$$

solving the $u$ equation gives

$$u = \frac{F(s)}{1 - \tau F(s)} \quad \text{or} \quad u(x,y) = \frac{F(x - y/2)}{1 - \frac{y}{2}F(x - y/2)} \ .$$

This expression is certainly defined for small enough values of $y$ (assuming $F$ is bounded). However, $u$ will blow up if $y$ becomes large enough for the denominator to vanish. So we only can expect a local solution to this semilinear problem near the $x$-axis. But this example shows our procedure works for semilinear problems.

**Summary:** From this section, know the classification of first order PDEs, and know how to solve Cauchy problems for such equations.

*Exercises*

1. $u_x + u_y = 1$, $u(x, 2x) = f(x)$, for any continuously differentiable function $f(x)$ on $\mathbb{R}$. (Notationally we would write for this statement $f \in C^1(\mathbb{R})$.)

2. $u_x - u_y + u = 1$, with $u(x, 0) = \sin(x)$.

3. $xu_x + yu_y = 3$, $u(1, y) = \ln(y)$, $y > 0$.

4. $x^2u_x - y^2u_y = 0$, $u(1, y) = f(y)$, $f \in C^1(\mathbb{R})$.

5. $yu_x + xu_y = 0$, $u(1, y) = y^2$.

6. $xu_y - yu_x = 0$, $u(x, 0) = x^2$. (Ans: $u = x^2 + y^2$.)

7. Suppose $N$ computers are infected with a virus and at $t = 0$. An antidote is sent that will cure the problem as soon as people open their inbox. Assume that if $n$ users are still infected at time $t$, then in the short time interval $(t, t + \delta t)$ one and only one user will log on. This probability is $\mu n p_n(t)$, where $p_n(t)$ = probability that there are still $n$ infected computers at time $t$; here $\mu > 0$ is a constant. By a conservation principle one can show that $\frac{dp_n}{dt}(t) = \mu(n+1)p_{n+1}(t) - \mu n p_n(t)$.[4] As in example 3 of this section, let $G(x, t) = \sum_{n=0}^{\infty} p_n(t)x^n$ be the generating function for $p_n(t)$. Show that $G$ satisfies the equation

$$\frac{\partial G}{\partial t} + \mu(x-1)\frac{\partial G}{\partial x} = 0 \ .$$

With $N$ victims initially, we have $p_N(0) = 1$, $p_j(0) = 0$ for $j \neq N$. Hence, $G(x, 0) = \sum_{n=0}^{\infty} p_n(0)x^2 = x^N$. Solve the Cauchy problem and show that

$$G(x, t) = (1 + (x-1)e^{-\mu t})^N \ .$$

Thus, show that the mean of the distribution, $\sum_{n=0}^{\infty} np_n(t)$, is given by $\partial G/\partial x(1, t) = Ne^{-\mu t}$.

8. Consider $u_x + yu_y = 0$, $u(x, 0) = f(x)$:

   (a) If $f(x) = x$, show no solution exists.

   (b) If $f(x) \equiv 1$, show there are many solutions.

[4] Another physical interpretation that gives you the same problem is to suppose $p_n(t)$ is the probability that after time $t$ spent proofreading a difficult text on differential equations, the draft still contains $n$ errors.