

LECTURE 2

First Order Linear PDEs

Let's look first at the simplest case of a first order linear PDE, as even the simplest cases can tell us something fundamental. Consider

$$\phi_x = 0 \quad \left(i.e. \quad \frac{\partial \phi}{\partial x} = 0 \right)$$

This says that the function ϕ does not change with the x coordinate is varied. The general solution is thus¹

$$\phi(x, y) = f(y) \quad , \quad f \text{ an arbitrary function of the "other" variable } y$$

Thus, in this simplest example we see already that the general solution to a PDE may involve arbitrary functions. We contrast this situation with the analogous ordinary differential equation (2) where general solution involved only a single arbitrary constant. In the present case, it will **not** suffice to specify that value of the solution at a point (x_0, y_0) to get a unique solution; because such a condition will fix only the value of $f(y)$ when $y = y_0$; it will not fix the values of f at other points y , and it so will not determine the function $f(y)$

As a second simple example consider

$$(1) \quad au_x + bu_y = 0$$

where a, b are constants not both zero.

0.1. Geometric construction of solution. Let \mathbf{v} be a vector in \mathbb{R}^n . Recall that the directional derivative $D_{\mathbf{v}}f(\mathbf{x})$ of a differentiable function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ at a point $\mathbf{x} \in \mathbb{R}^n$ is defined as

$$D_{\mathbf{v}}f(\mathbf{x}) = \lim_{\varepsilon \rightarrow 0} \frac{f(\mathbf{x} + \varepsilon \mathbf{v}) - f(\mathbf{x})}{\varepsilon}$$

and is computable as

$$D_{\mathbf{v}}f(\mathbf{x}) = \mathbf{v} \cdot \nabla f(\mathbf{x}) := v_1 \frac{\partial f}{\partial x_1}(\mathbf{x}) + v_2 \frac{\partial f}{\partial x_2}(\mathbf{x}) + \cdots + v_n \frac{\partial f}{\partial x_n}(\mathbf{x})$$

(i.e. $D_{\mathbf{v}}f(\mathbf{x})$ is the dot product of the gradient of f at \mathbf{x} with the vector \mathbf{v} .)

The PDE

$$au_x + bu_y = 0$$

thus says that the function $u(x, y)$ must be constant in the direction of $\mathbf{v} = (a, b)$. Put another way, any solution of (1) must be constant along any line parallel to the direction of (a, b) . Such lines are of the form

$$\ell = \{(x, y) \in \mathbb{R}^2 \mid (x, y) = (c_1, c_2) + t(a, b)\}$$

from which we may infer

$$\begin{aligned} x &= c_1 + ta \\ y &= c_2 + tb \end{aligned}$$

¹Actually, it is not clear from the way the PDE is written that there are only two underlying variables. For the expression $\phi_x = 0$ only indicates that x is an underlying variable of the function ϕ . In practice, the number of underlying variables will be clear from the context in which a PDE arises.

multiplying the first equation by $-b$ and the second equation by a and then adding we obtain

$$-bx + ay = -bc_1 + ac_2 = \text{some constant.}$$

Thus the lines parallel to (a, b) satisfy linear equations of the form

$$(2) \quad -bx + ay = \text{constant}$$

Since along such lines a solution of (1) is constant, it follows that the value of a solution $u(x, y)$ is determined by which of these lines parallel to \mathbf{v} , the point (x, y) sits on; which is in turn determined by the value of the constant on the right hand side of (2). Thus, we can infer that

$$u(x, y) = f(-bx + ay)$$

for some arbitrary function f .

0.2. Digression: Chain Rule for Functions of Several Variables. In order to make a change of variables or to look at the behaviour of the solution of a PDE along a particular curve, it is important to understand how to implement the Chain Rule for functions of several variables. Let me take a minute to explain this formula in a fairly general setting.

Suppose you have a function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ of n variables (x_1, \dots, x_n) and a map

$$f : \mathbb{R}^m \rightarrow \mathbb{R}^n \quad : \quad (y_1, \dots, y_m) \mapsto (x_1(y_1, \dots, y_m), x_2(y_1, \dots, y_m), \dots, x_n(y_1, \dots, y_m))$$

mapping the points of an m -dimensional space to the points in \mathbb{R}^n . We can then form the composed function

$$\Phi : \mathbb{R}^m \rightarrow \mathbb{R} \quad : \quad (y_1, \dots, y_m) \rightarrow \phi((x_1(y_1, \dots, y_m), x_2(y_1, \dots, y_m), \dots, x_n(y_1, \dots, y_m)))$$

The Chain Rule tells us how to compute the partial derivatives $\frac{\partial \Phi}{\partial y_i}$

$$\frac{\partial \Phi}{\partial y_i} = \sum_{j=1}^n \frac{\partial x_j}{\partial y_i} \frac{\partial \phi}{\partial x_j}$$

0.2.1. Application 1: Change of Variables. Suppose we have a function $\phi(x, y)$ and we make a change of variables; (for example, we might want to change from rectangular coordinates to polar coordinates). The Chain Rule tells us how to relate derivatives with respect to the old coordinates to derivatives with respect to the new coordinates. Suppose

$$\begin{aligned} x &= r \cos(\theta) & r &= \sqrt{x^2 + y^2} \\ y &= r \sin(\theta) & \theta &= \tan^{-1}\left(\frac{y}{x}\right) \end{aligned}$$

and we define

$$\Phi(r, \theta) = \phi(r \cos(\theta), r \sin(\theta))$$

Then the Chain rule says

$$\begin{aligned} \frac{\partial \Phi}{\partial r} &= \frac{\partial x}{\partial r} \frac{\partial \phi}{\partial x} + \frac{\partial y}{\partial r} \frac{\partial \phi}{\partial y} = \cos \theta \frac{\partial \phi}{\partial x} + \sin(\theta) \frac{\partial \phi}{\partial y} \\ \frac{\partial \Phi}{\partial \theta} &= \frac{\partial x}{\partial \theta} \frac{\partial \phi}{\partial x} + \frac{\partial y}{\partial \theta} \frac{\partial \phi}{\partial y} = -r \sin(\theta) \frac{\partial \phi}{\partial x} + r \cos \theta \frac{\partial \phi}{\partial y} \end{aligned}$$

You'll see another example of such a change of variables computation when we discuss the coordinate method for solving $a\phi_x + b\phi_y = 0$, just below.

0.2.2. *Application 2: Converting Differential Operators.* One reason for making a change of variables is that a differential operator may have a simpler expression in terms of a new set of variables. Consider the differential operator

$$a \frac{\partial}{\partial x} + b \frac{\partial}{\partial y}$$

If we introduce new variables

$$\begin{aligned} s &= ax + by \\ t &= -bx + ay \end{aligned}$$

then according to the Chain Rule (at the level of differential operators)

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial s}{\partial x} \frac{\partial}{\partial s} + \frac{\partial t}{\partial x} \frac{\partial}{\partial t} = a \frac{\partial}{\partial s} - b \frac{\partial}{\partial t} \\ \frac{\partial}{\partial y} &= \frac{\partial s}{\partial y} \frac{\partial}{\partial s} + \frac{\partial t}{\partial y} \frac{\partial}{\partial t} = b \frac{\partial}{\partial s} + a \frac{\partial}{\partial t} \end{aligned}$$

and so the differential operator with respect to s, t that is *equivalent* to $a \frac{\partial}{\partial x} + b \frac{\partial}{\partial y}$ will be

$$a \frac{\partial}{\partial x} + b \frac{\partial}{\partial y} = a \left(a \frac{\partial}{\partial s} - b \frac{\partial}{\partial t} \right) + b \left(b \frac{\partial}{\partial s} + a \frac{\partial}{\partial t} \right) = (a^2 + b^2) \frac{\partial}{\partial s} + (ab - ba) \frac{\partial}{\partial t} = (a^2 + b^2) \frac{\partial}{\partial s}$$

This means that if $\phi(x, y)$ is a solution of

$$\left(a \frac{\partial}{\partial x} + b \frac{\partial}{\partial y} \right) \phi(x, y) = 0$$

Then if we set

$$\tilde{\phi}(s, t) = \phi(x(s, t), y(s, t))$$

then $\tilde{\phi}(s, t)$ must satisfy

$$(a^2 + b^2) \frac{\partial \tilde{\phi}}{\partial s} = 0$$

And so the PDE becomes much simple in terms of the new variables s and t .

0.2.3. *Application 2: Derivative of a function along a curve.* Suppose we want to know the rate of change of a function $\phi(x, y)$ changes as we move along a particular curve $y = f(x)$. This could be computed by setting

$$\Phi(x) = \phi(x, f(x))$$

and then computing $\frac{d}{dx} \Phi(x)$. Here we regard $\Phi(x)$ as the function obtained by composing $g : \mathbb{R} \rightarrow \mathbb{R}^2 : t \rightarrow (t, f(t))$ with $\phi(x, y)$. The Chain Rule then tells us that

$$\frac{d\Phi}{dt} = \frac{dx}{dt} \frac{\partial \phi}{\partial x} + \frac{dy}{dt} \frac{\partial \phi}{\partial y} = \frac{dt}{dt} \frac{\partial \phi}{\partial x} + \frac{df}{dt} \frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial x} + f' \frac{\partial \phi}{\partial y}$$

We'll see an explicit application of this formula in the next lecture.

0.3. Digression: Two By Two Linear Systems. It will often happen in two variable situations that we'll need to solve a pair of linear equations for two unknowns. Here I'll just state a general formula for the solution to such a *two by two linear system*:

$$\left. \begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned} \right\} \Rightarrow \left\{ \begin{aligned} x &= \frac{1}{ad-bc} (de - bf) \\ y &= \frac{1}{ad-bc} (af - ce) \end{aligned} \right.$$

0.4. Change of Coordinates Method. Consider the vector $\mathbf{v}^\perp := (b, -a)$. We have

$$\mathbf{v} \cdot \mathbf{v}^\perp = ab - ba = 0$$

so \mathbf{v}^\perp is perpendicular to \mathbf{v} . The pair $\mathbf{v}, \mathbf{v}^\perp$ thus constitute a pair of orthogonal directions in the plane. Now set

$$\begin{aligned} s &= ax + by \\ t &= -bx + ay \end{aligned}$$

with inverse relations

$$\begin{aligned} x &= \frac{1}{a^2 + b^2} (bs + at) \\ y &= \frac{1}{a^2 + b^2} (as - bt) \end{aligned}$$

regarding s, t as a new set coordinates. Using the chain rule for functions of two variables we can express derivatives with respect to x, y in terms of derivatives with respect to the new variables s and t . Thus, if

$$\tilde{\phi}(s, t) \equiv \phi(x(s, t), y(s, t))$$

we will have

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{\partial s}{\partial x} \frac{\partial \tilde{\phi}}{\partial s} + \frac{\partial t}{\partial x} \frac{\partial \tilde{\phi}}{\partial t} = a \frac{\partial \tilde{\phi}}{\partial s} - b \frac{\partial \tilde{\phi}}{\partial t} \\ \frac{\partial \phi}{\partial y} &= \frac{\partial s}{\partial y} \frac{\partial \tilde{\phi}}{\partial s} + \frac{\partial t}{\partial y} \frac{\partial \tilde{\phi}}{\partial t} = b \frac{\partial \tilde{\phi}}{\partial s} + a \frac{\partial \tilde{\phi}}{\partial t} \end{aligned}$$

So if

$$0 = a\phi_x + b\phi_y$$

we must have

$$\begin{aligned} 0 &= a \left(a \frac{\partial \tilde{\phi}}{\partial s} - b \frac{\partial \tilde{\phi}}{\partial t} \right) + b \left(b \frac{\partial \tilde{\phi}}{\partial s} + a \frac{\partial \tilde{\phi}}{\partial t} \right) \\ &= a^2 \frac{\partial \tilde{\phi}}{\partial s} - ab \frac{\partial \tilde{\phi}}{\partial t} + b^2 \frac{\partial \tilde{\phi}}{\partial s} + ba \frac{\partial \tilde{\phi}}{\partial t} \\ &= (a^2 + b^2) \frac{\partial \tilde{\phi}}{\partial s} \end{aligned}$$

Dividing the extreme sides of this chain of equalities by $a^2 + b^2$ we get

$$0 = \frac{\partial \tilde{\phi}}{\partial s}$$

So the function $\tilde{\phi}(s, t)$ corresponding to a solution of $a\phi_x + b\phi_y = 0$ by a simple change of variables cannot depend on s : Thus

$$\tilde{\phi}(s, t) = f(t)$$

But then

$$\phi(x, y) = \tilde{\phi}(s(x, y), t(x, y)) = f(t(x, y)) = f(-bx + ay)$$

which is the same answer we got before.

EXAMPLE 2.1. Let's consider again the PDE

$$(3) \quad a\phi_x + b\phi_y = 0$$

but this time coupled with a boundary condition

$$(4) \quad \phi(0, y) = f(y)$$

We'll solve this PDE with boundary condition yet another way.

Consider the curves (actually a line) $\gamma_c : t \mapsto [x(t), y(t)] \in \mathbb{R}^2$ where

$$\begin{aligned}x(t) &= at \\ y(t) &= bt + c\end{aligned}$$

the curve γ_c is just a straight line with tangent vector

$$\frac{d\gamma_c}{dt} = \left[\frac{dx}{dt}, \frac{dy}{dt} \right] = [a, b]$$

Now suppose $\phi(x, y)$ is a solution of $a\phi_x + b\phi_y = 0$. How does the solution ϕ vary as we move along the curve γ_c ? Well, by the Chain Rule,

$$\frac{d}{dt}(\phi(\gamma_c(t))) = \frac{\partial \phi}{\partial x} \frac{dx}{dt} + \frac{\partial \phi}{\partial y} \frac{dy}{dt} = a\phi_x + b\phi_y = 0$$

In other words, a solution ϕ must remain constant as one moves along a curve $\gamma_c(t)$.

But what is the actual value of the solution at an arbitrary point (x, y) ? Well, the point (x, y) sits on one of the curves $\gamma_c(t)$ for some choice of c and t . Indeed, if we set

$$\begin{aligned}x &= at \\ y &= bt + c\end{aligned}$$

we can solve this pair of equations for t and c . One gets

$$\begin{aligned}t &= \frac{x}{a} \\ c &= y - \frac{b}{a}x\end{aligned}$$

Thus, the point (x, y) lies on the curve γ_c where $c = y - \frac{b}{a}x$. But the point $(0, c) = (0, y - \frac{b}{a}x)$ also lies on this same curve. So the value of the solution at (x, y) must be the same as its value at $(0, y - \frac{b}{a}x)$. Thus,

$$\phi(x, y) = \phi\left(0, y - \frac{b}{a}x\right) = f\left(y - \frac{b}{a}x\right)$$

Noting that there was nothing special about the way we chose the point (x, y) we can infer the value of the solution ϕ to (3) - (4) any point (x, y) is given by

$$\phi(x, y) = f\left(y - \frac{b}{a}x\right)$$

EXAMPLE 2.2. Find the solution of

$$(5) \quad \phi_x = x$$

$$(6) \quad \phi(0, y) = y^2$$

Here we've introduced two complications to our initial example: we've introduced an inhomogeneous term to the right hand side of the PDE and we've imposed a certain boundary condition along the y -axis.

Now PDE is again just telling us how a solution must change along the x -direction; it's just that now it's changing non-trivially in the x -direction. In fact, suppose we set Φ_{y_0} to be the restriction of the solution to the line $y = y_0$. This we can regard as a function of x alone

$$\Phi_{y_0}(x) = \phi(x, y_0)$$

And we have, by the chain rule

$$\frac{d}{dx}\Phi_{y_0}(x) = \frac{\partial \phi}{\partial x} \frac{dx}{dx} + \frac{\partial \phi}{\partial y} \frac{dy_0}{dx} = \frac{\partial \phi}{\partial x}$$

(because $\frac{dx}{dx} = 1$ and $\frac{dy_o}{dx} = 0$). The original PDE says that the right-hand side must be x . Hence, $\Phi_{y_o}(x)$ must satisfy

$$\frac{d\Phi_{y_o}}{dx} = x \implies \Phi_{y_o}(x) = \frac{1}{2}x^2 + C_{y_o}$$

Thus, we know

$$\phi(x, y_o) = \Phi_{y_o}(x) = \frac{1}{2}x^2 + C_{y_o}$$

i.e., we know the solution at all points (x, y_o) up to a term that depends only on y_o . We may as well write this as

$$\phi(x, y) = \frac{1}{2}x^2 + C(y)$$

To figure out $C(y)$ we just need to impose the boundary condition.

$$y^2 = \phi(0, y) = \frac{1}{2}(0)^2 + C(y) \implies C(y) = y^2$$

And so we have

$$\phi(x, y) = \frac{1}{2}x^2 + y^2$$

REMARK 2.3. The moral of the preceding example is that we can sometimes solve a first order linear PDE by interpreting it as telling us how the solution must change along a particular direction. In the next couple of examples, we'll generalize this idea and solve first order linear PDEs by interpreting them as telling us how solutions must vary along a particular curves.

EXAMPLE 2.4. Consider the following first order linear PDE

$$(7) \quad \phi_x + 2x\phi_y = y$$

subject to the boundary condition

$$(8) \quad \phi(0, y) = 1 + y^2 \quad , \quad \text{for } 1 < y < 2 \quad .$$

Suppose that $\phi(x, y)$ is a solution of this PDE/BVP. If we consider the values $\phi(x, y)$ along the curves of the form

$$y = f(x)$$

we have, via the Chain Rule,

$$\frac{d}{dx}\phi(x, f(x)) = \phi_x + f'(x)\phi_y \quad .$$

Thus, if we evaluate (7) along the particular curve

$$(9) \quad y = x^2 + C_1$$

we get

$$(10) \quad \frac{d}{dx}\phi(x, x^2 + C_1) = \phi_x + 2x\phi_y.$$

or, noting that the left hand side of (10) coincides with the right hand side of our original PDE, we get

$$(11) \quad \frac{d}{dx}\phi(x, x^2 + C_1) = x^2 + C_1$$

Let's set

$$\psi(x) := \phi(x, x^2 + C_1)$$

then (11) amounts to a ODE for $\psi(x)$; that is easily solved by integrating both sides:

$$\frac{d\psi}{dx} = x^2 + C_1 \implies \psi(x) = \frac{1}{3}x^3 + C_1x + C_2$$

We have thus shown that along the parabolas $y = C_1 + x^2$, the solutions of (7) must be have like

$$(12) \quad \phi(x, x^2 + C_1) = \frac{1}{3}x^3 + C_1x + C_2 \quad .$$

The constants C_1 and C_2 can be interpreted as follows. The curve (??) is a parabola that intersects the y -axis at the point C_1 . From (??) it is clear that C_2 corresponds to the value of ϕ at the point $(0, C_1)$. Note that this value is determined by the boundary condition (??).

In fact, let's determine C_2 straightaway. Taking the limit as $x \rightarrow 0$ on both sides of (12) we get

$$\phi(0, C_1) = C_2$$

Applying the boundary condition (8) we then have

$$1 + C_1^2 = \phi(0, C_1) = C_2 \quad \Rightarrow \quad C_2 = 1 + C_1^2$$

We can thus, by imposing (8), fix the constant of integration C_2 and get

$$(13) \quad \phi(x, x^2 + C_1) = \frac{1}{3}x^3 + C_1x + 1 + C_1^2$$

What remains is to organize what we have figured out into a formula for computing the value of ϕ at an arbitrary point (x, y) .

So consider a point (x, y) in the plane. We know from (13) how the solution $\phi(x, y)$ must vary along the parabola $y = x^2 + C_1$. We also know that if (x, y) lies on this parabola, then necessarily

$$C_1 = y - x^2$$

We thus have

$$\begin{aligned} \phi(x, y) &= \frac{1}{3}x^3 + C_1x + C_2 \\ &= \frac{1}{3}x^3 + (y - x^2)x + 1 + (y - x^2)^2 \\ &= \frac{1}{3}x^3 + (y - x^2)(x + y - x^2) + 1 \end{aligned}$$

The solution to (7) and (8) is thus

$$\phi(x, y) = \frac{1}{3}x^3 + (y - x^2)(x + y - x^2) + 1$$

EXAMPLE 2.5. Consider now the following PDE/BVP:

$$(14) \quad x\phi_x + y\phi_y = 1 + y$$

$$(15) \quad \phi(x, 1) = x + 1$$

In this example, we could divide through by x , to get

$$\phi_x + \frac{y}{x}\phi_y = \frac{1}{x} + \frac{y}{x}$$

and then try to construct solutions along curves $y = f(x)$ with

$$f'(x) = \frac{y}{x} \quad .$$

However, such a formulation would introduce singularities at $x = 0$ which could be avoided.

So instead, consider a curve in the (x, y) -plane defined by some function $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2; t \mapsto (x(t), y(t))$. If

$$(16) \quad \frac{dx}{dt} = x \quad , \quad \frac{dy}{dt} = y \quad ,$$

then

$$(17) \quad \frac{d}{dt}\phi(\gamma(t)) = \frac{\partial\phi}{\partial x}\frac{dx}{dt} + \frac{\partial\phi}{\partial y}\frac{dy}{dt} = x\phi_x + y\phi_y \quad .$$

Thus, along the curve γ , the PDE in (14) becomes the ODE

$$\frac{d}{dt}\phi(\gamma(t)) = 1 + y(t).$$

Solving the differential equations (16) for $x(t)$ and $y(t)$, we can make this equation for $\phi(\gamma(t))$ even more explicit:

$$\begin{aligned}\frac{dx}{dt} &= x &\Rightarrow & x = C_1 e^t \\ \frac{dy}{dt} &= y &\Rightarrow & y = C_2 e^t\end{aligned}$$

so

$$(18) \quad \frac{d}{dt}(\phi \circ \gamma) = 1 + C_2 e^t \quad .$$

Integrating (18) produces

$$\phi \circ \gamma(t) = t + C_2 e^t + C_3$$

and so

$$(19) \quad \phi(x, y) = \phi(C_1 e^t, C_2 e^t) = t + C_2 e^{2t} + C_3 \quad .$$

Now the solutions to (16) actually correspond to large family of curves. To choose a specific one, we need to fix some initial condition. Thus, without loss of generality, we can assume that the curve γ crosses the line $y = 1$ when $t = 0$ and that when it does, it crosses the line $y = 1$ at the point x_o . In other words, we stipulate

$$[x_o, 1] = \gamma(0) = [C_1 e^0, C_2 e^0] = [C_1, C_2]$$

or,

$$(20) \quad C_1 = x_o \quad , \quad C_2 = 1 \quad .$$

Now consider an arbitrary point $P = (x, y)$ in the first quadrant. Suppose γ passes through P , then

$$(21) \quad \begin{aligned}x &= x_o e^t \\ y &= e^t\end{aligned}$$

for some t . Solving (21) for x_o and t we get

$$(22) \quad \begin{aligned}t &= \ln |y| \\ x_o &= \frac{x}{y} \quad .\end{aligned}$$

We have thus figured out for an arbitrary point $(x, y) \in \mathbb{R}^2$, which of the curves γ it lies on (the various curves are distinguished by the initial condition $\gamma(t) = (x_o, 1)$) and how far along the curve (i.e. the value of t) it lies.

Let's now impose the boundary condition (15). Evaluating the right hand side of (19) at $t = 0$ yields

$$0 + 1 + C_3 \quad ,$$

On the other hand, the left hand side of (19) evaluates to

$$\phi(C_1 e^0, C_2 e^0) = \phi(x_o, 1) = x_o + 1$$

by virtue of the boundary condition (15). And so we have

$$1 + C_3 = x_o + 1$$

or

$$(23) \quad C_3 = x_o \quad .$$

We've now figured out the appropriate choices for all three constants of integration, enabling us to rewrite (19) as

$$\begin{aligned}\phi(x, y) &= t + C_2 e^t + C_3 \\ &= \ln |y| + (e^{\ln |y|}) + x_o \\ &= \ln |y| + y + \frac{x}{y} \quad .\end{aligned}$$

We conclude that the solution of (14)–(15) is

$$\phi(x,y) = \ln[y] + y + \frac{x}{y}$$