# 3 Solving first order linear PDE

## 3.1 Algorithm to solve first order linear PDE

In this lecture I consider a general linear first order PDE of the form

$$a(x,y)u_x + b(x,y)u_y = c(x,y)u + d(x,y), \tag{3.1}$$

where $a, b, c, d \in \mathcal{C}^{(1)}(\mathbf{R}^2;\mathbf{R})$ are given functions. I also consider an initial condition

$$u(x,y)|_{(x,y)\in\Gamma} = g(x,y), \tag{3.2}$$

which says that the initial condition is prescribed along some arbitrary (well, not totally, see below) curve $\Gamma$ on the plane $(x,y)$. The main difference with the previous section is that I stop treating one variable as "time" and allow to have rather general initial conditions contrary to the fixed value $t = 0$. I also use the variables $x$ and $y$ for the independent variables to emphasize that while it is important to keep in mind the physical description of the problem, mathematically for us both variables $t$ (time) and $x$ (space) are equally important and sometimes better make them indistinguishable. Even more importantly, a lot of first order PDE appear naturally in geometric rather than physical problems, and for this setting $x$ and $y$ are our familiar Cartesian coordinates.

**Remark 3.1.** All I am going to present is almost equally valid for a *semi-linear* first order equation

$$a(x,y)u_x + b(x,y)u_y = f(x,y,u), \tag{3.3}$$

where $f$ is some, generally nonlinear, function.

Let me consider the following system of ordinary differential equations

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}\tau} &= a(x,y), \\ \frac{\mathrm{d}y}{\mathrm{d}\tau} &= b(x,y). \end{aligned} \tag{3.4}$$

Its solutions $(x(\tau), y(\tau))$ form a family of curves on the plane, parameterized by the variable $\tau$. These curves (which is a basic fact from ODE theory) do not intersect and called (hopefully, not surprisingly) the *characteristics* or *characteristic curves*. The key fact is that

along the characteristics problem (3.1) (or (3.3)) becomes an ordinary differential equation.

Indeed, consider the solution $u$ to (3.1) along the parametric curve $(x(\tau), y(\tau))$. It becomes just the function of $\tau$ alone: $v(\tau) = u(x(\tau), y(\tau))$ (here I picked a different letter to emphasize that $v$ depends only on $\tau$). Now take the derivative with respect to $\tau$:

$$\frac{\mathrm{d}v}{\mathrm{d}\tau} = u_x x'_\tau + u_y y'_\tau = a(\tau)u_x + b(\tau)u_y = c(\tau)v + d(\tau),$$

which is a linear first order ODE. To get the initial condition for this ODE I will use (3.2).

Math 483/683: Partial Differential Equations by Artem Novozhilov
e-mail: artem.novozhilov@ndsu.edu. Spring 2023

In general (several examples are given below), to solve the initial value problem (3.1)-(3.2) I proceed in the following way. I consider a parametrization of the initial curve $\Gamma$:

$$\Gamma\colon x(\xi),\, y(\xi),$$

along which my initial condition becomes just a function of $\xi$: $g(\xi)$.

Now, for each fixed $\xi$ I solve problem (3.4) with the initial condition $x(\xi), y(\xi)$, my unique (due to ODE theory) solution will be

$$(x(\tau,\xi), y(\tau,\xi)).$$

Along this curve, as showed above, my PDE becomes the ODE

$$\dot{v} = c(\tau)v + d(\tau)$$

with the initial condition

$$v(0,\xi) = g(\xi).$$

The unique solution to this problem is $(\tau,\xi) \mapsto v(\tau,\xi)$ and I found a parametric representation of the solution to (3.1)-(3.2):

$$(\tau,\xi) \mapsto \big(x(\tau,\xi), y(\tau,\xi), v(\tau,\xi)\big).$$

If I am able to express $\tau$ and $\xi$ from the first two functions in the vector above then I will finally get the unique solution

$$u(x,y) = v(\tau(x,y), \xi(x,y)).$$

There is still the question whether I will always be able to do it, but I will postpone the general disussion and instead consider a few examples.

**Example 3.2.** Solve

$$\begin{aligned} xu_x + yu_y &= u, \\ u(x,1) &= g(x). \end{aligned} \tag{3.5}$$

The curve of the initial conditions is given simply as $y = 1$. In this case I can always take

$$x = \xi, \quad y = 1.$$

Hence I have for my characteristics

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}\tau} &= x, \quad x(0,\xi) = \xi, \\ \frac{\mathrm{d}y}{\mathrm{d}\tau} &= y, \quad y(0,\xi) = 1, \end{aligned}$$

which immediately implies that

$$x(\tau,\xi) = \xi e^{\tau}, \quad y(\tau,\xi) = e^{\tau}.$$

From the last expression I also have that

$$\xi = \frac{x}{y}\,.$$

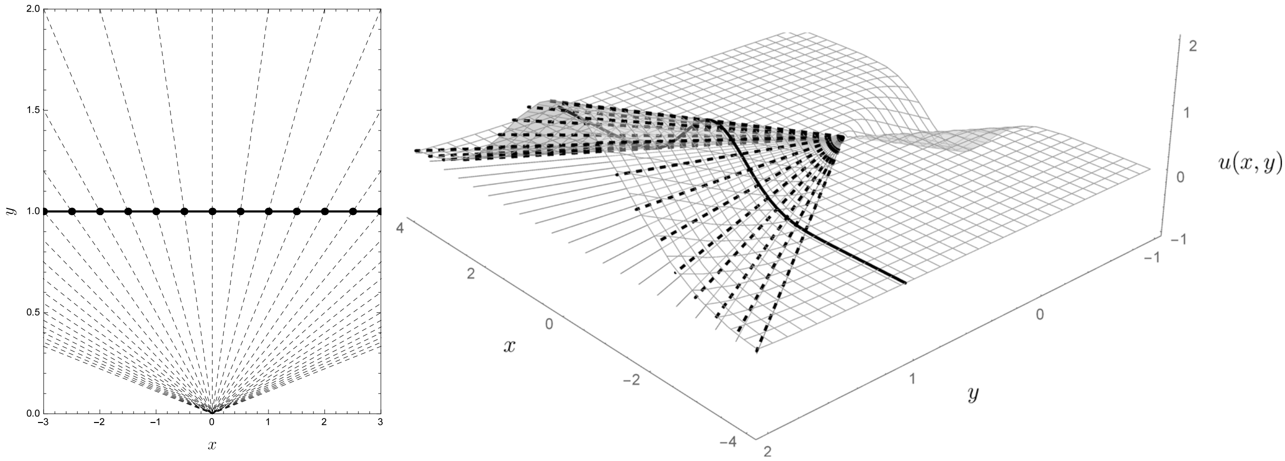



Figure 1: Left panel: The bold solid line is the curve of the initial conditions. The dotted lines are the characteristics. The points show where the initial condition for each characteristic is given. Note that the characteristics are defined only for $y > 0$. Right panel: The surface of the solution along with the initial condition (bold curve) and solutions of the corresponding ODE along the characteristics.

Along my characteristics I have (note the initial condition)

$$\frac{\mathrm{d}v}{\mathrm{d}\tau} = v, \quad v(0, \xi) = g(\xi) \implies v(\tau, \xi) = g(\xi)e^{\tau}.$$

Finally, returning to the initial variables $(x, y)$, I have

$$u(x, y) = g\left(\frac{x}{y}\right) y.$$

Note that my solution is not defined at $y = 0$.

To present graphs (see Fig. 1), I will use

$$g(x) = e^{-x^2},$$

and hence my solution is

$$u(x, y) = e^{-(x/y)^2} y,$$

which is valid only for $y > 0$.

**Example 3.3.** Solve

$$\begin{aligned} yu_x - xu_y &= 0, \\ u(x, 0) &= g(x), \quad x > 0. \end{aligned} \tag{3.6}$$

The reason why I define the initial condition only for $x > 0$ will be given below.

The system for characteristics is given by

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}\tau} &= y, \quad x(0, \xi) = \xi, \\ \frac{\mathrm{d}y}{\mathrm{d}\tau} &= -x, \quad y(0, \xi) = 0. \end{aligned}$$

Probably the easiest way to solve it is to reduce this system to one second order ODE. Denoting with prime the derivative with respect to $\tau$ I have

$$x'' = y' = -x \implies x'' + x = 0.$$

This is the equation for the harmonic oscillator, its general solution is

$$x(\tau) = C_1 \cos\tau + C_2 \sin\tau.$$

Using the initial conditions $x(0) = \xi, x'(0) = 0$ I get

$$x(\tau) = \xi\cos\tau, \quad y(\tau) = -\xi\sin\tau.$$

Squaring and adding the equations together I will find

$$x^2 + y^2 = \xi^2,$$

hence my characteristics are circles of radius $\xi$. As a side remark I note that the same result can be obtained by reducing the system to just one equation:

$$\frac{\frac{\mathrm{d}x}{\mathrm{d}\tau}}{\frac{\mathrm{d}y}{\mathrm{d}\tau}} = -\frac{x}{y}$$

and integrating this separable equation.

Along the characteristics I have

$$\frac{\mathrm{d}v}{\mathrm{d}\tau} = 0, \quad v(0,\xi) = g(\xi) \implies v(\tau,\xi) = g(\xi).$$

Returning to the original coordinates, I have

$$u(x,y) = g\left(\sqrt{x^2+y^2}\right),$$

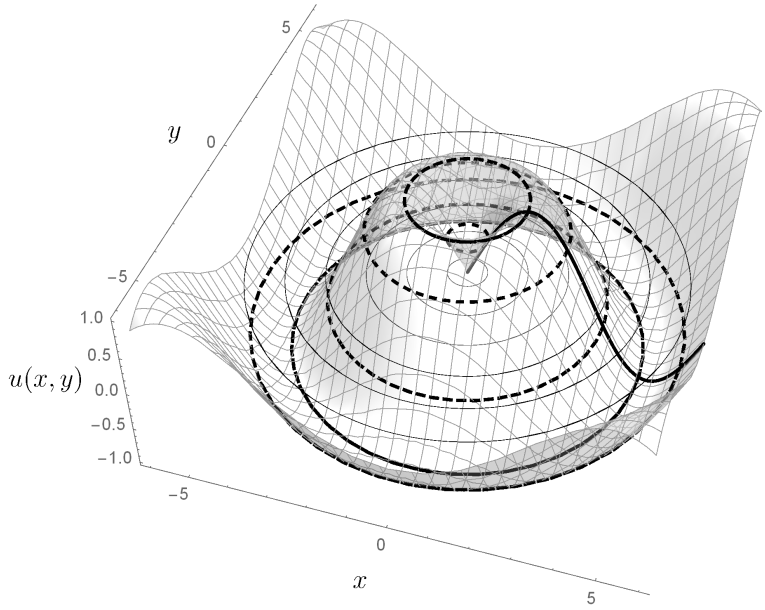
which gives me the solution to my problem. If I take $g(x) = \sin x$, then the solution is drawn in Fig. 2. Now we can see why the initial condition was prescribed only for $x > 0$. Since the characteristics are circles in this problem, if my initial condition was given for $-\infty < x < \infty$ then each characteristic would intersect it as two points. Therefore, for each ODE along this characteristic I would have *two* initial conditions, which yields a contradiction (nonexistence of solution in all but very special cases).

To conclude these examples I must decide when I actually can express my two parameters $\tau$ and $\xi$ as functions of $x, y$. It turns out (this is usually not covered in Calc III, but a curious student can look up *the inverse function theorem*) that it is always true if

> the curve of the initial conditions is not tangent to any characteristic.

Summarizing,

**Proposition 3.4.** *Problem* (3.1)-(3.2) *has a unique solution, which can in general be defined on some subset of* $\mathbf{R}^2$*, if the curve* $\Gamma$ *on which the initial conditions are given is not tangent to a characteristic, and if the characteristics do not intersect* $\Gamma$ *at more than one point.*




]

Figure 2: The surface of the solution along with the initial condition (bold curve) and solutions of the corresponding ODE (bold dashed curves) along the characteristics (thin solid curves).

**Example 3.5.** Now I would like to reconcile the theory I outlined above and the approach in the previous sections, for which I consider a more general transport equation with a non-constant velocity:

$$u_t + c(x)u_x = 0,$$

with the initial condition

$$u(0, x) = g(x).$$

Using the parametrization as in the previous examples, I will find that

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = 1, \quad t(0, \xi) = 0 \implies t = \tau,$$

and hence I will only need one parameter $\xi$, which is introduced in the equation for the characteristics (note I use $t$ instead of $\tau$)

$$\frac{\mathrm{d}x}{\mathrm{d}t} = c(x), \quad x(0) = \xi.$$

Along the curve defined by this equation I have an ODE

$$\frac{\mathrm{d}v}{\mathrm{d}t} = 0, \quad v(0, \xi) = g(\xi) \implies v(t, \xi) = g(\xi).$$

Hence, if I can express $\xi$ from the equation of the characteristic, then I will have my unique solution $u(t, x) = g(\xi(t, x))$.

To illustrate, let me take

$$c(x) = -x\,.$$

In this case the characteristics are the solutions to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -x\,, x(0) = \xi \implies x = \xi e^{-t}.$$

If the initial condition is given by

$$u(0,x) = \frac{1}{x^2+1}\,,$$

then my unique solution is hence

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

To conclude, in the last two lectures I considered the so-called *method of characteristics* to solve an initial value problem for a linear (or semi-linear) first order PDE, where the unknown function depends on two independent variables. The key fact is that along the special curves, called the characteristic curves or characteristics, these PDE turn into ODE, for which an extensive theory exists (from a physical point of view this is a manifestation of particle-wave duality, when the system can be either described by the positions of discrete particles or using a continuous representation of a force field, as explained in the book by V.I. Arnold, which I cited in the first lecture). This method can be immediately generalized to linear first order PDE with more than two independent variables and also, with some modifications, to nonlinear equations (as it was mentioned earlier, if the equation is semi-linear, then the method of characteristics is pretty much exactly the same). I will only touch on a more general nonlinear equation in the following lecture.

**Exercise 1.** Solve

$$xu_x + yu_y + zu_z = 1$$

with the condition

$$u = 0 \quad \text{on} \quad x + y + z = 1.$$

**Exercise 2.** Solve

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

## 3.2 Test yourself

3.1. Solve

$$u_x + u_y = 1, \quad u = 0 \quad \text{on} \quad x + y = 1.$$

3.2. Let $u$ be a smooth function of two variables $x, y$. We know that $u$ is constant along each curve $y = x^2 + 2x + \xi$ (but certainly $u$ can be different from one curve to another). Which PDE does $u$ solve?

### 3.3 Solutions to the exercises

*Exercise 1.* The only difference here while solving first order linear PDE with more than two independent variables is the lack of possibility to give a simple geometric illustration. In this particular example the solution $u$ is a hyper-surface in 4-dimensional space, and hence no drawing can be easily made. The characteristics live in 3-dimensional space and hence can be drawn, but this picture would be less informative compared to the case covered in the rest of the notes. So let me proceed purely algebraically, following the same steps as in the examples above.

First, I need to parameterize the initial condition. I choose (note that now I need two parameters)

$$x = \xi, \quad y = \zeta, \quad z = 1 - \xi - \zeta.$$

I have

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}\tau} &= x, \quad x(0,\xi,\zeta) = \xi,\\ \frac{\mathrm{d}y}{\mathrm{d}\tau} &= y, \quad y(0,\xi,\zeta) = \zeta,\\ \frac{\mathrm{d}z}{\mathrm{d}\tau} &= z, \quad z(0,\xi,\zeta) = 1 - \xi - \zeta,\end{aligned}$$

and hence

$$x(\tau,\xi,\zeta) = \xi e^{\tau}, \quad y(\tau,\xi,\zeta) = \zeta e^{\tau}, \quad z(\tau,\xi,\zeta) = (1 - \xi - \zeta)e^{\tau}$$

are my characteristics.

The equation along the characteristics takes the form

$$\frac{\mathrm{d}v}{\mathrm{d}\tau} = 1, \quad v(0,\xi,\zeta) = 0,$$

which yields

$$v(\tau,\xi,\zeta) = \tau.$$

Since from the equations for the characteristics I have that $x + y + z = e^{\tau}$, I conclude that

$$u(x,y,z) = \log(x + y + z),$$

which can be checked directly.

From the final answer I see that my solutions are defined only for $x + y + z > 0$. ■

*Exercise 2.* This equation is nonlinear, but all the nonlinearity is kept away from the partial derivatives. Such equations are called semi-linear because they behave *almost* like linear equations in some respect. Let me just follow exactly the same steps as I would do for a linear equation.

The characteristics are $x(t,\xi) = t+\xi$ (I did it multiple times above, so here I just write the answer). Along the characteristics I have the ODE

$$\frac{\mathrm{d}v}{\mathrm{d}t} = v^2, \quad v(0,\xi) = e^{-\xi^2},$$

which integrates to

$$v(t,\xi) = \frac{1}{e^{\xi^2} - t}.$$

Therefore, my final solution is

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

So, what is the difference? Now just looking at the characteristics I cannot conclude for which $t, x$ my solution is defined (how I did for the linear equations). In this particular example the characteristics carry the information throughout the whole plane. The solution, however, has a possibility *to blow up* in the case when $e^{(x-t)^2} = t$, which is a typically nonlinear effect. ■