3 Solving first order linear PDE

In this lecture I will deviate slightly from the textbook and consider a general linear first order PDE of the form

\[ a(x, y)u_x + b(x, y)u_y = c(x, y)u + d(x, y), \]  

(3.1)

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

\[ u(x, y)|_{(x, y) \in \Gamma} = g(x, y), \]  

(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 from the textbook is that I am allowing to have rather general initial conditions contrary to the fixed value \( t = 0 \) in the textbook. 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

\[ a(x, y)u_x + b(x, y)u_y = f(x, y, u), \]  

(3.3)

where \( f \) is some, generally nonlinear, function.

Let me consider the system of ordinary differential equations

\[
\begin{align*}
\frac{dx}{d\tau} &= a(x, y), \\
\frac{dy}{d\tau} &= b(x, y).
\end{align*}
\]

(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 \( u(x, y) \) to (3.1) along \((x(\tau), y(\tau))\). 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 \( \tau \):

\[
\frac{dv}{d\tau} = u_x x' + u_y y' = a(\tau)u_x + b(\tau)u_y = c(\tau)v + d(\tau),
\]

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

\[ \Gamma: x(\xi), y(\xi), \]

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):

\[ (\tau, \xi) \mapsto (x(\tau, \xi), y(\tau, \xi), v(\tau, \xi)). \]

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 discussion and instead consider a few examples.

**Example 3.2.** Solve

\[ xu_x + yu_y = u, \]
\[ u(x, 1) = g(x). \] (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

\[ \frac{dx}{d\tau} = x, \quad x(0, \xi) = \xi, \]
\[ \frac{dy}{d\tau} = y, \quad y(0, \xi) = 1, \]

which immediately implies that

\[ x(\tau, \xi) = \xi e^\tau, \quad y(\tau, \xi) = e^\tau. \]

From the last expression I also have that

\[ \xi = \frac{x}{y}. \]
Along my characteristics I have (note the initial condition)
\[
\frac{dv}{d\tau} = v, \quad v(0, \xi) = g(\xi) \implies v(\tau, \xi) = g(\xi)e^{\tau}.
\]
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
\[u(x, y) = e^{-(x/y)^2}y.
\]

**Example 3.3.** Solve
\[
yu_x - xu_y = 0, \quad u(x,0) = g(x), \quad x > 0.
\]
\hspace{1cm} (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
\[
\frac{dx}{d\tau} = y, \quad x(0, \xi) = \xi, \\
\frac{dy}{d\tau} = -x, \quad y(0, \xi) = 0.
\]
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:

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

and integrating this separable equation.

Along the characteristics I have

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

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 $\mathbb{R}^2$, 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.
Example 3.5. Now I would like to reconcile the theory I outlined above and the approach in the textbook, where the transport equation with a non-constant velocity is given:

\[ u_t + c(x)u_x = 0, \]

with the initial condition

\[ u(0, x) = g(x). \]

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

\[ \frac{dt}{d\tau} = 1, \quad t(0, \xi) = 0 \implies t = \tau, \]

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

\[ \frac{dx}{dt} = c(x), \quad x(0) = \xi. \]

Along the curve defined by this equation I have an ODE

\[ \frac{dv}{dt} = 0, \quad v(0, \xi) = g(\xi) \implies v(t, \xi) = g(\xi). \]

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 (Problem 2.2.17), let me take
\[ c(x) = -x. \]

In this case the characteristics are the solutions to
\[ \frac{dx}{dt} = -x, \quad x(0) = \xi \implies x = \xi e^{-t}. \]

If the initial condition is given by
\[ u(0, x) = \frac{1}{x^2 + 1}, \]
then my unique solution is hence
\[ u(t, x) = \frac{1}{x^2 e^{2t} + 1}. \]

More details can be found in the textbook and in the homework problems.

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). 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. I will only touch on the latter topic in the following lecture.