Partial Differential Equations (Week 2)
First Order PDEs

Gustav Holzegel

January 24, 2019

1 Introduction

Consider the following *transport equation*

$$u_t + cu_x = 0 \tag{1}$$

for a function $u : \mathbb{R}^2 \to \mathbb{R}$ depending on x and t (time) and $c > 0$. As a first idea, you may try to solve the above using a power series approach. Given $u(x, 0) = h(x)$ smooth for $t = 0$, one can indeed compute all partial derivatives of u on $t = 0$ from the equation (note all ∂_x^k-derivatives are already given by specifying $h(x)$):

$$\partial_t^j \partial_x^k u(x, 0) = (-c)^j \left(\frac{d}{dx}\right)^{j+k} h(x) \ .$$

The hope then is that

$$u(x, t) =^! \sum_{j,k} \frac{(-c)^j h^{(j+k)}(0)}{j!k!} t^j x^k \tag{2}$$

converges in a neighborhood of $(0, 0)$.

Exercise 1.1. *Show that if h is real analytic and converges for $|x| < R$, then the series on the right indeed converges to a solution of the PDE in a neighborhood of $(0, 0)$. [Hint: You may want to use the multinomial identity $(x_1 + \ldots x_n)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} x^\alpha$.]*

We will talk more about this power series approach when we discuss the Cauchy-Kovalevskaya theorem for general PDEs (Week 3). Note already, however, that if $h \in C^\infty$ and all derivatives vanish at 0 but not in a neighborhood of 0 (so h is *not* analytic), then the series (2) converges, but not to a solution.

In any case, there is a much simpler way of understanding (1). Suppose we have a classical (i.e. C^1) solution $u(x,t)$ of (1). We observe that along the lines L_a defined by $t = \frac{1}{c}x + a$ (and parametrized by a) the solution is constant:

$$\frac{d}{ds} u(c(s-a), s) = u_t + u_x \cdot c = 0 .$$

This tells us that any C^1 solution $u(x,t)$ will be uniquely determined by specifying a C^1 function $h(x)$ at $t = 0$ (or, more generally, on any line which is not one of the L_a). Indeed, the solution is simply going to get transported along the *characteristic direction* L_a of the PDE. In particular, the *domain of influence* of a point is precisely given by the characteristic line through that point. We summarize this as

Proposition 1.2. *The unique classical solution of (1) with initial data a C^1 function $h(x)$ prescribed at $t = 0$ is given by $u(x,t) = h(x - ct)$.*

Remark 1.3. *An interesting question which we will touch below is what happens for initial functions $h(x)$ which are not C^1 (propagation of singularities along characteristics).*

2 Quasi-linear case

Consider the general quasi-linear first order PDE

$$a(x,y,u)\, u_x + b(x,y,u)\, u_y = c(x,y,u) , \tag{3}$$

for C^1 functions $a, b, c : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$. For simplicity, we will discuss this *two-dimensional* case, the higher dimensional generalization being straightforward (see Exercise 2 below).

Definition 2.1. *Let $z = u(x,y)$ be a C^1-surface in $\mathbb{R}^3$. If $z = u(x,y)$ solves (3), we will call the surface an integral surface of the PDE.*

Suppose we are given an integral surface S, $z = u(x,y)$. Then, the normal of the tangent plane at any point of S is proportional to the vector $(u_x, u_y, -1)^T$ by standard multi-variable calculus. If we also interpret $(a(x,y,z), b(x,y,z), c(x,y,z))^T$ as a vector-field in $\mathbb{R}^3$ (defining the *characteristic direction* of the PDE at each point), then the PDE (3) says precisely that the characteristic direction at each point has to be perpendicular to the normal of of the tangent plane. In other words, the characteristic direction has to lie in the tangent plane to the integral surface. This suggest to obtain an integral surface via the integral curves along the characteristic direction.

Suppose we are at a point $(x_0, y_0, z_0) \in S$. Given the characteristic direction, we can define an integral curve through (x_0, y_0, z_0), by solving the ODE-system

$$\begin{aligned}
\frac{dx}{dt} &= a(x,y,z) \qquad & x(0) &= x_0 \\
\frac{dy}{dt} &= b(x,y,z) \qquad & y(0) &= y_0 \\
\frac{dz}{dt} &= c(x,y,z) \qquad & z(0) &= z_0
\end{aligned}$$

The corresponding curve, which exists by the ODE theory we developed (Exercise: Check!) is called *the characteristic* curve.

Lemma 2.2. *If a C^1-surface S is a union of characteristic curves, then it is an integral surface.*

Proof. Through any point P we have a characteristic curve lying in S. Since the tangent to the characteristic curve is everywhere equal to the characteristic direction, the latter lies in the tangent plane to S. This means that the normal to the tangent plane is perpendicular to the characteristic direction, which implies that the PDE (3) is satisfied. □

Proposition 2.3. *Given an integral surface S of (3) defined by $z = u(x,y)$ and a point $P = (x_0, y_0, z_0)$ together with the characteristic curve γ through P, then γ lies entirely in S.*

Proof. Let $\gamma(t) = (x(t), y(t), z(t))$ be the characteristic curve going through P at $t = 0$ and consider the quantity

$$U(t) = z(t) - u(x(t), y(t))$$

Clearly $U(t=0) = 0$. If we can show that $U = 0$ for all t, then this clearly implies that γ lies in S. Differentiating U and using that the characteristic system of ODEs is satisfied, we find

$$\frac{dU}{dt} = c(x(t), y(t), z(t)) - u_x \cdot a(x(t), y(t), z(t)) - u_y \cdot b(x(t), y(t), z(t))$$

which we may write as

$$\begin{aligned}
\frac{dU}{dt} &= c(x(t), y(t), U(t) + u(x(t), y(t))) \\
&- u_x \cdot a(x(t), y(t), U(t) + u(x(t), y(t))) \\
&- u_y \cdot b(x(t), y(t), U(t) + u(x(t), y(t))) \ .
\end{aligned}$$

This is an ODE for $U(t)$, all other quantities being given by the characteristic curve and the value of u along it. This ODE has the solution $U = 0$ identically as is easily checked. However, by the uniqueness theorem for solutions to ODEs, it is also the only solution. □

We conclude that the integral surface is the union of characteristic curves.

Now we want to formulate the appropriate Cauchy-problem inspired by our initial example of the linear transport equation. The idea is to specify the solution along a (non-characteristic) curve Γ and then solve the characteristic ODEs which should provide a unique solution at least in a neighborhood of a point.

Let the C^1-curve Γ be represented parametrically by

$$x = f(s) \qquad\qquad y = g(s) \qquad\qquad z = h(s) \tag{4}$$

and let $x_0 = f(s_0)$, $y_0 = g(s_0)$, $z_0 = h(s_0)$. We are looking for a solution of the PDE (3) such that we have

$$h(s) = u(f(s), g(s))$$

for all s – this is our initial condition. (Note that in the simplest case (the curve Γ being "time 0") we have the familiar $x = 0, y = s, z = h(s) = h(y)$ and hence $h(y) = u(0, y)$ initially.) We would like to construct a local solution near $x_0 = f(s_0)$ and $y_0 = g(s_0)$, parametrized locally by $x = X(s,t), y = Y(s,t), z = Z(s,t)$ where X, Y, Z satisfy the characteristic ODEs

$$\begin{aligned} X_t &= a(X(s,t), Y,(s,t), Z(s,t)) & \qquad X(s,0) &= f(s) \\ Y_t &= b(X(s,t), Y,(s,t), Z(s,t)) & \qquad Y(s,0) &= g(s) \\ Z_t &= c(X(s,t), Y,(s,t), Z(s,t)) & \qquad Z(s,0) &= h(s) \end{aligned} \tag{5}$$

By existence and uniqueness and continuous dependence on the parameters of the solution to the ODE system, we obtain a unique solution of class C^1 for (s,t) near $(s_0, 0)$.

Exercise 2.4. *(Week 1 refresher) Show explicitly that the solution depends continuously on the parameter s.*

In order to represent the solution as an integral surface $z = u(x,y)$ we need to be able to invert the equations $x = X(s,t), y = Y(s,t)$ to $s = S(x,y)$ and $t = T(x,y)$, since then $u(x,y) = Z(S(x,y), T(x,y))$ is an integral surface. The implicit function theorem tells us that this is possible locally near $(s_0, 0)$, provided the determinant of the differential

$$\begin{bmatrix} \frac{\partial X}{\partial s}(s,t) & \frac{\partial X}{\partial t}(s,t) \\ \frac{\partial Y}{\partial s}(s,t) & \frac{\partial Y}{\partial t}(s,t) \end{bmatrix} \tag{6}$$

is non-zero at $(s_0, 0)$, which yields the condition

$$f'(s_0) \cdot b(x_0, y_0, z_0) - g'(s_0) \cdot a(x_0, y_0, z_0) \neq 0 \tag{7}$$

Exercise 2.5. *Show that with this condition being satisfied, the function $u(x,y) = Z(S(x,y), T(x,y))$ indeed solves the PDE (3). [Hint: This follows either from carefully applying the chain rule or from observing that by (5), the parametrized surface has the characteristic direction in its tangent space.]*

The previous exercise ensures *existence* of a solution to the PDE near (x_0, y_0). This solution is also the *unique* C^1 solution realizing the "data" prescribed on Γ: Indeed, in view of Proposition 2.3, any integral surface through Γ has to contain all characteristic curves through Γ and hence locally agrees with the integral surface we constructed.

Let us finally investigate what happens if condition (7) is violated. In the latter case we have

$$f'b - g'a = 0\,, \qquad h' = u_x f' + g' u_y\,, \qquad c = au_x + bu_y \tag{8}$$

at $(s_0, 0)$, which implies that the vectors (f', g', h') and (a, b, c) are parallel (Exercise), so Γ has to be characteristic at $s = s_0$ in this case.

In summary, we have proven the following theorem

Theorem 2.6. *Consider the PDE (3), a point $P = (x_0, y_0, z_0 = u\,(x_0, y_0))$ in $\mathbb{R}^3$ and a parametrized C^1 curve $\Gamma : s \to (f\,(s)\,, g\,(s)\,, h\,(s))$ going through P at s_0, which is also non-characteristic at P (i.e. (7) holds). Then, there exists a small neighborhood $\mathcal{U} \subset \mathbb{R}^2$ of (x_0, y_0) and a unique C^1 function $u : \mathcal{U} \to \mathbb{R}$ which solves the PDE (3) in $\mathcal{U}$ and satisfies $h\,(s) = u\,(f\,(s)\,, g\,(s))$ along Γ.*

Remark 2.7. *It is important to note that in the quasi-linear case, the condition (7) of being non-characteristic at a point depends on the value of the solution at that point.*

3 Examples

3.1 The transport equation revisited

Let us go back to (1) with C^1 initial data $h\,(x)$ prescribed at $t = 0$. The initial data on the curve Γ is then given by $t = 0$, $x = s$, $z = h\,(s) = h\,(x)$. The characteristic equations are

$$\frac{dt}{d\tau} = 1 \qquad , \qquad \frac{dx}{d\tau} = c \qquad , \qquad \frac{dz}{d\tau} = 0 \tag{9}$$

with initial conditions $t\,(s, 0) = 0$ (hence $t\,(s, \tau) = \tau$), $x\,(s, 0) = s$ (hence $x\,(s, \tau) = c \cdot \tau + s$) and $z\,(s, 0) = h\,(s)$ (hence $z\,(s, \tau) = h\,(s)$). From this we easily obtain $u\,(x, t) = h\,(x - c \cdot t)$ in agreement with what we found earlier.

3.2 Inviscid Burger's equation

Consider the non-linear first order PDE

$$u_t + uu_x = u_t + \left(\frac{u^2}{2}\right)_x = 0 \tag{10}$$

with initial data prescribed at $t = 0$. This equation arises in fluid dynamics and is the simplest model exhibiting *shock formation.* Recall that the equation describing a free particle in Newtonian mechanics is $\ddot{x}(t) = 0$. Defining the velocity field $u(x(t), t) = \dot{x}(t)$, the chain rule implies that $u(t, x(t))$ satisfies Burger's equation. We will elucidate this relation further once we gained insight on the structure of solutions to (10).

In the notation of the previous section, the initial data is prescribed on the curve Γ given by $t = 0$, $x = s$, $z = h(s) = h(x)$ (say again with h being C^1). The characteristic equations are

$$\frac{dt}{d\tau} = 1 \qquad , \qquad \frac{dx}{d\tau} = z \qquad , \qquad \frac{dz}{d\tau} = 0 \tag{11}$$

with initial conditions $t(s,0) = 0$ (hence $t(s,\tau) = \tau$), $z(s,0) = h(s)$ (hence $z(s,\tau) = h(s)$ and $x(s,0) = s$ (hence $x(s,\tau) = z \cdot \tau + s$ noting that $\frac{dz}{d\tau} = 0$). From this we obtain the implicit formula $u(x,t) = h(x - u(x,t) \cdot t)$.

To understand what general solutions look like, we depict the projected characteristics of the solution in the (x,t) plane. Fixing a point $(s,0)$, where $u(s,0) = h(s)$ we have that the solution is constant (equal to $h(s)$) along the line $x = h(s) \cdot t + s$ (Exercise: Why?). In other words, $h(s)$ (or rather $\frac{1}{h(s)}$) determines the slope of the characteristic going through each point. From the Newtonian particle point of view, we can think of $h(x)$ as the initial velocity distribution of particles. Every particle moves along its characteristic. It is already clear graphically (see the figures below) that faster particles will eventually overtake slower ones and hence that the evolution will break down (unless $h(s)$ is monotonically increasing). More precisely, two characteristics C_1 and C_2 will intersect at some (x,t) where

$$t = -\frac{s_2 - s_1}{h(s_2) - h(s_1)} \, . \tag{12}$$

At the point of intersection, the solution u can no-longer be uni-valued. To visualize what happens it is instructive to plot different snapshots of a solution arising from data of compact support (Exercise). This breakdown of the solution in finite time is characteristic of non-linear PDEs. Here, the geometry of the break-up is simple enough to be understood.

It is clear that if we want to continue the solution beyond the shock, we have to give up our requirement that the solution should be C^1, i.e. classical. However, if we do this, then, how should the PDE (10) be understood? The key is the introduction of the notion of a so-called *weak solution* (below called an *integral solution*) of (10). This will concern us in greater detail later in the course once we have developed some distribution theory. For now we will only provide a first taste by discussing the present example

The basic idea is simple. Suppose first that we have a C^1 solution of (10) with initial datum $u(x,0) = h(x)$. Multiply the PDE by a smooth test function $v : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ of compact support and integrate over spacetime. This produces the identity

$$\int_0^\infty dt \int_{-\infty}^\infty dx \left[u v_t + \frac{u^2}{2} v_x \right] + \int_{-\infty}^\infty h v dx \Bigg|_{t=0} = 0 \, . \tag{13}$$

Although we arrived at (13) assuming that u was classical, this identity makes sense for much less regular u (and data $h(x)$). Therefore, we define

Definition 3.1. *We say that* $u \in L^{\infty}(\mathbb{R} \times (0,\infty))$ *is an integral solution of (10) with data* $u(x,0) = h(x)$ *provided that (13) holds for any smooth test function* $v : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ *of compact support.*

Exercise 3.2. *Show that if* u *is an integral solution that is also* C^1*, then it is a classical solution.*

Of course, the difficult question to answer is: Does there always exist a global weak solution of (10) and if so, is it unique?

We will not be able to answer this question in fully generality but will gain further insight by studying a particular class of integral solutions. Let us investigate what happens for a piecewise smooth solution u given in some connected region $V \subset \mathbb{R} \times (0,\infty)$ which is discontinuous along a smooth curve C separating V into V_l and V_r. We assume u is an integral solution and that moreover u and its first derivatives are uniformly continuous in V_l and V_r.

Choosing a test function compactly supported in V_l (or V_r respectively) we easily find that we must have

$$u_t + u u_x = 0$$

in both V_l and V_r (recall Exercise 3.2). However, if the test function doesn't vanish on the curve (but say still vanishes near the boundary of V) we will find (with $\nu = \left(\nu^1, \nu^2\right)$ denoting the unit-normal to C pointing from V_l into V_r)

$$\iint_{V_l} u v_t + \frac{u^2}{2} v_x \, dx dt = \int_C \left(u_l \nu^2 + \frac{u_l^2}{2} \nu^1 \right) v \; dl$$

$$\iint_{V_r} u v_t + \frac{u^2}{2} v_x \, dx dt = -\int_C \left(u_r \nu^2 + \frac{u_r^2}{2} \nu^1 \right) v \; dl$$

where u_l and u_r denote the limits from the left and the right respectively. Since u is assumed to be an integral solution we must have

$$\int_C \left((u_l - u_r) \nu^2 + \frac{u_l^2 - u_r^2}{2} \nu^1 \right) v \; dl = 0 \tag{14}$$

for all test functions, and hence

$$(u_l - u_r) \nu^2 + \frac{u_l^2 - u_r^2}{2} \nu^1 = 0$$

along C. Suppose that the curve C is parametrized by $\gamma(t) = (x = s(t), t)$ for some smooth function $s : [0,\infty) \to \mathbb{R}$. The tangent is then proportional to $(\dot{s}, 1)$ and hence for the normal we obtain $\nu = \left(\nu^1, \nu^2\right) = \frac{1}{\sqrt{1+\dot{s}^2}} (1, -\dot{s})$. Therefore, we have

$$u_l^2 - u_r^2 = 2\dot{s} (u_l - u_r) \; . \tag{15}$$

This is the *Rankine-Hugoniot condition* relating the jump in u to the *speed* $\dot{s}$. Note that u_l, u_r and $\dot{s}$ all change along C, however, the relation above must always hold.

Exercise 3.3. *Repeat the above computations for the more general PDE*

$$u_t + (F(u))_x = 0$$

to conclude the Rankine-Hugoniot condition

$$F(u_l) - F(u_r) = \dot{s}(u_l - u_r) \, . \tag{16}$$

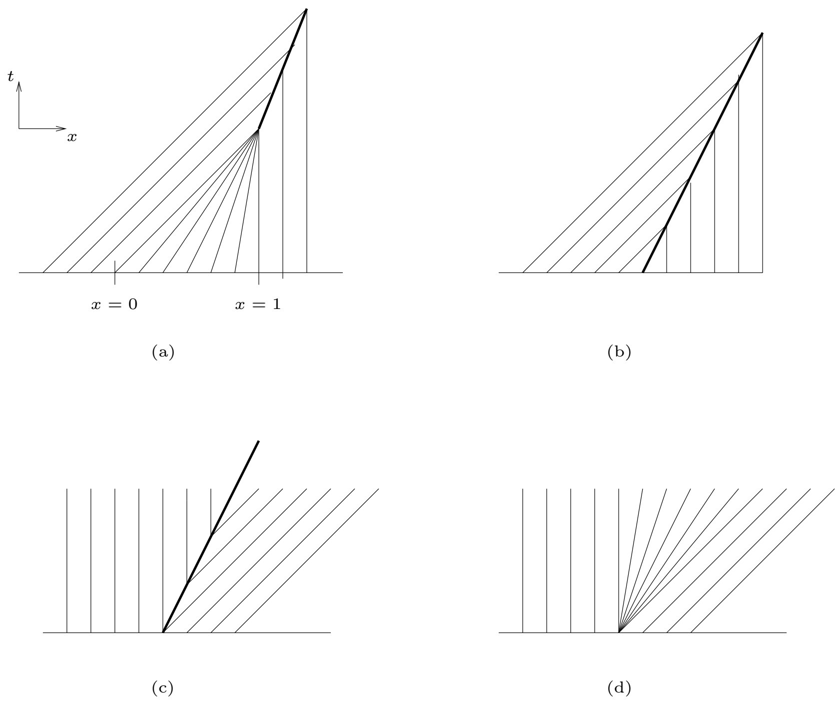
We will illustrate the above computation by examples. We consider the evolution of the three initial-data sets

$$h_1(x) = \begin{cases} 1 & \text{if } x \le 0, \\ 1 - x & \text{if } 0 \le x \le 1 \\ 0 & \text{if } x > 1. \end{cases} \tag{17}$$

$$h_2(x) = \begin{cases} 1 & \text{if } x < 0, \\ 0 & \text{if } x \ge 0. \end{cases} \tag{18}$$

$$h_3(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \ge 0. \end{cases} \tag{19}$$

For h_1 (piecewise smooth), the method of characteristics yields the solution depicted below in (a).



(a) (b) (c) (d)

Indeed, up to time $t < 1$ we have

$$u_1(x,t) = \begin{cases} 1 & \text{if } x \leq t \text{ and } \ 0 \leq t < 1, \\ \frac{1-x}{1-t} & \text{if } t \leq x \leq 1 \text{ and } \ 0 \leq t < 1 \\ 0 & \text{if } x \geq 1 \text{ and } \ 0 \leq t < 1. \end{cases} \tag{20}$$

How should u_1 be defined for $t \geq 1$? The RH-condition suggests that $\dot{s} = \frac{1}{2}$ should hold along the curve separating the two regions where the solution is still smooth. Hence we define $s(t) = \frac{1+t}{2}$ and

$$u_1(x,t) = \begin{cases} 1 & \text{if } x < s(t) \ \text{ and } \ t \geq 1, \\ 0 & \text{if } x > s(t) \ \text{ and } \ t \geq 1. \end{cases} \tag{21}$$

We have thus found a global integral solution to our PDE and it also turns out to be unique. In understanding the problem for h_1, we have understood it for h_2 because the situation at $t = 1$ for h_1 is precisely the situation for h_2 at time zero (only shifted).

The evolution of h_3 is more tricky because of a failure of uniqueness. Note that a-priori, the solution is not determined by the characteristics in a small wedge opening up at zero. It is easy to check that both

$$u_{3c}(x,t) = \begin{cases} 0 & \text{if } x < \frac{t}{2}, \\ 1 & \text{if } x > \frac{t}{2} \end{cases}, \tag{22}$$

depicted in (c) above, and

$$u_{3d}(x) = \begin{cases} 1 & \text{if } x > t, \\ \frac{x}{t} & \text{if } 0 \leq t \leq x \\ 0 & \text{if } x < 0. \end{cases} \tag{23}$$

depicted in (d), define an integral solution in $\mathbb{R} \times [0, \infty)$ to (10) with data $h_3(x)$. This example shows that we can, in general, not expect to find a *unique* weak solution of (10) unless we prescribe additional information (in the above case, the rarefaction wave (d) turns out to be physically correct).

The additional information ensuring the uniqueness of weak solutions is typically given by so-called *entropy conditions*. Suppose we are in the class of *piecewise-smooth* integral solutions. We may hope that starting from a arbitrary point P in the (t,x)-plane, we will not see any crossing characteristics when moving towards the *past* along the characteristic through that point. This requires $u_l > \dot{s} > u_r$ to hold along a curve of discontinuities. It turns out (as suggested by the examples) that the above condition along any shock curve suffices to ensure uniqueness of weak solutions. For our example, it singles out the "rarefaction wave" u_{3d} as the unique weak solution.[1]

We will not go into the existence theory and uniqueness theory of global weak solutions at this point. For further discussion, see Chapter 3 of Evans.

[1] This is also the solution one obtains from the *viscous* Burger's equation in the zero viscosity limit.

4 The fully non-linear case

The geometry of fully non-linear first order PDEs is considerably more complicated but remarkably, the local problem still reduces to that of solving ODEs. I will first (Section 4.1) explain the geometric intuition leading to the system of characteristic ODEs (30)–(32), (36), (37). The impatient reader may jump immediately to this ODE system and – after doing the computation of Lemma 4.1 – continue with Section 4.2. See the book of Fritz John for a more detailed treatment.

4.1 Heuristic derivation of the ODE system

We start with the PDE

$$F\left(x, y, z, p=u_x, q=u_y\right)=0 \tag{24}$$

with (say) F being smooth. At a given point $P=\left(x_0, y_0, z_0\right)$, the tangent plane to an integral surface $z=u\left(x, y\right)$ is given by

$$z-z_0=p\left(x-x_0\right)+q\left(y-y_0\right)$$

with p and q satisfying $F\left(x_0, y_0, z_0, p, q\right)=0$. Suppose that we found an explicit $\left(p_0, q_0\right)$ with $F\left(x_0, y_0, z_0, p_0, q_0\right)=0$. Then, provided $F_q\left(x_0, y_0, z_0, p_0, q_0\right) \neq 0$, we expect that for p near p_0 we can solve for $q\left(p\right)$ and thereby obtain an entire family of possible tangent planes at P, parametrized by p.

How do we, from this, single out a *characteristic direction*? The key idea is to use the envelope of the family of planes (i.e. the surface which is tangent to all members of the family). The envelope touches each plane in a certain generator; therefore, once we have chosen p and q at the point P (hence a particular tangent plane), this determines a particular direction: the characteristic direction.[2] The envelope is typically a cone and one speaks of the field of Monge-cones associated with the PDE (24).

Let us improve the above heuristics by some formulae. The envelope is defined as follows. If

$$G\left(x, y, z, p\right):=p\left(x-x_0\right)+q\left(p\right)\left(y-y_0\right)-\left(z-z_0\right)=0 \tag{25}$$

is our family of tangent planes, then its envelope is obtained by solving

$$\partial_p G\left(x, y, z, p\right)=0 \tag{26}$$

for $p=p\left(x, y, z\right)$ and plugging it back into (25). The generator of intersection of the envelope and the plane is given by solving (25) and (26) for *fixed* p. It is convenient to work in the language of differential forms. For fixed p we have

$$dz=p dx+q\left(p\right) dy \tag{27}$$

[2]For the quasi-linear problem, this generator will be the same for any plane, as the envelope degenerates to a line in this case, as can be checked by a quick computation.

describing the plane, while condition (26) yields

$$dx + \frac{dq}{dp} dy = 0\,. \tag{28}$$

Finally, from differentiating $F(x_0, y_0, z_0, p, q) = 0$ we obtain

$$F_p + \frac{dq}{dp} F_q = 0\,. \tag{29}$$

This suggests that we should solve the system of ODEs

$$\frac{dx}{dt} = F_p(x, y, z, p, q)\ , \tag{30}$$

$$\frac{dy}{dt} = F_q(x, y, z, p, q)\ , \tag{31}$$

$$\frac{dz}{dt} = pF_p(x, y, z, p, q) + qF_q(x, y, z, p, q)\ . \tag{32}$$

However, this is not a closed system unless we are in the quasi-linear case! In general, only if we are *given* the integral surface we can determine $p = u_x$ and $q = u_y$ and hence solve the above ODEs. Remarkably, we can actually complement the above system by ODEs for p and q to obtain a closed system.

Here is how. Differentiate the PDE with respect to x and y to produce

$$\begin{aligned} &F_x + F_z p + F_p u_{xx} + F_q u_{xy} = 0\,, \\ &F_y + F_z q + F_p u_{xy} + F_q u_{yy} = 0\,. \end{aligned} \tag{33}$$

On the other hand, along a characteristic we have

$$\frac{dp}{dt} = \frac{d}{dt}\left(u_x\left(x\left(t\right), y\left(t\right)\right)\right) = u_{xx} F_p + u_{xy} F_q\,, \tag{34}$$

$$\frac{dq}{dt} = \frac{d}{dt}\left(u_y\left(x\left(t\right), y\left(t\right)\right)\right) = u_{yx} F_p + u_{yy} F_q\,. \tag{35}$$

Combining these equations it becomes apparent that we should complement the ODE system (30)–(32) above with the equations

$$\frac{dp}{dt} = -F_x(x, y, z, p, q) - pF_z(x, y, z, p, q)\ , \tag{36}$$

$$\frac{dq}{dt} = -F_y(x, y, z, p, q) - qF_z(x, y, z, p, q)\ , \tag{37}$$

thereby achieving *closure*. We will call the system (30)–(32), (36), (37) the system of characteristic ODEs. The following is easy to check:

Lemma 4.1. *The function $F(x, y, z, p, q)$ is an integral of the system of characteristic ODEs (30)–(32), (36), (37) in that*

$$\frac{d}{dt}\left[F\left(x\left(t\right), y\left(t\right), z\left(t\right), p\left(t\right), q\left(t\right)\right)\right] = 0$$

holds along a solution.

Proof. Straightforward computation. □

Geometrically, compared with the quasi-linear case, we are now propagating *plane elements* (x, y, z, p, q) in $\mathbb{R}^5$ (i.e. a point in $\mathbb{R}^3$ and a tangent plane) according to the system of ODEs. We will call $(x(t), y(t), z(t), p(t), q(t))$ a *characteristic strip* if the tangent to the curve defined by $(x(t), y(t), z(t))$ lies in the tangent plane determined by the plane element. This is indeed ensured if the characteristic ODEs are satisfied, as they imply that $\frac{dz}{dt} = p\frac{dx}{dt} + q\frac{dy}{dt}$ holds.

4.2 The Cauchy problem

We proceed with the Cauchy problem. As in the quasi-linear case, we would like to specify data on a curve Γ parametrized by $x = f(s)$, $y = g(s)$, $z = h(s)$ with functions of class C^1 near s_0. However, to obtain appropriate data for the characteristic ODEs, we now need to complete Γ into a strip, i.e. we need to determine functions $\phi(s)$ and $\psi(s)$ satisfying

$$h'(s) = \phi(s) f'(s) + \psi(s) g'(s) \,, \tag{38}$$

$$0 = F(f(s), g(s), h(s), \phi(s), \psi(s)) \,. \tag{39}$$

In the quasi-linear case, these equations are *linear* in $\phi(s)$ and $\psi(s)$; therefore, $\phi(s)$ and $\psi(s)$ are determined uniquely near s_0 (from f, g and h) in view of the non-degeneracy condition (7). In the present fully non-linear case, the equations (38) and (39) may have no, one or many solutions. To ensure uniqueness, we will need to invoke the implicit function theorem as follows:

Lemma 4.2. *Given C^1 functions $f(s)$, $g(s)$, $h(s)$ of class C^1 near s_0 and $p_0, q_0 \in \mathbb{R}$ such that*

$$h'(s_0) = p_0 f'(s_0) + q_0 g'(s_0) \tag{40}$$

$$0 = F(f(s_0), g(s_0), h(s_0), p_0, q_0) \tag{41}$$

and also

$$det \begin{bmatrix} f'(s_0) & g'(s_0) \\ F_p(f(s_0), g(s_0), h(s_0), p_0, q_0) & F_q(f(s_0), g(s_0), h(s_0), p_0, q_0) \end{bmatrix} \neq 0 \tag{42}$$

holds. Then, there exist unique C^1-functions $\phi(s)$ and $\psi(s)$ with $\phi(s_0) = p_0$ and $\psi(s_0) = q_0$ satisfying both (38) and (39).

Proof. Standard application of the implicit function theorem. □

We will call a quintuple $(f(s), g(s), h(s), p_0, q_0)$ such that the conditions (40)–(42) of Lemma 4.2 are satisfied an *admissible data set* of the PDE (24). By the Lemma, an admissible data set determines uniquely a quintuple of functions $(f(s), g(s), h(s), \phi(s), \psi(s))$ of class C^1 near s_0 which serves as a one-parameter initial data set for the characteristic system of ODEs. Solving the characteristic ODEs with this data we obtain unique C^1 functions

$$x = X(s,t)\ ,\quad y = Y(s,t)\ ,\quad z = Z(s,t)\ ,\quad p = P(s,t)\ ,\quad q = Q(s,t)\ . \tag{43}$$

Since $F = 0$ holds identically along Γ and moreover $\frac{d}{dt}F = 0$ along the solutions of the system of ODEs (cf. Lemma 4.1) we have that $F(X(s,t), ..., Q(s,t)) = 0$ in a neighborhood of $(s_0, 0)$. As any integral surface through Γ (if it exists) contains its characteristic curves through Γ, it must must locally agree with the surface we constructed (*uniqueness*).[3]

Conversely, (to show *existence*) let us show that the parametrized surface (43) (obtained from an admissible data set) represents a solution of the Cauchy problem for the PDE in a neighborhood of (x_0, y_0, z_0). To do this, note first that can invert the relation $X(s,t)$ and $Y(s,t)$ locally near $(s_0, 0)$ in view of the non-degeneracy condition (42) to obtain

$$s = S(x,y) \qquad t = T(x,y)\ . \tag{44}$$

Then

$$u(x,y) := Z(S(x,y), T(x,y)) \tag{45}$$

defines a function which is C^1 is a neighborhood of $(x_0 = f(s_0), y_0 = g(s_0))$ and agrees with the data prescribed on Γ. Observe that equivalently one may write

$$Z(s,t) = u(X(s,t), Y(s,t))\ . \tag{46}$$

We want to show that the u defined above actually solves the PDE. We claim that for this it is sufficient to establish the two identities $u_x(X(s,t), Y(s,t)) = P(s,t)$ and $u_y(X(s,t), Y(s,t)) = Q(s,t)$, where the right hand side is defined by the solution of the characteristic ODEs. Indeed, given these identities, the known (from solving the ODE system) fact that

$$F(X(s,t), Y(s,t), Z(s,t), P(s,t), Q(s,t)) = 0 \quad \text{holds near } (s_0, 0) \tag{47}$$

is equivalent to

$$F(x, y, u(x,y), u_x(x,y), u_y(x,y)) = 0 \tag{48}$$

holding in a neighborhood of $(x_0 = f(s_0), y_0 = g(s_0))$. To finally prove the desired identites note that the chain rule applied to (46) yields

$$Z_s(s,t) = u_x(X(s,t), Y(s,t)) \cdot X_s(s,t) + u_y(X(s,t), Y(s,t)) \cdot Y_s(s,t) \tag{49}$$

$$Z_t(s,t) = u_x(X(s,t), Y(s,t)) \cdot X_t(s,t) + u_y(X(s,t), Y(s,t)) \cdot Y_t(s,t) \tag{50}$$

[3] If this is not clear, repeat the proof of Proposition 2.3 adapted to the fully non-linear case.

for all (s,t) in a neighborhood of $(s_0, 0)$. We are done if we can also establish the identities (why?)

$$Z_s\,(s,t) = P\,(s,t)\,X_s\,(s,t) + Q\,(s,t)\,Y_s\,(s,t)\ , \tag{51}$$
$$Z_t\,(s,t) = P\,(s,t)\,X_t\,(s,t) + Q\,(s,t)\,Y_t\,(s,t)\ . \tag{52}$$

Now (52) follows directly from the characteristic system of ODEs. For (51), we use a familiar trick: Defining $A = Z_s - PX_s - QY_s$ we derive an ODE for A_t to conclude that $A = 0$ on Γ implies $A = 0$ for all times (Exercise). Summarizing, we have proven

Theorem 4.3. *Let $F\,(x, y, u, u_x, u_y) = 0$ be a PDE for a smooth function $F : \mathbb{R}^5 \to \mathbb{R}$. Let $(f\,(s)\,, g\,(s)\,, h\,(s)\,, p_0, q_0)$ be an admissible data set for the PDE (i.e. the assumptions of Lemma 4.2 are satisfied). Then there exists a unique solution $z = u\,(x, y)$ of the PDE in a neighborhood of $(x_0 = f\,(s_0)\,, y_0 = g\,(s_0))$.*

Of course, we have also obtained an explicit representation for the solution in the above proof.

4.3 Example: The Eikonal Equation

As an interesting application of the above techniques, we will solve the eikonal equation, which is central in geometric optics describing the motion of wave fronts.

$$c^2\left(u_x^2 + u_y^2\right) = 1 \tag{53}$$

with $c > 0$ a constant, which can be interpreted as the speed. Geometrically, the family of possible tangent planes at each point envelopes a cone with opening angle $2\theta_0 = \arctan c$ (Exercise: Why?). From

$$F\,(x, y, z, p, q) = \frac{1}{2}\left(c^2p^2 + c^2q^2 - 1\right) \tag{54}$$

we obtain the characteristic ODEs

$$\frac{dx}{dt} = c^2p\,,\ \frac{dy}{dt} = c^2q\,,\ \frac{dz}{dt} = c^2p + c^2q = 1\,,\ \frac{dp}{dt} = 0\,,\ \frac{dq}{dt} = 0\,. \tag{55}$$

Given a curve Γ parametrized by $x = f\,(s)$, $y = g\,(s)$ and $z = h\,(s)$, we need to find $\phi\,(s)$, $\psi\,(s)$ satisfying (cf. Lemma 4.2)

$$h'\,(s) = \phi\,(s)\,f'\,(s) + \psi\,(s)\,g'\,(s) \qquad \textit{and} \qquad \phi^2\,(s) + \psi^2\,(s) = c^{-2} \tag{56}$$

It is easy to see that there can be no solutions if $(f')^2 + (g')^2 < c^2(h')^2$ (corresponding to Γ making an angle $\theta < \theta_0$ with the z-axis, "timelike") and that in the case $(f')^2 + (g')^2 > c^2(h')^2$ there are two solutions differing in sign.

In the special case that $x = f\,(s)$, $y = g\,(s)$ and $z = 0$ (prescribing the intersection of the integral surface with the xy-plane) we obtain

$$x\,(s,t) = f\,(s) + c^2\phi\,(s)\,t\,, \qquad y\,(s,t) = g\,(s) + c^2\psi\,(s)\,t\,, \qquad z\,(t) = t\,, \tag{57}$$

while $p(s,t) = \phi(s)$ and $q(s,t) = \psi(s)$. It is instructive to draw some pictures of the evolution. In particular, we can illustrate Huygen's principle by depicting the level surfaces of constant u (= constant t) in the xy-plane (Exercise). This is an instructive example of the duality between a particle description (ODEs, characteristics) and a wave description (PDE for u).

5 Exercises

1. Consider the PDE $u_t - iu_x = 0$ for $u(t,x) \in \mathbb{C}$. Identifying the (t,x) plane appropriately with $\mathbb{C}$, show that the solution u has to be holomorphic. Conclude that the initial value problem can only be solved for *analytic* data. Compare and contrast with the transport equation.

2. Generalize the quasi-linear theory developed for in two dimensions to any dimension. In particular, state an appropriate Cauchy problem (cf. Theorem 2.6 and Theorem 4.3).

3. (F. John's PDE book; Picone) Let u be a solution of

$$a(x,y)\,u_x + b(x,y)\,u_y = -u$$

   of class C^1 in the closed unit disk Ω in the xy-plane. Let $a(x,y)\,x + b(x,y)\,y > 0$ hold along the boundary $\partial\Omega$. Prove that u vanishes identically. (Hint: Show that $\max_\Omega u \le 0$ and $\min_\Omega u \ge 0$.)

4. (F. John's PDE book) Show that the function $u(x,t)$ defined for $t \ge 0$ by

$$u = -\frac{2}{3}\left(t + \sqrt{3x + t^2}\right) \qquad \text{for } 4x + t^2 > 0$$
$$u = 0 \qquad \text{for } 4x + t^2 < 0$$

   is an integral solution of (10).

5. Find a global weak solution of Burger's equation for $t > 0$ arising from the initial data

$$h(x) = \begin{cases} 0 & \text{if } x \le 0, \\ 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{if } x > 1. \end{cases} \tag{58}$$

   prescribed at $t = 0$.

6. (F. John's PDE book) Solve the initial value problem for the PDE

$$u_x + u_y = u^2$$

   with initial data $u(x,0) = h(x)$.

7. (F. John's PDE book) Solve the initial value problem for the PDE

$$u_x^2 + u_y^2 = u^2$$

with the following initial data prescribed on the unit circle in the xy-plane: $u(\cos s, \sin s) = 1$ for all $s \in [0, 2\pi]$. Is the solution unique?
[Answer: $u(x, y) = \exp\left(\pm\left(1 - \sqrt{x^2 + y^2}\right)\right)$.]