How to solve quasi linear first order PDE

A quasi linear PDE is an equation of the form

$$v(u,x) \cdot Du(x) = f(u,x) \text{ on } U \subset \mathbb{R}^n \ , \tag{1}$$

subject to the initial condition

$$u = g \text{ on } \Gamma \ , \tag{2}$$

where Γ is a hypersurface in U. We assume again that the field $v(g(x_0), x_0)$ is not zero for $x_0 \in \Gamma$. The only difference when compared with linear first order PDE is that the vector field v and the inhomogeneity may depend on the unknown function u.

We prove the

Theorem

Assume that there exists a point $x_0 \in \Gamma$ such that that the surface Γ is not characteristic at that point, i.e., the vector $v(g(x_0), x_0)$ is not tangent to Γ at x_0. Then, in a vicinity of the point x_0, the PDE (1) with the initial condition (2) has a unique solution.

One can look at this problem in many ways and one of them is the following. The graph of the solution forms an n dimensional surface in $\mathbb{R}^{n+1}$. This surface can be described locally around the point x_0 as the solution of an equation of the form $U(u(x), x) = 0$ for some function $U = U(x^0, x)$ of $n+1$ variables. A sufficient condition for solving for $u(x)$ is that $U_{x^0}(g(x_0), x_0)$ does not vanish. The implicit function theorem then garantees the existence of a solution $u(x)$ in the vicinity of the point x_0. This function U should satisfy the 'initial conditions' $U(g(x), x) = 0$ for all $x \in \Gamma$ in a vicinity of the point x_0 which also sits on Γ. Note that this amounts to an initial condition on a surface of dimension $n-1$ in $\mathbb{R}^{n+1}$.

Next, one writes an equation for U. We shall show later the connection with the original equation (1). Introduce the new vector field

$$V(x^0, x) = (f(x^0, x), v(x^0, x))$$

and denote the gradient with respect to all variables (x^0, x) by ∇. Abreviate (x^0, x) by y, and onsider the PDE

$$V(y) \cdot \nabla U = 0 \text{ in } U \times \mathbb{R} \subset \mathbb{R}^{n+1} \tag{3}$$

subject to the initial conditions

$$U(g(x), x) = 0 \text{ for } x \in \Gamma \ . \tag{4}$$

We want to add as another condition that

$$U_{x^0}(g(x_0), x_0) \neq 0 \ . \tag{5}$$

Recall that this condition is important for solving for $u(x)$.

Notice that there exist n independent integrals for the equation

$$y' = V(y) \ , \tag{6}$$

but the intitial condition requires that the function U be constant on a 'surface' of dimension $n-1$. Thus, the solution of the problem (3),(4) will be, in general, not be unique. E.g., the zero function will be certainly a solution satisfying (3) and (4) but does not deliver us any reasonable candidate for a solution to (1) and (2). Certainly, one can extend the initial condition in the vicinity of x_0 to an $n-1$ dimensional surface, but keeping of course the condition (4). This can be done by extending the condition (4) to the condition

$$U(x^0, x) = x_0 - g(x) \text{ for } x \in \Gamma \ . \tag{7}$$

This is now an initial condition on the hypersurface γ, that is parallel to the x^0-axis and passes through the 'curve' Γ. Another way to express this is, that the projection of γ along the x^0 axis yields Γ. This surface γ contains also the 'curve' $(g(x), x), x \in \Gamma$, in particular the point $(g(x_0), x_0)$ is in γ. Note that $U(g(x), x) = 0$.

First we show that the initial data of this new, extended problem is non characteristic. The vectors tangent to γ at the point $(g(x_0), x_0)$ are of the form

$$\begin{bmatrix} a \\ \vec{\tau} \end{bmatrix} \ ,$$

where $\vec{\tau}$ is any unit vector, tangent to Γ at the point x_0. Clearly, the vector V at the point $(g(x_0), x_0)$ is not tangent to the surface γ at $(g(x_0), x_0)$ since v is not tangent to Γ and hence not parallel to any of the vectors $\vec{\tau}$.

Thus, the initial conditions are non characteristic and there exists a unique solution for the extended problem.

Next, we note that $U_{x^0}(g(x_0), x_0) = 1$ by (7). This shows (5).

Thus, since $U_{x_0}(g(x_0), x_0)$ does not vanish, the implicit function theorem garantees the existence of a unique solution to the equation

$$U(u(x), x) = 0$$

in the vicinity of x_0 which satisfies $u(x) = g(x)$ for $x \in \Gamma$.

Next we check that the function $u(x)$ defined implicitely by the equation $U(u(x), x) = 0$ solves our original PDE.

Certainly, in a vicinity of x_0

$$Du = -\frac{DU}{U_{x^0}}$$

where DU denotes the part of ∇U coming form the x variables. Thus,

$$v \cdot Du = -\frac{v \cdot DU}{U_{x^0}} = f$$

using (3).

Note that in this picture, solving the quasi linear equation looks quite simple. The vector field V has n independent first integrals, call them $I^1(y), \ldots, I^n(y)$, recalling that $y = (x^0, x)$. Do not forget that these integrals are functions of $n+1$ variables. Since the surface Γ is non characteristic we can, at least in a small neighborhood of x_0, choose these integrals as coordinates for γ. In other words, every point $y = (x^0, x)$ in this neighborhood can be written as $y = f(I)$, i.e., $x^0 = f^0(I)$ and $x = f(I)$ for some functions f^0 and f. Clearly, $f^0(I) - g(f(I))$ is the solution we are looking for since on the surface γ it reduces to $x^0 - g(x)$. The function $u(x)$ is found by solving the equation

$$f^0(I(u(x), x)) - g(f(I(u(x), x)) = 0 \ .$$

In other words, the graph of u is the set of all points where $f^0(I(y)) - g(f(I(y))$ vanishes.

Example 1

$$u_x + uu_y = 0 \ , \ u(0, y) = g(y) \ .$$

Extend this problem to

$$U_x + zU_y = 0 \ , \ U(0, y, z) = z - g(y) \ .$$

The characteristic equations are

$$x' = 1 \ , y' = z \ , z' = 0 \ ,$$

with solutions

$$x = t + x_0 \ , z = z_0 \ , y = z_0 t + y_0 \ .$$

The two independent integrals are

$$z \ , y - zx \ .$$

Thus,

$$U(x, y, z) = f(z, y - zx) \ ,$$

and

$$f(z, y) = z - g(y) \ .$$

Therefore, our solution U is given by

$$U(x, y, z) = z - g(y - zx) \ .$$

Now, $U_z(x,y,z) = 1 + xg'(y - zx)$ and hence for $x = 0$ $U_z(0,y,z) = 1$ and by the implicit function theorem there is always a solution of our original initial value problem, at least locally.

Pick as an example $g(y) = y^2$. Solving the equation $z = (z - zx)^2$ for z yields the solution

$$u(x,y) = \frac{1 + 2xy - \sqrt{1+4xy}}{2x^2}$$

for $x > 0$ and $1 + 4xy > 0$. Of course

$$\lim_{x \to 0} u(x,y) = y^2 \ .$$

In this picture, uniqueness is not so easy to see. After all the process of finding the solution depends on an extension of the problem that is quite arbitrary.

The following point of view makes up for that. Implicitely, in solving (3) the characteristic equation (6) has been used. All the first integrals refer to that equation, which when written in detail is

$$(d/dt)x^0 = f(x^0, x) \ , (d/dt)x = v(x^0, x) \ . \tag{8}$$

Given a solution u of (1). Consider a point on the graph of u, i.e., a point of the form $(u(y), y)$. Solve equation (8) for this initial condition which yields the solution $x^0(t), x(t)$.

The claim is, *that the solution $x^0(t), x(t)$ stays on the graph of u, i.e., $x^0(t) = u(x(t))$.*

To see that define $z(t)$ by solving the differential equation

$$(d/dt)z = v(u(z), z) \ , z(0) = y \ . \tag{9}$$

Define $z^0(t) := u(z(t))$. Certainly, $z^0(0) = u(y)$. Thus $(z^0(t), z(t))$ has the same initial condition as $(x^0(t), x(t))$. Next,

$$(d/dt)z^0 = (d/dt)z \cdot Du(z) = v(u(z), z) \cdot Du(z) \ ,$$

and hence $(z^0(t), z(t))$ satisfies also the same ODE, and hence, by uniqueness of the solution $x^0(t) = z^0(t) = u(z(t))$ and $x(t) = z(t)$ which proves our claim.

The solution $u(x)$ of the initial value problem (1),(2) can be understood in the following way. Start with any initial condition of the form $(g(y), y)$ which is on the graph of the solution, and solve the characteristic equation. The solution curve will be on the the graph of $u(x)$. By considering all initial conditions of the form $(g(x), x)$ the corresponding solution curves trace out a surface that coincides with the graph of u in the vicinity of the point $(g(x_0), x_0)$. For that it is important that the initial condition is noncharacteristic! This procedure produces a unique surface and hence the solution of the initial value problem must be unique at least in a vicinity of x_0.

This procedure, however, does not deliver the solution explicitely. For this purpose the one using first integrals is more useful. Often, however, there is no way that the solution can be explicitely given, and then one has to rely on qualitative reasoning for which the second picture is very well suited. We describe this for the same equation as in Example 1.

Example 2

Consider again the equation

$$u_t + uu_x = 0 \tag{10}$$

with the initial condition

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

The characteristic equations are

$$\dot{t} = 1 \ , \dot{x} = z \ , \dot{z} = 0 \ .$$

Solving them with the initial conditions $x(0) = x_0, t(0) = 0$, and $z(0) = g(x_0)$ yields

$$t(s) = s \ , z(s) = g(x_0) \ , x(s) = g(x_0)s + x_0 \ .$$

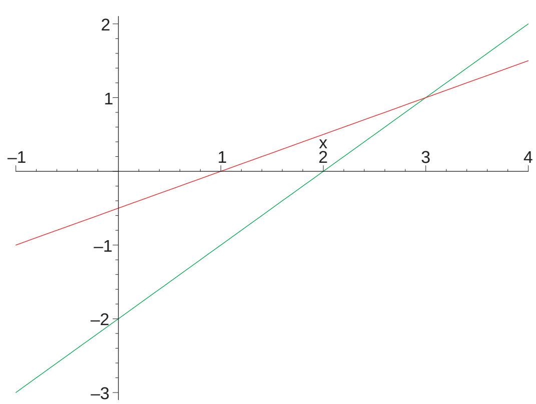
Recall that the geometric curve defined by these equations sits on the graph of the solution $u(t,x)$. Hence we learn that the solution is constant along the line defined by the equation

$$x = g(x_0)t + x_0 \ .$$

From this one can see immediately where the problems arise. Pick an initial condition g so that $g(1) = 2$ and $g(2) = 1$. Obviously the (projected) characteristic lines intersect at the point $x = 3$ and $t = 1$. Hence, the solution cannot exist beyond the time $t = 1$.

\>

\>



If one thinks a bit how the equation (10) has been derived from physics, the loss of existence of solutions does not come as a great surprise. Consider a gas of particles in one dimension. These particles do not interact with each other and hence their individiual momenta and energies are conserved. Now, suppose that the only quantity of interest is the distribution on their velocities as a function of space. Thus, $u(t,x)$ tells the velocity of the particles that sit at x at time t. The initial condition is $g(x)$, the velocitiy distribution at time zero. How does $u(t,x)$ evolve in time? At the point x at time t sits the particle that had its position at $x - tu(t,x)$ since the particles move with constant velocity. At time $t = 0$ the velocity of the particle at the point $x - tu(t,x)$ is given by $g(x - tu(t,x))$ which is the same as the velocity at the point x at time t. Thus,

$$u(t,x) = g(x - tu(t,x)) \ .$$

Now,

$$u_t(t,x) = -g'(x - tu(t,x))[u(t,x) + tu_t(t,x)] \ ,$$

and

$$u_x(t,x) = g'(x - tu(t,x))[1 - tu_x(t,x)] \ .$$

Combining these two equations yields

$$u_t + uu_x = -g'(x - tu)[u + tu_t - u + tuu_x] \ ,$$

or

$$[u_t + uu_x][1 + tg'(x - tu)] = 0 \ ,$$

and hence for t small equation (10) results.

Armed with this physical insight it is now obious what is happening, namely if the initial velocity distribution is decrasing on some interval, then in the course of the motion some of the faster particles will overtake the slower ones and the velocity distribution will 'tilt' over, i.e., it ceases to be a function.

One can take the attitude that this is a bad model and that it should be discarded. However, the assumption that particles are non interacting is for many practical purposes not such a bad assumption. The equation (10) can be viewed as a good approximation in a regime where the interactions can be neglected but has to be modified when the density gets large and hence the interactions become important. This modification leads to the shock wave theory and we pick up that subject at some later moment.