First order PDE: The Methods of Characteristics.

John Andersson, johnan@kth.se

September 29, 2015

1 Briefly about the course.

This course consists of three parts and these notes are only the theoretical aspects of the first part. But since these notes introduce the first part it might be in order to briefly describe the course. The three parts of the course forms a progression.

In **part 1** we study mostly partial differential equations where we have some hope to write down explicit solutions. We will in particular study first order quasilinear equations (quasilinear means that the equation is linear in it's highest order derivative). For these equations one may write down explicit solutions, by using undergraduate calculus, in some easy cases. However, it will soon be clear that one needs abstract theory in order to analyze the equations. Tis is the first baptism of abstract theory in the course.

In **part 2** of the course you will study second order linear equations. There are three types of second order equations that serves as models for most partial differential equations. These are the elliptic equations (represented by the laplace equation), the parabolic equations (represented by the heat equation) and hyperbolic equations (represented by the wave equation). This is the heart of the course and many of the standard theorems for these three equations will be covered in this part of the course. The theory developed will to a large extent be based on representation formulas. This is somewhat disingenuous sinc ewe can only write down these formulas in very simple settings (say for a very nice PDE such as the laplace equation in a very nice domain such as the ball). But the material is standard for any PDE course at masters level and it is a very nice introduction to semi-abstract theory.

In the final part, **part 3**, of the course we will study the obstacle problem. This is meant as an introduction to modern mathematics. The obstacle problem is, in the setting we will study it, a non-linear problem based on the laplacian. For this problem we have no representation formulas and we will therefore be forced to develop an abstract theory. The progression of the course is therefore from partial differential equations of first order that can be approached by first year calculus to the obstacle problem where we are close to modern research and we will have to work with abstract theories. In between you will get the foundation of classical PDE theory.

In this first part of the course we will mimic the greater story of the course. We will start by calculating solutions and then, when calculations fails, move toward an abstract theory. The questions that are most important to us when studying PDE are: *"Does solutions exist?"*, *"If they do, can we calculate or construct them?"*, *"What properties does solutions have? Are they bounded?, Will they be defined everywhere?" et.c.* and *"If we cannot calculate the solutions can we at least estimate them (meaning controlling some norm of the solutions)?"*. In the next few weeks we will encounter all these questions and different answers. In that sense the first part of the course will be a microcosmos of the entire course and PDE theory in general.

2 Introduction to the method of Characteristics.

Consider the following equation:

$$\begin{array}{ll} a(x,y,u(x,y))\frac{\partial u(x,y)}{\partial x} + b(x,y,u(x,y))\frac{\partial u(x,y)}{\partial y} = c(x,y,u(x,y)) & \text{for } (x,y) \in \Omega \\ u(x,y) = h(x,y) & \text{on the curve } \Gamma \subset \overline{\Omega}. \end{array} \tag{1}$$

We assume that $\Omega \subset \mathbb{R}^2$ is a given open set, and that $a(x,y,z), b(x,y,z), c(x,y,z) \in C(\Omega \times \mathbb{R})$ are given continuous functions and $h(x) \in C^1(\Gamma \cap \overline{\Omega})$. The curve Γ will be any differentiable one dimensional curve in $\mathbb{R}^2$, say

$$\Gamma = \{(f(s), g(s));\ s \in [0,1]\}$$

for some functions $f, g \in C^1([0,1])$ such that $|f'(s)| + |g'(s)| \neq 0$.

We are interested in whether there exists a function $u(x,y) \in C^1(\Omega)$ satisfying (1). If such a function exists, is it unique? Can we calculate them?

Without any information about a, b, c and f it is impossible to say much about the equations. But, in order to increase our intuition about the equations, we will try to understand them informally. The equations are of the form

$$a\frac{\partial u(x,y)}{\partial x} + b\frac{\partial u(x,y)}{\partial y} = (a,b) \cdot \nabla u = c. \tag{2}$$

In particular, (2) shows that the partial differential equation determines the value of the derivative of u in the direction (a,b). If (a,b) are known functions, say that a and b only depend on x and y: $a = a(x,y)$ and $b = b(x,y)$ then we should be able to choose a coordinates (s,t) such that curves s =constant are tangential to $(a(x,y), b(x,y))$ at every point.

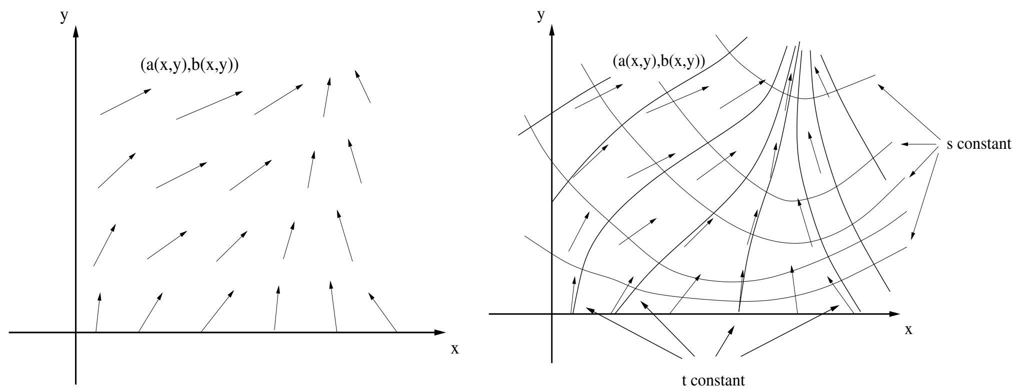


Figure 1: *The vector field (a, b) (left) and new curvilinear coordinates such that the lines t =const. are tangential to (a, b).*

For each (s, t), the crossing of two lines in the left graph of figure 1, there exists a point (x, y). We may write that point $(x(t; s), y(t; s))$.[1] The tangent of the curve $(x(t; s), y(t; s))$ for a fixed s is given by

$$\left(\frac{dx(t;s)}{dt}, \frac{dy(t;s)}{dt}\right). \tag{3}$$

But if the curve $(x(t; s), y(t; s))$, s being fixed, is tangential to $(a(x, y), b(x, y))$ only if the tangent (3) is proportional to $\big(a(x(t; s), y(t; s)), b(x(t; s), y(t; s))\big)$. In particular if

$$\left(\frac{dx(t;s)}{dt}, \frac{dy(t;s)}{dt}\right) = (a(x(t;s), y(t;s))), b(x(t;s), y(t;s)))$$

or equivalently:

$$\frac{dx(t;s)}{dt} = a(x(t;s), y(t;s))) \text{ and } \frac{dy(t;s)}{dt} = b(x(t;s), y(t;s)) \tag{4}$$

then $(a(x, y), b(x, y))$ is a tangent vector field to the curves $(x(t; s), y(t; s))$.

Assuming that we can solve the ordinary differential equations (4) then, according to the chain rule, we may write, for s fixed,

$$\frac{du(x(t;s), y(t;s))}{dt} = \frac{dx(t;s)}{dt}\frac{\partial u(x,y)}{\partial x} + \frac{dy(t;s)}{dt}\frac{\partial u(x,y)}{\partial x} =$$

$$= \left\{ \begin{array}{c} \text{using} \\ (4) \end{array} \right\} = a(x,y)\frac{\partial u(x,y)}{\partial x} + b(x,y)\frac{\partial u(x,y)}{\partial y}.$$

Thus, *if* we can solve (4), then the system (1) can be reduced to

$$\frac{du(x(t;s))}{dt} = c(x(t;s), y(t;s), u(x(t;s), y(t;s))).$$

[1] The notation might seem to be backwards. But later we will think of $(x(t; s), y(t; s))$ as functions of t with s being a parameter.

We have effectively reduced the partial differential equation (1) to a system of ordinary differential equations:

$$\frac{dx(t;s)}{dt} = a(x(t;s), y(t;s))),$$

$$\frac{dy(t;s)}{dt} = b(x(t;s), y(t;s))$$

and

$$\frac{du(x(t;s))}{dt} = c(x(t;s), y(t;s), u(x(t;s), y(t;s))).$$

This is a beautiful idea since it reduces a difficult problem, that is a partial differential equation, to a simpler problem, a system of ordinary differential equations, that we can solve. In the first part of this course we will pursue this idea in order to solve first order partial differential equations.

In the next section we will discuss the idea in further detail as well as pointing out some limits and difficulties with the method. We will then develop some theory for the solution of these equations and try to analyze some of the difficulties that we point out in the next section.

3 The Method of Characteristics.

We are interested in solutions to (1) which we restate here for convenience:

$$\begin{array}{ll} a(x,y,u(x,y))\frac{\partial u(x,y)}{\partial x} + b(x,y,u(x,y))\frac{\partial u(x,y)}{\partial y} = c(x,y,u(x,y)) & \text{for } (x,y) \in \Omega \\ u(x,y) = h(x,y) & \text{on the curve } \Gamma \subset \overline{\Omega}, \end{array} \tag{5}$$

where h is a given function on the curve $\Gamma = \{(f(s), g(s));\ s \in [0,1]\}$.

In this section we will **assume** that we have a solution $u(x, y)$ to the partial differential equation (5). By the analysis in the previous section we know that changing the parametrization of the plane and writing $x = x(t;s)$ and $y = y(t;s)$ we might reduce the PDE (5) into a system of ordinary differential equations. In particular, if we **assume** that we can solve the ordinary differential equations, for each fixed s,

$$\frac{dx(t;s)}{dt} = a(x(t;s), y(t;s)), u(x(t;s), y(t;s))), \tag{6}$$

$$\frac{dy(t;s)}{dt} = b(x(t;s), y(t;s)), u(x(t;s), y(t;s))) \tag{7}$$

then $u(x(t;s), y(t;s))$ will solve (just as in the previous section)

$$\frac{du(x(t;s), y(t;s))}{dt} = \frac{dx(t;s)}{dt}\frac{\partial u(x(t;s), y(t;s))}{\partial x} + \frac{dy(t;s)}{dt}\frac{\partial u(x(t;s), y(t;s))}{\partial y} = \tag{8}$$

$$= a(x, y, u(x, y))\frac{\partial u(x(t; s), y(t; s))}{\partial x} + b(x, y, u(x, y))\frac{\partial u(x, y)}{\partial y} = c(x, y, u(x, y)), \tag{9}$$

where we used the *assumptions* (6) and (7) in the second equality and the *assumption* that u solves the PDE (5) in the final equality. Notice that (8)-(9) together becomes

$$\frac{du(x(t; s), y(t; s))}{dt} = c(x, y, u(x, y)), \tag{10}$$

which again is an ordinary differential equation.

This shows that *if* we have a solution $u(x, y)$ and *if* we can solve (6)-(7) then, $x(t; s)$, $y(t; s)$ and $u(t; s) = u(x(t; s), y(t; s))$ will solve the ordinary differential equations (6), (7) and (10).

The PDE (5) also contains some boundary data

$$u(x, y) = h(x, y) \quad \text{on the curve } \Gamma \subset \overline{\Omega} \tag{11}$$

where h is a given function on the curve $\Gamma = \{(f(s), g(s));\ s \in [0, 1]\}$. Notice that if we impose that

$$\begin{aligned} &x(0; s) = f(s) \\ &y(0; s) = g(s) \text{ and} \\ &u(x(0; s), y(0; s)) = h(s) \end{aligned} \tag{12}$$

then it follows that (11) is satisfied.

To summarize, if we can solve (6)-(7) with the first two boundary conditions in (12) and if $u(x, y)$ is a solution to (5) then $u(x(t; s), y(t; s))$ solves the ODE (10) and $u(x(t; s), y(t; s))$ satisfies the third boundary condition in (12).

This is somewhat backwards (and not at all how mathematics is supposed to be taught!). But it leads us to the following conjecture.

Conjecture 1. *Assume that $x(t; s)$, $y(t; s)$ and $z(t; s)$ solves the following system of ordinary differential equations*

$$\begin{aligned} \tfrac{dx(t;s)}{dt} &= a(x, y, z) \quad & x(0; s) = f(s) \\ \tfrac{dy(t;s)}{dt} &= b(x, y, z) \quad & y(0; s) = g(s) \\ \tfrac{dz(t;s)}{dt} &= c(x, y, z) \quad & z(0; s) = h(s), \end{aligned} \tag{13}$$

for each $s \in [0, 1]$.

Remark: *This conjecture illustrates is a typical trick in mathematics. We want to solve a certain problem that is very difficult. Instead we transform the problem to something different that we can solve and work with that problem instead. Since we can solve (13), as we will see later, we get a problem where we can apply the strong calculational tools of mathematics to gain more understanding of the problem than we could by just using our intuition.*

Notice that we made some daring assumptions, in particular that we already had a solution to the PDE, in order to arrive at the conjecture. The conjecture *itself is however independent of these assumption and will prove itself if it leads to interesting results.*

Then the function $u(x(t;s), y(t;s)) = z(t;s)$ *will solve the partial differential equation (5).*

We will call the equations (13) for the *characteristic equations*. The method of solving the PDE (5) by solving the characteristic equations is called *the method of characteristics*. Moreover, the curves in $\mathbb{R}^3$

$$\{(x(t;s), y(t;s), z(t;s));\ t \in \mathbb{R},\ s \text{ fixed}\}$$

are called the characteristic curves. And the projections of the characteristic curves into $\mathbb{R}^2$

$$\{(x(t;s), y(t;s));\ t \in \mathbb{R},\ s \text{ fixed}\} \subset \mathbb{R}^2$$

are called the *characteristic projections*.

This conjecture gives us a strategy for solving the PDE (5). In particular, as we have remarked before, it is in general much easier to solve an ordinary differential equation than it is to solve a partial differential equation. So if the conjecture turns out to be true[2] then we have reduced a more difficult problem to a simpler problem. Before we start to state theorems and try to prove the conjecture we will make some remarks about it:.

Existence of Solutions to ODEs. The first step in the analysis of a partial differential equation is usually to show that a solution exists. The reduction of the PDE (5) to the ODE (13) would be of little help to do this unless we can actually show that solutions to the ODE exists. So the first order of business will be to show that we can actually find solutions to the ODE (13).

We will prove that solutions the ODE (13) exists, under certain (mild) assumptions on the functions a, b and c, in Theorem 1 in the next section. In the process of proving the existence we will see some interesting geometrical situations where we have no solutions.

However we will, in general, not be able to write down explicit solutions to the equations. This is not strange or unusual in higher mathematics and we will have to get used to not being able to calculate solutions (except in the most trivial circumstances) when we work with PDE.

Invertability of the map $(t;s) \mapsto (x,y)$**.** The second problem we will encounter is that the solution to the ODE is given in the parameters $(t;s)$. In particular, if we solve the ODE (13) and if the conjecture is true, then we will be given the solution $u(x(t;s), y(t;s)) = z(t;s)$ which is a function in $(t;s)$. Therefore we would like to invert the relationship $(t;s) \mapsto (x,y)$ and define $(t;s)$ as functions of (x,y). If that is possible then we may write $u(x,y) = z(t(x,y); s(x,y))$ which is a function of (x,y).

It is however, in general, not possible to find an explicit formula for the inversion of the function $(t;s) \mapsto (x,y)$. At times we will be able to express

[2] As we will see later, the conjecture will be true with certain modifications.

$u(x, y)$ in some implicit form $u = f(x, y, u)$ (see example 1.6 in *"Applied Partial Differential Equations"*). This provides much information regarding u. But naturally it is far from optimal.

We will use the implicit function theorem to show that, in theory, we can invert the relation locally around certain points.

Domain of definition. The next problem is that, in the best case scenario, the solution $u(x, y)$ will only be defined on the set

$$\{(x(t; s), y(t; s));\ s \in [0, 1] \text{ and } t \in \mathbb{R}\}.$$

This set is determined by the solutions $x(t; s)$ and $y(t; s)$. In particular, the domain of definition of $u(x, y)$ is determined by the initial conditions. This means that the initial data doesn't provide enough information for the solution to be defined in the entire set Ω.

This is not a problem with the method. It is rather a problem with our formulation of the problem. We simply can not decide in which set Ω the solution will be defined - the mathematics itself chooses where the solution is defined. The situation is a little bit similar to analytic continuation in complex analysis. In complex analysis an analytic function defined in some disc D may be extend to a certain set (maybe on a Riemann surface) but the set is determined by the values of the function on D.

For examples of this problem[3] see example 1.6 in *"Applied Partial Differential Equations"* or below on the discussion on shocks and rarefacation.

Shocks and Rarefacation. This problem is a little similar to the problem of the domain of definition and it stems from the fact that the characteristic projections $\{(x(t; s), y(t; s));\ t \in \mathbb{R}\}$ are determined by the equations and may cross, which causes shocks, or diverge, which causes rarefacation. It will be easier to see this by means of examples.

EXAMPLE 1 [RAREFARCATIONS]: Consider the PDE, defined for $y > 0$,

$$\begin{aligned} &u(x, y)u_x(x, y) + u_y(x, y) = 0 \quad \text{in } \mathbb{R}^2 \\ &u(x, 0) = \begin{cases} 1 & \text{if } x < 0 \\ 2 & \text{if } x \geq 0 \end{cases} \end{aligned} \tag{14}$$

The characteristic equations are, with the prime indicating the derivative with respect to t: $u' = \frac{\partial x}{\partial t}$,

$$\begin{array}{ll} x'(t; s) = z(t; s) & x(0; s) = s \\ y'(t; s) = 1 & y(0; s) = 0 \\ z'(t; s) = 0 & z(0; s) = \begin{cases} 1 & \text{if } s < 0 \\ 2 & \text{if } s \geq 0 \end{cases} \end{array}$$

[3] Or maybe it isn't a problem but one of these beautiful instances where mathematics gently guides us towards a the right conclusion whether that happens to be the conclusion we wanted or expected?

The third equation implies that $z(t;s)$ is independent of t and thus

$$z(t;s) = \begin{cases} 1 & \text{if } s < 0 \\ 2 & \text{if } s \geq 0 \end{cases}$$

The ordinary differential equation for y implies that $y(t;s) = t$ The equation for x becomes, after substituting our expression of z,

$$x'(t;s) = \begin{cases} 1 & \text{if } s < 0 \\ 2 & \text{if } s \geq 0 \end{cases} \qquad \text{and } x(0;s) = s.$$

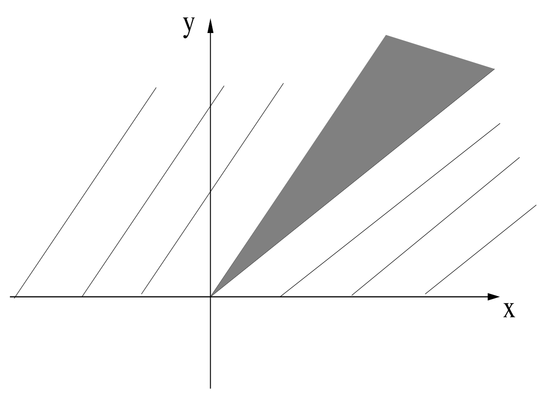
This clearly implies that

$$x(t;s) = \begin{cases} s + t & \text{if } s < 0 \\ s + 2t & \text{if } s \geq 0. \end{cases}$$

The projected characteristics are therefore given by the curves (lines actually)

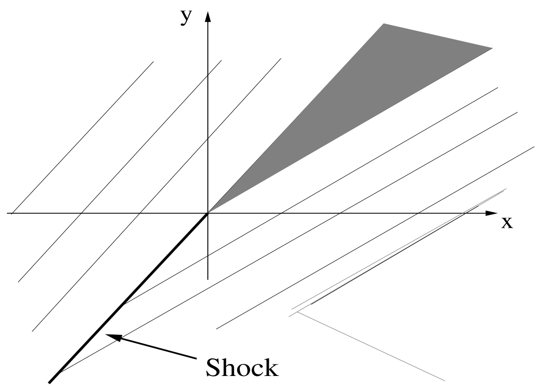
$$PC_s = \left\{ (x,y);\ x = \begin{cases} s + t & \text{if } s < 0 \\ s + 2t & \text{if } s \geq 0. \end{cases},\ y = t,\ t > 0 \right\}.$$

If we plot some of these lines we see that they are given by lines. The lines that have slope 1/2 if they intersect positive $x-$axis and slope 1 if they intersect the negative $x-$axis.



The geometry of the projected characteristics is depicted above. The gray area indicates a region where no projected characteristics enter. This means that the initial data does not specify the solution u in this region. This is logically the same as the domain of definition of the solution is $\mathbb{R}^2 \cap \{y > 0\}$ minus the gray region and this can thus be viewed as a problem with the domain of definition.

However, if we extend the solution to negative y then the projected characteristics will start to cross. In the picture below we can see the geometry of the situation.



We will call the bold line in the picture *a shock*, that is a line where the solution becomes discontinuous. In this case it is easy to see that the solution $u = 2$ to the right of the shock and $u = 1$ to the left of the shock.

The appearance of shocks is not really a problem with the method of characteristics. As a matter of fact we can not define a solution to (14) that is continuous in $\mathbb{R}^2 \cap \{y < 0\}$.

One might suspect that the appearance of shocks and rarefacation is due to the discontinuity of the boundary data. Later we will see that this is locally true under some additional assumptions. By *"locally true"* we mean that we can exclude shocks/rarefacation close to the curve where the boundary data is given.

An extreme case of the problem with the domain of definition. In the extreme case the initial line $(f(s), g(s))$ is a characteristic line. In this case we prescribe $u(f(s), g(s)) = h(s)$ but simultaneously u should solve an ordinary differential equation. This is of course only possible for very special functions $h(s)$. This means that there might not be any solutions.

Blow-ups. The final problem is also not related to the method of characteristics but to the partial differential equation itself. In particular, the solution might *blow-up*. That means that the solution may develop singularities beyond which we cannot define it. This is easiest seen by means of an easy example.

EXAMPLE 2 [BLOW-UPS]: Consider the simple PDE

$$\begin{aligned} &u_y(x, y) = u(x, y)^2 \quad \text{in } \mathbb{R}^2 \\ &u(x, 0) = 1. \end{aligned} \tag{15}$$

This equation is independent of x so we may treat it as a ordinary differential equation directly: $u' = -u^2$ where $u' = \frac{\partial u}{\partial y}$. This ODE is separable and the general solution is $u(x, y) = -\frac{1}{y+c}$. Choosing $c = -1$, that is $u(x, y) = \frac{1}{1-y}$, assures that $u(x, y)$ satisfies the boundary data. This implies that $\lim_{y \to 1^-} u(x, y) = \infty$. The solution blows up at the line $y = 1$ and we may not extend the solution in a continuous way beyond that line.

For another example of blow-up behavior see example 1.7 in *"Applied Partial Differential Equations"*.

4 Some Theory for ODEs.

We have conjectured that solving a first order PDE,

$$\begin{array}{l} a(x,y,u)\frac{\partial u}{\partial x} + b(x,y,u)\frac{\partial u}{\partial y} = c(x,y,u) \\ u(f(s),g(s)) = h(s) \quad \text{ for } s \in (-\alpha,\alpha), \end{array} \tag{16}$$

can be reduced to solving the Characteristic equations:

$$\begin{array}{ll} \frac{dx(t;s)}{dt} = a(x,y,z) & x(0;s) = f(s) \\ \frac{dy(t;s)}{dt} = b(x,y,z) & y(0;s) = g(s) \\ \frac{dz(t;s)}{dt} = c(x,y,z) & z(0;s) = h(s). \end{array} \tag{17}$$

We will see that this conjecture is essentially true but in order to benefit from the reduction of (16) to the system (17) we need to show that we can find solutions to (17).

In the following Theorem we show that there exist a solution to (17) under some mild conditions on a, b and c. In the Theorem we use the, standard, notation $B_K(p_0)$ for the open ball with radius K and center p_0.

Theorem 1. *Let $K > 0$ and $F(t,p) : (-\alpha,\alpha) \times B_K(p_0) \to \mathbb{R}^n$ be a continuous function satisfying the following Lipschitz condition*[4]

$$|F(t,p_1) - F(t,p_2)| \leq L|p_1 - p_2| \tag{18}$$

for all $t \in (-\alpha,\alpha)$ and $p_1, p_2 \in B_K(p_0)$.

Then there exist an $\epsilon > 0$ such that the following ODE:

$$\begin{array}{l} \frac{dp(t)}{dt} = F(t,p(t)) \\ p(0) = p_0 \end{array} \tag{19}$$

has a solution $p(t) \in C^1(-\epsilon,\epsilon)$.

Remark: Notice that p and p_0 are vectors in the theorem and that F is vector valued. If we set, for a fixed s,

$$p(t) = \begin{bmatrix} x(t;s) \\ y(t;s) \\ z(t;s) \end{bmatrix} \text{ and } F(t,p) = \begin{bmatrix} a(x(t;s),y(t;s),z(t;s)) \\ b(x(t;s),y(t;s),z(t;s)) \\ c(x(t;s),y(t;s),z(t;s)) \end{bmatrix}$$

then the solution to (19) is a solution to (17).

Before we prove the Theorem we need to prove a simple but powerful result from functional analysis.

[4]Remember that a function $f(x)$ is called Lipschitz in x with Lipschitz constant L if $|f(x) - f(y)| \leq L|x-y|$ for all x and y in f's domain of definition. The Lipschitz condition is used in order to be able to use the following form of the fundamental Theorem of calculus $\int_a^b f'(x)dx = f(b) - f(a)$ for absolutely continuous, and therefore Lipschitz, functions f. This result is very profound and might not be well known. If you don't know that the fundamental Theorem of calculus applies to Lipschitz functions feel free to assume that F is continuously differentiable in p and $|\nabla_p F(t,p)| \leq L$. This is just a slight relaxation of the Lipschitz condition.

Lemma 1. *Let X be a complete metric space. Also let $T : X \to X$ be a mapping satisfying*

$$\textit{dist}(Tx, Ty) \leq r\,\textit{dist}(x, y) \tag{20}$$

for some $r < 1$ and all $x, y \in X$. Then T has a unique fixed point $x \in X$. That is, there exist a unique $x \in X$ such that $Tx = x$.

Proof: We begin to show uniqueness. Assume that we have two fixed points x and y. Then form (20) we have

$$r\text{dist}(x, y) \geq \text{dist}(Tx, Ty) = \text{dist}(x, y), \tag{21}$$

where we used that, by assumption, $Tx = x$ and $Ty = y$ in the last equality. Since $r < 1$ equation (21) implies that $\text{dist}(x, y) = 0$ which implies that $x = y$.

To show existence we pick any $x_0 \in X$ and define

$$x_k = T^k(x_0),$$

where $T^k(x_0) = T(T(T(...T(x_0))))$ is T applied to x_0 k times. Then by the triangle inequality

$$\text{dist}(x_{k+m}, x_k) \leq \text{dist}(x_{k+m}, x_{k+m-1}) + \text{dist}(x_{k+m-1}, x_{k+m-2}) + \tag{22}$$

$$+ ... + \text{dist}(x_{k+1}, x_k).$$

Notice that by (20) we have

$$\text{dist}(x_l, x_{l-1}) = \text{dist}(T(x_{l-1}), T(x_{l-2})) \leq r\text{dist}(x_{l-1}, x_{l-2}) \leq ... \leq r^l \text{dist}(x_1, x_0)$$

so equation (22) can be estimated

$$\text{dist}(x_{k+m}, x_k) \leq \left(\sum_{j=k}^{k+m-1} r^j \right) \text{dist}(x_1, x_0) \leq \frac{r^k}{1-r} \text{dist}(x_0, x_1).$$

Since $r < 1$ it follows that x_k is a Cauchy sequence and we may use that X is complete to conclude that $\lim_{k\to\infty} x_k = x$ for some x.

We need to show that $Tx = x$. Since $x_k \to x$ there exist an N_ϵ such that $\text{dist}(x, x_k) < \epsilon$ for all $k > N_\epsilon$. Form (20) it follows that

$$\text{dist}(T(x), T(x_k)) < r\epsilon$$

for $k > N_\epsilon$. In particular, by the triangle inequality, we have

$$\text{dist}(Tx, x) \leq \text{dist}(T(x), T(x_k)) + \text{dist}(T(x_k), x) < r\epsilon + \epsilon < 2\epsilon \tag{23}$$

for all $k > N_\epsilon$. We also used that $T(x_k) = x_{k+1}$.

Since ϵ is arbitrary we may, from (23), conclude that $Tx = x$. The Lemma follows. □

We are now ready to prove Theorem 1.

Proof of Theorem 1: We would like to apply Lemma 1. To that end we define T on the space X by

$$T(p)(t) = p_0 + \int_0^t F(s, p(s))ds$$

where

$$X = \left\{ u \in C\big((-\epsilon, \epsilon)\big);\ u(0) = p_0,\ \sup_{t \in (-\epsilon, \epsilon)} |u(t) - p_0| < K \big) \right\},$$

where ϵ is to be chosen later. We also define the distance on X according to

$$\text{dist}(p(t), q(t)) \equiv \|p(t) - q(t)\| \equiv \sup_{t \in (-\epsilon, \epsilon)} |p(t) - q(t)|.$$

In order to use Lemma 1 we need to show

1. that $T : X \to X$. In particular we need to show that $|T(p)(t) - p_0| \leq K$ for all $p \in X$ that $T(p)(0) = p_0$ and that $T(p)(t) \in C(-\epsilon, \epsilon)$. It is easy to verify that $T(p)(0) = p_0$ and that $T(p)(t) \in C(-\epsilon, \epsilon)$ and it is therefore left as an exercise.

2. That $\|T(p)(t) - T(q)(t)\| \leq r\|p(t) - q(t)\|$.

To show 1 we notice that

$$\sup_{\substack{s \in (-\epsilon, \epsilon) \\ \|p(t) - p_0\| \leq K}} \big|F(s, p(s))\big| \leq$$

$$\leq \sup_{\substack{s \in (-\epsilon, \epsilon) \\ \|p(t) - p_0\| \leq K}} \big|F(s, p(s)) - F(s, p_0) + F(s, p_0)\big| \leq$$

$$\leq \sup_{\substack{s \in (-\epsilon, \epsilon) \\ \|p(t) - p_0\| \leq K}} \big(\big|F(s, p_0)\big| + \big|F(s, p(s)) - F(s, p_0)\big|\big) \leq$$

$$\leq \sup_{\substack{s \in (-\epsilon, \epsilon) \\ \|p(t) - p_0\| \leq K}} \big(\big|F(s, p_0)\big| + L|p(s) - p_0|\big) \leq$$

$$\leq \sup_{s \in (-\epsilon, \epsilon)} \big(\big|F(s, p_0)\big|\big) + LK \leq M$$

Where M is a constant depending only on L, K, F and p_0. Most importantly, M is some finite constant.

In particular if $\epsilon < K/M$, $|t| < \epsilon$ and $p(t) \in X$ then

$$\big|T(p(t)) - p_0\big| \leq \left| \int_0^t F(s, p(s))ds \right| \leq \int_0^t |F(s, p(s))|ds \leq |t|M < K.$$

That is $T(p)(t) \in X$ if $p(t) \in X$ and $\epsilon < K/M$.

Next we need to show that 2 holds, that is that T is a contraction. For that we calculate

$$|T(p)(t) - T(q)(t)| = \left| \int_0^t F(s, p(s))ds - \int_0^t F(s, q(s))ds \right|$$

$$\leq \int_0^t |F(s, p(s) - F(s, q(s)))|\, ds$$

$$\leq \int_0^t L|p(s) - q(s)|ds \leq |t|L \sup_{s \in (0,t)} |p(s) - q(s)|,$$

here we also used (18). In particular it follows that

$$\|T(p)(t) - T(q)(t)\| \leq L|t|\|p - q\|.$$

So if $|t| < \epsilon \leq \frac{1}{2L}$ then we have

$$\|T(p)(t) - T(q)(t)\| < \frac{1}{2}\|p - q\|,$$

which is the same as 2 with $r = \frac{1}{2}$. So if we choose $\epsilon < \inf(1/(2L), K/M)$ may thus use Lemma 1 and conclude that T has a fixed point $p(t) \in X$. That is, there exists a $p(t) \in X$ such that

$$p(t) = p_0 + \int_0^t F(s, p(s))ds. \tag{24}$$

It is a simple exercise, using the fundamental theorem of calculus, to verify that the fixed point $p(t) \in C^1(-\epsilon, \epsilon)$ so we may differentiate (24) and conclude that

$$\frac{dp(t)}{dt} = F(t, p(t)).$$

Moreover, substituting $t = 0$ in (24) we see that

$$p(0) = p_0.$$

It follows that $p(t)$ is a solution to the initial value problem (19). □

The above theorem gives a solution to the ordinary differential equation that is continuously differentiable in t. We are however interested in the first order PDE so we will also need to show that the solution is continuously differentiable in s. We do that in the following theorem.

Theorem 2. *Let where F satisfy the condition in Theorem 1 and assume that $p(t; s)$ be a family (parametrized by $s \in (-\alpha, \alpha)$) of solutions, for $t \in (-\epsilon, \epsilon)$, to the following ODE*

$$\begin{aligned} \tfrac{dp(t;s)}{dt} &= F(t, p(t; s)) \\ p(0; s) &= p_0(s). \end{aligned} \tag{25}$$

1. *If* $p_0(s) \in C(-\alpha,\alpha)$ *then* $p(t;s) \in C((-\epsilon,\epsilon)\times(-\alpha,\alpha))$.
2. *If* $|p_0(s_0)-p_0(s_1)| \le K|s_0-s_1|$ *then* $|p(t;s_0)-p(t;s_1)| \le Ke^{Lt}|s_0-s_1|$ *with* L *as in Theorem 1.*
3. *If* $F \in C^1$ *and* $p_0(s) \in C^1(-\alpha,\alpha)$ *then* $p(t;s) \in C^1((-\epsilon,\epsilon)\times(-\alpha,\alpha))$.

In order to prove this theorem we need a simple Lemma. The Lemma is the first instance of an a priori estimate (see the inequality (26)). An a priori estimate is an inequality where we estimate th solution to a PDE, in this case we estimate the value of the solution but most often it is the maximal value or an integral of the derivatives that is estimated, by means of the coefficients of the equation and the initial data (and for PDE the geometry of the domain). It is almost impossible to overestimate the importance of a priori estimates in modern PDE theory. A priori estimates are used to show existence of solutions, symmetry properties of solutions, construction of counterexamples et.c. et.c.

Lemma 2. *Assume that* $u(t) = [u^1(t),u^2(t),...,u^n(t)]^T$ *is a solution to the following ordinary differential equation*

$$\begin{aligned} &u'(t) = F(t)u(t)+f(t) \qquad \text{for } t\in\mathbb{R} \\ &|u(0)| = \tau, \end{aligned}$$

where $F(t)$ *is a matrix valued function.*

Assume furthermore that $|F(t)v| \le M|v|$ *for* $t\in\mathbb{R}$ *and any vector* $v \in R^n$ *and* $|f(t)| \le c_f$. *Then*

$$|u(t)| \le e^{M|t|}\left[\frac{1}{M}-\frac{e^{-M|t|}}{M}\right]c_f + \tau e^{M|t|}. \tag{26}$$

Proof: Let us first prove the Lemma when $f=0$, we will also only prove the Lemma for $t>$, for $t<0$ the proof is analogous. To do the proof for $f=0$ separately is not necessary since we will prove the general case later. But the proof is somewhat clearer when $f(t)=0$ so we include it for pedagogical reasons. To that end we consider the function $v(t)=e^{-Mt}u(t)$, then the Lemma states that $\frac{\partial|v(t)|^2}{\partial t} \le 0$. To see this we just differentiate $|v|^2$:

$$\frac{\partial|v|^2}{\partial t} = 2v(t)\cdot v'(t) = 2e^{-2Mt}u(t)\cdot u'(t) - 2Me^{-2Mt}|u(t)|^2 =$$

$$= \left\{\ u' = Fu\ \right\} = 2e^{-2Mt}u(t)\langle F(t),u(t)\rangle - 2Me^{-2Mt}|u(t)|^2 \le 0,$$

since $u(t)\langle F(t),u(t)\rangle \le |u||F(t)u(t)| \le M|u|^2$.

For the general case we notice that if $w_\epsilon = (w_\epsilon^1, w_\epsilon^2, ..., w_\epsilon^n)$, where

$$w_\epsilon^i(t) = e^{Mt}\left[\frac{1}{M}-\frac{e^{-Mt}}{M}\right]\frac{c_f}{\sqrt{n}} + \frac{\tau+\epsilon}{\sqrt{n}}e^{Mt},$$

then

$$\frac{dw_\epsilon^i(t)}{dt} = Mw_\epsilon^i + \frac{c_f}{\sqrt{n}} \text{ and } w_\epsilon^i(0) > \frac{\tau}{\sqrt{n}}. \tag{27}$$

Therefore (27) implies that

$$\frac{d|w_\epsilon(t)|^2}{dt} = 2w_\epsilon(t)\cdot w_\epsilon'(t) = 2\Big(|w_\epsilon(t)|^2 + \underbrace{\sum_{i=1}^{n}\frac{c_f}{\sqrt{n}}w_\epsilon^i(t)}_{=c_f|w_\epsilon(t)|}\Big) >$$

$$> \left\{ \begin{array}{l} \text{if} \\ |w_\epsilon(t)|^2 > |u(t)| \end{array} \right\} > 2\left(M|u(t)|^2 + c_f|u(t)|\right) \geq$$

$$\geq 2\big(u(t)\cdot\langle F(t)\cdot u(t)\rangle + \underbrace{c_f|u(t)|}_{\geq f(t)\cdot u(t)}\big) \geq$$

$$\geq 2u(t)\cdot u'(t) = \frac{d|u(t)|^2}{dt},$$

where we used that $u' = F(t)\cdot u(t) + f(t)$ in the last equality. We have thus shown that if $|w_\epsilon(t)|^2 \geq |u(t)|$ then

$$\frac{d|w_\epsilon(t)|^2}{dt} \geq \frac{d|u(t)|^2}{dt}.$$

By assumption on u and construction of w_ϵ we know that $|w_\epsilon(0)|^2 = \tau + \epsilon > |u(0)|^2$. We may therefore conclude that $|w_\epsilon(t)|^2 \geq |u(t)|^2$ for any $\epsilon > 0$ and any $t > 0$. Sending $\epsilon \to 0$ we conclude that

$$|u(t)| \leq e^{Mt}\left[\frac{1}{M} - \frac{e^{-Mt}}{M}\right]c_f + \tau e^{Mt}.$$

□

Before we prove the Theorem 2 we need to make a clarification about the notation that we use. By ∇_p we will mean the operator

$$\nabla_p = \left(\frac{\partial}{\partial p_1}, \frac{\partial}{\partial p_2}, ..., \frac{\partial}{\partial p_n}\right),$$

in particular if $F(t,p) = [F_1, F_2, ..., F_n]^T$ is the vector valued function in Theorem 2 then

$$\nabla_p F(t,p(t;s)) = \begin{bmatrix} \frac{\partial F_1(t,p(t;s))}{\partial p_1} & \frac{\partial F_1(t,p(t;s))}{\partial p_2} & \cdots & \frac{\partial F_1(t,p(t;s))}{\partial p_n} \\ \frac{\partial F_2(t,p(t;s))}{\partial p_1} & \frac{\partial F_2(t,p(t;s))}{\partial p_2} & \cdots & \frac{\partial F_2(t,p(t;s))}{\partial p_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial F_n(t,p(t;s))}{\partial p_1} & \frac{\partial F_n(t,p(t;s))}{\partial p_2} & \cdots & \frac{\partial F_n(t,p(t;s))}{\partial p_n} \end{bmatrix}$$

Later on we will also use the notation $D_{(t,s)}$ for the total derivative

$$D_{(t,s)} = \left[\frac{d}{dt}, \frac{d}{ds}\right].$$

We are now ready to prove Theorem 2.

Proof of Theorem 2: For the first part of the theorem it is enough to show that $p(t;s)$ is continuous in s since, by assumption, $p(t;s)$ is differentiable in t, and thus continuous in t.

Let $s_0 \in (-\alpha, \alpha)$ be any point and define

$$v(t;s) = p(t;s) - p(t;s_0).$$

In order to show that $p(t;s)$ is continuous in s at s_0 we need to show that $v(t;s) \to 0$ as $s \to s_0$.

Notice that $v(t;s)$ solves the following differential equation, for each fixed s,

$$v'(t;s) = F(t, p(t;s)) - F(t, p(t;s_0)) \tag{28}$$

where we used v' to denote $\frac{dv(t;s)}{dt}$.

Next we, and here is the main point in the proof, notice that

$$v'(t;s) = \int_0^1 \left[\frac{d}{dr} F(t, (1-r)p(t;s_0) + rp(t;s)) \right] dr =$$

$$= \underbrace{\int_0^1 \nabla_p F(t;(1-r)p(t;s_0) + rp(t;s)) dr}_{=G(t;s)} \underbrace{(p(t;s) - p(t;s_0))}_{=v(t;s)}, \tag{29}$$

where we changed the derivative to a partial derivative and applied the chain rule in the last step. We also define $G(t;s)$ according to the formula above. Comparing this to (28) we see that

$$v'(t;s) = G(t;s)v(t;s).$$

Notice that since F is a Lipschitz function it follows that

$$|\nabla_p F(t,p)| \leq L \Rightarrow |G(t,p)| \leq L,$$

where L is the constant in Theorem 1. In particular, from Lemma 2, we see that

$$|v(t;s)| \leq |v(0;s)|e^{Lt}. \tag{30}$$

The inequality (30) shows that the value of $v(t;s)$ is controlled by

$$v(0;s) = p(0;s) - p(0;s_0). \tag{31}$$

But by assumption $p(0,s) = p_0(s) \in C(-\alpha, \alpha)$. This means that, for each $t \in \mathbb{R}$, there exists a $\delta_{\epsilon,t} > 0$ for each $\epsilon > 0$ such that

$$|p(0;s) - p(0;s_0)| < \epsilon e^{-Lt} \text{ for all } s \text{ such that } |s - s_0| < \delta_{\epsilon,t}. \tag{32}$$

This implies that for each t and each $\epsilon > 0$ there exists a $\delta_{\epsilon,t} > 0$ such that

$$|p(t;s) - p(t;s_0)| = \left\{ \text{ def. of } v \right\} = |v(t;s)| \leq \left\{ \text{ eq. (30) } \right\} \leq |v(0;s)|e^{Lt} \leq$$

$$\leq \{ \text{ eq. (31) } \} \leq |p(0;s) - p(0;s_0)|e^{Lt} < \{ \text{ eq. (32) } \} < \epsilon e^{-Lt} e^{Lt} = \epsilon \quad (33)$$

for all s such that $|s - s_0| < \delta_{\epsilon,t}$. This proves that $p(t;s)$ is continuous in s.

In order to show the second part we just notice that we may choose $\delta = |s - s_0|$ and use $|p_0(s_0) - p_0(s_1)| \leq K|s_0 - s_1|$ instead of (32) in the calculation (33) gives the estimate

$$|p(t;s_0) - p(t;s_1)| \leq K e^{Lt} |s_0 - s_1|$$

which is what we desire.

Next we show that if F and $p_0(s)$ are C^1 then $p(t;s)$ is C^1. In order to do that we will show that there exists a function $w(t;s_0)$ such that

$$|p(t;s) - p(t;s_0) - (s - s_0) w p_0'(s_0)(t;s_0)| = o(|s - s_0|)$$

for each t. Then it follows that $\frac{\partial p(t;s)}{\partial s}$ is continuous and thus uniformly continuous on compact sets. We choose $w(t;s_0)$ to be the solution to

$$w'(t;s_0) = \nabla_p F(t, p(t;s_0)) w(t;s_0) \quad \text{and } w(t;s_0) = 1, \quad (34)$$

notice that (34) is a linear ODE and therefore has a solution.

We also need to make a slightly more careful analysis of the function G introduced in (29) before we continue. Since F is continuously differentiable and, by the second statement of the theorem,[5] $|p(t;s) - p(t;s_0)| \leq K|s - s_0| e^{Lt}$ we can estimate

$$G(t;s) = \int_0^1 \nabla_p F(t; (1-r)p(t;s) + rp(t;s_0)) dr =$$

$$= \int_0^1 \left(\nabla_p F(t; p(t;s_0)) + \left(\nabla_p F(t; (1-r)p(t;s) + rp(t;s_0)) - \nabla_p F(t; p(t;s_0)) \right) \right) dr =$$

$$= \underbrace{\int_0^1 \nabla_p F(t; p(t;s_0)) dr}_{= \nabla_p F(t, p(t;s_0))} + \int_0^1 \underbrace{\nabla_p F(t; (1-r)p(t;s) + rp(t;s_0)) - \nabla_p F(t; p(t;s_0))}_{= o(|s-s_0|) \text{since } F \in C^1} dr =$$

$$= \nabla_p F(t, p(t;s_0)) + o(|s - s_0|)$$

where we used that $F \in C^1$ and thus

$$\nabla_p F(t; (1-r)p(t;s) + rp(t;s_0)) - \nabla_p F(t; p(t;s_0)) \leq$$

$$\leq o\left(\sup_{r \in [0,1]} |(1-r)p(t;s) + rp(t;s_0) - p(t;s_0)| \right) =$$

[5]Which is valid since if $p_0'(s)$ is continuous then $p'(s)$ is bounded on compact sets by some constant K. Therefore $|p_0(s) - p_0(s_0)| = |p_0'(s_1)||s - s_0| \leq K|s - s_0|$ by the mean value theorem for derivatives.

$$= o\left(\sup_{r\in[0,1]}(1-r)\,|p(t;s)+p(t;s_0)|\right) \leq o\left(Ke^{Lt}|s-s_0|\right) = o(|s-s_0|),$$

where we used the second statement of the proof and $|1-r| \leq 1$ in the last equality. We have therefore shown that

$$G(t;s) = \nabla_p F(t,p(t;s_0)) + o(|s-s_0|) \tag{35}$$

We continue to write down the differential equation for $v(t;s)-(s-s_0)p_0'(s_0)w(t;s_0)$, where we have used the notation $v(t;s) = p(t;s) - p(t;s_0)$,

$$\frac{\partial\left(v(t;s)-(s-s_0)p_0'(s_0)w(t;s_0)\right)}{\partial t} =$$

$$= G(t;s)v(t;s) + (s-s_0)p_0'(s_0)\nabla_p F(t,p(t;s_0))w(t;s_0) =$$

$$= \nabla_p F(t,p(t;s_0))\left(v(t;s)-(s-s_0)p_0'(s_0)w(t;s_0)\right) + \underbrace{o(|s-s_0|)v(t;s)}_{=o(|s-s_0|)},$$

where we also used (35) in the final step. Also, at $t=0$ we have, using that $w(0,s_0)=1$

$$\left|v(0;s)-(s-s_0)p_0'(s_0)w(0;s_0)\right| = \left|p_0(s)-p_0(s_0)-p_0'(s_0)(s-s_0)\right| = o(|s-s_0|),$$

where we have used that $p_0 \in C^1$ and Taylor's theorem in the last equality.

To summarize, we have shown that $\chi(t;s,s_0) = v(t;s)-(s-s_0)p_0'(s_0)w(t;s_0)$ satisfies

$$\frac{\partial\chi(t;s,s_0)}{\partial t} = \nabla_p F(t,p(t;s_0))\chi(t;s,s_0) + o(|s-s_0|)$$

and

$$\chi(t;s,s_0) = o(|s-s_0|).$$

Thus form Lemma 2 it follows that $\frac{|\chi(t;s,s_0)|}{|s-s_0|} \to 0$ as $|s-s_0| \to 0$. Writing this in terms of $p(t;s)$ we arrive at

$$\frac{|p(t;s)-p(t;s_0)-(s-s_0)p_0'(s_0)w(t;s_0)|}{|s-s_0|} \to 0 \text{ as } s \to s_0$$

which is the same as

$$\frac{\partial p(t;s)}{\partial s} = p_0'(s_0)w(t;s_0).$$

□

5 Existence and Uniqueness of Solutions.

So far we have show that we can find a C^1 solutions to the characteristic equations. As mentioned before there are some difficulties related to this method.

The first difficulty (that we will consider in this section) is that the characteristic lines might intersect and form a shock where the solution is not continuous - not to mention not differentiable. Or alternatively, diverge and create a region where the solution is not defined.

The second difficulty is that the initial line $\{(f(s), g(s)); s \in (-\alpha, \alpha)\}$ may be a characteristic. In this case we might have either infinitely many or no solutions.

The third difficulty is that the solution to the characteristic equations are $(x(t;s), y(t;s), z(t;s))$ where $z(t;s) = u(x(t;s), y(t;s))$. Given this expression for u is not clear what u is at a point (x, y) unless we can invert the map $(t;s) \to (x(t;s), y(t;s))$ and find an expression of $(t;s)$ in terms of (x, y).

In this section we will try to develop a theory to handle these problems. In particular, we will show that none of these bad things happens close to the initial curve if we have good C^1 boundary data $h(s)$ and the coefficients a and b and $f(s)$ and $g(s)$ and $h(s)$ satisfy certain compatibility conditions.

To derive the compatibility conditions we assume that we have a solution $u \in C^1$ to the PDE

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u) \tag{36}$$

$$u(f(s), g(s)) = h(s) \quad \text{for } s \in (-\alpha, \alpha). \tag{37}$$

Furthermore we assume that $f, g, h \in C^1$. We may differentiate (37) with respect to s and conclude that

$$f'(s)u_x(f(s), g(s)) + g'(s)u_y(f(s), g(s)) = h'(s). \tag{38}$$

Writing equations (36) and (38) as a system of equations we get

$$\begin{bmatrix} a\big(f(s), g(s), h(s)\big) & a\big(f(s), g(s), h(s)\big) \\ f'(s) & g'(s) \end{bmatrix} \begin{bmatrix} u_x\big(f(s), g(s), h(s)\big) \\ u_y\big(f(s), g(s), h(s)\big) \end{bmatrix} \tag{39}$$

$$= \begin{bmatrix} c\big(f(s), g(s), h(s)\big) \\ h'(s) \end{bmatrix}.$$

We immediately see that unless we can solve the system (39) we can not solve the PDE (36) with initial data (37).

Remember that we may always find a solution to (39) at a point s_0 if the determinant

$$\det\left(\begin{bmatrix} a\big(f(s), g(s), h(s)\big) & b\big(f(s), g(s), h(s)\big) \\ f'(s) & g'(s) \end{bmatrix}\right) \neq 0.$$

We formulate this insight as a Lemma.

Lemma 3. *A necessary condition in order to solve (36) with initial data (37) is that (39) is solvable at each point* $s \in (-\alpha, \alpha)$*.*

In order to find sufficient conditions, at least in a small neighborhood around a point, we will need the inverse function theorem which we state without proof.

Theorem 3. [THE INVERSE FUNCTION THEOREM.] *Suppose that* $\Psi \in C^1(D;\mathbb{R}^n)$ *for some open set* $D \subset \mathbb{R}^n$ *and that* $\Psi'(x_0)$ *is invertible for some* $x_0 \in D$ *and* $y_0 = \Psi(x_0)$*. Then*

1. *there exist open subsets* $U \subset \mathbb{R}^n$ *and* $V \subset \mathbb{R}^n$ *such that* $x_0 \in U$ *and* $y_0 \in V$ *and* Ψ *is one-to-one on* U *and* $\Psi(U) = V$*. In particular* Ψ *has an inverse* Ψ^{-1} *defined in* V*:* $\Psi^{-1}(\Psi(x)) = x$ *for all* $x \in U$*.*

2. *The inverse* Ψ^{-1} *of* Ψ *is a* $C^1(V;U)$ *function.*

With the inverse function Theorem in place we are able to prove the main existence Theorem for first order PDE

Theorem 4. *Let* $f, g, h \in C^1(-\alpha,\alpha)$ *and* $a, b, c \in C^1(\mathbb{R}^3;\mathbb{R})$ *be given functions and furthermore assume that there exist an* $s_0 \in (-\alpha,\alpha)$ *such that*

$$det\left(\begin{bmatrix} a\big(f(s_0),g(s_0),h(s_0)\big) & a\big(f(s_0),g(s_0),h(s_0)\big) \\ f'(s_0) & g'(s_0) \end{bmatrix}\right) \neq 0. \tag{40}$$

Then there exist an open neighbourhood in $\mathbb{R}^2$ *around* $(f(s_0), g(s_0))$ *where the following initial value problem*

$$\begin{array}{ll} a(x,y,u)\frac{\partial u}{\partial x} + b(x,y,u)\frac{\partial u}{\partial y} = c(x,y,u) & \\ u(f(s),g(s)) = h(s) & \text{for } s \in (-\alpha,\alpha), \end{array}$$

has a unique solution $u(x,y)$*.*

Before we prove the Theorem we remind ourselves of the notation we use (which is terribly confusing). We will denote the partial differentials by ∇:

$$\nabla_{(t,s)} = \left[\frac{\partial}{\partial t}, \frac{\partial}{\partial s}\right].$$

Use D for the total derivative

$$D_{(t,s)} = \left[\frac{d}{dt}, \frac{d}{ds}\right].$$

Proof of Theorem 4: From Theorem 1 we know that there exists a unique a solution to the characteristic equations

$$\begin{array}{ll} \frac{dx(t;s)}{dt} = a(x,y,z) & x(0;s) = f(s) \\ \frac{dy(t;s)}{dt} = b(x,y,z) & x(0;s) = g(s) \\ \frac{dz(t;s)}{dt} = c(x,y,z) & x(0;s) = h(s), \end{array} \tag{41}$$

for $t \in (-\epsilon,\epsilon)$. Moreover the solution $(x(t;s), y(t;s), z(t;s))$ is in C^1 by Theorem 2.

We may therefore define a C^1 mapping $\Psi(t;s) = (x(t;s), y(t;s))^T$. Since (x, y) solves the characteristic equations we have, with the notation $\Psi'(\cdot) = D_{(t;s)}\Psi(t;s)\lfloor_{(\cdot)}$,

$$\Psi'(0, s_0) = \begin{bmatrix} \frac{dx(0;s_0)}{dt} & \frac{dx(0;s_0)}{ds} \\ \frac{dy(0;s_0)}{dt} & \frac{dy(0;s_0)}{ds} \end{bmatrix} = \begin{bmatrix} a(f(s_0), g(s_0), h(s_0)) & f'(s_0) \\ b(f(s_0), g(s_0), h(s_0)) & g'(s_0) \end{bmatrix}$$

where we have used that $x(0; s_0) = f(s_0)$ and $y(0; s_0) = g(s_0)$. From condition (40) we know that

$$\det\big(\Psi'(0; s_0)\big) = \det\left(\begin{bmatrix} a(f(s_0), g(s_0), h(s_0)) & f'(s_0) \\ b(f(s_0), g(s_0), h(s_0)) & g'(s_0) \end{bmatrix}\right) \neq 0.$$

It follows, from the inverse function Theorem, that $\Psi(t; s)$ has a C^1 inverse Ψ^{-1} in some open neighbourhood U of $(0, s_0)$. In particular there exist a representation $(s(x, y), t(x, y)) = \Psi^{-1}(x, y)$ and we may write

$$u(x, y) = z(t(x, y); s(x, y)). \tag{42}$$

This shows that there is a C^1-function $u(x, y)$.

We still need to show that u is a solution to the PDE. To that end we need to calculate the derivatives of the inverse Ψ^{-1}.

If we denote by I the identity matrix then we have, by the chain rule,

$$I = D_{(t,s)}\left(\Psi^{-1}(\Psi(t; s))\right) = \left(\nabla_{(x,y)}\Psi^{-1}(\Psi(t; s))\right) D_{(t,s)}\Psi(t; s).$$

Multiplying both sides (from the left) by the inverse of $D\Psi(t; s)$ which we denote by $(D\Psi(t; s))^{-1}$ we get, after reversing the order of the equality,

$$\nabla_{(x,y)}\Psi^{-1}(\Psi(t; s)) = \left(D_{(t,s)}\Psi(t; s)\right)^{-1} = \frac{1}{a\frac{dy}{ds} - b\frac{dx}{ds}} \begin{bmatrix} \frac{dy}{ds} & -\frac{dx}{ds} \\ -b & a \end{bmatrix}. \tag{43}$$

But we may also calculate

$$\nabla_{(x,y)}\Psi^{-1}(\Psi(t; s)) = \begin{bmatrix} \frac{dt}{dx} & \frac{dt}{dy} \\ \frac{ds}{dx} & \frac{ds}{dy} \end{bmatrix} \Bigg\lfloor_{\Psi(t;s)}. \tag{44}$$

In order to verify that u is a solution we start to calculate

$$a(x, y, u)\frac{\partial u(x, y)}{\partial x} + b(x, y, u)\frac{\partial u(x, y)}{\partial y}$$

$$= \begin{bmatrix} \frac{\partial u(x,y)}{\partial x} & \frac{\partial u(x,y)}{\partial y} \end{bmatrix} \begin{bmatrix} a(x, y, u) \\ b(x, y, u) \end{bmatrix} = \left\{ \begin{array}{l} \text{to be} \\ \text{continued...} \end{array} \right\}. \tag{45}$$

Noticing that, by the chain rule and $u(x, y) = z(t(x, y); s(x, y))$,

$$\frac{\partial u(x, y)}{\partial x} = \frac{\partial z(t(x, y); s(x, y))}{\partial x} = \frac{dt}{dx}\frac{dz}{dt} + \frac{ds}{dx}\frac{dz}{ds}$$

and
$$\frac{\partial u(x,y)}{\partial y} = \frac{\partial z(t(x,y);s(x,y))}{\partial y} = \frac{dt}{dy}\frac{dz}{dt} + \frac{ds}{dy}\frac{dz}{ds}$$
we may continue (45)
$$= \begin{bmatrix} \frac{dz}{dt} & \frac{dz}{ds} \end{bmatrix} \begin{bmatrix} \frac{dt}{dx} & \frac{dt}{dy} \\ \frac{ds}{dx} & \frac{ds}{dy} \end{bmatrix} \begin{bmatrix} a(x,y,u) \\ b(x,y,u) \end{bmatrix} = \tag{46}$$
$$= \begin{bmatrix} \frac{dz}{dt} & \frac{dz}{ds} \end{bmatrix} \nabla_{(x,y)}\Psi^{-1}(\Psi(t;s)) \begin{bmatrix} a(x,y,u) \\ b(x,y,u) \end{bmatrix} \tag{47}$$
where we have used the second representation formula for $D\Psi^{-1}$ (44). We may continue (47) by substituting $\frac{dz}{dt} = c(x,y,z)$ and the use formula (43)
$$= \frac{1}{a\frac{dy}{ds} - b\frac{dx}{ds}} \begin{bmatrix} \frac{dz}{dt} & \frac{dz}{ds} \end{bmatrix} \begin{bmatrix} \frac{dy}{ds} & -\frac{dx}{ds} \\ -b & a \end{bmatrix} \begin{bmatrix} a(x,y,u) \\ b(x,y,u) \end{bmatrix} = c(x,y,u), \tag{48}$$
where we used that $\frac{dz}{dt} = c$ in the last trivial calculation. Putting (45), (47) and (48) together we see that
$$a(x,y,u)\frac{\partial u(x,y)}{\partial x} + b(x,y,u)\frac{\partial u(x,y)}{\partial y} = c(x,y,u),$$
so u is indeed a solution to the PDE. □

The above proof is rather straightforward and we define $u(x,y)$ in (42) and then we verify that $u(x,y)$ is a solution. The verification itself is however very nasty and depends on the calculation of the derivatives of the inverse and several rather nasty formula manipulations. Therefore I thought it necessary to provide a different proof.

Alternetive Proof of Theorem 4: Up to equation (42) the previous proof is rather straightforward. Therefore we will only change the details after that equation. In particular, we will provide another argument for the fact that $u(x,y)$ solves the PDE in some small neighborhood V around the point $(f(s_0), g(s_0)))$. We need to show that $u(x,y)$ is a solution at any $(\hat{x},\hat{y}) \in V$.

We know that $\Psi : U \mapsto V$ has a C^1 inverse Ψ^{-1}. Therefore $(\hat{t},\hat{s}) = \Psi^{-1}(\hat{x},\hat{y})$ is some well defined point in the $(t;s)-$space. Moreover, by definition
$$u(x(\hat{t};\hat{s}), y(\hat{t};\hat{s})) = u(\hat{x},\hat{y}) = z(\hat{t};\hat{s}).$$
If we differentiate $u(x(t;s),y(t;s)) = z(t;s)$ by t we arrive at
$$\frac{du(x(t;s),y(t;s))}{dt} = \frac{dz(t;s)}{dt} = c(x(t;s),y(t;s),z(t;s)) = \tag{49}$$
$$= c(x(t;s),y(t;s),u(x(t;s),y(t;s)))$$

but, by the chain rule,

$$\frac{du(x(t;s),y(t;s))}{dt} = \frac{dx(t;s)}{dt}\frac{\partial u(x(t;s),y(t;s))}{\partial x} + \frac{dy(t;s)}{dt}\frac{\partial u(x(t;s),y(t;s))}{\partial y} =$$

$$= a(x(t;s),y(t;s),z(t;s))\frac{\partial u(x,y)}{\partial x} + b(x(t;s),y(t;s),z(t;s))\frac{\partial u(x,y)}{\partial y} = \quad (50)$$
$$= a(x(t;s),y(t;s),u(x,y))\frac{\partial u(x,y)}{\partial x} + b(x,y,u(x,y))\frac{\partial u(x(t;s),y(t;s))}{\partial y},$$

where we used that $x(t;s)$ and $y(t;s)$ solves the characteristic equations (41) in the second equality. We also remark that $(x,y) = (x(t;s),y(t;s))$ throughout equation (50).

Equations (49) and (50) together implies that

$$a(x(t;s),y(t;s),u(x,y))\frac{\partial u(x,y)}{\partial x} + b(x,y,u(x,y))\frac{\partial u(x(t;s),y(t;s))}{\partial y} = \quad (51)$$
$$= c(x(t;s),y(t;s),u(x(t;s),y(t;s))),$$

where again $(x,y) = (x(t;s),y(t;s))$. Notice that (51) is an equality between functions.[6] We may therefore substitute $(\hat{t};\hat{s})$ for $(t;s)$ throughout (51). But if we immediately write $(\hat{x},\hat{y})$ for $(x(\hat{t};\hat{s}),y(\hat{t};\hat{s}))$ we arrive at

$$a(\hat{x},\hat{y},u(\hat{x},\hat{y}))\frac{\partial u(\hat{x},\hat{y})}{\partial x} + b(\hat{x},\hat{y},u(\hat{x},\hat{y}))\frac{\partial u(\hat{x},\hat{y})}{\partial y} = c(\hat{x},\hat{y},u(\hat{x},\hat{y})). \quad (52)$$

But $(\hat{x},\hat{y}) \in V$ where arbitrary so (52) proves that $u(x,y)$ satisfies the PDE at an arbitrary point in V. This finishes the proof. □

We have therefore shown that we may solve first order PDEs by the method of characteristics, at least locally (that is for small t). To actually write down a solution in an explicit form is usually much harder. This is just the nature of advanced mathematics - that we cannot in general write down explicit solutions. However, the method shows that solutions exists and that they are well behaved for small t. To actually solve the PDE most people would use numerical methods.

In the beginning of this section we listed three difficulties with the method of characteristics. The preceding theorem shows that if (40) is satisfied at a point s_0 then there is a small neighbourhood where the initial line is not a characteristic and that the characteristics do not intersect in this neighbourhood - that is we have no shocks close to C^1 initial data.

In relation to the third difficulty we can only show that we can invert the relation $(x(t;s),y(t;s))$ - to actually calculate $(s(x,y),t(x,y))$ might be very difficult. But the deeper we submerge in the theory of PDE the more we will have to rely on abstract theorems so we might as well get used to it.

[6] All the derivatives are partial.