Chapter 1

Quasilinear first order PDEs

In this section of the course we will discuss quasilinear first order PDE. We will be able to give a fairly complete account of the solvability of such equations because, as we shall see, we will be able to use the theory of ODEs to construct solutions.

1.1 Introduction

For simplicity we'll stick to functions defined on a subset of $\mathbb{R}^2$, but the methods are applicable for functions u defined on higher dimensional sets. The general quasilinear first order PDE in two dimensions has the form:

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u) \tag{1.1}$$

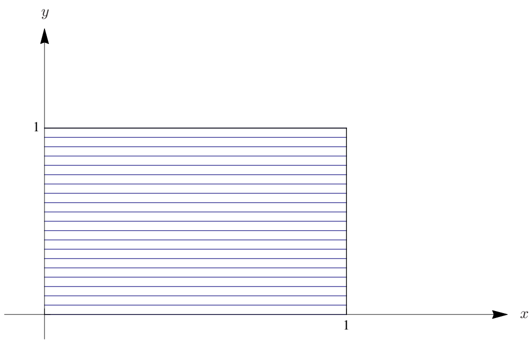

Figure 1.1 The domain $\Omega = [0, 1] \times [0, 1]$

Example. Let $\Omega \subset \mathbb{R}^2$ and consider the very simple quasilinear first order PDE:

$$\frac{\partial u}{\partial x}(x, y) = 0 \quad \text{in } \Omega$$

The equation means that u does not change if we move along a curve of constant y. If $\Omega = [0, 1] \times [0, 1]$, the value of u on $\{0\} \times [0, 1]$ determines u everywhere on Ω. The domain, and the lines of constant y are shown in Figure 1.1. In fact, if $u(0,y) = h(y)$ with h continuous on $[0,1]$ we can write

$$u(x,y) = u(0,y) + \int_0^x \frac{\partial u}{\partial x}(s,y)ds = u(0,y) = h(y).$$

We say that the value of u propagates along lines of constant y. What about other domains Ω? When is specifying the value of u on $\{0\} \times [0,1]$ enough to allow us to find u everywhere? Obviously for this approach to work we must be able to connect every point in Ω to $\{0\} \times [0,1]$ by a horizontal line which lies in Ω. An example where this is possible and where it's not possible is given in Figure 1.2

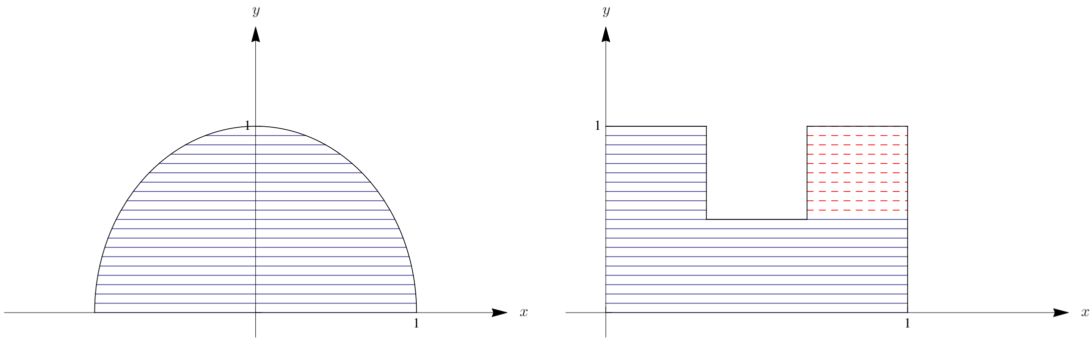

Figure 1.2 *[L] A domain Ω on which $u_x = 0$ is solvable with data on $\{0\} \times [0,1]$. [R] A domain Ω on which $u_x = 0$ is not solvable with data on $\{0\} \times [0,1]$*

Notice that while for a classical solution of a first order equation we would require $u \in C^1(\Omega)$, the procedure above in fact makes sense even if $h(y)$ is only continuous, so that u_y need not be C^0. This is a hint that we will need to consider a notion of solution for some PDE which is weaker than the classical notion.

1.2 The method of characteristics

To deal with the general case, we take the same approach. We identify some curves along which the function u has nice behaviour (in particular it satisfies an ODE) and attempt to propagate u along these curves. This requires specifying the value of u on some initial data curve (in our example above, the segment $\{0\} \times [0,1]$).

Suppose then we are given a curve $\gamma(s) = (f(s), g(s))$, $s \in (\alpha, \beta)$. The problem we wish to solve then is to find a function u satisfying

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

such that for $s \in (\alpha, \beta)$ we have $[u \circ \gamma](s) = u(f(s), g(s)) = h(s)$ for some given function $h : (\alpha, \beta) \to \mathbb{R}$. In other words, we specify the value of u along γ.

The main idea we shall pursue is to find the graph of u over some domain Ω, by showing that we can foliate it by curves along which the value of u is determined by ODEs. Recall that the graph of u over Ω is the surface in $\mathbb{R}^3$ given by

$$\text{Graph}(u) = \{(x, y, u(x,y)) : (x,y) \in \Omega\}$$

We know from the initial conditions that the curve Γ, given by $(f(s), g(s), h(s))$, $s \in (\alpha, \beta)$ must belong to $\text{Graph}(u)$, so we want to write

$$\text{Graph}(u) \supset \bigcup_{s\in(\alpha,\beta)} C_s$$

where

$$C_s = \{((x(t;s), y(t;s), z(t;s)) : t \in (-\epsilon_s, \epsilon_s)\}, \quad \epsilon_s > 0$$

with $x(0;s) = f(s)$, $y(0;s) = g(s)$ and $z(0;s) = g(s)$. In other words, through each point of the initial data curve $(f(s), g(s), h(s))$ we would like to find another curve which lies in the graph of u.

Now, the tangent vector to C_s, given by $\boldsymbol{t} := \left(\frac{dx}{dt}(t;s), \frac{dy}{dt}(t;s), \frac{dz}{dt}(t;s)\right)$ lies in the surface $\text{Graph}(u)$, so it must be orthogonal to any vector which is normal to $\text{Graph}(u)$. The map

$$(x, y) \mapsto \boldsymbol{u}(x, y) := (x, y, u(x, y))$$

is a parameterisation of the surface $\text{Graph}(u)$, so we can easily write down a vector normal to $\text{Graph}(u)$:

$$\begin{aligned} \boldsymbol{N} &= \boldsymbol{u}_x \wedge \boldsymbol{u}_y \\ &= \begin{pmatrix} 1 \\ 0 \\ u_x \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 1 \\ u_y \end{pmatrix} \\ &= \begin{pmatrix} -u_x \\ -u_y \\ 1 \end{pmatrix} \end{aligned}$$

Thus we see that C_s lies in the surface $\text{Graph}(u)$ if and only if $\boldsymbol{t} \cdot \boldsymbol{N} = 0$, i.e.:

$$\frac{dx}{dt}u_x + \frac{dy}{dt}u_y - \frac{dz}{dt} = 0$$

should hold at each point on C_s. On its own, this is not much use to us since we do not know the functions u_x or u_y. However, recall that our PDE is

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u).$$

which tells us that we do know a certain linear combination of u_x and u_y. Comparing the two, we see that if $x(t;s), y(t;s), z(t;s)$ is the unique solution of the system of ODEs:

$$\begin{aligned} \frac{dx}{dt} &= a(x, y, z), \qquad x(0;s) = f(s) \\ \frac{dy}{dt} &= b(x, y, z), \qquad y(0;s) = g(s) \\ \frac{dz}{dt} &= c(x, y, z), \qquad z(0;s) = h(s) \end{aligned} \tag{1.2}$$

then the curve C_s will lie in Graph(u). The Picard–Lindelöf theorem[1] implies that provided a, b, c are sufficiently well behaved, there exists a unique solution to (1.2) on some interval $t \in (-\epsilon_s, \epsilon_s)$, and that moreover the solution depends continuously on s. Later we'll state more precisely some sufficient conditions on a, b, c such that we can solve (1.2).

The equations (1.2) are known as the *characteristic equations* and the corresponding solution curves $(x(t;s), y(t;s), z(t,s))$ are the *characteristic curves* or simply *characteristics.* The union of the characteristics forms a surface:

$$\bigcup_{s\in(\alpha,\beta)} C_s = \{(x(t;s), y(t;s), z(t,s)) : t \in (-\epsilon_s, \epsilon_s), s \in (\alpha, \beta)\}$$

which is a portion of the graph of u that includes the curve Γ. We want to write this as a graph of the form $\{(x, y, u(x, y)) : (x, y) \in \Omega\}$. We can do this, provided we can invert the map $(t, s) \mapsto (x(t; s), y(t; s))$ to give $t(x, y), s(x, y)$. We then finally have $u(x, y) = z(t(x, y); s(x, y))$.

Example 11 (Constant speed transport equation)**.** Let $c \in \mathbb{R}$ and consider the problem

$$cu_x + u_y = 0, \qquad u(x, 0) = h(x). \tag{1.3}$$

Solution. The characteristic equations reduce to:

$$\begin{aligned} \frac{dx}{dt} &= c, & x(0;s) &= s \\ \frac{dy}{dt} &= 1, & y(0;s) &= 0 \\ \frac{dz}{dt} &= 0, & z(0;s) &= h(s) \end{aligned}$$

We can solve these easily to give:

$$x(t;s) = ct + s, \qquad y(t;s) = t, \qquad z(t;s) = h(s)$$

Now the map $(t, s) \mapsto (x, y)$ can be inverted to give

$$t(x, y) = y, \qquad s(x, y) = x - cy$$

So that the solution to (1.3) is simply

$$u(x, y) = z(t(x, y); s(x, y)) = h(x - cy).$$

The curve Γ is given by $(s, 0, h(s))$, $s \in \mathbb{R}$ and Graph(u) is the set

$$\begin{aligned} \text{Graph}(u) &= \{(ct + s, t, h(s)), (t, s) \in \mathbb{R}^2\} \\ &= \{(x, y, h(x - cy)), (x, y) \in \mathbb{R}^2\}. \end{aligned}$$

[1] Also known as Picard's existence theorem or the Cauchy–Lipschitz theorem

Exercise 1.2. Solve the problem

$$u_x + u_y = u, \qquad u(x,0) = \cos x.$$

Exercise 1.3. Solve the problem

$$u_x + xu_y = y, \qquad u(0,y) = \cos y.$$

Our method does not require that the equations are linear. The next example is a quasilinear equation.

Example 12 (Burger's equation)**.** This quasilinear PDE provides a simple model for gas dynamics and traffic flow. The problem we wish to solve is:

$$uu_x + u_y = 0, \qquad u(x,0) = h(x), \qquad x, y \in \mathbb{R},\ y \geq 0$$

Solution. The characteristic equations reduce to:

$$\begin{aligned} \frac{dx}{dt} &= z, & x(0;s) &= s \\ \frac{dy}{dt} &= 1, & y(0;s) &= 0 \\ \frac{dz}{dt} &= 0, & z(0;s) &= h(s) \end{aligned}$$

We can solve the second two equations directly to deduce

$$y(t;s) = t, \qquad z(t;s) = h(s)$$

Putting this solution back into the first equation we have

$$x(t;s) = h(s)t + s.$$

Now, this time we cannot explicitly invert the relation $(t,s) \mapsto (x,y)$. We do know that $t(x,y) = y$ and that $s(x,y) = x - h(s(x,y))y$. From here we have

$$u(x,y) = z(t(x,y); s(x,y)) = h(s(x,y)) = h(x - h(s(x,y)))y) = h(x - u(x,y)y),$$

so that the solution obeys the implicit equation

$$u(x,y) = h(x - u(x,y)y).$$

The solution is a travelling wave, but one whose speed depends on the value of u. One case where we are able to solve the implicit equation is for $h(x) = \sqrt{x^2+1}$. Then we find

$$u = \sqrt{(x-uy)^2+1}, \quad \Longleftrightarrow \quad (y^2-1)u^2 - 2xyu + x^2 + 1 = 0$$

so that

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

Notice that this solution is only valid for $|y| < 1$. At this point, we have that u blows up.

To explore blow-up in more generality, let us return to the characteristics for Burger's equation, which are given by

$$x(t;s) = h(s)t + s, \qquad y(t;s) = t, \qquad z(t;s) = h(s).$$

Notice that at fixed s the characteristic is simply a straight line parallel to the $x - y$ plane:

$$y = \frac{x-s}{h(s)}, \qquad z = h(s).$$

This means that $u(x,y) = z$ remains constant on each of the lines $y = \frac{1}{h(s)}(x - s)$. Put

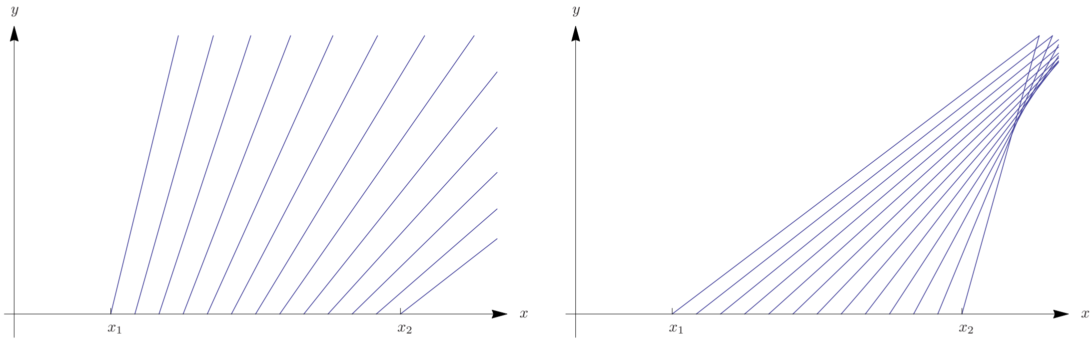

Figure 1.3 *Projected characteristics for Burger's equation with $h(s)$ increasing [L] and $h(s)$ decreasing [R].*

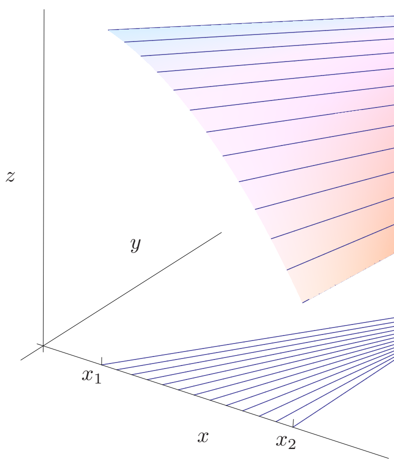

Figure 1.4 The full characteristic surface for the solution with characteristic projections shown in Figure 1.3 [R]

another way, through each point $(x_0, 0)$ on the $x-$axis there is a line $y = \frac{1}{h(x_0)}(x - x_0)$ on which the function u is constant and equal to $\frac{1}{h(x_0)}$. We show two examples of this in Figure 1.3. In the first, we suppose that $h(s)$ is increasing on the interval $[x_1, x_2]$. This means that the lines of constant u are 'fanning out' and do not meet. In the second we assume that $h(s)$ is decreasing on the interval $[x_1, x_2]$. We see then that the lines of constant u must meet, since the line emanating from x_2 will be steeper than the line emanating from x_1. At such a point, our solution is no longer admissible since the two characteristics are trying to assign two different values to u. This discontinuity is called a *shock.* In figure 1.4 we show the full characteristic surface for this data. Notice that the characteristic surface stays smooth at the shock, but it ceases to be a graph over the plane $z = 0$.

To discuss the region on which the solution is valid, we introduce

Definition. Let $(x(t, s), y(t, s), z(t, s))$ be a characteristic curve for the PDE $au_x+bu_y = c$. Then the curve

$$\sigma_s(t) = \{x(t, s), y(t, s)\}$$

is called a *characteristic projection.*

As the name suggests, a characteristic projection is simply the projection of a characteristic onto the $x - y$ plane. The PDE propagates information along characteristic projections.

A shock has formed if two characteristic projections meet at a point and try to assign different values of u at that point. In other words, if there exist $(s, t), (s', t')$ such that $x(s, t) = x(s', t'), y(s, t) = y(s', t')$ and $z(s, t) \neq z(s', t')$. This is another way of saying that the union of the characteristic curves has failed to be a graph over the $z = 0$ plane.

In the case of Burger's equation above, the projected characteristics emanating from $(s_1, 0)$ and $(s_2, 0)$ meet at (X, Y) with

$$Y = \frac{1}{h(s_1)}(X - s_1), \qquad Y = \frac{1}{h(s_2)}(X - s_2)$$

so eliminating X we find

$$Y = \frac{s_2 - s_1}{h(s_1) - h(s_2)}$$

If h is increasing on $[s_1, s_2]$, then the projected characteristics do not meet in the region $y \geq 0$ so no shock forms. If h is decreasing then we find that a shock must form in the region $y \geq 0$.

Exercise 1.4. Consider Burger's equation

$$uu_x + u_y = 0, \qquad u(x, 0) = h(x), \qquad x, y \in \mathbb{R},\ y \geq 0$$

Suppose that $h \in C^1(\mathbb{R})$, h bounded and decreasing, $h' < 0$, and assume that h' is bounded. Show that no shock forms in the region

$$y < \frac{1}{\sup_{\mathbb{R}} |h'|},$$

and that for any $y > \frac{1}{\sup_{\mathbb{R}} |h'|}$ a shock has formed for some x.

1.2.1 Compatibility of the initial condition

Having seen some examples where the method of characteristics allows us to solve the PDE, we'll now try and formalise the circumstances under which we can expect the method to work. One obstruction to the method working is the requirement that the initial condition is *compatible* with the PDE. Recall the initial condition can be written in the form

$$u(f(s), g(s)) = h(s) \tag{1.4}$$

where $\gamma : s \mapsto (f(s), g(s))$ is the curve on which we specify the initial data and h represents the value of u along this curve. We can differentiate this relation as a function of s to obtain

$$u_x|_{(f(s),g(s))} f'(s) + u_y|_{(f(s),g(s))} g'(s) = h'(s) \tag{1.5}$$

on the other hand, we expect the PDE

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u) \tag{1.6}$$

to hold for our solution, so in particular it holds when $x = f(s), y = g(s), u = h(s)$. Inserting these values into (1.6), we find

$$u_x|_{(f(s),g(s))}\, a + u_y|_{(f(s),g(s))}\, b = c \tag{1.7}$$

where $a = a(f(s), g(s), h(s))$, etc. The two equations (1.5), (1.7) comprise a linear system to determine $u_x|_{(f(s),g(s))}$, $u_y|_{(f(s),g(s))}$. Now if $ag' - bf' \neq 0$ then this system admits a *unique solution*. In this case, we can expect that the method of characteristics will work. In this situation we say that the initial conditions are *compatible* with the PDE.

If, on the other hand, $ag' - bf' = 0$, then either the initial conditions are *incompatible* with the PDE, or they are *redundant*.

Exercise(∗). Show that $ag' - bf' = 0$ holds for some s if and only if the curve $\gamma(s) = (f(s), g(s))$ is tangent to the projection of the characteristic through $(f(s), g(s), h(s))$ at s. This means that the characteristics do not propagate information away from γ.

We give three simple examples to show the three possibilities (compatible, incompatible and redundant).

Example (Compatible initial conditions)**.** Consider

$$u_x = 0, \qquad u(0, y) = h(y)$$

Here $\gamma(s) = (0, s)$ is not tangent to any projected characteristic, so we get a unique solution $u(x, y) = h(y)$.

Example (Incompatible initial conditions)**.** Let $h(x)$ be non-constant and consider

$$u_x = 0, \qquad u(x, 0) = h(x)$$

Differentiating in x we find $h'(x) \equiv 0$, a contradiction, so there exists no solution. Here $\gamma(s) = (s, 0)$ is tangent to the projected characteristic through $(s, 0, h(s))$.

Example (Redundant initial conditions)**.** Fix $c \in \mathbb{R}$ and consider

$$u_x = 0, \qquad u(x,0) = c$$

Differentiating in x we find $h'(x) \equiv 0$, which is no contradiction. There exist infinitely many solutions since any u of the form $u(x,y) = w(y)$ with $w(0) = c$ is a solution.

1.2.2 Formalising the method of characteristics

Remember we want to solve the PDE

$$a(x,y,u)u_x + b(x,y,u)u_y = c(x,y,u) \qquad \text{on } \Omega$$

with initial data given along a curve $\gamma : s \mapsto (f(s), g(s)) \in \Omega$ such that $u(f(s), g(s)) = h(s)$. We'll now show that the only obstruction to finding a solution in a neighbourhood of γ is the compatibility of the initial conditions. Recall that we found the characteristics by solving the ODEs

$$\begin{aligned} \frac{dx}{dt} &= a(x,y,z), \qquad x(0;s) = f(s) \\ \frac{dy}{dt} &= b(x,y,z), \qquad y(0;s) = g(s) \\ \frac{dz}{dt} &= c(x,y,z), \qquad z(0;s) = h(s) \end{aligned} \tag{1.8}$$

Provided that a, b, c are functions of class C^1 on $\Omega \times \mathbb{R}$ and f, g, h are of class C^1 from (α, β) into $\mathbb{R}$ we can solve these equations uniquely for each $s \in (\alpha, \beta)$ and for $t \in (-\epsilon_s, \epsilon_s)$ and the solution will be of class C^1. The delicate point now is to invert the map $(t, s) \mapsto (x, y)$. Let's define $\Psi(t,s) = (x(t,s), y(t,s))$. Ψ is of class C^1 in a neighbourhood of γ and

$$D\Psi(0,s) = \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial s} \end{pmatrix}(0,s) = \begin{pmatrix} a & f'(s) \\ b & g'(s) \end{pmatrix}.$$

Now we assume that $a(f(s), g(s), h(s))g'(s) - b(f(s), g(s), h(s)f'(s) \neq 0$ for all $s \in (\alpha, \beta)$, so that $D\Psi(0,s)$ is invertible. By the inverse function theorem Ψ^{-1} exists in a neighbourhood of $f(s), g(s)$. Furthermore, we have

$$\begin{pmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \end{pmatrix}(\Psi(t,s)) = (D\Psi^{-1})(\Psi(t,s)) = (D\Psi(t,s))^{-1} = \frac{1}{a\frac{\partial y}{\partial s} - b\frac{\partial x}{\partial s}} \begin{pmatrix} \frac{\partial y}{\partial s} & -\frac{\partial x}{\partial s} \\ -b & a \end{pmatrix}. \tag{1.9}$$

Now by definition we have $u(x,y) = z(t(x,y), s(x,y))$ so by the chain rule

$$Du((x(t,s), y(t,s))) = \begin{pmatrix} u_x & u_y \end{pmatrix} = \begin{pmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial s} \end{pmatrix} \begin{pmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \end{pmatrix}$$

From here, we immediately see

$$\begin{aligned} au_x + bu_y &= \begin{pmatrix} u_x & u_y \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial s} \end{pmatrix} \begin{pmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial s} \end{pmatrix} (D\Psi^{-1}) \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c & \frac{\partial z}{\partial s} \end{pmatrix} (D\Psi^{-1}) \begin{pmatrix} a \\ b \end{pmatrix}. \end{aligned}$$

From (1.9) we have an expression for $(D\Psi^{-1})$. Inserting this, we calculate

$$\begin{aligned} au_x + bu_y &= \frac{1}{a\frac{\partial y}{\partial s} - b\frac{\partial x}{\partial s}} \begin{pmatrix} c & \frac{\partial z}{\partial s} \end{pmatrix} \begin{pmatrix} \frac{\partial y}{\partial s} & -\frac{\partial x}{\partial s} \\ -b & a \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{1}{a\frac{\partial y}{\partial s} - b\frac{\partial x}{\partial s}} \begin{pmatrix} c & \frac{\partial z}{\partial s} \end{pmatrix} \begin{pmatrix} a\frac{\partial y}{\partial s} - b\frac{\partial x}{\partial s} \\ 0 \end{pmatrix} \\ &= c \end{aligned}$$

so the equation is indeed satisfied. We summarise this section in the following theorem:

Theorem 1. *Let $a(x,y,u)$, $b(x,y,u)$ and $c(x,y,u)$ be functions of class C^1 on $\Omega \times \mathbb{R}$. Suppose that $f(s), g(s), h(s)$ are of class C^1 from (α, β) into $\mathbb{R}$, and that $\gamma(s) := (f(s), g(s)) \in \Omega$ is injective. Suppose that $a(\Gamma(s))g'(s) - b(\Gamma(s))f'(s) \neq 0$ for all $s \in (\alpha, \beta)$ where $\Gamma(s) = (f(s), g(s), h(s))$. Then, given $K \subset (\alpha, \beta)$ a closed (and hence compact) interval, there exists a neighbourhood U of $\gamma(K)$ such that*

$$au_x + bu_y = c$$

has a unique solution u (of class C^1) such that $u(\gamma(s)) = h(s)$.

Remark. Uniqueness follows from the fact that the characteristic through a point (x_0, y_0, z_0) is uniquely determined and moreover lies completely in Graph(u) if (x_0, y_0, z_0) lies in Graph(u). Thus any solution would have to include all of the characteristics through Γ, and hence would have to locally agree with the solution we construct above.

Exercise 1.5. a) Solve the problem

$$u_x + 3x^2 u_y = 1, \qquad u(x,0) = h(x), \quad (x,y) \in \mathbb{R}^2$$

for $h \in C^1(\mathbb{R})$.

b) Show and explain why the problem

$$u_x + 3x^2 u_y = 1, \qquad u(s, s^3) = 1, \quad (x,y) \in \mathbb{R}^2$$

has no solution $u \in C^1(\mathbb{R}^2)$.

c) Show and explain why the problem

$$u_x + 3x^2 u_y = 1, \qquad u(s, s^3) = s - 1, \quad (x, y) \in \mathbb{R}^2$$

has infinitely many solutions $u \in C^1(\mathbb{R}^2)$.

Exercise(∗). Go through the conditions of Theorem 1 and check that you understand where each is used in the proof of the theorem.