Chapter 1

First Order PDEs

In this introductory chapter we first introduce partial differential equations and then consider first order partial differential equations. We shall see that they are simpler than higher order partial differential equations. In contrast to higher order partial differential equations these first order partial differential equations are similar to ordinary differential equations and can be solved by using the theory of ordinary differential equations. After this introductory chapter we shall focus on second order partial differential equations. Before we consider the three main examples of second order differential equations we introduce some general concepts in the next chapter. These general concepts are partially motivated by observations contained in the first chapter.

A partial differential equation is an equation on the partial derivatives of a function depending on at least two variables.

Definition 1.1. *A possibly vector valued equation of the following form*

$$F\left(D^k u(x), D^{k-1} u(x), \ldots, Du(x), u(x), x\right) = 0$$

is called partial differential equation of order k. Here F is a given function and u an unknown function. The expressions $D^k u$ denotes the vector of all partial derivatives of the function u of order k. The function u is called a solution of the differential equation, if u is k times differentiable and obeys the partial differential equation.

On open subsets $\Omega \subset \mathbb{R}^n$ we denote the partial derivatives of higher order by $\partial^\gamma = \prod_i \partial_i^{\gamma_i} = \prod_i (\frac{\partial}{\partial x_i})^{\gamma_i}$ with multi-indices $\gamma \in \mathbb{N}_0^n$ of length $|\gamma| = \sum_i \gamma_i$. The multi-indices are ordered by $\delta \leq \gamma \iff \delta_i \leq \gamma_i$ for $i = 1, \ldots, n$. The partial derivative acts only on the immediately following function; they only act on a product of functions if the product is grouped together in brackets.

1.1   Homogeneous Transport Equation

One of the simplest partial differential equations is the transport equation:

$$\dot{u} + b \cdot \nabla u = 0.$$

Here $\dot{u}$ denotes the partial derivative $\frac{\partial u}{\partial t}$ of the unknown function $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $b \in \mathbb{R}^n$ is a vector, and the product $b \cdot \nabla u$ denotes the scalar product of the vector b with the vector of the first partial derivatives of u with respect to x:

$$b \cdot \nabla u(x,t) = b_1 \frac{\partial u(x,t)}{\partial x_1} + \ldots + b_n \frac{\partial u(x,t)}{\partial x_n}.$$

Let us first assume that $u(x,t)$ is a differentiable solution of the transport equation. For all fixed $(x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ the function

$$z(s) = u(x_0 + s \cdot b, t_0 + s)$$

is a differentiable function on $s \in \mathbb{R}$, whose first derivative vanishes:

$$z'(s) = b \nabla u(x_0 + s \cdot b, t_0 + s) + \dot{u}(x_0 + s \cdot b, t_0 + s) = 0.$$

Therefore u is constant along all parallel straight lines in direction of $(b, 1)$. Furthermore, u is completely determined by the values on all these parallel straight lines.

Initial Value Problem 1.2. *We seek a solution $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ of the transport equation $\dot{u} + b \cdot \nabla u = 0$ with given $b \in \mathbb{R}^n$, which at $t_0 = 0$ is equal to some given function $g : \mathbb{R}^n \to \mathbb{R}$. We call this the Cauchy problem (or initial value problem) for the transport equation.*

With the additional initial data, we can now uniquely determine a solution. We set $t_0 = 0$. All parallel straight lines in direction of $(b, 1)$ intersect $\mathbb{R}^n \times \{0\}$ exactly once. The point (x, t) lies on the line $(x_0 + sb, s)$ with $s = t$ and $x = x_0 + tb$. Thus the value of $u(x, t)$ is determined by the initial condition at $u(x_0, 0) = g(x_0)$ with $x_0 = x - bt$. These lines are in general called characteristic curves. The solution has to be equal to

$$u(x,t) = u(x_0 + tb, t) = u(x_0, 0) = g(x_0) = g(x - tb).$$

If g is differentiable on $\mathbb{R}^n$, then this function indeed solves the transport equation. In this case the initial value problem has a unique solution. Otherwise, if g is not differentiable on $\mathbb{R}^n$, then the initial value problem does not have a solution. As we have seen above, whenever the initial value problem has a solution, then the function $u(x,t) = g(x - tb)$ is the unique solution. So it might be that this candidate is a solution in a more general sense.

1.2   Inhomogeneous Transport Equation

Now we consider the corresponding inhomogeneous transport equation:

$$\dot{u} + b \cdot \nabla u = f.$$

Again $b \in \mathbb{R}^n$ is a given vector, $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is a given function and $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is the unknown function.

Initial Value Problem 1.3. *Given a vector $b \in \mathbb{R}$, a function $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ and an initial value $g : \mathbb{R}^n \to \mathbb{R}$, we seek a solution to the Cauchy problem for the inhomogeneous transport equation: a function $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ that satisfies*

$$\dot{u} + b \cdot \nabla u = f \qquad with \qquad u(x, 0) = g(x).$$

Similar to the homogeneous case, we define for each $(x_0, 0) \in \mathbb{R}^n \times \mathbb{R}$ the function $z(s) = u(x_0 + sb, s)$ which solves

$$z'(s) = b \cdot \nabla u(x_0 + sb, s) + \dot{u}(x_0 + sb, s) = f(x_0 + sb, s).$$

Notice that the right hand side is only a function of s. Moreover $z(0) = u(x_0, 0) = g(x_0)$ is known. Thus we can integrate and determine $z(s)$ completely. This tells us the value of u and any point on the line $(x_0 + sb, s) \in \mathbb{R}^n \times \mathbb{R}$.

We can also gather this information into a formula for u. The point (x, t) lies on the line $(x_0 + sb, s)$ with $s = t$ and $x_0 = x - tb$. Therefore

$$u(x, t) = z(t) = z(0) + \int_0^t z'(s) \ \mathrm{d}s = g(x_0) + \int_0^t f(x_0 + sb, s) \ \mathrm{d}s$$

$$= g(x - tb) + \int_0^t f(x + (s - t)b, s) \ \mathrm{d}s.$$

We observe that this formula is analogous to the formula for solutions of inhomogeneous initial value problems of linear ODEs. The unique solution is the sum of the unique solution of the corresponding homogeneous initial value problem and the integral over solution of the homogeneous equation with the inhomogeneity as initial values. We obtained these solutions of the first order homogeneous and inhomogeneous transport equation by solving an ODE. We shall generalise this method in Section 1.6 and solve more general first order PDEs by solving an appropriate chosen system of first order ODEs.

1.3   Scalar Conservation Laws

In this section we consider the following class of non-linear first order differential equations

$$\dot{u}(x, t) + \frac{\partial f(u(x, t))}{\partial x} = \dot{u}(x, t) + f'(u(x, t)) \cdot \frac{\partial u(x, t)}{\partial x} = 0$$

for a smooth function $f : \mathbb{R} \to \mathbb{R}$. Here $u : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the unknown function. This equation is called a scalar conservation law and is a non-linear first order PDE. For any compact interval $[a, b]$ we calculate

$$\frac{d}{dt} \int_a^b u(x,t)dx = \int_a^b \dot{u}(x,t)dx = -\int_a^b \frac{\partial f(u(x,t))}{\partial x} dx = f(u(a,t)) - f(u(b,t)).$$

This is the meaning of a conservation law: the change of the integral of $u(\cdot, t)$ over $[a, b]$ is equal to the 'flux' of $f(u(x,t))$ through the 'boundary' $\partial[a, b] = \{a, b\}$.

Thinking of t as time, the natural boundary condition to consider is $u(x, 0) = g(x)$ for all $x \in \mathbb{R}$ with some given function $g : \mathbb{R} \to \mathbb{R}$. Let us try to apply the method of characteristics to these equations, namely we assume that there exists a solution u try to understand how the value of u changes along a curve $(x(s), s)$ in its domain. The difference to the transport equation is that we do not assume that the curves are straight lines; it remains to be seen which curves we should choose. Let $z(s) = u(x(s), t(s))$. The derivative is

$$z'(s) = \frac{\partial u(x(s), t(s))}{\partial x} x'(s) + \frac{\partial u(x(s), t(s))}{\partial t} t'(s)$$

Hence if we choose the curve $x(s)$ with the property that $x'(s) = f'(u(x(s), s))$ and $t(s)$ with the property that $t'(s) = 1$ then

$$z'(s) = \frac{\partial u(x(s), s)}{\partial x} f'(u(x(s), s)) + \dot{u}(x(s), s) = 0.$$

This shows that z is constant along these particular curves.

There remain two things to determine: what is the value of z and does there even exist a curve $x(s)$ with the required property? We make the assumption that the characteristic curve begins at the point $(x_0, 0)$. In other words $x(0) = x_0$. By the constancy of z and the initial conditions we have $z(s) = u(x(0), 0) = u(x_0, 0) = g(x_0)$. This answers the first question. The second question is now answerable too: the derivative of $x(s)$ is constant equal to

$$x'(s) = f'(u(x(s), s)) = f'(z(s)) = f'(g(x_0)).$$

The characteristic curve is therefore $x(s) = x_0 + sf'(g(x_0))$. Together this shows that the solution of the PDE is uniquely determine from the initial condition, if it exists.

Instead of thinking about a single characteristic curve and initial point, let us think about all characteristic curves. This point of view implies the solution obeys

$$u(x_0 + tf'(g(x_0)), t) = g(x_0) \quad \text{for all} \quad x_0, t \in \mathbb{R}.$$

The characteristic curves with initial points $x_1, x_2 \in \mathbb{R}$ with $g(x_1) \neq g(x_2)$ might intersect at $t \in \mathbb{R}^+$. In this case the method of characteristic implies $g(x_1) = u(x_1 + tf'(g(x_1)), t) = u(x_2 + tf'(g(x_2)), t) = g(x_2)$, which is impossible. This situation is called crossing characteristics. But otherwise the above implicit equation for u can be solved and defines a solution to the PDE.

Theorem 1.4. *If $f \in C^2(\mathbb{R}, \mathbb{R})$ and $g \in C^1(\mathbb{R}, \mathbb{R})$ with $f''(g(x))g'(x) > -\alpha$ for all $x \in \mathbb{R}$ and some $\alpha \geq 0$, then there is a unique C^1-solution of the initial value problem for the scalar conservation law*

$$\frac{\partial u(x,t)}{\partial t} + f'(u(x,t))\frac{\partial u(x,t)}{\partial x} = 0 \quad with \quad u(x,0) = g(x)$$

on $(x,t) \in \mathbb{R} \times [0, \alpha^{-1})$ for $\alpha > 0$ and on $(x,t) \in \mathbb{R} \times [0, \infty)$ for $\alpha = 0$.

Proof. By the method of characteristics the solution $u(x,t)$ is on the lines $x_0 + tf'(g(x_0))$ equal to $g(x_0)$. For all $t \geq 0$ with $1 - t\alpha > 0$ the derivative of $x_0 \mapsto x_0 + tf'(g(x_0))$ obeys

$$1 + tf''(g(x_0))g'(x_0) \geq 1 - t\alpha > 0.$$

Hence $x_0 + tf'(g(x_0))$ is a strictly increasing function of x_0 and therefore injective. Moreover $\lim_{x_0 \to \pm\infty} x_0 + tf'(g(x_0)) = \pm\infty$, because there is a minimum rate of growth. So $x_0 \mapsto x_0 + tf'(g(x_0))$ is a C^1-diffeomorphism from $\mathbb{R}$ onto $\mathbb{R}$. Therefore there exists for any $x \in \mathbb{R}$ a unique x_0 with $x_0 + tf'(g(x_0)) = x$. Then $u(x,t) = g(x_0)$ solves the initial value problem. $\qquad\square$

Example 1.5. *For $n = 1$ and $f(u) = \frac{1}{2}u^2$ we obtain Burgers equation:*

$$\dot{u}(x,t) + u(x,t)\frac{\partial u(x,t)}{\partial x} = 0.$$

The solutions of the corresponding characteristic equations are $x(t) = x_0 + g(x_0)t$. Therefore the solutions of the corresponding initial value problem obey

$$u(x + tg(x), t) = g(x).$$

If g is continuously differentiable and monotonic increasing, then for all $t \in [0, \infty)$ the map $x \mapsto x + tg(x)$ is a C^1-diffeomorphism from $\mathbb{R}$ onto $\mathbb{R}$ and there is a unique C^1-solution on $\mathbb{R} \times [0, \infty)$. More generally, if $g'(x) > -\alpha$ with $\alpha \geq 0$, then there is a unique C^1-solution on $\mathbb{R} \times [0, \alpha^{-1})$ for $\alpha > 0$ and $(x,t) \in \mathbb{R} \times [0, \infty)$ for $\alpha = 0$.

1.4 Weak Solutions

In the first few sections there were situations with no solutions, or the method of characteristics gave a 'solution' that was not differentiable. In this section we take a scalar conservation law and look for more general notions of solutions which allow us to extend solutions across the crossing characteristics by allowing a limited amount of non-differentiability. But if we don't have differentiability, what does it meant to satisfy a PDE? For this purpose we use the conserved integrals. Since we will restrict ourselves to

the one-dimensional situation for the moment, the natural domains are intervals $\Omega = [a, b]$ with $a < b \in \mathbb{R}$. In this case the conservation law implies

$$\frac{d}{dt} \int_a^b u(x, t)dx = f(u(a, t)) - f(u(b, t)).$$

Now we look for functions u with discontinuities along the graph $\{(x, t) \mid x = y(t)\}$ of a C^1-function y. In the case that $y(t)$ belongs to $[a, b]$, we split the integral over $[a, b]$ into the integrals over $[a, b] = [a, y(t)] \cup [y(t), b]$. In such a case let us calculate the derivative of the integral over $[a, b]$:

$$\frac{d}{dt} \int_a^b u(x, t)dx = \frac{d}{dt} \int_a^{y(t)} u(x, t)dx + \frac{d}{dt} \int_{y(t)}^b u(x, t)dx =$$

$$= \dot{y}(t) \lim_{x \uparrow y(t)} u(x, t) + \int_a^{y(t)} \dot{u}(x, t)dx - \dot{y}(t) \lim_{x \downarrow y(t)} u(x, t) + \int_{y(t)}^b \dot{u}(x, t)dx.$$

We abbreviate $\lim_{x \uparrow y(t)} u(x, t)$ as $u^l(y(t), t)$ and $\lim_{x \downarrow y(t)} u(x, t)$ as $u^r(y(t), t)$ and assume that on both sides of the graph of y the function u is a classical solution of the conservation law:

$$\frac{d}{dt} \int_a^b u(x, t)dx$$

$$= \dot{y}(t)(u^l(y(t), t) - u^r(y(t), t)) - \int_a^{y(t)} \frac{d}{dx} f(u(x, t))dx - \int_{y(t)}^b \frac{d}{dx} f(u(x, t))dx$$

$$= \dot{y}(t)(u^l(y(t), t) - u^r(y(t), t)) + f(u(a, t)) - f(u(b, t)) + f(u^r(y(t), t)) - f(u^l(y(t), t)).$$

Hence the integrated version of the conservation law still holds, if the following Rankine-Hugonoit condition is fulfilled:

$$\dot{y}(t) = \frac{f(u^r(y, t)) - f(u^l(y, t))}{u^r(y, t) - u^l(y, t)}.$$

Example 1.6. *We consider Burgers equation* $\dot{u}(x, t) + u(x, t)\frac{\partial u}{\partial x}(x, t) = 0$ *for* $(x, t) \in \mathbb{R} \times \mathbb{R}^+$ *with the following continuous initial values* $u(x, 0) = g(x)$ *and*

$$g(x) = \begin{cases} 1 & \text{for } x \leq 0, \\ 1 - x & \text{for } 0 \leq x < 1 \\ 0 & \text{for } 1 \leq x. \end{cases}$$

The first crossing of characteristics happens for $t = 1$:

$$x = x_0 + tg(x_0) = \begin{cases} x_0 + t & \text{for } x_0 \leq 0, \\ x_0 + t(1 - x_0) & \text{for } 0 < x_0 < 1, \\ x_0 & \text{for } 1 \leq x_0. \end{cases}$$

For $t < 1$ the evaluation at t is a homeomorphism from $\mathbb{R}$ onto itself with inverse

$$x \mapsto \begin{cases} x - t & \text{for } x \leq t, \\ \frac{x-t}{1-t} & \text{for } t < x < 1, \\ x & \text{for } 1 \leq x. \end{cases}$$

Therefore the solution is for $0 < t < 1$ equal to

$$u(x,t) = \begin{cases} 1 & \text{for } x < t, \\ \frac{x-1}{t-1} & \text{for } t < x < 1, \\ 0 & \text{for } 1 \leq x. \end{cases}$$

At $t = 1$ the solutions of the characteristic equations starting at $x \in [0,1]$ all meet at $x = 1$. For $t > 1$ there exists a unique solution satisfying the Rankine-Hugonoit condition, which is 1 on some interval $(\infty, y(t))$ and 0 on the interval $(y(t), \infty)$. The corresponding regions have to be separated by a path with velocity $\frac{1}{2}$ which starts at $(x,t) = (1,1)$. This gives $y(t) = 1 + \frac{t-1}{2}$. For $t \geq 1$ this solution is equal to

$$u(x,t) = \begin{cases} 1 & \text{for } x < 1 + \frac{t-1}{2}, \\ 0 & \text{for } 1 + \frac{t-1}{2} < x. \end{cases}$$

The second initial value problem is not continuous but monotonic increasing. For continuous monotonic increasing functions g the evaluation at t of the solutions of the characteristic equation would be a homeomorphism for all $t > 0$. Therefore in such cases there exists a unique continuous solution for all $t > 0$. But for non-continuous initial values this is not the case.

Example 1.7. *We again consider Burgers equation $\dot{u}(x,t) + u(x,t)\frac{\partial u}{\partial x}(x,t) = 0$ for $(x,t) \in \mathbb{R} \times \mathbb{R}^+$ with the following non-continuous initial values $u(x,0) = g(x)$ and*

$$g(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1 & \text{for } 0 < x. \end{cases}$$

Again there is a unique discontinuous solution which is 0 on some interval $(-\infty, y(t))$ and 1 on the interval $(y(t), \infty)$. By the Rankine-Hugonoit condition both regions are separated by a path with velocity $\frac{1}{2}$. This solution is equal to

$$u(x,t) = \begin{cases} 0 & \text{for } x < \frac{t}{2}, \\ 1 & \text{for } \frac{1}{2} < x. \end{cases}$$

But there exists another continuous solution, which clearly also satisfies the Rankine-Hugonoit condition:

$$u(x,t) = \begin{cases} 0 & \text{for } x \leq 0, \\ \frac{x}{t} & \text{for } 0 < x < t, \\ 1 & \text{for } t \leq x. \end{cases}$$

These solutions are constant along the lines $x = ct$ for $c \in [0, 1]$. These lines all intersect in the discontinuity at $(x, t) = (0, 0)$. Besides these two extreme cases there exists infinitely many other solutions with several regions of discontinuity, which all satisfy the Rankine-Hugonoit condition.

These examples show that such weak solutions exists for all $t \geq 0$ but are not unique. We now restrict the space of weak solutions such that they have a unique solution for all $t \geq 0$. Since we want to maximise the regularity we only accept discontinuities if there are no continuous solutions. In the last example we prefer the continuous solution. So for Burgers equation this means we only accept discontinuous solutions that take larger values for smaller x and smaller values for larger x.

Definition 1.8 (Lax Entropy condition). *A discontinuity of a weak solution along a C^1-path $t \mapsto y(t)$ satisfies the Lax entropy condition, if along the path the following inequality is fulfilled:*

$$f'(u^l(y, t)) > \dot{y}(t) > f'(u^r(y, t)).$$

A weak solutions with discontinuities along C^1-paths is called an admissible solution, if along the path both the Rankine-Hugonoit condition and the Lax Entropy condition are satisfied.

There is a justification of the entropy condition on physical grounds in Evans' book p. 142-3.

For continuous g there is a crossing of characteristics if $f'(g(x_1)) > f'(g(x_2))$ for $x_1 < x_2$. So this condition ensures that discontinuities can only show up if we cannot avoid a crossing of characteristics.

Theorem 1.9. *Let $f \in C^1(\mathbb{R}, \mathbb{R})$ be convex and u and v two admissible solutions of*

$$\dot{u}(x, t) + f'(u(x, t))\frac{\partial u}{\partial x}(x, t) = 0.$$

in $L^1(\mathbb{R})$. Then $t \mapsto \|u(\cdot, t) - v(\cdot, t)\|_{L^1(\mathbb{R})}$ is monotonically decreasing.

Proof. We divide $\mathbb{R}$ into maximal intervals $I = [a(t), b(t)]$ with the property that either $u(x, t) > v(x, t)$ or $v(x, t) > u(x, t)$ for all $x \in (a(t), b(t))$. This means that either $x \mapsto u - v$ vanishes at the boundary, or is discontinuous and changes sign at the boundary. We claim that the boundaries $a(t)$ and $b(t)$ of these maximal intervals are differentiable. We prove this only for $a(t)$. For $b(t)$ the proof is analogous. If either $u(\cdot, t)$ or $v(\cdot, t)$ is discontinuous at a, then by definition of an admissible solution the locus of the discontinuity a is differentiable with respect to t. On the other hand, suppose u and v are both continuously differentiable at $(a(t_1), t_1)$ with $u(a(t_1), t_1) = v(a(t_1), t_1)$. Then we know that u and v have a common characteristic through this point $s \mapsto (a(t_1) + sf'(u(a(t_1), t_1)), t_1 + s)$,

and moreover they are equal along this characteristic. Hence the line of equality is given by $a(t) = a(t_1) + (t - t_1)f'(u(a(t_1), t_1))$.

To simplify notation we will sometimes write a and b instead of $a(t)$ and $b(t)$. Additionally, we only consider intervals on whose interior $u > v$. On the other intervals these arguments apply with interchanged u and v. Now we calculate

$$\frac{d}{dt} \int_{a(t)}^{b(t)} (u(x,t) - v(x,t))dx$$

$$= \int_{a(t)}^{b(t)} \Big(\dot{u}(x,t) - \dot{v}(x,t)\Big)dx + \dot{b}(t)\Big(u(b,t) - v(b,t)\Big) - \dot{a}(t)\Big(u(a,t) - v(a,t)\Big)$$

$$= \int_{a(t)}^{b(t)} \frac{d}{dx}\Big(f(v(x,t)) - f(u(x,t))\Big)dx + \dot{b}(t)\Big(u(b,t) - v(b,t)\Big) - \dot{a}(t)\Big(u(a,t) - v(a,t)\Big)$$

$$= f(v(b,t)) - f(u(b,t)) + \dot{b}(t)\Big(u(b,t) - v(b,t)\Big)$$

$$+ f(u(a,t)) - f(v(a,t)) + \dot{a}(t)\Big(v(a,t) - u(a,t)\Big).$$

If u and v are both differentiable at (a,t), then they take the same values at (a,t) and the corresponding terms in the last line vanishes. The same holds, if u and v are both differentiable at (b,t). For convex f the derivative f' is monotonically increasing and the Lax-Entropy condition implies at all discontinuities y of $u(\cdot,t)$ and $v(\cdot,t)$

$$u^l(y,t) > u^r(y,t), \qquad\qquad v^l(y,t) > v^r(y,t),$$

respectively. If one of the two solutions u and v is at the boundary of I continuous and the other is non-continuous, then the value of the continuous solution belongs to the closed interval between the limits of the non-continuous solution, because at the boundary either $u - v$ becomes zero or changes sign. For v being continuous and u being discontinuous at a we would have $u^l(a,t) \leq v(a,t) \leq u^r(a,t)$ by $u > v$ on (a,b) in contradiction to the former inequality. So either $u(\cdot,t)$ is continuous and differentiable at a and $v(\cdot,t)$ is discontinuous at $a(t)$ and analogously u is discontinuous at b and v is continuous and differentiable at b. The Rankine Hugonoit condition determines $\dot{a}(t)$ and $\dot{b}(t)$. At $a(t)$ the corresponding contribution to $\frac{d}{dt}\|u(\cdot,t) - v(\cdot,t)\|_1$ is

$$f(u(a,t)) - f(v^r(a,t)) + \dot{a}(t)(v^r(a,t) - u(a,t)) =$$

$$= f(u(a,t)) - f(v^r(a,t)) + \frac{f(v^r(a,t)) - f(v^l(a,t))}{v^r(a,t) - v^l(a,t)}(v^r(a,t) - u(a,t))$$

$$= f(u(a,t)) - \left(f(v^r(a,t))\frac{v^l(a,t) - u(a,t)}{v^l(a,t) - v^r(a,t)} + f(v^l(a,t))\frac{u(a,t) - v^r(a,t)}{v^l(a,t) - v^r(a,t)}\right).$$

Since f is convex the secant lies above the graph of f. Since $u(a,t) \in [v^r(a,t), v^l(a,t)]$

this expression is non-positive. At $b(t)$ this contribution is

$$f(v(b,t)) - f(u^l(b,t)) + \dot{b}(t)\left(u^l(b,t) - v(b,t)\right) =$$
$$= f(v(b,t)) - f(u^l(b,t)) + \frac{f(u^r(b,t)) - f(u^l(b,t))}{u^r(b,t) - u^l(b,t)}\left(u^l(b,t) - v(b,t)\right)$$
$$= f(v(b,t)) - \left(f(u^r(b,t))\frac{u^l(b,t) - v(b,t)}{u^l(b,t) - u^r(b,t)} + f(u^l(b,t))\frac{v(b,t) - u^r(b,t)}{u^l(b,t) - u^r(b,t)}\right).$$

Again due to $v(b,t) \in [u^r(b,t), u^l(b,t)]$ this expression is non-positive.

If finally both solutions are discontinuous at $a(t)$ or $b(t)$. Since $u(\cdot,t) - v(\cdot,t)$ is positive on I, the Lax Entropy condition implies $[u^r(a,t), u^l, (a,t)] \subset [v^r(a,t), v^l(a,t)]$ and $[v^r(b,t), v^l(b,t)] \subset [u^r(b,t), u^l(b,t)]$, respectively. The corresponding contributions to $\frac{d}{dt}\|u(\cdot,t) - v(\cdot,t)\|_1$ are again non-positive:

$$f(u^r(a,t)) - f(v^r(a,t)) + \dot{a}(t)\left(v^r(a,t) - u^r(a,t)\right) =$$
$$= f(u^r(a,t)) - f(v^r(a,t)) + \frac{f(v^r(a,t)) - f(v^l(a,t))}{v^r(a,t) - v^l(a,t)}\left(v^r(a,t) - u^r(a,t)\right)$$
$$= f(u^r(a,t)) - \left(f(v^r(a,t))\frac{v^l(a,t) - u^r(a,t)}{v^l(a,t) - v^r(a,t)} + f(v^l(a,t))\frac{u^r(a,t) - v^r(a,t)}{v^l(a,t) - v^r(a,t)}\right).$$

$$f(v^l(b,t)) - f(u^l(b,t)) + \dot{b}(t)\left(u^l(b,t) - v^l(b,t)\right) =$$
$$= f(v^l(b,t)) - f(u^l(b,t)) + \frac{f(u^r(b,t)) - f(u^l(b,t))}{u^r(b,t) - u^l(b,t)}\left(u^l(b,t) - v^l(b,t)\right)$$
$$= f(v^l(b,t)) - \left(f(u^r(b,t))\frac{u^l(b,t) - v^l(b,t)}{u^l(b,t) - u^r(b,t)} + f(u^l(b,t))\frac{v^l(b,t) - u^r(b,t)}{u^l(b,t) - u^r(b,t)}\right).$$

Hence the contributions to $\frac{d}{dt}\|u(\cdot,t) - v(\cdot,t)\|_1$ of all intervals are non-positive. $\quad\square$

This implies that admissible solutions to an IVP are unique, if they exist. By utilising an explicit formula for admissible solutions one can also prove the existence of admissible solutions. The following theorem is Theorem 10.3 in the lecture notes "Hyperbolic Partial Differential Equations" by Peter Lax, Courant Lecture Notes in Mathematics **14**, American Mathematical Society (2006), which also supplies a proof.

Theorem 1.10. *For $f \in C^2(\mathbb{R}, \mathbb{R})$ is strictly convex and $g \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ there exists an unique admissible solution $u(x,t)$ of*

$$\dot{u}(x,t) + f'(u(x,t))\frac{\partial u}{\partial x}(x,t) = 0 \quad and \quad u(x,0) = g(x) \quad for\ all\ x \in \mathbb{R}.$$

1.5 Noncharacteristic Hypersurfaces

Until now we have only considered specific PDEs where one variable was labelled 'time' and the initial conditions was $t = 0$. In this section we shall consider boundary conditions for the general first order PDE:

$$F(\nabla u(x), u(x), x) = 0$$

on the domain $\Omega \subseteq \mathbb{R}^n$ with the boundary condition $u(y) = g(y)$ for all $y \in \Sigma$. Here u is a real unknown function on an open domain $\Omega \subset \mathbb{R}^n$ and F is a real function on an open subset of $W \subset \mathbb{R}^n \times \mathbb{R} \times \Omega$. For the boundary condition we assume that $\Sigma = \{x \in \Omega \mid \varphi(x) = \varphi(x_0)\}$ is the level-set of the function φ, which we call a hypersurface.

We will first show that locally every Cauchy problem can be brought into the following form:

$$u(y) = g(y) \text{ for all } y \in \Omega \cap H \text{ with } H = \{x \in \mathbb{R}^n \mid x \cdot e_n = x_0 \cdot e_n\}.$$

Here $e_n = (0, \ldots, 0, 1)$ denotes the n-th element of the canonical basis and H the unique hyperplane through $x_0 \in \Omega$ orthogonal to e_n. If $\nabla \varphi(x_0) \neq 0$ we may assume without loss of generality that $\frac{\partial \varphi}{\partial x_n}(x_0) \neq 0$ (relabel the variables if necessary). Then we apply the inverse function theorem to $x \mapsto \Phi(x) = (x_1, \ldots, x_{n-1}, \varphi(x))$ to get a continuously differentiable coordinate transformation $x = \Phi^{-1}(y)$ in a neighbourhood of x_0. This coordinate change has the property that $\varphi(x) = \varphi(x_0)$ if and only if $y \cdot e_n = y_n = \varphi(x_0)$. We say that the boundary has been straighten at x_0. Then by the chain rule the composition $u = v \circ \Phi$ of a function $v : \Omega' \to \mathbb{R}$ with Φ obeys

$$\nabla u(x) = \nabla v(\Phi(x)) \cdot \Phi'(x) = \nabla v(y) \cdot \Phi'\left(\Phi^{-1}(y)\right).$$

Here ∇v and ∇u are row vectors and $\Phi'(x)$ the Jacobi matrix. Hence u solves the PDE

$$F(\nabla u(x), u(x), x) = 0$$

if and only if v solves the PDE

$$G(\nabla v(y), v(y), y) := F\left(\nabla v(y) \cdot \Phi'\left(\Phi^{-1}(y)\right), v(y), \Phi^{-1}(y)\right) = 0.$$

Thus we can indeed assume locally (the coordinate change is only guaranteed to exist in a neighbourhood of x_0) that the boundary is a hyperplane, at the cost of changing the form of the PDE.

Next we ask the question: given the values of u on the hypersurface H is there anything else we can determine about u on the hypersurface? Can we determine the value of its derivatives for example, or can we see immediately that there is no possible u (like for some situations of Burgers' equation)?

We can compute the partial derivatives in most directions at $x_0 \in H$. Observe

$$\frac{\partial u(x_0)}{\partial x_1} = \lim_{h \to 0} \frac{u(x_0 + he_1) - u(x_0)}{h} = \lim_{h \to 0} \frac{g(x_0 + he_1) - g(x_0)}{h} = \frac{\partial g(x_0)}{\partial x_1}.$$

This also works for the directions $x_2, \ldots, x_{n-1}$ which lie in the hyperplane. This idea does not determine $\frac{\partial u(x_0)}{\partial x_n}$, but we have not used the PDE yet. If we substitute all the values we know, there is only one free variable in the PDE:

$$F(\nabla u(x_0), u(x_0), x_0) = F\left(\frac{\partial g(x_0)}{\partial x_1}, \ldots, \frac{\partial g(x_0)}{\partial x_{n-1}}, p_n, g(x_0), x_0\right) = 0.$$

Whether or not this has a solution depends on both the PDE F and the initial condition g. However, if there does exist a solution then there is a simple criterion depending only on F that ensures further that it is solvable in a neighbourhood of x_0.

Definition 1.11. *Consider the PDE as a function of $2n + 1$ variables $F(p, z, x) = 0$ and suppose that there is a solution (p_0, z_0, x_0). The hyperplane $H = \{x_n = x_{0,n}\}$ is called noncharacteristic at x_0 if*

$$\frac{\partial F}{\partial p_n}(p_0, z_0, x_0) \neq 0.$$

To understand the name 'noncharacteristic' let us consider the example

$$\frac{\partial u}{\partial x_1} = 0, \qquad u(x_1, 0) = g(x_1).$$

The PDE in this case is $F(p_1, p_2, z, x_1, x_2) = p_1$, which clearly does not enjoy the non-characteristic property. We see that the initial condition is fighting against the PDE; they are only compatible if g is constant. And even if they happen to be compatible then the initial condition does not determine $\frac{\partial u}{\partial x_2}$ on $H = \{x_2 = 0\}$. If we apply the method of characteristics to this PDE, we must try to find a curve $(x_1(s), x_2(s))$ along which $z(s) = u(x_1(s), x_2(s))$ is nicely behaved. Differentiating z gives

$$z' = \frac{\partial u}{\partial x_1} x_1' + \frac{\partial u}{\partial x_2} x_2',$$

which 'aligns' with the PDE if we choose $x_1' = 1$ and $x_2' = 0$. However this choice of characteristics gives $x_1(s) = x_{0,1} + s, x_s(s) = x_{0,2}$, which lies in the hyperplane. The method fails to be useful because no points in the domain can be reached by characteristics starting on the hyperplane.

Lemma 1.12. *Let $F : W \to \mathbb{R}$ and $g : H \to \mathbb{R}$ be continuously differentiable, $x_0 \in \Omega \cap H$, $z_0 = g(x_0)$ and $p_{0,1} = \frac{\partial g(x_0)}{\partial x_1}, \ldots, p_{0,n-1} = \frac{\partial g(x_0)}{\partial x_{n-1}}$. If there exists $p_{0,n}$ with $F(p_0, z_0, x_0) = 0$ and H is noncharacteristic at x_0 then in an open neighbourhood $\Omega_{x_0} \subset \Omega$ of x_0 there exists for $x \in \Omega_{x_0} \cap H$ a unique solution q of*

$$F(q(x), g(x), x) = 0, \qquad q_i(x) = \frac{\partial g(x)}{\partial x_i} \text{ for } i = 1, \ldots, n-1 \qquad and \qquad q(x_0) = p_0.$$

Proof. Consider the function $(x, q_n) \mapsto F(q_1(x), \ldots, q_{n-1}(x), q_n, g(x), x)$. This takes the value 0 at $(x_0, p_{0,n})$. The noncharacteristic assumption means that we can apply the implicit function theorem to define q_n as a unique function of x in a neighbourhood of x_0. $\qquad\square$

1.6 Method of Characteristics

In this section continue to consider the general first order PDE and try to formalise the method of characteristics, which thus far we have developed only ad hoc. We try to obtain the solution to the PDE by understanding the function u along a curve in the domain. For a clever choice of the curve this reduces to the solution of an appropriate system of first order ODEs. So let $x(s)$ be a curve in the domain of the PDE and $z(s) = u(x(s))$ be the value of u along the curve. The new ingredient is that we must also consider $p(s) = \nabla u(x(s))$, the gradient of u along this curve. But how should be choose the curve $s \mapsto x(s)$? For this purpose we first differentiate

$$p_i'(s) = \frac{d}{ds}\frac{\partial u(x(s))}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial^2 u(x(s))}{\partial x_j \partial x_i} x_j'(s).$$

The total derivative of $F(\nabla u(x), u(x), x) = 0$ with respect to x_i gives

$$0 = \frac{dF(\nabla u(x), u(x), x)}{dx_i} =$$
$$= \sum_{j=1}^{n} \frac{\partial F(\nabla u(x), u(x), x)}{\partial p_j}\frac{\partial^2 u(x)}{\partial x_i \partial x_j} + \frac{\partial F(\nabla u(x), u(x), x)}{\partial z}\frac{\partial u(x)}{\partial x_i} + \frac{\partial F(\nabla u(x), u(x), x)}{\partial x_i}.$$

Due to the commutativity $\partial_i \partial_j u = \partial_j \partial_i u$ of the second partial derivatives we obtain

$$\sum_{j=1}^{n} \frac{\partial F(p(s), z(s), x(s))}{\partial p_j}\frac{\partial^2 u(x(s))}{\partial x_j \partial x_i} = -\frac{\partial F(p(s), z(s), x(s))}{\partial z}p_i(s) - \frac{\partial F(p(s), z(s), x(s))}{\partial x_i}.$$

We want to eliminate the explicit dependence on u from all our equations. If we compare this equation with the derivative of p_i we see that we should choose the vector field for the characteristic curves as

$$x_j'(s) = \frac{\partial F(p(s), z(s), x(s))}{\partial p_j}.$$

This choice allows us to rewrite the equation above for p' as

$$p_i'(s) = \sum_{j=1}^{n} \frac{\partial^2 u(x(s))}{\partial x_j \partial x_i}\frac{\partial F(p(s), z(s), x(s))}{\partial p_j}$$
$$= -\frac{\partial F(p(s), z(s), x(s))}{\partial z}p_i(s) - \frac{\partial F(p(s), z(s), x(s))}{\partial x_i}.$$

Finally we differentiate

$$z'(s) = \frac{d}{ds}u(x(s)) = \sum_{j=1}^{n} \frac{\partial u(x(s))}{\partial x_j} x_j'(s) = \sum_{j=1}^{h} p_j(s) \frac{\partial F(p(s), z(s), x(s))}{\partial p_j}.$$

In this way we indeed obtain the following system of first order ODEs:

$$x_i'(s) = \frac{\partial F(p(s), z(s), x(s))}{\partial p_i}$$

$$p_i'(s) = -\frac{\partial F(p(s), z(s), x(s))}{\partial x_i} - \frac{\partial F(p(s), z(s), x(s))}{\partial z} p_i(s)$$

$$z'(s) = \sum_{j=1}^{n} \frac{\partial F(p(s), z(s), x(s))}{\partial p_j} p_j(s).$$

This is a system of first order ODEs with $2n + 1$ unknown real functions. Importantly this is a 'closed' system; it only depends on these $2n + 1$ functions, not on any other information from u. This is a little surprising, particularly that p', which is effectively a certain second derivative of u, only depends on the location x, the value z, and the first derivatives p. The fact that this idea of characteristics leads to a finite system of ODEs is what makes this an effective method. Let us summarise these calculations in the following theorem:

Theorem 1.13. *Let F be a real differentiable function on an open subset $W \subset \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$ and $u : \Omega \to \mathbb{R}$ a twice differentiable solution on an open subset $\Omega \subset \mathbb{R}^n$ of the first order PDE $F(\nabla u(x), u(x), x) = 0$. For every solution $s \mapsto x(s)$ of the ODE*

$$x_i'(s) = \frac{\partial F}{\partial p_i}(\nabla u(x(s)), u(x(s)), x(s))$$

the functions $p(s) = \nabla u(x(s))$ and $z(s) = u(x(s))$ solve the ODEs

$$p_i'(s) = -\frac{\partial F(p(s), z(s), x(s))}{\partial x_i} - \frac{\partial F(p(s), z(s), x(s))}{\partial z} p_i(s) \ \ and$$

$$z'(s) = \sum_{j=1}^{n} \frac{\partial F(p(s), z(s), x(s))}{\partial p_j} p_j(s). \qquad \square$$

This theorem can be used to address the uniqueness of the solution of PDE, reducing it to the question of uniqueness of the solution of this system of ODEs. This is useful because we have many theorems that tell us when a system of ODEs is unique. For example, the Picard-Lindelöf theorem tells us the solution is uniquely determined by initial conditions if the right hand side is Lipschitz.

We must also pay attention to the logical structure of this theorem. It says *if a solution to the PDE exists* then it solves the ODE; it tells us where to look for potential solutions.

But that was not the task we set for ourselves at the outset of this section. We want to prove that a solution of the PDE does in fact exist. We have seen that global solutions may not exist due to crossing characteristics, so the best we can hope for is a local existence result. This takes a little work but is achieved in the following theorem.

Theorem 1.14. *Let $F : W \to \mathbb{R}$ and $g : H \to \mathbb{R}$ be three times differentiable functions. Suppose we have a point $(p_0, z_0, x_0) \in W$ with*

$$F(p_0, z_0, x_0) = 0, \qquad z_0 = g(x_0), \qquad p_{0,1} = \tfrac{\partial g(x_0)}{\partial x_1}, \ \ldots, \ p_{0,n-1} = \tfrac{\partial g(x_0)}{\partial x_{n-1}}.$$

Furthermore, assume that H is noncharacteristic at x_0. Then in a neighbourhood $\Omega_{x_0} \subset \Omega$ of x_0 there exists a unique solution of the boundary value problem

$$F(\nabla u(x), u(x), x) = 0 \quad for \quad x \in \Omega_{x_0} \quad and \quad u(y) = g(y) \quad for \quad y \in \Omega_{x_0} \cap H.$$

Proof. The strategy of this proof is to solve the system of ODEs given by the method of characteristics and show that it does solve the PDE and the initial conditions. First we need to translate the initial conditions of the PDE to initial conditions for the ODEs. By Lemma 1.12 there exists a solution q on an open neighbourhood of x_0 in H of the following equations

$$F(q(y), g(y), y) = 0, \qquad q_i(y) = \tfrac{\partial g(y)}{\partial x_i} \text{ for } i = 1, \ldots, n-1 \quad and \quad q(x_0) = p_0.$$

If F is twice and g are three times differentiable then the implicit function theorem yields a twice differentiable solution. The Picard-Lindelöf theorem shows that the following initial value problem has for all y in the intersection of an open neighbourhood of x_0 with H a unique solution:

$$x_i'(s) = \frac{\partial F}{\partial p_i}(p(s), z(s), x(s)) \qquad\qquad\qquad\qquad \text{with} \qquad x(0) = y$$

$$p_i'(s) = -\frac{\partial F}{\partial x_i}(p(s), z(s), x(s)) - \frac{\partial F}{\partial z}(p(s), z(s), x(s))p_i(s) \qquad \text{with} \qquad p(0) = q(y)$$

$$z'(s) = \sum_{j=1}^{n} \frac{\partial F}{\partial p_j}(p(s), z(s), x(s))p_j(s) \qquad\qquad\qquad \text{with} \qquad z(0) = g(y).$$

We denote the family of solutions by $(x(y, s), p(y, s), z(y, s))$. For a neighbourhood $\Omega_{x_0} \ni x_0$ there exists an $\epsilon > 0$ such that these solutions are uniquely defined on $(y, s) \in (\Omega \cap H) \times (-\epsilon, \epsilon)$. This is a local proof so let us just write Ω instead of Ω_{x_0}. Since F and g are three times differentiable all coefficients and initial values are twice differentiable. The theorem on the dependence of solutions of ODEs on the initial values gives that $(y, s) \mapsto (x(y, s), p(y, s), z(y, s))$ is on $(\Omega \cap H) \times (-\epsilon, \epsilon)$ twice differentiable.

Now let us examine the characteristic curves in more detail. The function $(y, s) \mapsto x(y, s)$ on $(\Omega \cap H) \times (-\epsilon, \epsilon) \to \mathbb{R}^n$ has at $(y, s) = (x_0, 0)$ the Jacobi matrix

$$
\begin{pmatrix}
1 & 0 & \cdots & 0 & \frac{\partial F(p_0, z_0, x_0)}{\partial p_1} \\
& & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & \frac{\partial F(p_0, z_0, x_0)}{\partial p_{n-1}} \\
0 & 0 & \cdots & 0 & \frac{\partial F(p_0, z_0, x_0)}{\partial p_n}
\end{pmatrix}.
$$

Since $\frac{\partial F(p_0, z_0, x_0)}{\partial p_n} \neq 0$ this matrix is invertible. The inverse function theorem implies that on the (possibly diminished) neighbourhood Ω of x_0 and suitable $\epsilon > 0$ this map is a twice differentiable homeomorphism $(\Omega \cap H) \times (-\epsilon, \epsilon) \to \Omega$ with twice differentiable inverse mapping. Because we know that the inverse mapping exists, the function $u : \Omega \to \mathbb{R}$ defined in implicit form by

$$
u(x(y, s)) = z(y, s) \text{ for all } (y, s) \in (\Omega \cap H) \times (-\epsilon, \epsilon)
$$

is well-defined.

This function u satisfies the initial conditions of the PDE: we have $x(y, 0) = y$ and so

$$
u(y) = u(x(y, 0)) = z(y, 0) = g(y)
$$

for all $y \in \Omega \cap H$. It remains to show that u solves the PDE $F(\nabla u(x), u(x), x) = 0$. Observe that the ODEs imply

$$
\frac{d}{ds} F(p(y, s), z(y, s), x(y, s)) = 0.
$$

Since $F(q(y), g(y), y)$ vanishes for all $y \in \Omega \cap H$ we conclude

$$
F(p(y, s), z(y, s), x(y, s)) = 0 \text{ for all } (y, s) \in (\Omega \cap H) \times (-\epsilon, \epsilon).
$$

Hence to show that u solves the PDE it suffices to show $p(y, s) = \nabla u(x(y, s))$ for all $(y, s) \in (\Omega \cap H) \times (-\epsilon, \epsilon)$.

To this end, we need to establish the following equalities

$$
\frac{\partial z(y, s)}{\partial s} = \sum_{j=1}^{n} p_j(y, s) \frac{\partial x_j(y, s)}{\partial s} \quad \text{and} \quad \frac{\partial z(y, s)}{\partial y_i} = \sum_{j=1}^{n} p_j(y, s) \frac{\partial x_j(y, s)}{\partial y_i}
$$

for all $(y, s) \in (\Omega \cap H) \times (-\epsilon, \epsilon)$ and all $i = 1, \ldots, n-1$. The first equation follows from the ODE for $x(y, s)$ and $z(y, s)$. For $s = 0$ the second equation follows from the initial conditions for $z(y, s)$, $p(y, s)$ and $x(y, s)$. For $s \neq 0$, let us use $v(y, s)$ for the difference between the left and right hand sides of the second equation:

$$
v(y, s) := \frac{\partial z(y, s)}{\partial y_i} - \sum_{j=1}^{n} p_j(y, s) \frac{\partial x_j(y, s)}{\partial y_i}.
$$

We need to show that v is always zero. The derivative of the first equation with respect to y_i yields

$$\frac{\partial^2 z(y,s)}{\partial y_i \partial s} = \sum_{j=1}^{n} \left(\frac{\partial p_j(y,s)}{\partial y_i} \frac{\partial x_j(y,s)}{\partial s} + p_j(y,s) \frac{\partial^2 x_j(y,s)}{\partial y_i \partial s} \right).$$

By the commutativity of the second partial derivatives we obtain

$$\frac{\partial}{\partial s} v(y,s) = \frac{\partial^2 z(y,s)}{\partial s \partial y_i} - \sum_{j=1}^{n} \frac{\partial p_j(y,s)}{\partial s} \frac{\partial x_j(y,s)}{\partial y_i} - \sum_{j=1}^{n} p_j(y,s) \frac{\partial^2 x_j(y,s)}{\partial s \partial y_i}$$

$$= \sum_{j=1}^{n} \left(\frac{\partial p_j(y,s)}{\partial y_i} \frac{\partial x_j(y,s)}{\partial s} - \frac{\partial p_j(y,s)}{\partial s} \frac{\partial x_j(y,s)}{\partial y_i} \right)$$

$$= \sum_{j=1}^{n} \frac{\partial p_j(y,s)}{\partial y_i} \frac{\partial F(p(y,s), z(y,s), x(y,s))}{\partial p_j}$$

$$+ \sum_{j=1}^{n} \left(\frac{\partial F(p(y,s), z(y,s), x(y,s))}{\partial x_j} + \frac{\partial F(p(y,s), z(y,s), x(y,s)) p_j(y,s)}{\partial z} \right) \frac{\partial x_j(y,s)}{\partial y_i}$$

$$= \frac{\partial}{\partial y_i} F(p(y,s), z(y,s), x(y,s))$$

$$- \frac{\partial F(p(y,s), z(y,s), x(y,s))}{\partial z} \left(\frac{\partial z(y,s)}{\partial y_i} - \sum_{j=1}^{n} p_j(y,s) \frac{\partial x_j(y,s)}{\partial y_i} \right).$$

Notice that the bracketed expression is exactly v. Inserting $F(p(y,s), z(y,s), x(y,s)) = 0$ we obtain

$$\frac{\partial}{\partial s} v(y,s) = -\frac{\partial F(p(y,s), z(y,s), x(y,s))}{\partial z} v(y,s).$$

For each y this is a linear homogeneous ODE for $v(y,s)$ in the variable s with initial value 0 at $s = 0$. The unique solution is $v(y,s) \equiv 0$. This implies the second equation for all y and s:

$$\frac{\partial z(y,s)}{\partial y_i} = \sum_{j=1}^{n} p_j(y,s) \frac{\partial x_j(y,s)}{\partial y_i}.$$

Now that we have established the two equalities, we demonstrate that they are not only necessary but also sufficient for the conclusion $p(y,s) = \nabla u(x(y,s))$ for all $(y,s) \in (\Omega \cap H) \times (-\epsilon, \epsilon)$. The solution u is defined as the composition of the inverse of $(y,s) \mapsto x(y,s)$ with $(y,s) \mapsto z(y,s)$. The chain rule implies

$$\frac{\partial u}{\partial x_j} = \frac{\partial z}{\partial s} \frac{\partial s}{\partial x_j} + \sum_{i=1}^{n-1} \frac{\partial z}{\partial y_i} \frac{\partial y_i}{\partial x_j} = \left(\sum_{k=1}^{n} p_k \frac{\partial x_k}{\partial s} \right) \frac{\partial s}{\partial x_j} + \sum_{i=1}^{n-1} \left(\sum_{k=1}^{n} p_k \frac{\partial x_k}{\partial y_i} \right) \frac{\partial y_i}{\partial x_j}$$

$$= \sum_{k=1}^{n} p_k \left(\frac{\partial x_k}{\partial s} \frac{\partial s}{\partial x_j} + \sum_{i=1}^{n-1} \frac{\partial x_k}{\partial y_i} \frac{\partial y_i}{\partial x_j} \right) = \sum_{k=1}^{n} p_k \frac{\partial x_k}{\partial x_j} = p_j.$$

Thus we have shown that the function u, which was constructed from the method of characteristics, solves the PDE.

Theorem 1.13 and the theorem of Picard-Lindelöf imply the uniqueness of the solutions.
$\square$

The relation between the method of characteristics as explained in this section and the ad hoc versions we used in previous sections will be explored now. The important point is they are really the same method, but in many cases the system decouples and the ODEs for x' and z' do not depend on p. This is a nice simplification because it makes solving the p' equations redundant. We solved the Cauchy problem by solving a family of ODEs. In the case of the inhomogeneous transport equation, we combine the coordinates x and t to one coordinate (x,t) and calculate the corresponding function F. Consequently we write

$$F(p, z, (x, t)) = \tilde{F}(p, x, t) = b_1 p_1 + \ldots + b_n p_n + p_{n+1} - f(x, t).$$

We use the equation $F(p, z, (x, t)) = 0$ and rewrite the ODE for z. Then the ODE becomes independent of p and we can solve $x(s)$, $t(s)$ and $z(s)$ separately:

$$x'(s) = b \qquad t'(s) = 1, \qquad z'(s) = \tilde{F}(p(s), x(s), t(s)) + f(x(s), t(s)) = f(x(s), t(s)).$$

Whenever the function F is a first order polynomial with respect to p, then the functions

$$\frac{\partial F(p, z, x)}{\partial p_i} \text{ for } i = 1, \ldots, n, \text{ and } \sum_{j=1}^{n} \frac{\partial F(p, z, x)}{\partial p_j} p_j - F(p, z, x)$$

do not depend on p. Therefore the ODE system for the functions $x(s)$ and $z(s)$ becomes independent of $p(s)$, and the components $x(s)$ and $z(s)$ can be solved independently of $p(s)$. This situation describes the transport equation with vector b depending on z, x and t. For the solution of this equation we do not need to introduce the function $p(s) = \nabla u(x(s))$. Another example is the scalar conservation law in the general form for unknown function $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$:

$$\dot{u}(x, t) + \nabla f(u(x, t)) = \dot{u}(x, t) + f'(u(x, t)) \cdot \nabla u(x, t) = 0$$

with a continuously differentiable function $f : \mathbb{R} \to \mathbb{R}^n$. Again we impose the initial values $u(x, 0) = g(x)$ for all $x \in \mathbb{R}^n$ and some given function $g : \mathbb{R}^n \to \mathbb{R}$. If $x_{n+1} = t$ then the corresponding function F is indeed linear in p:

$$F(p, z, (x, t)) = f'(z) \cdot (p_1, \ldots, p_n) + p_{n+1}.$$

So the corresponding ODE is independent of p

$$x'(s) = f'(z(s)) \qquad t'(s) = 1 \qquad z'(s) = F(p, z(s), (x(s), t(s))) = 0.$$

For any $x \in \mathbb{R}^n$ the unique solution is $x(s) = x + s f'(g(x))$, $t(s) = s$ and $z(s) = g(x)$. So we recover in this more general situation the implicit equation from Section 1.3:

$$u(x + t f'(g(x)), t) = g(x) \quad \text{for all} \quad (x, t) \in \mathbb{R}^n \times \mathbb{R}.$$