Chapter 2

First-order PDEs

In this Chapter, we will investigate what are called **transport equations**. In practice, we will be interested in cases where the motion is directed. Examples of applications of this work can be found in modelling:

- **Pedestrian dynamics** in corridors;
- **Traffic flow** on highways;
- **Water flow** in pipes and hoses;
- **Gas dynamics**;
- **Transport of blood cells** in vessels.

While the underlying physics of each of these systems is quite different, they all involve the motion of individual constituents (e.g. water molecules, cars, pedestrians) along what is effectively a one-dimensional path.

Figure 2.1 Example of applications of transport equations.

We thus aim at developing a modelling strategy which can deal with very different systems. We must then decide on the scale we will use to characterize the motion. In all cases, we will assume that the objects we are describing are numerous enough that following the track of all of them individually is a lost cause and rather use averaged physical quantities. What does this mean in practice?

It means that the equations we will develop will inform us on lengthscale which are large compared to the typical size of a single constituent. We will see that similar equations will be used to model fluid and traffic flow: our equations will be able to describe the flow of fluids at scales above (say) millimeters but will fail to describe nanofluidics; similarly, our equations will inform us about the flow of cars on the scale of highways (kilometers) of but not the scale of individual cars (meters). In most of the applications we will introduce in this Chapter, we will characterize our system by its density $u(x,t)$. In our traffic flow example, the density is the number of cars per unit length.

2.1 Conservation laws

For the sake of simplicity, we will consider in this first chapter one-dimensional problems. Throughout this chapter, we will consider equations of the general form

$$\frac{\partial u}{\partial t} + \frac{\partial q(u)}{\partial x} = 0, \quad x \in \mathbb{R}, t > 0 \tag{2.1}$$

2.1.1 What is a conservation law?

In general, $u(x,t)$ will represent the density or concentration of a physical quantity Q and $q(u)$ its flux function. If for instance you picture the mass concentration of a chemical species in a long and thin water channel (aligned with the x-axis), the dimensions of u are given by

$$[u] = M \cdot L^{-1} \tag{2.2}$$

(It is sometimes called a linear density) and the dimensions of q are

$$[q] = M \cdot T^{-1} \tag{2.3}$$

Equation (2.1) expresses what is a called a scalar conservation law. We can understand this terminology the following way. Consider the interval $[a,b]$,

$$\int_a^b u(x,t)dx \tag{2.4}$$

is the total amount of Q in the interval $[a,b]$ at time t.

A **conservation law** states that: *without sources or sinks, the rate of change of Q in the interior of $[a,b]$ is equal to the net flux through the boundaries of the interval.* If the flux is given by the function $q = q(u)$, we can write this statement mathematically as

$$\frac{d}{dt}\int_a^b u(x,t)dx = -q(u(b,t)) + q(u(a,t)) \tag{2.5}$$

where we have assumed that $q > 0$ (respectively, $q < 0$) if the flux is oriented towards the positive (respectively, negative) direction of the x-axis. If u and q are smooth functions, we can rewrite the previous expression as

$$\int_a^b \left[\frac{\partial u}{\partial t} + \frac{\partial q(u)}{\partial x}\right] dx = 0 \tag{2.6}$$

where we have used the Fundamental Theorem of Calculus. Finally, as the interval we chose is arbitrary, we obtain

$$\frac{\partial u}{\partial t} + \frac{\partial q(u)}{\partial x} = 0 \tag{2.7}$$

Remark. *While here we have used the ideas behind the conservation of mass to derive the general form of a scalar conservation law, note that these concepts are much more general; we will see for instance in the next chapter that the derivation of the heat equation relies on the idea of conservation of energy instead.*

2.1.2 Pollutant in a channel

Equation (2.7) provides a link between the density and the flux but remains very general until we have postulated the form of the flux function. At this point, we need to decide what kind of mass flux we are dealing with. This is called determining a **constitutive relation for** q. There are several possibilities. Here, we will introduce them via an example.

Consider a pollutant on the surface of a narrow channel. A water stream at constant speed v transports the pollutant along the channel (in the positive direction of the x-axis, i.e. $v > 0$). As long as one considers that the pollutant remains at the surface of the water and that the channel is narrow enough, this problem can be considered as one-dimensional.

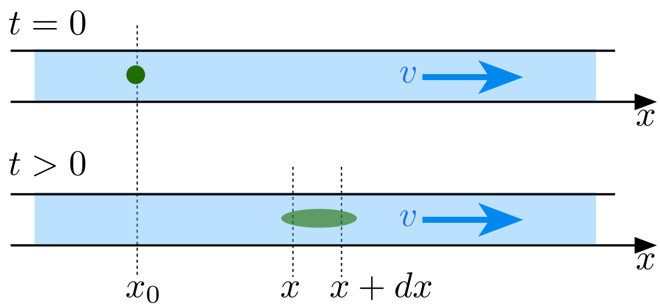


Figure 2.2 Evolution of a pollutant concentration in a narrow channel

The flux function q can have different physical origins. Imagine the time evolution of an initial drop of pollutant (at $t = 0$, see Fig. 2.2), we expect that over time the drop of pollutant will move along the channel and expand in size (i.e. the pollutant gets diluted). This comes from:

- **advection**: the flux is determined by the water stream only, the drop of pollutant moves in bulk, without deformation of expansion. Mathematically, this leads to a flux function
$$q(x,t) = vu(x,t) \tag{2.8}$$
- **diffusion**: the drop of pollutant expands from higher concentration regions to lower concentration regions and gets diluted. Mathematically, the flux function is then proportional to the gradient of density
$$q(x,t) = -D\frac{\partial u}{\partial x} \tag{2.9}$$
where D is called the diffusion coefficient and has for dimensions $[D] = L^2 \cdot T^{-1}$. Note that this is called Fick's law, it will be central in the derivation of the diffusion equation in the next chapter.

More complicated scenarios are obviously possible, but these two effects already lead to very interesting behavior as we shall see. Going back to our example, advection and diffusion will, in general, be present and so we superpose both effects and write the flux function as
$$q(x,t) = vu(x,t) - D\frac{\partial u}{\partial x} \tag{2.10}$$
Finally, we obtain the following closed form evolution equation for the concentration of pollutant in the channel
$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} - v\frac{\partial u}{\partial x} \tag{2.11}$$
This is called an **advection-diffusion (or convection-diffusion) equation**.

2.1.3 Linear transport equation

Equation (2.11) is formed of two terms and it may be useful to look separately at their effect:

- In the limit of no stream $v \to 0$, (2.11) becomes

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{2.12}$$

  which is called a **diffusion equation** and will be the focus of the first part of Chapter 3. We will see that the typical effects of a diffusion term are to spread and smooth.

- Conversely, in the limit where $D \to 0$, (2.11) becomes

$$\frac{\partial u}{\partial t} = -v\frac{\partial u}{\partial x} \tag{2.13}$$

  which is a pure **transport equation** in which v is the **advection speed**. Diffusion can be neglected for instance if the fluid in which the pollutant is transported is very viscous.

As v is a constant, Equation (2.13) is more specifically a **linear** transport equation. In the next section, we will introduce a model for traffic flow and shall introduce an advection speed depending on density $v(u)$. But before we move on, it is natural to ask what sort of solutions (2.13) admits. We combine (2.13) and the following initial condition

$$u(x,0) = f(x) \tag{2.14}$$

to form what is called an **initlal-value problem**. In the previous example, the initial condition determines the profile of concentration of the pollutant at $t = 0$ in the whole channel.

We can rewrite (2.13) as follows

$$\left(\frac{\partial}{\partial t} + v\frac{\partial}{\partial x}\right) u = 0 \tag{2.15}$$

If one can proceed to a change of variables from (x,t) to (r,s) in such a way that the derivatives transform as

$$\frac{\partial}{\partial r} = \frac{\partial}{\partial t} + v\frac{\partial}{\partial x}, \tag{2.16}$$

in those new variables, the equation becomes $\partial u/\partial r = 0$ and is then very easy to solve. With this in mind, we write $x = x(r,s)$ and $t = t(r,s)$ and use the chain rule to obtain

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial t}{\partial r}\frac{\partial}{\partial t} \tag{2.17}$$

Comparing (2.16) and (2.17), we thus require that $\partial x/\partial r = v$ and $\partial t/\partial r = 1$. By integration, these constraints yield

$$x = vr + p(s) \quad \text{and} \quad t = r + q(s) \tag{2.18}$$

We need to determine $p(s)$ and $q(s)$. At this point, it is useful to think about the initial conditions. The initial condition specifies the solution along the x-axis, i.e. for $t = 0$. To facilitate the application of the initial condition, it is convenient to make sure that in our new variables the initial conditions also specify the solution along, e.g. the s-axis corresponding to $r = 0$. In other words, we require that $r = 0$ implies $t = 0$ and $x = s$.

Setting $r = t = 0$, we conclude that $p(s) = s$ and $q(s) = 0$. Finally, we can write that our variable transformation reads:

$$r = t \quad \text{and} \quad s = x - vt \tag{2.19}$$

In these new variables, our equation is written

$$\frac{\partial u}{\partial r} = 0 \tag{2.20}$$

which means that

$$u = u(s) = u(x - vt) \tag{2.21}$$

The solution is only a function of the reduced variable $x - vt$. Thanks to our clever choice of new variables, it is easy to apply the initial condition and find that the solution of the problem is

$$\boxed{u(x,t) = f(x - vt)} \tag{2.22}$$

We can conclude that the profile of concentration set by the initial condition is transported (or advected) along the channel by the water stream at speed v without expansion or deformation. Note that the form of the solution could have been guessed from physical insights about this process. We established earlier that the advection term led to the pollutant moving in bulk, i.e. that all molecules composing the initial drop of pollutant are transported along the channel at the same speed v. Finally, note also that the initial condition need not be continuous for this solution to exist.

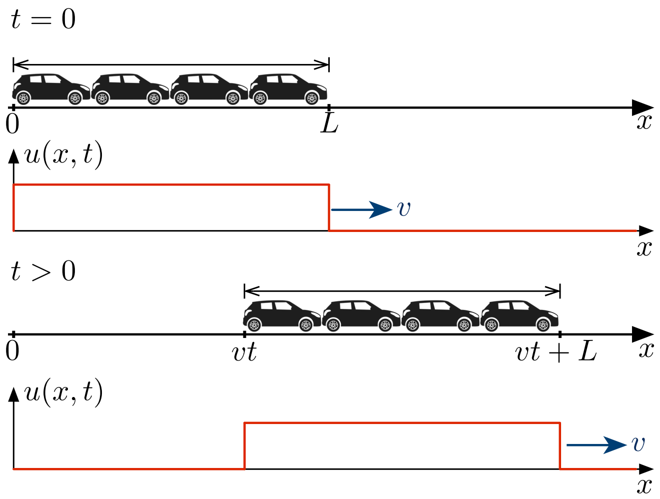


Figure 2.3 Schematic of a simple traffic flow example in which a group of cars all move at the same speed v. For each situation, we also represent the associated density profile.

Example

Here, we consider a first very elementary traffic flow example. Consider a group of cars uniformly spaced on a single lane road (as shown on Fig. 2.3). We also assume that the cars all move at the same speed v along the x-axis. As a consequence, cars move as a unit and do not get closer or further from each other. This implies that the density profile (a square bump) does not change over time. Mathematically, this can expressed as initial conditions given by

$$f(x) = \begin{cases} c, & \text{if } 0 < x < L \\ 0, & \text{otherwise} \end{cases} \tag{2.23}$$

where c is a constant density (which depends on the car spacing). From (2.22), the solution is then given for all $t > 0$

$$u(x,t) = \begin{cases} c, & \text{if } 0 < x - vt < L \\ 0, & \text{otherwise} \end{cases} \Rightarrow u(x,t) = \begin{cases} c, & \text{if } vt < x < vt + L \\ 0, & \text{otherwise} \end{cases} \tag{2.24}$$

2.2 The method of characteristics

In this section, we will introduce a powerful method of resolution for PDEs: the **method of characteristics**. We will introduce this method via examples starting off with the simplest possible case, the linear advection equation.

2.2.1 Linear transport equation with constant coefficients

We have just seen that the following initial value problem

$$\frac{\partial u}{\partial t} + v\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.25}$$

$$u(x,0) = u_0(x), \quad x \in \mathbb{R} \tag{2.26}$$

with v a constant has for solution

$$u(x,t) = u_0(x - vt) \tag{2.27}$$

If $v > 0$, it means physically that the original profile $u_0(x)$ moves to the right with constant speed as is show in Fig. 2.4. This is called a **travelling wave** solution.

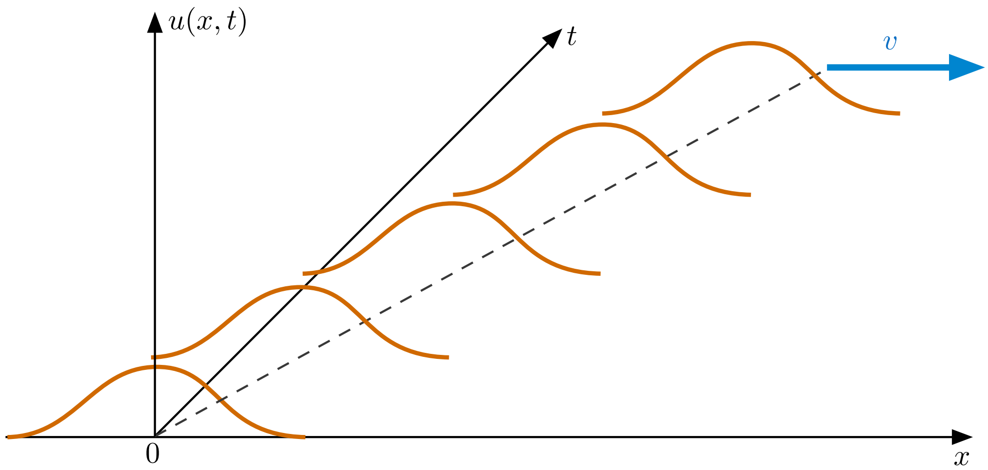


Figure 2.4

Here is a powerful way of constructing solutions to both linear and nonlinear equations. The solution to the PDE is a function of x and t, i.e. $u \equiv u(x,t)$.

Consider the value of the function on a curve $x = x(t)$ in the (x,t)-plane. Along this curve, the solution is given by $u = u(x(t),t)$. One may ask: what is du/dt?

The chain rule yields

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\frac{dx}{dt} \tag{2.28}$$

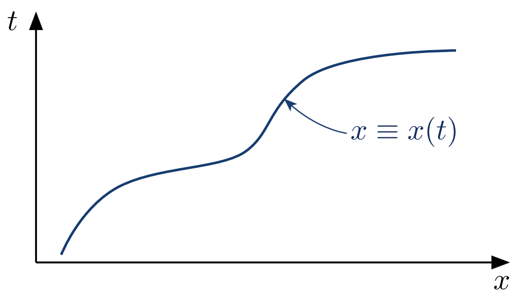


Figure 2.5

Using (2.25), the expression we obtained becomes particularly useful if we identify **special paths** $x(t)$ that will make the analysis easy. In particular, we easily realize that if $x(t)$ is **chosen** so that

$$\frac{dx}{dt} = v \tag{2.29}$$

where we recall that v is constant, then we have

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + v\frac{\partial u}{\partial x} = 0 \Rightarrow u = \text{const.} \tag{2.30}$$

We thus conclude that

$$\boxed{\text{On the curve defined by } \frac{dx}{dt} = v, \quad u = \text{const.}} \tag{2.31}$$

Integrating dx/dt, we obtain that $x(t) = \xi + vt$, where we have chosen $x(0) = \xi \in \mathbb{R}$. The only thing left to do is to find the constant value of u which must depend on our choice of ξ. This comes from the initial condition (2.26) which reads

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R} \tag{2.32}$$

which we may as well write

$$u(\xi, 0) = u_0(\xi), \quad \xi \in \mathbb{R} \tag{2.33}$$

Finally, as ξ is the value of $x(0)$ of our curve $x(t)$, we conclude that

$$u = u_0(\xi) \quad \text{on} \quad x = \xi + vt \tag{2.34}$$

algebraically, this gives $\xi = x - vt$ and so

$$u(x, t) = u_0(\xi) = u_0(x - vt). \tag{2.35}$$

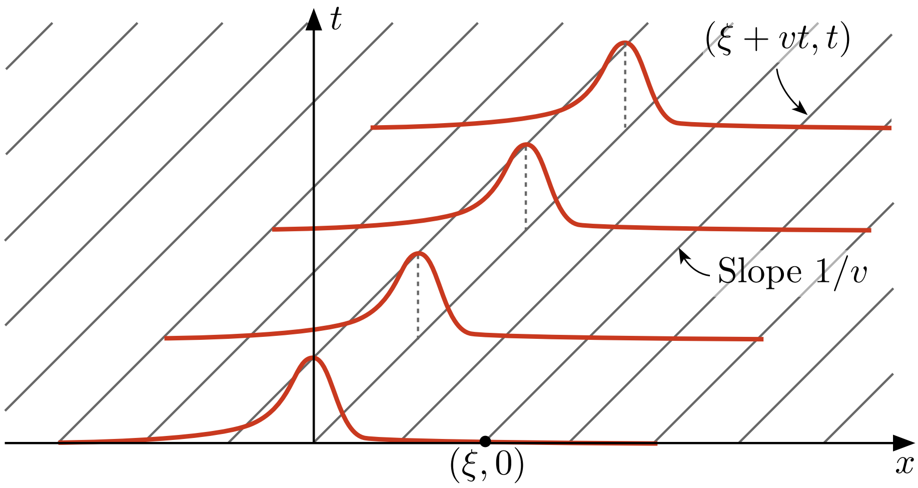


Figure 2.6 Solution of the linear advection equation in the (x, t)-plane. Grey lines show the characteristics of the linear advection equation with slope $1/v$; each characteristic curve correspond to a different parameter ξ. The initial condition $u_0(x)$ is transported without deformation at constant velocity v.

Definition 2.2.1: Characteristics

The geometrical construction we just introduced is **very important**. The special paths we introduced are called the **characteristic curves** (or **characteristics**) of the PDE. The union of all of them for $\xi \in \mathbb{R}$ is called the **set of characteristics**.

For the linear advection equation, the characteristics are straight lines $x = \xi + vt$ (see Fig. 2.6). The solution all along the characteristic with parameter ξ is $u_0(\xi)$. To find the solution at neighboring points, one would change ξ and draw the associated characteristics. For the linear advection problem, the characteristics are parallel lines all with slope $1/v$.

2.2.2 Linear transport equation with variable coefficients

Let us now consider a slightly more complicated case: the case of a linear advection equation but with variable coefficient, i.e. $v = v(x,t)$ is not a constant anymore. This may correspond for example to the case where the stream speed may not be constant in the channel or more generally, the medium in which the travelling wave is propagating is not homogeneous. The initial value problem, we are trying to solve, is

$$\frac{\partial u}{\partial t} + v(x,t)\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.36}$$

$$u(x,0) = u_0(x), \quad x \in \mathbb{R} \tag{2.37}$$

with $v(x,t)$ a given continuous function. In this case, we do not expect the characteristics to be straight lines any more. In fact, we know that

$$\frac{du}{dt} = 0 \quad \text{on} \quad \frac{dx}{dt} = v(x,t) \tag{2.38}$$

To find the solution, we thus need to solve

$$\frac{dx}{dt} = v(x,t) \tag{2.39}$$

$$x(0) = \xi \tag{2.40}$$

The solution to the problem (2.36)-(2.37) is then constant $u = u_0(\xi)$ along the curve $x \equiv x(t)$ solution of the problem (2.39)-(2.40).

Example

To go further, we need to provide an explicit expression for $v(x,t)$. Here, we consider the following problem

$$\frac{\partial u}{\partial t} - xt\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.41}$$

$$u(x,0) = u_0(x), \quad x \in \mathbb{R} \tag{2.42}$$

The characteristics are here solutions of the following equation

$$\frac{dx}{dt} = -xt \tag{2.43}$$

where the initial condition is set to be $x(0) = \xi$. By integration, we obtain

$$\log x = -\frac{t^2}{2} + C \tag{2.44}$$

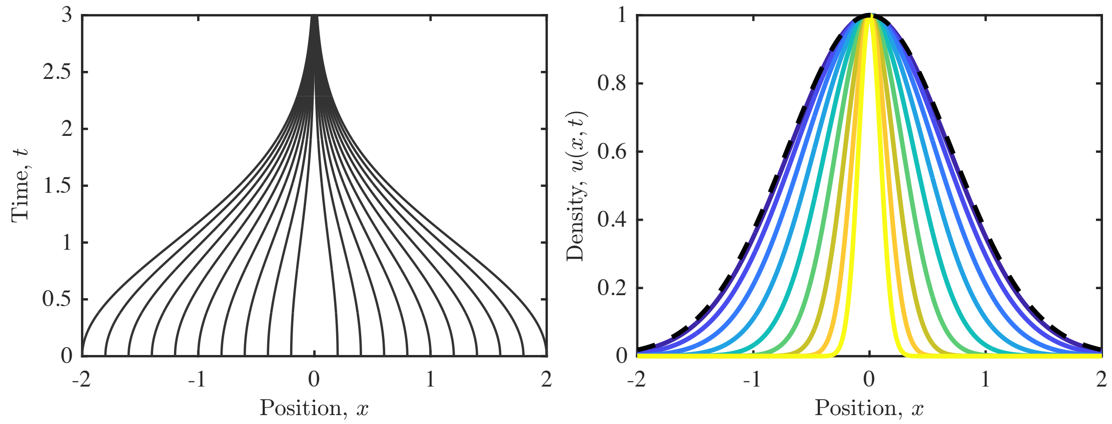


Figure 2.7 **(a)** Characteristic curves for equation (2.41) for various ξ, the characteristics are going inwards towards $x = 0$. **(b)** Density profiles $u(x,t)$ for various times (with time increasing from blue to yellow). Initial conditions $u_0(x) = e^{-x^2}$ are shown as dashed black lines.

where C is a constant to be determined. Using the initial condition, we finally obtain

$$x(t) = \xi e^{-t^2/2} \tag{2.45}$$

Fig. 2.7**a** provides a sketch of the characteristics and how the solution develops in the (x,t)-plane. From the characteristics, we can conclude that the solution will get focused around $x = 0$ as time increases. Indeed, this can be seen from the characteristics going inwards towards $x = 0$. Recall that $u(x,t)$ is constant along the characteristics.

Finally, we can express the solution in terms of (x,t). To do so, we need to eliminate the variable ξ. Now since

$$\xi = x\exp(t^2/2) \tag{2.46}$$

we directly have that

$$u(x,t) = u_0(xe^{t^2/2}) \tag{2.47}$$

If for instance, we set the initial profile to be $u_0(x) = e^{-x^2}$ (i.e. a gaussian profile), then the solution to this particular problem will be given by

$$u(x,t) = e^{-x^2e^{t^2}} \tag{2.48}$$

The profiles of density $u(x,t)$ are shown on Fig. 2.7**b** for different times.

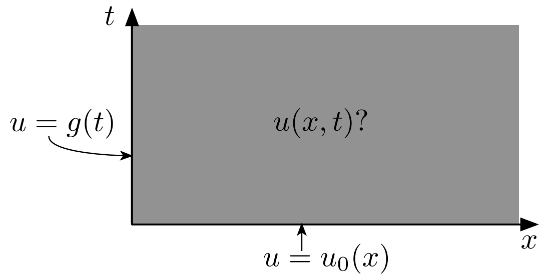
2.2.3 Boundary value problems

Consider the problem of transport of a pollutant in a semi-infinite channel ($x \geq 0$), where the density of pollutant is controlled at all times at the boundary and equal to $g(t)$. This type of problem is called a boundary-value problems; it takes the following form

$$\frac{\partial u}{\partial t} + v\frac{\partial u}{\partial x} = 0, \quad x > 0,\ t > 0 \tag{2.49}$$

$$u(0,t) = g(t), \quad t > 0 \tag{2.50}$$

$$u(x,0) = u_0(x), \quad x > 0 \tag{2.51}$$



How can the method of characteristics help us solve this problem? We want to obtain the solution $u(x, t)$ in the $x > 0$ and $t > 0$ quadrant. Just as we parametrized along the x-axis before with ξ, we will do a parametrization along the t-axis (using τ). The easiest is to look at an example: we will consider the case where the water channel does not originally contain any pollutant but where the density of pollutant is set to be first increasing and then decreasing.

Example

Consider the following boundary value problem

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad x > 0,\ t > 0 \tag{2.52}$$

$$u(0, t) = te^{-t}, \quad t > 0 \tag{2.53}$$

$$u(x, 0) = 0, \quad x > 0 \tag{2.54}$$

The characteristic curves associated to (2.52) are given by

$$\frac{du}{dt} = 0 \quad \text{on} \quad \frac{dx}{dt} = 1 \Rightarrow u = \xi_1 \quad \text{on} \quad x = t + \xi_2 \tag{2.55}$$

The characteristics are here (as expected) straight lines with slope 1. Clearly, the solution is given by $u(x, t) = 0$ for $x > t$ (see Fig. 2.8)

To find the solution in the sector $x < t$, we parametrize with τ along the t-axis ($x = 0$). Along the t-axis, we have

$$t = \tau, \quad u = \tau e^{-\tau} \quad \text{and} \quad x = 0 \Rightarrow \xi_1 = \tau e^{-\tau} \quad \text{and} \quad \xi_2 = -\tau \tag{2.56}$$

The solution to this problem is then given by

$$u = \tau e^{-\tau} \quad \text{on} \quad x = t - \tau \Rightarrow u(x, t) = (t - x)e^{x-t}, \quad \text{for } x < t \tag{2.57}$$

We show in Fig. 2.8 the density profiles for various times.

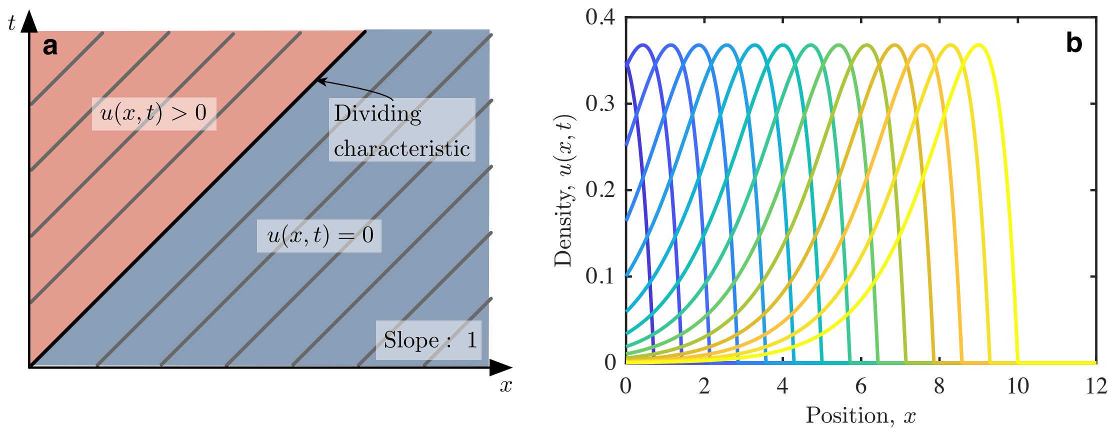


Figure 2.8 **(a)** Characteristic curves associated to (2.52). **(b)** Density profiles for various times (time increases from blue to yellow)

2.2.4 Method of characteristics for quasilinear first-order PDEs

So far, we have only used the method of characteristics to solve linear PDE problems. **Does this method work for nonlinear equations as well?**

The method of characteristics can be extended to nonlinear first-order PDEs. The method of characteristics is a geometrical method, let us try to understand how we can construct the solutions of quasilinear first-order PDEs.

The most general quasilinear first-order equation is written:

$$A(x,y,u)\frac{\partial u}{\partial x} + B(x,y,u)\frac{\partial u}{\partial y} = C(x,y,u) \tag{2.58}$$

In addition to the PDE itself, we will assume that the PDE is subject to the boundary condition that $u(x,y)$ is specified along some curve γ in the (x,y)-plane.

Example

> For instance, in the following conditions:
>
> $$u(x,0) = x^2 \qquad \text{...}\gamma \text{ is the } x\text{-axis} \tag{2.59}$$
> $$u(0,y) = ye^y \qquad \text{...}\gamma \text{ is the } y\text{-axis} \tag{2.60}$$
> $$u(x,x^3-x) = \sin(x) \qquad \text{...}\gamma \text{ is the graph of } y = x^3 - x \tag{2.61}$$

The idea of the method of characteristics is to reduce a quasilinear PDE to a system of ODEs. Let us think geometrically! We identify the solution $u(x,y)$ with its graph, i.e. with the surface $z = u(x,y)$ in the (x,y,z)-space. Note that specifying the solution along γ gives us a space curve Γ which must lie on the graph; this curve is called the initial curve. We can then use the PDE to build the remainder of the surface (i.e. build the solution) as a collection of space curves which emanate from Γ.

The normal vector to the surface $z = u(x,y)$ is given by

$$\mathbf{N} = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, -1\right) \tag{2.62}$$

Let us denote $\mathbf{F}$ the vector field

$$\mathbf{F} = \Big(A(x,y,u), B(x,y,u), C(x,y,u)\Big) \tag{2.63}$$

whose components are the coefficient functions of the PDE itself. We can easily see that if $u(x,y)$ is a solution to the PDE, then on the surface $z = u(x,y)$, we have

$$\mathbf{F}\cdot\mathbf{N} = 0 \tag{2.64}$$

which means that $\mathbf{F}$ is perpendicular to $\mathbf{N}$. Said differently, $\mathbf{F}$ is **tangent to the graph** $z = u(x,y)$. So the graph of the solution is made of streamlines of the vector field $\mathbf{F}$. **So we can construct the graph of the solution to the PDE by finding the streamlines of $\mathbf{F}$ that pass through the initial curve** Γ.

First, let's parametrize the initial curve Γ and write

$$\Gamma : \begin{cases} x = x_0(t) \\ y = y_0(t) \\ z = z_0(t) \end{cases} \tag{2.65}$$

For each value of the parameter t, the streamline of $\mathbf{F}$ that passes through $\Gamma(t)$ is given by the solution to the following coupled ODEs

$$\frac{dx}{ds} = A(x,y,z), \quad \frac{dy}{ds} = B(x,y,z), \quad \frac{dz}{ds} = C(x,y,z), \tag{2.66}$$

subject to the following initial conditions

$$x(0) = x_0(t), \quad y(0) = y_0(t), \quad z(0) = z_0(t) \tag{2.67}$$

These are called the **characteristic equations** of the PDE.

Remark. *Note that here t does not stand for time, it is a parameter which dictates our location on the initial curve Γ.*

Granted that we can solve this system of ODEs (which may be very hard), we would obtain solutions in terms of the parameter s and t:

$$x = X(t, s) \tag{2.68}$$
$$y = Y(t, s) \tag{2.69}$$
$$z = Z(t, s) \tag{2.70}$$

giving us a **parametric representation** of the solution surface. We would need to invert (2.68) and (2.69) and express (s, t) in terms of (x, y) as in

$$t = T(x, y), \quad s = S(x, y) \tag{2.71}$$

to finally obtain the solution to the PDE in the form:

$$u(x, y) = Z(T(x, y), S(x, y)) \tag{2.72}$$

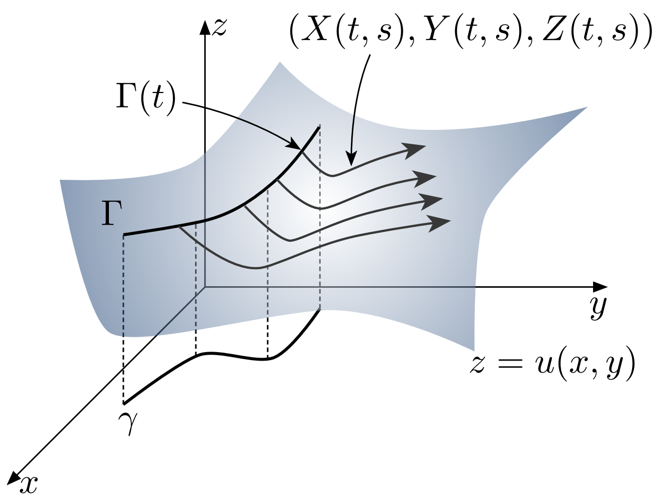


Figure 2.9 Schematic of the graph of the solution $z = u(x, y)$ (blue surface), initial curve $\Gamma(t)$ and the characteristics flowing out of the initial curve (grey lines).

There is nothing better to understand abstract concepts than to have a look at a concrete example!

Example

In this example, we want to find the solution to

$$x\frac{\partial u}{\partial x} - 2y\frac{\partial u}{\partial y} = u^2 \tag{2.73}$$

that satisfies $u(x, x) = x^3$. This is a quasilinear equation with the following coefficient functions:

$$A(x, y, u) = x, \quad B(x, y, u) = -2y, \quad C(x, y, u) = u^2 \tag{2.74}$$

The initial curve Γ can be parametrized as follows:

$$x = t, \quad y = t, \quad z = t^3 \tag{2.75}$$

The characteristic equations are then given by

$$\frac{dx}{ds} = x, \quad \frac{dy}{ds} = -2y, \quad \frac{dz}{ds} = z^2, \tag{2.76}$$

with the initial conditions

$$x(0) = t, \quad y(0) = t, \quad z(0) = t^3 \tag{2.77}$$

The characteristic equations can be integrated and the solutions are written

$$x(s) = te^s \tag{2.78}$$
$$y(s) = te^{-2s} \tag{2.79}$$
$$z(s) = \frac{t^3}{1 - t^3 s} \tag{2.80}$$

We can now invert (2.78) and (2.79) and obtain

$$y = te^{-2s} = t(e^s)^{-2} = t(x/t)^{-2} = t^3/x^2 \Rightarrow t^3 = x^2 y \tag{2.81}$$
$$\Rightarrow t = x^{2/3} y^{1/3} \tag{2.82}$$

and

$$x = te^s \Rightarrow e^s = x/t \tag{2.83}$$
$$\Rightarrow e^s = (x/y)^{1/3} \tag{2.84}$$
$$\Rightarrow s = \frac{1}{3}\ln(x/y) \tag{2.85}$$

Substituting these into (2.80), we find that

$$u(x, y) = \frac{x^2 y}{1 - \frac{1}{3}x^2 y \ln(x/y)} \tag{2.86}$$

Remark. *Note that we derived here the method of characteristics for the most general quasilinear equation. The characteristic equations we obtained were given by*

$$\frac{dx}{ds} = A(x, y, z), \quad \frac{dy}{ds} = B(x, y, z), \quad \frac{dz}{ds} = C(x, y, z), \tag{2.87}$$

To recover the formulation of the method of characteristics we used earlier, one can combine the ODEs and write

$$\frac{dy}{dx} = \frac{B(x, y, u)}{A(x, y, u)} \quad \text{and} \quad \frac{du}{dx} = \frac{C(x, y, u)}{A(x, y, u)} \tag{2.88}$$

In the case of an homogeneous equation, we would have $C(x, y, u) = 0$.

2.3 Nonlinear first-order PDEs

So far, we have mainly looked at linear scalar conservation laws. In this section, we explore the solutions of nonlinear first-order PDEs. As much as possible, we will introduce important concepts via examples; this section may appear a bit more abstract but it is a necessary step before we can talk further about traffic flow modelling. Indeed, nonlinear equations are a lot subtler than linear equations and require us to introduce some new important ideas.

2.3.1 Kinematic Wave Equation

First, we will consider a general kinematic wave equation, i.e. an equation of the form

$$\frac{\partial u}{\partial t} + c(u)\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.89}$$

with the initial condition $u(x,0) = u_0(x),\ x \in \mathbb{R}$. Here, we will assume that $c(u)$ is a given **smooth function** of u, e. g. if $c(u) = u$, one obtains the so-called inviscid Burgers equation. We also assume that $u_0(x)$ is smooth, in particular, the initial conditions do not exhibit any discontinuity.

The natural question one may ask is: **starting from smooth initial conditions, are we guaranteed to have smooth solutions for all time?** We will see that **sometimes, yes** (rarefaction waves) but **sometimes, no** (shock waves).

Let us assume that the solution and its first derivatives are continuous, i.e. $u(x,t) \in \mathcal{C}^1$. To simplify our problem, we can proceed as in the linear case and look for characteristics. By the chain rule on a special path $x(t)$, we have

$$\frac{du}{dt} = \frac{d}{dt}\left[u(x(t),t)\right] = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\frac{dx}{dt} = 0 \quad \text{if} \quad \frac{dx}{dt} = c(u) \tag{2.90}$$

in particular, we obtain that

$$u = u_0(\xi) \quad \text{on} \quad \frac{dx}{dt} = c(u),\ x(0) = \xi \tag{2.91}$$

Now, since $u = u_0(\xi)$ is constant along the characteristics so is $c(u) = c(u_0(\xi))$. To find the equation of the characteristics, we integrate and obtain

$$x = \xi + c(u_0(\xi))t \tag{2.92}$$

At this point, there are a few comments to make:

1. The characteristics are **straight lines**;

2. The **slope of the characteristics is not constant** as in the linear problems, but depends on ξ;

3. The expression in (2.92) is an **implicit relation for** ξ as a function of (x,t), i.e. $\xi(x,t)$. This means that given a point in the (x,t)-plane, the solution at this point is given by $u(x,t) = u_0(\xi(x,t))$, where ξ is the solution (which we assume to be unique) of (2.92). Thus, if given (x,t), one must first find the ξ it came from at $t = 0$ (i.e. solve (2.92)) to then obtain a formula for the solution. The problem with nonlinear equations is that this can rarely be done analytically.

Let us check that we have indeed found a solution in

$$u = u_0(\xi) \quad \text{on} \quad x = \xi + c(u_0(\xi))t \tag{2.93}$$

It may be easier to see this as

$$u(x,t) = u_0(\xi(x,t)) \tag{2.94}$$

In particular, we have that

$$\frac{\partial u}{\partial t} = u_0'(\xi)\frac{\partial \xi}{\partial t}, \quad \frac{\partial u}{\partial x} = u_0'(\xi)\frac{\partial \xi}{\partial x} \tag{2.95}$$

where the prime denote a derivative with respect to ξ. By implicit differentiation of (2.92) with respect to t, we obtain

$$0 = \frac{\partial \xi}{\partial t} + c(u_0(\xi)) + c'(u_0(\xi))u_0'(\xi)\frac{\partial \xi}{\partial t}t \Rightarrow \frac{\partial \xi}{\partial t} = -\frac{c(u_0(\xi))}{1 + c'(u_0(\xi))u_0'(\xi)t} \tag{2.96}$$

and by implicit differentiation with respect to x this time

$$1 = \frac{\partial \xi}{\partial x} + c'(u_0(\xi))u_0'(\xi)\frac{\partial \xi}{\partial x}t \Rightarrow \frac{\partial \xi}{\partial x} = \frac{1}{1 + c'(u_0(\xi))u_0'(\xi)t} \tag{2.97}$$

Clearly by combining (2.95), (2.96) and (2.97), we have

$$\frac{\partial u}{\partial t} + c(u)\frac{\partial u}{\partial x} = 0 \tag{2.98}$$

as expected. In our algebra, a crucial quantity emerged

$$\Lambda \equiv 1 + u_0'(\xi)c'(u_0(\xi))t \tag{2.99}$$

Indeed, for instance, if there exists a time $t > 0$ for which $\Lambda = 0$, then by (2.95), we can see that

$$\frac{\partial u}{\partial x} \to \infty \tag{2.100}$$

This is known **shock formation**; the profile $u(x,t)$ develops a discontinuity in space. On the other hand, if $\Lambda \neq 0$ for all time $t > 0$, then **a smooth solution exists for all times**. This will happen is $c'(u_0(\xi))$ and L have the same sign. This leads to the following theorem.

Theorem 2.3.1

Suppose that c and u_0 are continuously differentiable functions, i.e. $c, u_0 \in \mathcal{C}^1(\mathbb{R})$, and that they are either both nonincreasing or nondecreasing on $\mathbb{R}$. Then, the nonlinear initial value problem

$$\begin{aligned} \frac{\partial u}{\partial t} + c(u)\frac{\partial u}{\partial x} &= 0, \quad x \in \mathbb{R},\ t > 0 \\ u(x,0) = u_0(x), &\quad x \in \mathbb{R} \end{aligned}$$

has a unique solution for all time given implicitly by

$$u(x,t) = u_0(\xi), \quad x = \xi + c(u_0(\xi))t$$

Example

Consider the following nonlinear initial value problem

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.101}$$

$$u(x,0) = \tanh(x),\ x \in \mathbb{R} \tag{2.102}$$

We have in this problem

$$c(u) = u \Rightarrow c' = 1 > 0 \tag{2.103}$$

and

$$u_0(\xi) = \tanh(\xi) \Rightarrow u_0' = \operatorname{sech}^2 \xi > 0 \tag{2.104}$$

The characteristics are given by

$$x = \xi + \tanh(\xi)t \tag{2.105}$$

and the solution is given by

$$u(x,t) = \tanh(\xi(x,t)) \tag{2.106}$$

As can be seen in Fig. 2.10, the characteristics associated to this problem fan out and the solution spreads out from its initial conditions.

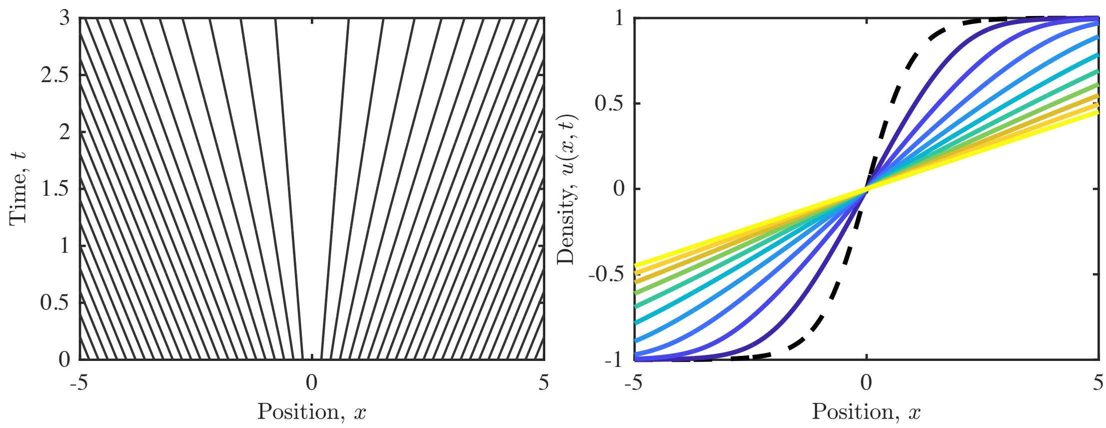


Figure 2.10 *(Left)* Characteristic curves associated to (2.101). *(Right)* Density profiles for various times (time increases from blue to yellow)

Another way to visualize the problem of shock formation is by using the characteristics themselves! For instance, consider the case where $c(u) = u$. Equation (2.89) is then called the inviscid Burgers equation. Now the initial condition $u(x,0) = u_0(x)$ not only specifies the solution on the line $t = 0$, it also specifies the slope of the characteristics. The characteristic line passing through $(x_1, 0)$ will have slope $c(u_0(x_1))$ and the characteristic line passing through $(x_2, 0)$ will have slope $c(u_0(x_2))$. If the slopes are such that the two lines intersect (see Fig. 2.11), then we are in trouble! Indeed, on one line $u = u_0(x_1)$ while on the other $u = u_0(x_2)$, but for the lines to cross, we needed the lines to have different slopes and so $c(u_0(x_1)) \neq c(u_0(x_2))$ and the solution is constant on the characteristics!

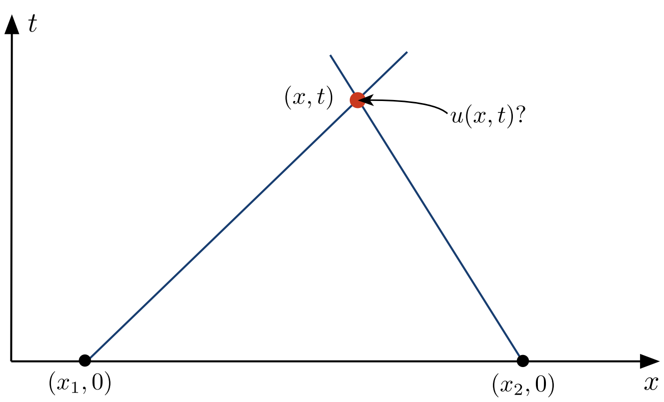


Figure 2.11

Geometrically, we can say that in the case where no pair of characteristic lines intersect in the half-plane $t > 0$, there exists a solution $u(x,t)$ throughout this half-plane. And we saw that this can only happen if the slope of the characteristics is increasing as a function of their intercept:

$$c(u_0(x)) \leq c(u_0(y)), \quad \text{for } x \leq y \tag{2.107}$$

In this case, the characteristic lines spread out for $t > 0$. Such a solution is called an **expansive wave** or **rarefaction wave**.

2.3.2 Shock formation

In the general case, some characteristics will cross and the solution will exist only up to the time of the crossing. **What does this look like qualitatively?** The wave speed is given by $c(u)$. As it depends on u, some parts of the wave will be faster than others. If you consider the first profile showed in Fig, 2.12, the bigger u the faster the wave moves, so the bigger, faster part of the wave overtakes the smaller and slower part of the wave. This leads to a wave breaking when the crest of the wave takes over leading to a triple-valued ”solution”. Shock waves occur in many contexts including in explosions, traffic flow, supernovae, airplanes breaking the sound barrier, and so on.

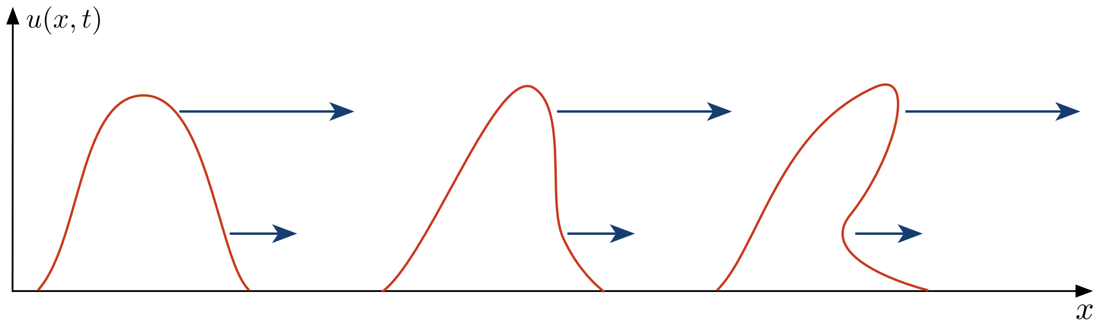


Figure 2.12

Now, let us study the formation of shocks from smooth initial data quantitatively. We saw that the formation of a shock is associated with a singularity in the first spatial derivative of u. A natural question is whether we can find the time when a singularity in $\partial u/\partial x$ first forms. We focus on the equation

$$\frac{\partial u}{\partial t} + c(u)\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.108}$$

with the initial condition $u(x,0) = u_0(x),\ x \in \mathbb{R}$. We place ourselves in the case where $c(u) > 0$, $c'(u) > 0$ and $u_0 \in \mathcal{C}^1(\mathbb{R})$. We have just seen that provided these assumptions on $c(u)$:

- If $u_0(x)$ is nondecreasing then the solution will be smooth for all times;
- If $u_0(x)$ is strictly decreasing on some interval $\mathcal{I} \subset \mathbb{R}$, then a shock will form.

To fix things, we can take $u_0(x) > 0$ and $u_0'(x) < 0$ in $\mathbb{R}$. The characteristic equations are given by

$$u(x,t) = \text{const.} \quad \text{on} \quad \frac{dx}{dt} = c(u) \tag{2.109}$$

Consider two characteristics emanating from the points $\xi_1 < \xi_2$. The respective speeds are $c(u_0(\xi_1))$ and $c(u_0(\xi_2))$, as u_0 is decreasing, c is increasing and so we conclude that

$$c(u_0(\xi_1)) > c(u_0(\xi_2)) \tag{2.110}$$

In the (x,t)-plane, these characteristics are straight lines with slopes set by $c(u_0(\xi))$. In particular, we have

$$\frac{dt}{dx} = \frac{1}{c(u)} \tag{2.111}$$

so the slope of the ξ_1 characteristic is smaller than the slope of the ξ_2 characteristic. As shown on Fig. 2.13, the characteristics cross at the point P. As we hinted at before, this is a contradiction! At the point P, the solution is non-unique. The values $u_0(\xi_1) \neq u_0(\xi_2)$ which are carried by the ξ_1 and ξ_2 characteristics respectively both end up at P. We conclude that **a smooth solution cannot exist for all times**.

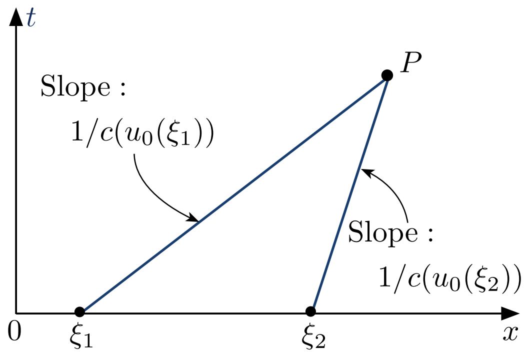


Figure 2.13

The breaking time can be found by considering the first time **any two characteristic cross**. One way to go about finding this time is to find $\partial u/\partial x$ and check where it blows up. We will take a characteristic starting at $x(0) = \xi$ and calculate $\partial u/\partial x$ along it.

The equation of the characteristic is given by

$$x = c(u_0(\xi))t + \xi \tag{2.112}$$

We want to calculate $\frac{\partial}{\partial x}u(x(t),t)$, with $x(t)$ given by (2.112). Let $g(t) = \frac{\partial u}{\partial x}$, then we can write

$$\frac{dg}{dt} = \frac{\partial^2 u}{\partial t \partial x} + \frac{\partial^2 u}{\partial x^2}\frac{dx}{dt} = \frac{\partial^2 u}{\partial t \partial x} + c(u)\frac{\partial^2 u}{\partial x^2} \tag{2.113}$$

By differentiating the original PDE (2.108) with respect to x, we obtain

$$\frac{\partial^2 u}{\partial x \partial t} + c(u)\frac{\partial^2 u}{\partial x^2} + c'(u)\left(\frac{\partial u}{\partial x}\right)^2 = 0 \tag{2.114}$$

Comparing these two equations, we find that

$$\frac{dg}{dt} + c'(u)g^2 = 0 \tag{2.115}$$

along the characteristic, i.e.

$$\frac{dg}{dt} + c'(u_0(\xi))g^2 = 0 \tag{2.116}$$

We can solve this separable ODE to find that

$$g(t) = \frac{g(0)}{1 + g(0)c'(u_0(\xi))t} \tag{2.117}$$

where $g(0) = \frac{\partial u_0}{\partial x}(\xi)$ for the characteristic we considered. So finally, we conclude that

$$\frac{\partial u}{\partial x} = \frac{u_0'(\xi)}{1 + u_0'(\xi)c'(u_0(\xi))t} \tag{2.118}$$

Now by assumption, $u_0' < 0$ and $c' > 0$, hence the first spatial derivative $\partial u/\partial x$ will become **infinite at a finite time for each characteristic** labelled by ξ. We need to find the characteristic that minimizes the breaking time t_B. From (2.118), we find that the time of singularity for the ξ characteristic, i.e. the time at which $\partial u/\partial x$ tends to infinity, is given by

$$t_{\text{sing}} = -\frac{1}{u_0'(\xi)c'(u_0(\xi))} \tag{2.119}$$

Hence, we need to maximize the following quantity $|u_0'(\xi)c'(u_0(\xi))|$ over ξ. Assume that this function is maximized for $\xi = \xi_B$, the breaking time is then given by

$$t_{\text{B}} = -\frac{1}{u_0'(\xi_B)c'(u_0(\xi_B))} \tag{2.120}$$

Note that if $u_0(x)$ is non-monotonous then the breaking will occur first on the characteristic $\xi = \xi_B$ for which we have both $u_0'(\xi_B)c'(u_0(\xi_B)) < 0$ and $|u_0'(\xi)c'(u_0(\xi))|$ is maximal.

Let us put his into practice and have a look at an example!

Example

Consider the following equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.121}$$

This is the inviscid Burgers equation or 1D Euler equation (without source term); this equation can be used to model the flow of a fluid with no viscosity, in which case u represents the momentum in the fluid. We assume that the initial conditions are given by

$$u(x,0) = e^{-x^2}, \quad x \in \mathbb{R} \tag{2.122}$$

In particular, this Gaussian profile is monotonically decreasing in $x > 0$, so we expect a shock to form. Given a label ξ, we found from (2.119) that the singular time is given by

$$t_{\text{sing}} = -\frac{1}{u_0'(\xi)c'(u_0(\xi))} \tag{2.123}$$

Here, we have $u_0(x) = e^{-\xi^2}$ and $c(u) = u$, so the singular time reads

$$t_{\text{sing}} = \frac{1}{2\xi e^{-\xi^2}} \tag{2.124}$$

To minimize this time, we need to maximize the function $\phi(\xi) = \xi e^{-\xi^2}$,

$$\phi'(\xi) = e^{-\xi^2} - 2\xi^2 e^{-\xi^2} = 0 \Rightarrow \xi_B = \frac{1}{\sqrt{2}} \tag{2.125}$$

This leads to the following breaking time

$$t_B = \frac{1}{\sqrt{2}e^{-1/2}} = \sqrt{\frac{e}{2}} \approx 1.16 \tag{2.126}$$

The implicit solution for $t < t_B$ is given by

$$u(x,t) = e^{-\xi^2} \quad \text{on} \quad x = e^{-\xi^2}t + \xi \tag{2.127}$$

Fig. 2.14 shows both the characteristics (crossing) and the profiles $u(x,t)$ for various times.

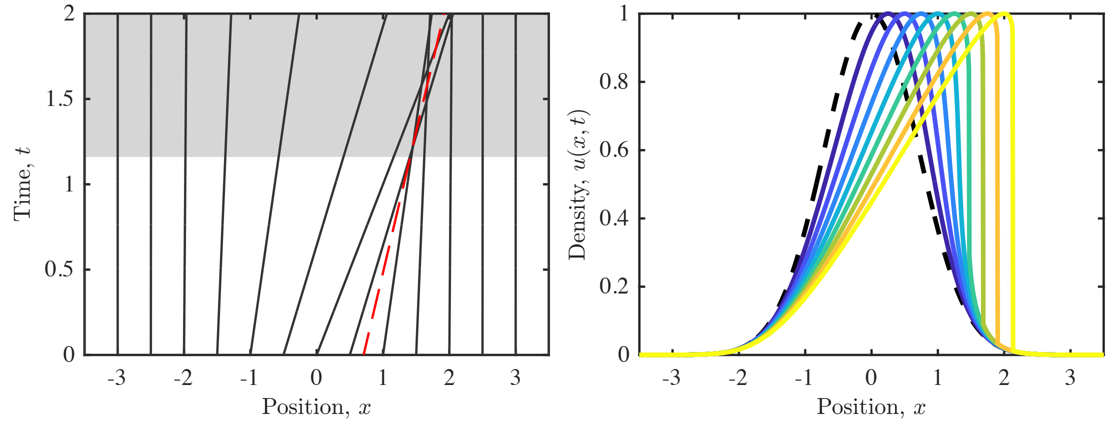


Figure 2.14 *(Left)* Characteristic curves associated with the above example.The characteristic associated to $\xi_B = 1/\sqrt{2}$ is given by the red dashed line. The greyed out region is the region where we do not know yet how to construct a solution. *(Right)* Density profiles for various times (time increases from blue to yellow). The initial profile is given by the black dashed line.

Remark. *Another way of seeing the condition for breakup is to consider when it is possible to solve for $\xi(x,t)$ in the characteristic equation*

$$x - \xi = c(u_0(\xi))t \tag{2.128}$$

so that we can find $u(x,t) = u_0(\xi(x,t))$. Let

$$J(x,t,\xi) = x - \xi - c(u_0(\xi))t \tag{2.129}$$

We know that (2.129) can be solved locally for ξ as long as $\frac{\partial J}{\partial \xi} \neq 0$ (so that the **implicit function theorem** *can be used). Taking the derivative with respect to ξ, we write*

$$\frac{\partial J}{\partial x} = -1 - c'(u_0(\xi))u_0'(\xi)t \tag{2.130}$$

So given a point (x,t) in the (x,t)-plane, the condition to be be able to find a unique $(\xi,0)$ from where (x,t) stems from is given by

$$-1 - c'(u_0(\xi))u_0'(\xi)t \neq 0 \tag{2.131}$$

which is the same condition as (2.119).

2.3.3 Discontinuous solutions – shock propagation

Let us start by summarizing our findings so far:

1. The **linear advection equation** $u_t + c_0 u_x = 0$ propagates **any** initial condition $u(x,0) = u_0(x)$ without deformation of the profile, i.e. giving a solution $u(x,t) = u_0(x - c_0 t)$.

2. For **nonlinear problems** of the form $u_t + c(u)u_x = 0$ where $x \in \mathbb{R}$ and $t > 0$, with **smooth initial conditions** $u(x,0) = u_0(x)$, the solution will (a) remain smooth for all times **or** (b) develop an infinite slope singularity (i.e. a shock).

3. For nonlinear problems of the form $u_t + c(u)u_x = 0$ where $x \in \mathbb{R}$ and $t > 0$, where the initial conditions $u(x, 0) = u_0(x)$ is **continuous** but has **discontinuous derivatives** at a finite number of points, it can be shown (not done here) that the discontinuity travels along the characteristic emanating from where it is at $t > 0$.

Naturally, these results beg **some related questions**:

(i) What happens if the initial condition $u_0(x)$ is discontinuous?

(ii) What happens to a shock solution after the shock forms?

Riemann problem

For that, we will consider a paradigmatic problem: the so-called **Riemann problem**. The Riemann problem is an initial value problem composed of a conservation equation and a piecewise constant initial data (which presents a single discontinuity). Here, we will be interested in the following

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.132}$$

$$u(x, 0) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{if } x > 0 \end{cases} \tag{2.133}$$

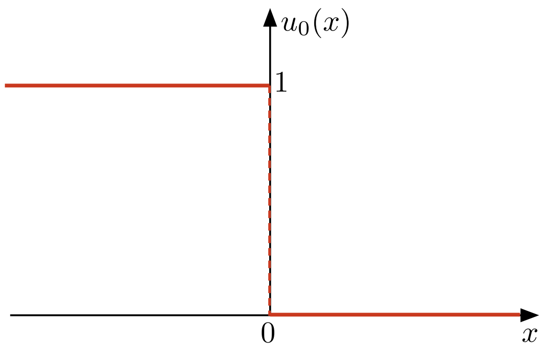


Figure 2.15

For this problem, the characteristics are given by

$$\frac{du}{dt} = 0 \quad \text{on} \quad \frac{dx}{dt} = u \tag{2.134}$$

so once again, we have

$$\frac{dx}{dt} = u_0(\xi), \quad \xi \in \mathbb{R} \tag{2.135}$$

which leads to

$$x(t) = u_0(\xi)t + \xi = \begin{cases} \xi \text{ (const.)}, & \xi > 0 \\ t + \xi, & \xi < 0 \end{cases} \tag{2.136}$$

We thus conclude that we have a serious problem! Indeed, for **any** $t > 0$, however small, characteristics from $\xi > 0$ cross with the characteristic emanating from $\xi < 0$. Fig. 2.16 shows the characteristics in the (x, t)-plane.

It is natural to think that the discontinuity will be carried in the (x, t)-plane along some curve. If, for instance, we assume that this curve is given by $x = s_0 t$, then any choice of s_0 provides a shock solution; the associated characteristics are shown in Fig. 2.17. We now have a **non-uniqueness problem**! Indeed, this construction holds for **any** s_0.

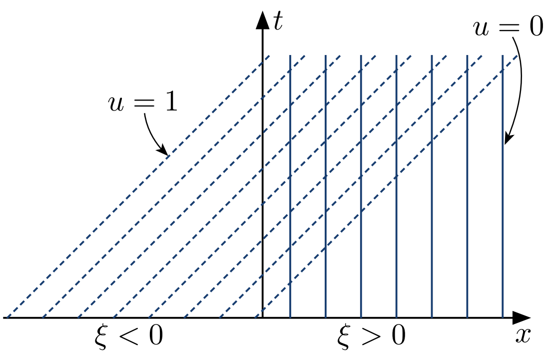


Figure 2.16

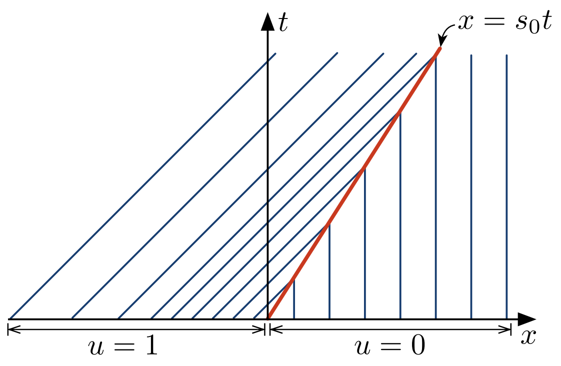


Figure 2.17

Jump condition

A natural question is then: how do we fix the shock speed? (i.e. how do we select s_0) To answer this question, we need to go back to the conservation law in integral form to derive what are called **jump conditions**.

Recall that our conservation equations came from the conservation principle. To derive it, we considered an interval $[a, b]$ and wrote the conservation of a quantity (e.g. mass, momentum, energy) on this interval; mathematically, this took the following form

$$\frac{d}{dt}\int_a^b (\text{Quantity})dx = [\text{Flux}]_a - [\text{Flux}]_b \tag{2.137}$$

In our example, we wrote the conservation law (here, the inviscid Burgers equation) in differential form as follows

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \tag{2.138}$$

or equivalently:

$$\frac{\partial u}{\partial t} + \frac{\partial q(u)}{\partial x} = 0 \quad \text{with} \quad q(u) = \frac{1}{2}u^2 \tag{2.139}$$

As a reminder, we can integrate this equation over the interval $[a, b]$ to obtain

$$\int_a^b \left[\frac{\partial u}{\partial t} + \frac{\partial q(u)}{\partial x}\right] = 0 \Rightarrow \frac{d}{dt}\int_a^b u(x,t)dx = \int_b^a \frac{\partial q}{\partial x}dx \tag{2.140}$$

Finally, this equation in integral form then reads

$$\frac{d}{dt}\int_a^b u(x,t)dx = q(a,t) - q(b,t) \tag{2.141}$$

It is necessary to go back to (2.141) to analyse the motion of shocks as (2.138) becomes singular. We will construct what are called **weak solutions**.

Assume that the shock path in the (x, t)-plane is given by $x = s(t)$. Across the shock, the solution is discontinuous but the one-sided limits of u and its derivatives exist, i.e. for any time t

$$\lim_{x\to s(t)^+} u, \quad \lim_{x\to s(t)^+} \frac{\partial u}{\partial x}, \quad \lim_{x\to s(t)^-} u, \quad \text{and} \quad \lim_{x\to s(t)^-} \frac{\partial u}{\partial x} \quad \text{are } \textbf{all finite} \tag{2.142}$$

Let us pick two arbitrary points (a, b) such that $a < s(t)$ and $b > s(t)$; the interval (a, b) thus contains the shock. The conservation law given in (2.141) becomes

$$\frac{d}{dt}\left[\int_a^{s(t)} u(x,t)dx\right] + \frac{d}{dt}\left[\int_{s(t)}^b u(x,t)dx\right] = q(a,t) - q(b,t) \tag{2.143}$$

Recall that Leibniz's rule for differentiating under the integral gives that

$$\frac{d}{dt}\left(\int_{a(t)}^{b(t)} f(x,t)dx\right) = f(b(t),t)\frac{d}{dt}b(t) - f(a(t),t)\frac{d}{dt}a(t) + \int_{a(t)}^{b(t)} \frac{\partial}{\partial t}f(x,t)dx \quad (2.144)$$

for $a(x) > -\infty$ and $b(x) < \infty$. Using Leibniz's rule, we can differentiate (2.143) under the integral to obtain

$$u(s^-,t)\frac{ds}{dt} + \int_a^{s(t)} \frac{\partial}{\partial t}u(x,t)dx - u(s^+,t)\frac{ds}{dt} + \int_{s(t)}^b \frac{\partial}{\partial t}u(x,t)dx = q(a,t) - q(b,t) \quad (2.145)$$

If we now take the limits $a \to s(t)^-$ and $b \to s(t)^+$, the integral terms become zero (i.e. the one-sided derivatives exist) and we obtain

$$\left[u(s^-,t) - u(s^+,t)\right]\frac{ds}{dt} = \left[q(s^-,t) - q(s^+,t)\right] \iff \frac{ds}{dt}\left[u(x,t)\right]_+^- = \left[q(u)\right]_+^- \quad (2.146)$$

or more compactly

$$[u(x,t)]\, s'(t) = [q(u)] \quad (2.147)$$

where the jump notation $[\cdots]$ is taken to implicitly always mean the value **before** the shock minus the value **after** the shock. The condition in (2.147) is called the **Rankine-Hugoniot** condition. It relates the speed of the shock $s'(t)$ to the state of the system before and after the shock; finally, $[u]$ is called the shock strength.

Remark. *The notation* $[\cdots]$ *is not to be confused with the notation for the dimension of a physical quantity. You should be able to know from context whether we are using a dimension of a jump notation. If you were to need to use both in the same problem, you should be very clear about what is meant by* $[\cdots]$ *each time it is used.*

Back to the Riemann problem

Recall that the PDE in differential form that we are considering here is

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \Rightarrow \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) = 0 \quad (2.148)$$

So the shock propagation speed is given by the following Rankine-Hugoniot condition

$$-[u]s'(t) + \left[\frac{1}{2}u^2\right] = 0 \quad (2.149)$$

which gives here

$$-[1-0]s'(t) + \left[\frac{1}{2} - 0\right] = 0 \Rightarrow s'(t) = 1/2 \quad (2.150)$$

Note that in the general case where

$$u_0(x) = \begin{cases} u^- & x < 0 \\ u^+ & x > 0 \end{cases} \quad (2.151)$$

one can easily show that for the inviscid Burgers equation, the shock propagation speed is given by

$$s'(t) = \frac{1}{2}(u^+ + u^-) \quad (2.152)$$

i.e. an average of the two states.

Finally, we conclude that our Riemann problem has for solution

$$u(x,t) = \begin{cases} u^+ & x > (u^+ + u^-)t/2 \\ u^- & x < (u^+ + u^-)t/2 \end{cases} \quad (2.153)$$

Rarefaction waves

For a moment, let us change the initial conditions of the Riemann problem to be

$$u_0(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases} \tag{2.154}$$

The characteristics for this new problem are shown in Fig. 2.19.

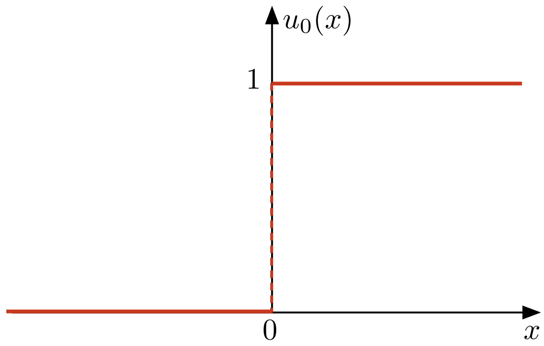


Figure 2.18

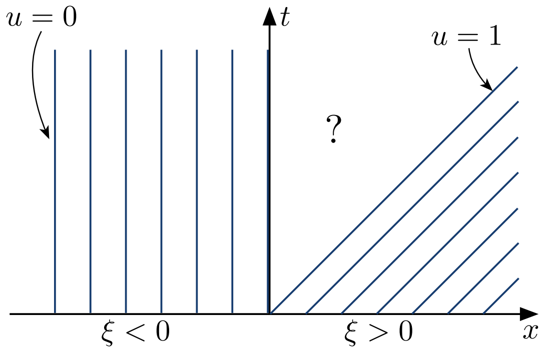


Figure 2.19

What is happening then in between the vertical characteristics ($\xi < 0$) and the characteristics labelled by $\xi > 0$. At any fixed time, we have

$$u(x,t) = \begin{cases} 0 & x < 0 \\ 1 & x > t \end{cases} \tag{2.155}$$

However, in the interval $(0, t)$, we need to have a linear variation connecting the two states, i.e.

$$u(x,t) = \frac{x}{t}, \quad 0 < x < t \tag{2.156}$$

This is then equivalent to having a fan of characteristics all emanating from $x = 0$ at $t = 0$. The true diagram of characteristics for this **rarefaction wave** problem is given in Fig. 2.20

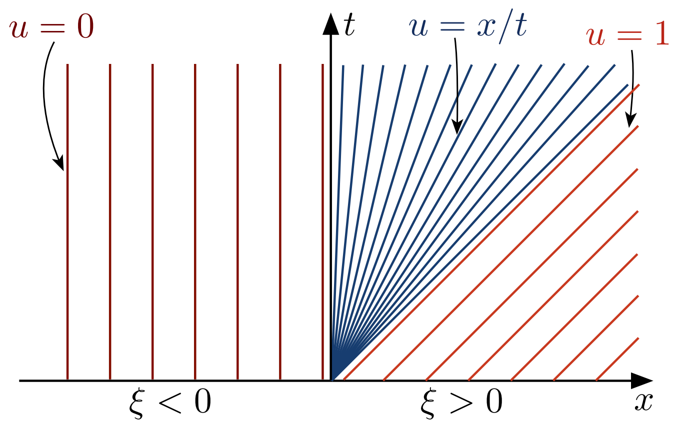


Figure 2.20 Characteristics for the rarefaction wave problem, showing a fan of characteristics.

Shock fitting

We will finish this section by looking at a slightly more complicated example, an example which involves what is called **shock fitting**. We consider the following PDE problem

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.157}$$

$$u(x,0) = \begin{cases} 1 & \text{if } x < 0 \\ -1 & \text{if } 0 < x < 1 \\ 0 & \text{if } x > 1 \end{cases} \tag{2.158}$$

The initial conditions are shown in Fig. 2.21. For this problem, we can easily draw the characteristics emanating from each region. In particular, we have

$$\frac{dx}{dt} = 1, \quad x < 0 \tag{2.159}$$

$$\frac{dx}{dt} = -1, \quad 0 < x < 1 \tag{2.160}$$

$$\frac{dx}{dt} = 0, \quad x > 1 \tag{2.161}$$

leading to the intersecting characteristics shown in Fig. 2.22.

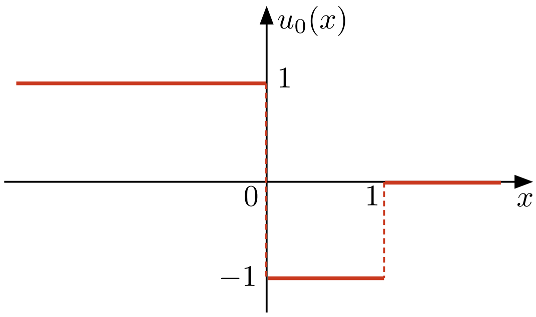


Figure 2.21

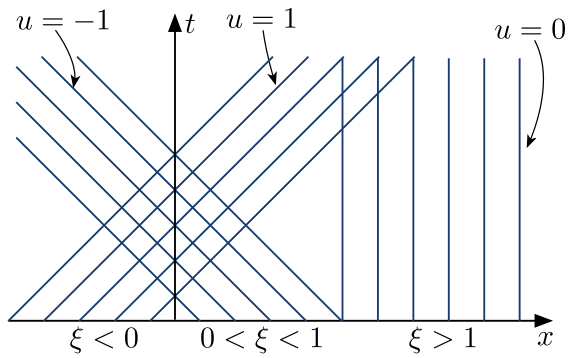


Figure 2.22

We can see that in this problem, we expect two shocks to propagate, initially at $x = 0$ and $x = 1$. Thanks to the Rankine-Hugoniot condition, we can obtain the shock speeds. The $x = 0$ discontinuity propagates with speed

$$s'(t) = \frac{1}{2}(u^+ + u^-) = \frac{1}{2}(-1+1) = 0 \tag{2.162}$$

The shock initially located at $x = 0$ will not propagate and will remain at $x = 0$. We can draw up what the shock will do until $t = 1$; this is shown in Fig. 2.23.

We could only go up to $t = 1$ with this construction because beyond $t = 1$, **we do not know the solution ahead of the shock**. In particular, we need an expansion fan in the void, i.e. a wave joining $u = 0$ for $x > 1$ with $u = -1$ for $x < 1 - t$. The equations of the characteristics of the fan are given by

$$t = -k(x-1) \tag{2.163}$$

where k is a positive constant or equivalently

$$\frac{1-x}{t} = \text{const.} \tag{2.164}$$

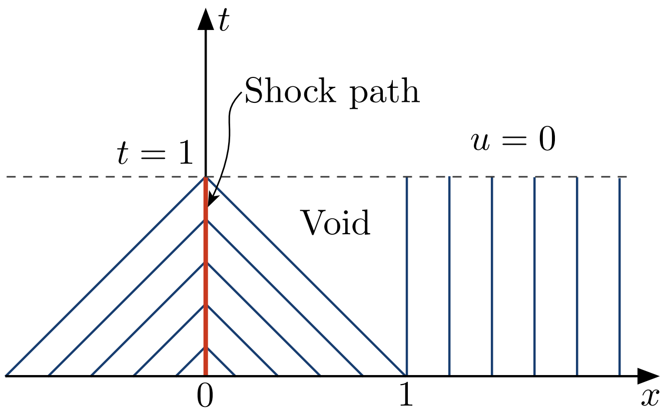


Figure 2.23

The solution in the fan region varies linearly from $u = -1$ at $x = 1 - t$ to $u = 0$ at $x = 1$, so it is given by

$$u(x,t) = \frac{x-1}{t}, \quad 1 - t < x < 1 \tag{2.165}$$

(To convince yourself, it is useful to draw the characteristics in the (x,t)-plane and the profile in the (x,u)-plane).

Now the next question is: what happens to the shock after $t = 1$? Drawing the characteristics close to $t = 1$ will be informative; those are shown in Fig. 2.24.

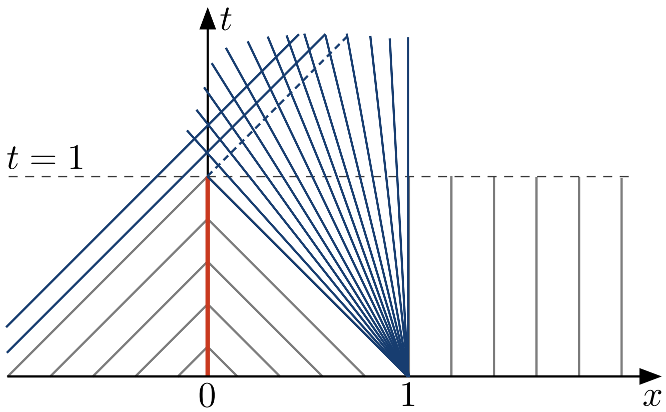


Figure 2.24

Beyond $t = 1$, the shock will connect two states:

- Ahead of the shock, a state with $u^+ = (x-1)/t$
- Behind the shock, a state with $u^- = 1$

The Rankine-Hugoniot condition gives that

$$\frac{ds}{dt} = \frac{1}{2}(u^+ + u^-) = \frac{1}{2}\left[\frac{x-1}{t} + 1\right] \tag{2.166}$$

We can see that here the **shock path will be curved.** Now, since $ds/dt = dx/dt$, we have the following ODE

$$\frac{dx}{dt} - \frac{1}{2t}x = \frac{1}{2}\left(1 - \frac{1}{t}\right) \tag{2.167}$$

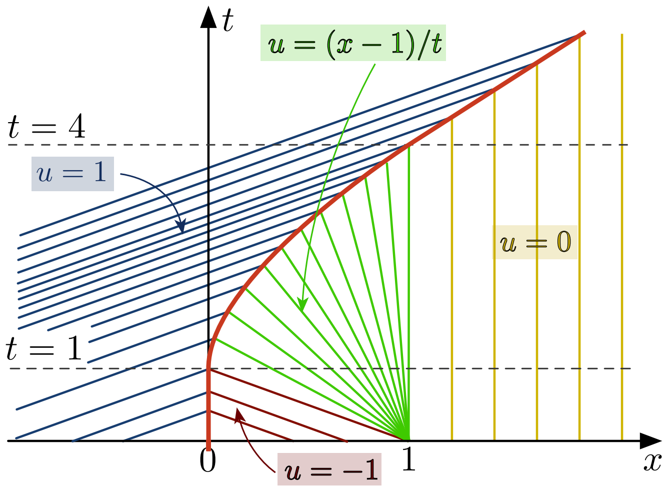


Figure 2.25 Shock fitting problem

subject to $x = 0$ at $t = 1$. This ODE can be solved using an integrating factor and we obtain

$$x = s(t) = t + 1 - 2t^{1/2}, \quad 1 \leq t \leq 4 \tag{2.168}$$

This shock path solution is indeed only valid up to $t = 4$, as at that time, the shock reaches $x = 1$ and hence the solution ahead of the shock changes and is then $u = 0$. What happens then beyond $t = 4$?

The principle is the same again: we have ahead of the shock $u = 0$ and behind of the shock $u = 1$. The speed of the shock is then given by $s'(t) = 1/2$. So the shock path is a solution of the following ODE

$$\frac{dx}{dt} = \frac{1}{2} \tag{2.169}$$

subject to the following initial condition $x = 1$ at $t = 4$, i.e.

$$x(t) = s(t) = \frac{1}{2}t - 1 \tag{2.170}$$

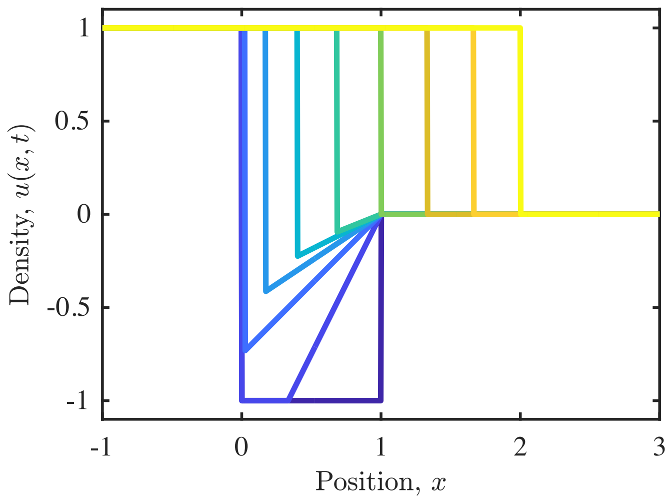


Figure 2.26 Shock fitting problem solution profiles (with time increasing from blue to yellow)

At this point, we can write that the solution of the PDE for $t > 4$ is given by

$$u(x,t) = \begin{cases} 0 & x > \frac{1}{2}t - 1 \\ 1 & x < \frac{1}{2}t - 1 \end{cases} \tag{2.171}$$

We summarize all our findings in Fig. 2.25. From the diagram of characteristics, we can determine the solution at any point (x, t); solution profiles for various times are given in Fig. 2.26.

2.4 Shock fitting and asymptotic behavior of shocks

Let us have a look at how we can obtain the shape of shock solutions in more complicated examples. For this, consider the now infamous inviscid Burgers equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \quad x \in \mathbb{R},\ t > 0 \tag{2.172}$$

Assume that the initial conditions are given by the following "tent" function

$$u(x,0) = f(x) = \begin{cases} u_0(1-x), & 0 < x < 1 \\ u_0(1+x), & -1 < x < 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.173}$$

We will first show that a shock will form and then we will find an explicit solution for this problem. Let us have a look at the characteristics

$$u = f(\xi) \quad \text{on} \quad \frac{dx}{dt} = f(\xi), \quad x(0) = \xi \tag{2.174}$$

Thus, we find that the equation of the characteristics is given by

$$x(t) = f(\xi)t + \xi \tag{2.175}$$

We draw the diagram of characteristics in Fig. 2.28. We see that characteristics emanating from the $0 \leq \xi \leq 1$ interval are crossing the vertical characteristics from $\xi > 1$. In particular, for $\xi = 0$ and $\xi = 1$, the characteristics are given respectively by

$$\xi = 0: \quad x = u_0 t \tag{2.176}$$
$$\xi = 1: \quad x = 1 \tag{2.177}$$

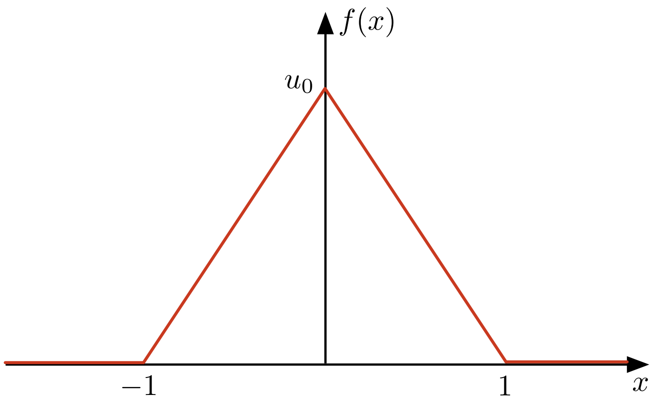


Figure 2.27 "Tent" initial conditions for the shock fitting problem

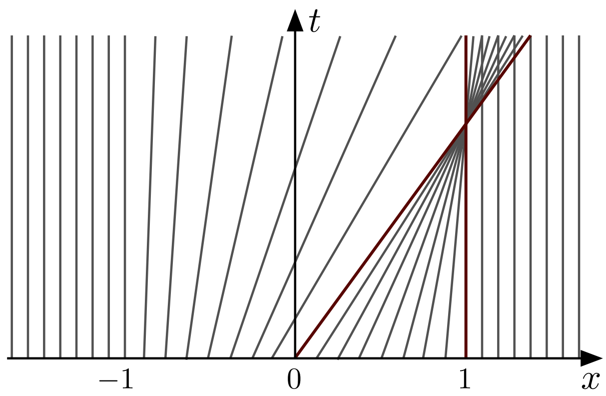


Figure 2.28 Diagram of characteristics for the shock fitting problem.

so these two characteristics cross at $t_s = 1/u_0$.

Now take $0 < \xi < 1$, the equation of the characteristics in this parameter range is written

$$x = u_0(1-\xi)t + \xi \tag{2.178}$$

We could look for the singular time for each of these characteristics by looking at when $|\partial u/\partial x| \to \infty$ like was done in the previous sections. But we can also realize that these characteristics cross when

$$\frac{\partial x}{\partial \xi} = 0 \tag{2.179}$$

Hence, these characteristics all cross when

$$\frac{\partial x}{\partial \xi} = -u_0 t + 1 = 0 \Rightarrow t_s = \frac{1}{u_0} \tag{2.180}$$

We can substitute this value of time back into the equation of the characteristics to obtain that $x_s = u_0(1-\xi)t_s + \xi = 1$. So we conclude that all the characteristics emanating from the interval $0 \leq \xi \leq 1$ cross at $x_s = 1$ and $t_s = 1/u_0$. For $t > 1/u_0$, the solution is triple-valued.

An explicit solution can be found up to $t = t_s$. From (2.175), we find that the implicit solution is given by

$$u(x,t) = f(x - u(x,t)t) \tag{2.181}$$

for an arbitrary function f given by the initial conditions. In this ”tent” example, recall that

$$f(x) = \begin{cases} u_0(1-x), & 0 < x < 1 \\ u_0(1+x), & -1 < x < 0 \\ 0, & |x| > 1 \end{cases} \tag{2.182}$$

so we find that

$$u(x,t) = f(x-u(x,t)t) = \begin{cases} u_0(1-(x-ut)), & 0 < x-ut < 1 \\ u_0(1+(x-ut)), & -1 < x-ut < 0 \\ 0, & |x-ut| > 1 \end{cases} \tag{2.183}$$

We can now solve for u in each case to obtain the explicit solution

$$u(x,t) = \begin{cases} u_0\left(\frac{1-x}{1-u_0 t}\right), & u_0 t < x < 1 \\ u_0\left(\frac{1+x}{1+u_0 t}\right), & -1 < x < u_0 t \\ 0, & \text{otherwise} \end{cases} \tag{2.184}$$

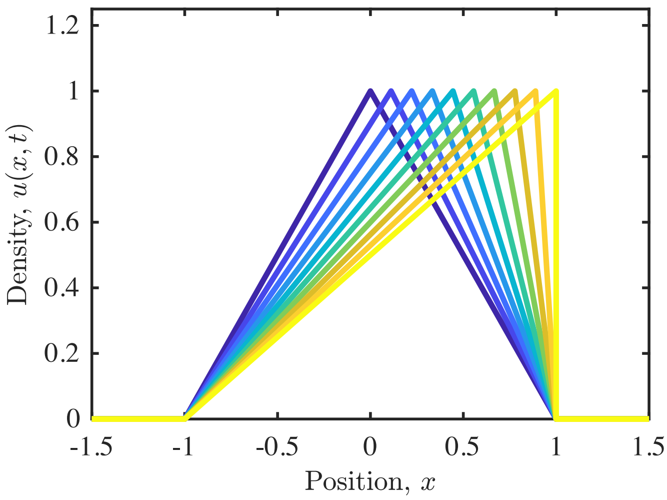


Figure 2.29 Solution profiles until the shock forms (time increases from blue to yellow).

We represent these profiles for various times in Fig. 2.29.

Clearly, this solution goes awry when $t = t_s = 1/u_0$. Indeed, if we ignore that a shock has formed and calculate the solution according to (2.184), we can easily see, as shown on Fig. 2.30, that the solution becomes triple valued.

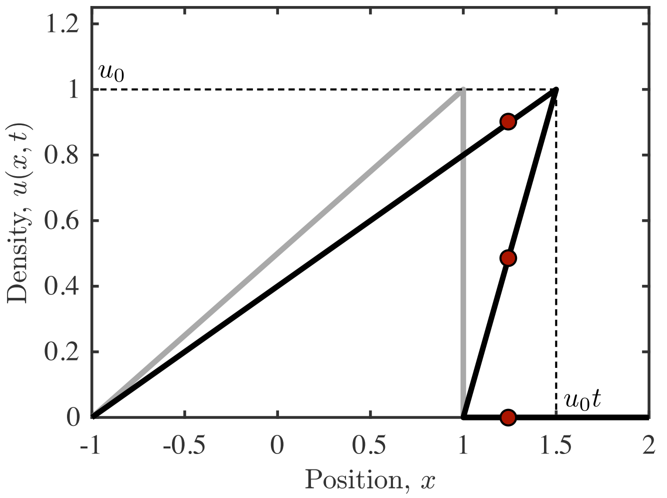


Figure 2.30 Triple-valued solution beyond $t = t_s$.

This solution is thus unphysical but we are going to see now that nevertheless it is connected to the shock path and strength. We have shown above that a shock form at $t_s = 1/u_0$ and propagates afterwards. As we were able to obtain the solution behind and ahead of the shock, using the Ranking-Hugoniot condition we can obtain the shock speed. Recall that it is given by

$$\frac{ds}{dt} = \frac{[u^2/2]}{[u]} = \frac{\frac{1}{2}u_-^2 - 0}{u_- - 0} = \frac{1}{2}u_- \tag{2.185}$$

where u_- is the solution right behind the shock. Now from (2.184), we can write that

$$u_- = u_0 \frac{1+s}{1+u_0 t} \tag{2.186}$$

which lead to the following ODE

$$\frac{ds}{dt} = \frac{1}{2}\frac{u_0}{1+u_0 t} + \frac{1}{2}\left(\frac{u_0}{1+u_0 t}\right) s \Rightarrow \frac{ds}{dt} - \frac{1}{2}\left(\frac{u_0}{1+u_0 t}\right) s = \frac{1}{2}\frac{u_0}{1+u_0 t} \tag{2.187}$$

We can integrate this ODE using an integrating factor. As we have

$$\int \frac{u_0}{1+u_0 t} dt = \ln(1+u_0 t) \tag{2.188}$$

the integrating factor can be written as

$$\exp\left[-\frac{1}{2}\ln(1+u_0 t)\right] = \frac{1}{(1+u_0 t)^{1/2}} \tag{2.189}$$

leading to

$$\frac{d}{dt}\left(\frac{s}{\sqrt{1+u_0 t}}\right) = \frac{1}{2}\frac{u_0}{(1+u_0 t)^{3/2}} \Rightarrow s(t) = -\frac{(1+u_0 t)^{1/2}}{(1+u_0 t)^{1/2}} + K(1+u_0 t)^{1/2} \tag{2.190}$$

But we know that at $t = t_s = 1/u_0$, $s(t_s) = 1$, which gives us that $K = \sqrt{2}$. Finally, we find that the shock path is given by

$$s(t) = -1 + \sqrt{2}(1+u_0 t)^{1/2}, \quad \text{for} \quad t \geq \frac{1}{u_0} \tag{2.191}$$

Given the shock path, we can find the maximum value $u_{\max}$ that the solution takes. Indeed, we have that

$$u_{\max} = u_-|_{x=s(t)} = u_0\left(\frac{1+s}{1+u_0 t}\right) = \frac{\sqrt{2}u_0}{(1+u_0 t)^{1/2}} \underset{t\to\infty}{\to} 0 \tag{2.192}$$

So the shock strength vanishes when $t \to \infty$. Solution profiles beyond $t = t_s$ are given in Fig. 2.31.

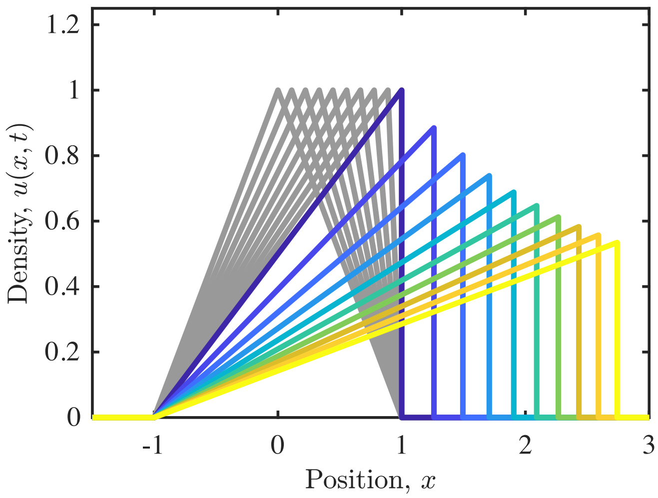


Figure 2.31 Solution before shock formation (grey lines) and after shock formation (time increases from blue to yellow).

If $u(x,t)$ is a mass density, we know that the area under the profile (i.e. the integral over space) is the total mass. Before shock formation, the mass is conserved (remember that we are dealing with a conservation law here!). But what about beyond the singular time?

Considering the shape of the profile, the area under the curve is given by

$$\mathcal{A} = \frac{1}{2}(1+s(t))u_{\max} = \frac{1}{2}\sqrt{2}(1+u_0t)^{1/2}\frac{\sqrt{2}u_0}{(1+u_0t)^{1/2}} = u_0 \tag{2.193}$$

So we conclude that the mass is conserved even after the shock forms. This result is not surprising as the solution has a compact support. Indeed, recall that the conservation law gives

$$\frac{d}{dt}\int_a^b u(x,t)dx = q(a,t) - q(b,t) \tag{2.194}$$

where the flux is defined here by $q(u) = \frac{1}{2}u^2$. If we take the limits $a \to -\infty$ and $b \to \infty$, then $q(\pm\infty) = 0$ and we obtain

$$\frac{d}{dt}\int_{-\infty}^{\infty} u(x,t)dx = 0 \Rightarrow \int_{-\infty}^{\infty} u(x,t)dx = u_0(\text{const.}) \tag{2.195}$$

While we have derived it here for the tent profile, this argument will work for all profiles with compact support. We can generally say that for solutions with compact support, the area under the profile is conserved even after the shock. This leads to a geometrical construction of the shock profile after $t = t_s$, using what is called **Whitham's equal area rule**.

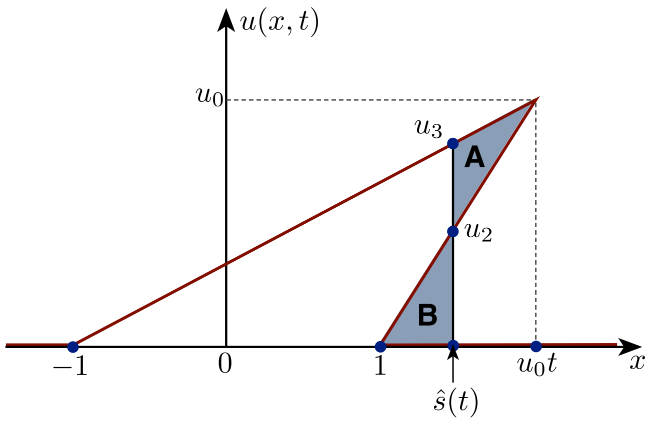


Figure 2.32 Whitham's equal area rule.

Consider the physically inadmissible solution after the shock has formed, i.e. after $t = t_s$ (see Fig. 2.32). We will fit the shock position at $x = \hat{s}(t)$ by picking $\hat{s}(t)$ so that the area of the region A is equal to that of the region B. Mathematically, we can write that

$$\text{Area of } A = \frac{1}{2}(u_3 - u_2)(u_0t - \hat{s}(t)) \tag{2.196}$$

$$\text{Area of } B = \frac{1}{2}u_2(\hat{s}(t) - 1) \tag{2.197}$$

Now, as u_2 and u_3 are known in terms of $\hat{s}(t)$ from the explicit solution (2.184), we can solve the equation

$$\text{Area of } A = \text{Area of } B \tag{2.198}$$

which is a quadratic equation in $\hat{s}(t)$. You are encouraged to do this calculation fully and check the result against our previous result. This equal area rule is very general and works for other initial conditions, e.g. smooth initial conditions with compact support (see Fig. 2.33).

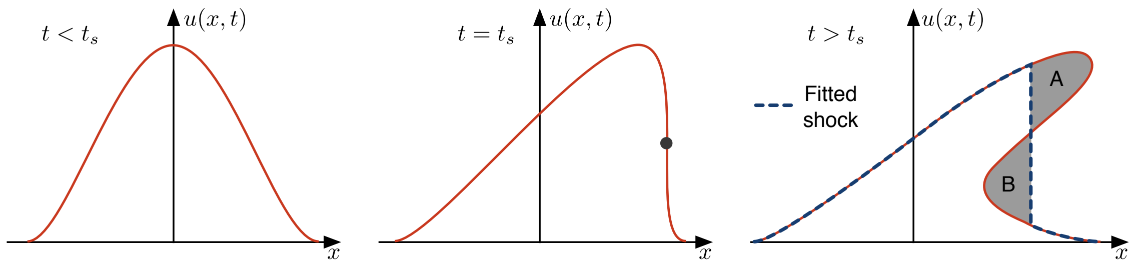


Figure 2.33 Example of shock fitting with smooth initial conditions.

2.5 Traffic flow

In the previous section, we have reported general results for quasilinear first-order equations. There are many examples of applications of this type of nonlinear conservation laws. Here, we will use what we just learnt to model **traffic flow**.

2.5.1 Conservation law

Here, we will model the one-dimensional flow of cars on a long road. We will assume that the road is fairly congested. Let us introduce some terminology. We will denote $\rho(x,t)$ the linear density of cars, i.e. the number of cars per unit length (e.g. measured in number of cars per km). Further, the **traffic flow or flux**, i.e. the number of cars per unit time passing a given fixed point x, will be denoted $q(x,t)$; it will be measured in number of cars per hour, for instance.

If we assume that there are no entrances or exits on the highway and pick two arbitrary points a and b, the conservation principle tells us that the **rate of change of cars in the interval $[a,b]$ is equal to the number of cars entering per unit time at $x = a$ minus the number of cars leaving per unit time at $x = b$.** Mathematically, this statement reads

$$\frac{d}{dt}\int_a^b \rho(x,t)dx = q(a,t)-q(b,t) \Rightarrow \frac{d}{dt}\int_a^b \rho(x,t)dx = -\int_a^b \frac{\partial q}{\partial x}dx \Rightarrow \int_a^b \left[\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x}\right] dx = 0 \tag{2.199}$$

It follows from (2.199) that if the integrand is continuous, then it must be zero since a and b are arbitrary. The density of cars on the highway is thus governed by the following conservation of cars equation

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{2.200}$$

2.5.2 Empirical models for the flux function

To complete this equation, we need to specify the flux function. A physical observation can help us for that: the number of cars passing a given point x per unit time must be equal to the density of cars times their velocity at that point and time, i.e.

$$q = \rho v \tag{2.201}$$

where $v(x,t)$ is the car velocity. We can easily confirm to be dimensionally correct.

Guided by our own experience of road trips, we assume that the car velocity depends on the density and that it decreases as the density increases (cars slow down in heavy traffic). Mathematically, we have assumed

$$v \equiv v(\rho) \quad \text{and} \quad \frac{dv}{d\rho} \leq 0 \tag{2.202}$$

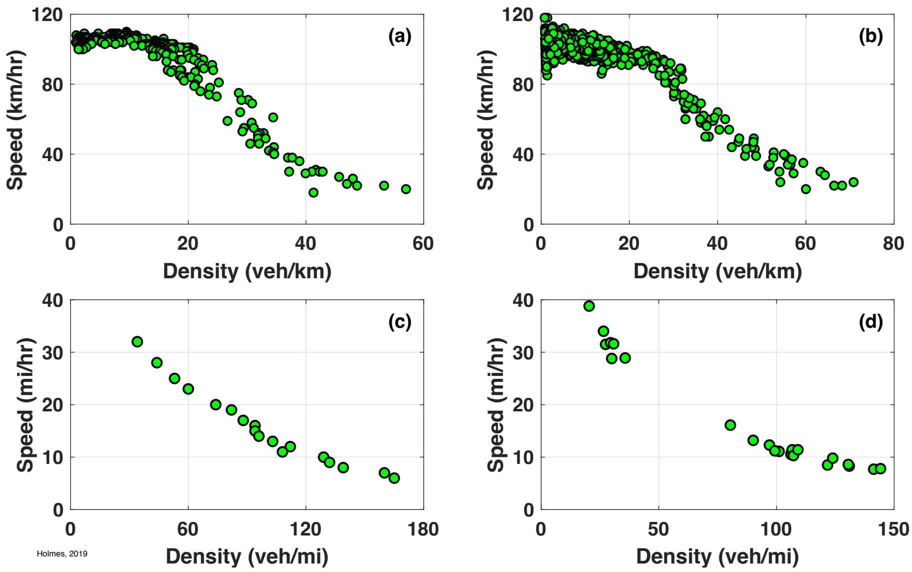


Figure 2.34 Velocity as a function of the density measured for different roadways: (a) highway near Toronto, (b) a highway near Amsterdam, (c) the Lincoln Tunnel (New York City) and (d) the Merritt Parkway (NY-CT, USA). Adapted from (Holmes, 2019).

Experimental observations confirm this intuition. Fig. 2.34 reports scatter plots of the observed car velocities in function of the car density for various roadways. The question which remains is what function best describes this data. Obviously, the simplest assumption would be to consider a constant velocity but it does not seem very realistic.

In the general case, the flux will be written $q = \rho v(\rho)$; if one assumes that v is a smooth function of ρ, by the chain rule, we write

$$\frac{\partial q}{\partial x} = q'(\rho)\frac{\partial \rho}{\partial x} \tag{2.203}$$

so the general form of the conservation law is given by

$$\frac{\partial \rho}{\partial t} + c(\rho)\frac{\partial \rho}{\partial x} = 0 \tag{2.204}$$

where

$$c(\rho) \equiv q'(\rho) = v(\rho) + \rho v'(\rho) \tag{2.205}$$

is known as the **wave velocity**. When deciding then, what functional form $v(\rho)$ should take, one should consider the experimental evidence but also whether the resulting problem has a solution. It will be of no use to derive the most accurate model of traffic flow, if this model is unsolvable! In modelling, one should try to extract the main features in the otherwise noisy experimental data. Here, the velocity decays monotonically with the density.

The simplest constitutive law which is not a constant would then be linear. For the traffic problem, this means that we would assume $v(\rho) = a - b\rho$, where a, b are constants. The most widely used constitutive law in traffic modelling is called the **Greenshields model** and it is written

$$v(\rho) = v_m\left(1 - \frac{\rho}{\rho_m}\right) \tag{2.206}$$

where v_m and ρ_m are the maximum velocity and density, respectively. The value of these constants can be obtained by fitting the model to the experimental data. As shown on

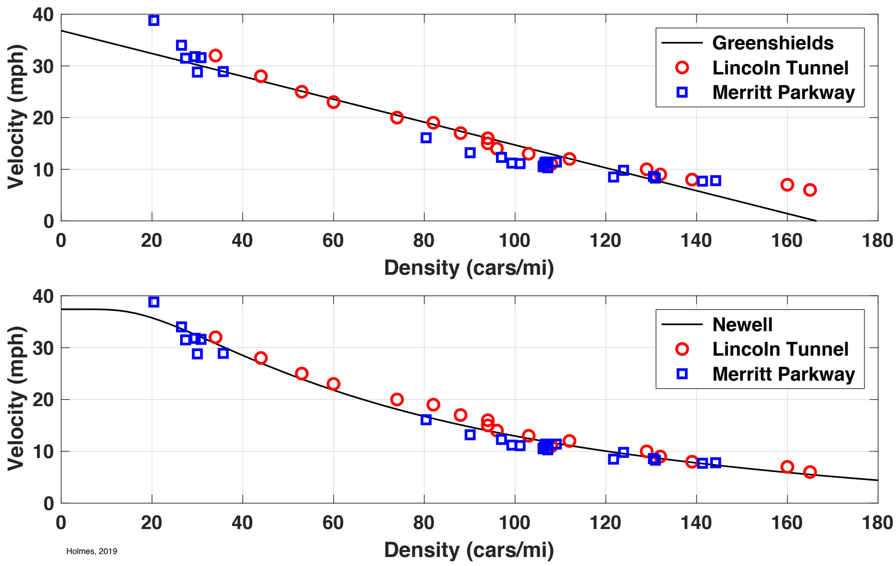


Figure 2.35 (*Top*)Greenshields law and (*Bottom*) Newell law fitted to the the Lincoln tunnel and Merritt parkway data. Adapted from (Holmes, 2019).

Fig. 2.35, we can see that, while not perfect, the linear constitutive law reproduces relatively well the Lincoln tunnel and Merritt parkway data. Strikingly, it will fail at reproducing the plateau observed in the velocity near $\rho = 0$ in the Toronto and Amsterdam data. Nevertheless, it does have the most important features of the system:

- Velocity monotonically decays with density;
- Velocity is maximal when $\rho = 0$;
- Velocity goes to zero at some maximum jamming density ρ_m (i.e. bumper-to-bumper traffic).

As far as we are concerned, it sounds like an acceptable approximation which finally leads to the following traffic flow equation

$$\frac{\partial \rho}{\partial t} + c(\rho)\frac{\partial \rho}{\partial x} = 0 \quad \text{with} \quad c(\rho) = v_m\left(1 - \frac{2\rho}{\rho_m}\right) \tag{2.207}$$

a quasilinear PDE.

This model leads to the car velocity, flux and wave velocity dependence with car density shown in Fig. 2.36. More realistic models (those fitting the experimental data more

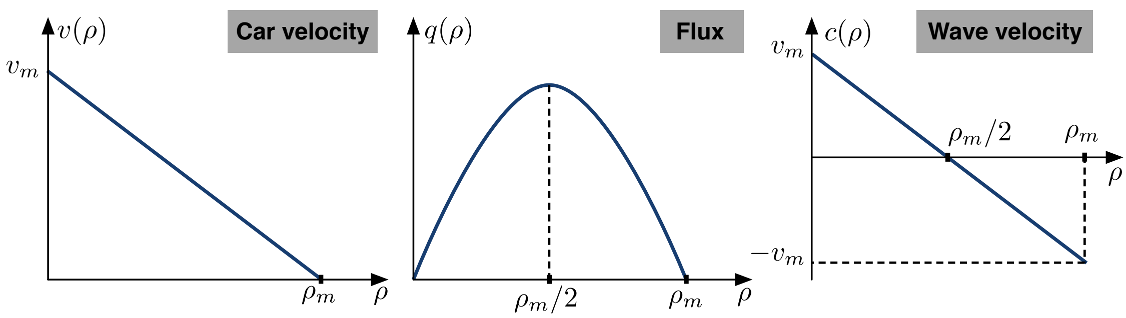


Figure 2.36

faithfully) would have a more asymmetric flux function than that in Fig. 2.36 while preserving necessary features, i.e. a zero flux at both $\rho = 0$ and $\rho = \rho_m$. Using Greenshield's law to fit the Lincoln tunnel and Merritt parkway data, one obtains a maximum car density $\rho_m = 166.4$ cars/mile and a maximum velocity $v_m = 36.8$ mph. Finally, the maximum flux would then be $q_m \approx 1530$ cars/hour and a velocity giving the maximum flux given by $v^* = 18.4$ mph.

Remark. *Another constitutive relation was proposed in 1961 by Newell to fit the velocity-density data. Newell law is written*

$$v = v_m \left(1 - e^{-\lambda(1/\rho - 1/\rho_m)}\right) \tag{2.208}$$

While this proposed constitutive law may fit better the data (see Fig. 2.35), for instance it does reproduce the plateau observed close to $\rho = 0$ that is seen in the Toronto and Amsterdam data, it comes at the cost of a much more complicated wave velocity

$$c(\rho) = v_m \left[1 - \left(1 + \frac{\lambda}{\rho}\right) e^{-\lambda(1/\rho - 1/\rho_m)}\right] \tag{2.209}$$

It is then natural to ask yourself whether the increased complexity in the traffic flow equation resulting from this more faithful constitutive law is worthwhile. The answer is not easy and may depend on the question you are trying to answer.

In what follows, we will have a look at two canonical problems.

2.5.3 Red light turning green

Let us use our new shiny traffic flow model (2.207) to try to understand what happens when a red light turns green. Our traffic flow equation reads

$$\frac{\partial \rho}{\partial t} + v_m \left(1 - \frac{2\rho}{\rho_m}\right) \frac{\partial \rho}{\partial x} = 0 \tag{2.210}$$

To start, suppose that a bumper-to-bumper traffic is standing at a red light, placed at $x = 0$, while the road ahead is empty. The initial density profile is then given by

$$f(x_0) = \begin{cases} \rho_m, & \text{for } x_0 \leq 0 \\ 0, & \text{for } x_0 > 0 \end{cases} \tag{2.211}$$

At time $t = 0$, the traffic light turns green. What is the evolution of the car density for $t > 0$? At first, only the cars close to the traffic light start moving while most remain standing. In this model, the local wave speed is given by

$$c(f(x_0)) = \begin{cases} -v_m, & \text{for } x_0 \leq 0 \\ v_m, & \text{for } x_0 > 0 \end{cases} \tag{2.212}$$

and so the characteristics are the straight lines given by

$$x = -v_m t + x_0, \quad \text{if } x_0 \leq 0 \tag{2.213}$$

$$x = v_m t + x_0, \quad \text{if } x_0 > 0 \tag{2.214}$$

The characteristics are shown in Fig. 2.37. One can see that the lines $x = v_m t$ and $x = -v_m t$ are partitioning the half-plane in three regions L, C and R.

In the L sector of the plane, $\rho(x,t) = \rho_m$, while inside the R sector, we have $\rho(x,t) = 0$. At a given time $t = t^*$, we can conclude that the front car is located at the point

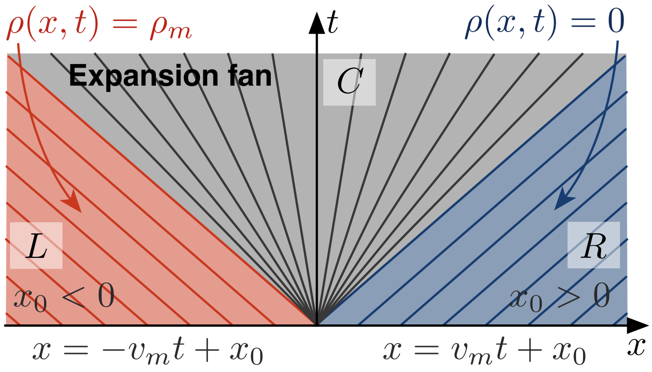


Figure 2.37

$x^* = v_m t^*$, so the traffic has not yet arrived in the sector R. The front car moves at maximum speed since the road ahead is empty. Similarly, the cars placed at the points $(x, t^*) \in L$ are still standing. The first car that starts moving at time $t = t^*$ is at the point $x^* = -v_m t^*$. So we can see that **the green light signal propagates back through the traffic at the speed** v_m. You are commonly experiencing this when at a traffic light; indeed when the traffic light turns green, you need to wait some time until you can start moving and this time is longer the further away from the traffic light you are.

What is the density in the C sector? As we did before in similar cases, we will place in this sector a **rarefaction fan**. In the fan region, the characteristics in this problem are written in the form

$$x = v_m t \left(1 - \frac{2\rho(x,t)}{\rho_m}\right) \tag{2.215}$$

They all pass through the origin and are parametrized by ρ that varies between $\rho = 0$ at the front and $\rho = \rho_m$ at the back of the fan.

The solution can then be found by solving (2.215) for ρ given x and t. The solution is thus easily found to be

$$\rho(x,t) = \frac{\rho_m}{2}\left(1 - \frac{x}{v_m t}\right) \tag{2.216}$$

Note that since $c(\rho) = v_m\left(1 - 2\frac{\rho}{\rho_m}\right)$, we can write the solution in the expansion fan as

$$\rho(x,t) = r(x/t) \tag{2.217}$$

where $r = c^{-1}$ is the inverse function of c the wave speed. This is the general form of a rarefaction wave centered at the origin for a conservation law.

We represent car density profiles for various times in Fig. 2.38.

Remark. *Some of you may wonder how we can justify the existence of the expansion fan in the region C. After all, up until now, I have just told you that this was the solution and asked you to accept it. For the curious among you, here is a possible justification.*

To find the solution in the region C, here is a strategy which should give us a reasonable answer:

1. *first, approximate the initial data by a continuous function f_ε, which converges to f as $\varepsilon \to 0$ at every point x, except $x = 0$;*
2. *second, construct the solution ρ_ε of the ε-problem by the method of characteristics;*
3. *third, let $\varepsilon \to 0$ and check that the limit of ρ_ε is a solution of the original problem.*

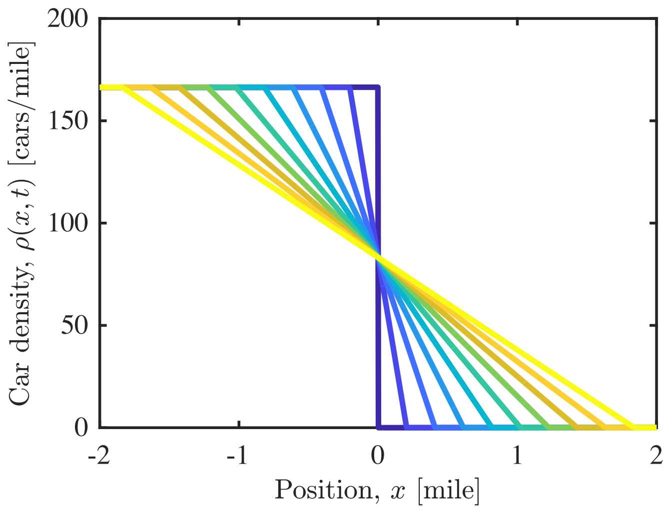


Figure 2.38 Car density in the green light problem for various times; time increases from $t = 0$ min (blue) to $t = 3$ min (yellow). Note that here we have used the Lincoln tunnel data, i.e. $\rho_m = 166.4$ cars/mile and $v_m = 36.8$ mph. The front car has driven 1.85 miles in these 3 min.

So first, let us choose f_ε as the function given by

$$f_\varepsilon(x) = \begin{cases} \rho_m, & x \le 0 \\ \rho_m (1 - x/\varepsilon), & 0 < x < \varepsilon \\ 0, & x \ge \varepsilon \end{cases} \tag{2.218}$$

i.e. that this function linear connects the states at ρ_m and 0 over the range ε. When $\varepsilon \to 0$, $f_\varepsilon(x) \to f(x)$ for all $x \neq 0$. Now, the characteristics for the ε-problem are given by

$$x = -v_m t + x_0, \qquad \textit{if} \quad x_0 < 0 \tag{2.219}$$

$$x = -v_m \left(1 - 2\frac{x_0}{\varepsilon}\right) t + x_0, \qquad \textit{if} \quad 0 \le x_0 < \varepsilon \tag{2.220}$$

$$x = v_m t + x_0, \qquad \textit{if} \quad x_0 \ge \varepsilon \tag{2.221}$$

Indeed, for $0 \le x_0 < \varepsilon$, we have that

$$c(f_\varepsilon(x_0)) = v_m \left(1 - \frac{2 f_\varepsilon(x_0)}{\rho_m}\right) = -v_m \left(1 - 2\frac{x_0}{\varepsilon}\right) \tag{2.222}$$

The characteristics in the region $-v_m t < x < v_m t + \varepsilon$ form a rarefaction fan. Now clearly, we have that $\rho_\varepsilon(x,t) = 0$ for $x \ge v_m t + \varepsilon$ and $\rho_\varepsilon(x,t) = \rho_m$ for $x \le -v_m t$. In the region $-v_m t < x < v_m t + \varepsilon$, we can solve for x_0 in the characteristic equation to obtain

$$x_0 = \varepsilon \frac{x + v_m t}{2 v_m t + \varepsilon} \tag{2.223}$$

Then, we obtain that

$$\rho_\varepsilon(x,t) = f_\varepsilon(x_0) = \rho_m (1 - \frac{x_0}{\varepsilon}) = \rho_m \left(1 - \frac{x + v_m t}{2 v_m t + \varepsilon}\right) \tag{2.224}$$

Finally, we can let $\varepsilon \to 0$ and we obtain

$$\rho(x,t) = \begin{cases} \rho_m, & x \le -v_m t \\ \frac{\rho_m}{2}\left(1 - \frac{x}{v_m t}\right), & -v_m t < x < v_m t \\ 0, & x \ge v_m t \end{cases} \tag{2.225}$$

It is then easy to check that ρ is a solution of the PDE problem in the regions L, C and R.

2.5.4 Green light turning red

Consider now the case where we have a simple situation with uniform density $\rho_0 < \rho_m$. All cars moving on the roadway are thus moving with velocity

$$v(\rho_0) = v_m \left(1 - \frac{\rho_0}{\rho_m}\right) < v_m \tag{2.226}$$

Consider that at $x = 0$, there is a traffic light; at $t = 0$, the **green light turns red**! Let us assume that the car density is governed by (2.210). In this problem initial and boundary conditions are given by

$$\rho(x,0) = \rho_0, \quad \text{for} \quad x < 0 \tag{2.227}$$

$$\rho(0,t) = \rho_m, \quad \text{for} \quad t > 0 \tag{2.228}$$

Indeed, we are not interested in what happens ahead of the light. The fact that the light turned red is **irrelevant to the cars that made it through**! The **characteristics** of this problem are solutions to the following equation

$$\frac{dx}{dt} = v_m \left(1 - 2\frac{\rho}{\rho_m}\right) \tag{2.229}$$

For $x < 0$, we obtain

$$\frac{dx}{dt} = v_m \left(1 - 2\frac{\rho_0}{\rho_m}\right) \Rightarrow x = v_m \left(1 - 2\frac{\rho_0}{\rho_m}\right) t + x_0, \quad x_0 < 0 \tag{2.230}$$

The slope of these characteristics can be positive or negative depending on whether $\rho < \rho_m/2$ or $\rho > \rho_m/2$, respectively. The characteristics from $x = 0$ are given by

$$\frac{dx}{dt} = v_m \left(1 - 2\frac{\rho_m}{\rho_m}\right) = -v_m \Rightarrow x = -v_m t + v_m t_0, \quad t_0 > 0 \tag{2.231}$$

The slope of these characteristics is $-1/v_m$; hence, if $c(\rho_0) > 0$, the characteristics will certainly cross and a shock will form.

What if $\rho_0 > \rho_m/2$, the slope of the characteristics for $x_0 < 0$ is given by

$$\frac{1}{v_m(1 - 2\rho_0/\rho_m)} = -\frac{1}{v_m(2\rho_0/\rho_m - 1)} < -\frac{1}{v_m} \tag{2.232}$$

and so the characteristics would still cross and a shock would form. We provide in Fig. 2.39 the diagram of characteristics in both cases.

We just concluded that **in all cases a shock will form**. Let us calculate the location of the shock. The **Rankine-Hugoniot condition** here gives us that

$$\frac{ds}{dt} = \frac{[q]}{[\rho]} \tag{2.233}$$

where we have used the jump notation. Now let's recall that the flux is defined as

$$q(\rho) = \rho v(\rho) = v_m \rho \left(1 - \frac{\rho}{\rho_m}\right) \tag{2.234}$$

Given the state of the system behind and ahead of the shock, we obtain

$$\frac{ds}{dt} = \frac{q(\rho_0) - q(\rho_m)}{\rho_0 - \rho_m} = -\frac{q(\rho_0)}{\rho_m - \rho_0} < 0 \tag{2.235}$$

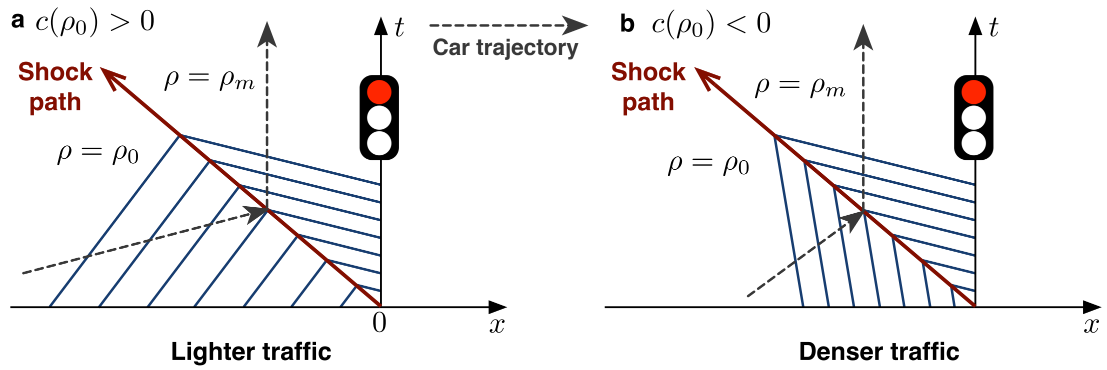


Figure 2.39 Diagram of characteristics for **(a)** the case where $c(\rho_0) > 0$ (lighter initial traffic) and **(b)** the case where $c(\rho_0) < 0$ (denser initial traffic)

We can then see that the shock propagates backwards in the traffic flow. As we know that the shock passes through $x = 0$ at $t = 0$, we can then obtain the shock path and write

$$x_s(t) = -\frac{q(\rho_0)}{\rho_m - \rho_0} t = \frac{\rho_0 v_m}{\rho_m - \rho_0} \left(1 - \frac{\rho_0}{\rho_m}\right) t = -\frac{\rho_0}{\rho_m} v_m t \tag{2.236}$$

We then conclude that the solution is given by

$$\rho(x,t) = \begin{cases} \rho_0, & x < x_s(t) \\ \rho_m, & x > x_s(t) \end{cases} \tag{2.237}$$

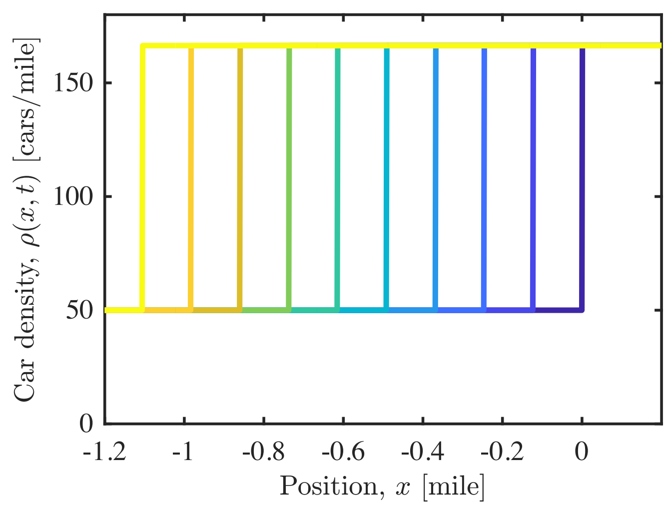


Figure 2.40 Solution to the green light turning red problem.

2.5.5 Towards a more realistic model

In this section, we have made a very simple modelling assumption. We have stated that the velocity of a car was only dependent on the local density $v \equiv v(\rho)$. However, it is natural to think that drivers in our traffic flow would slow down when they see increased (relative) density ahead of them; mathematically, this means that an appropriate model for the velocity would be given by

$$\tilde{v}\left(\rho, \frac{\partial \rho}{\partial x}\right) \equiv v(\rho) - \frac{\nu}{\rho}\frac{\partial \rho}{\partial x} \tag{2.238}$$

This new model for the velocity leads to a new flow rate of cars

$$\tilde{q}\left(\rho, \frac{\partial \rho}{\partial x}\right) \equiv \rho v(\rho) - \nu \frac{\partial \rho}{\partial x} = q(\rho) - \nu \frac{\partial \rho}{\partial x} \tag{2.239}$$

Finally, this means that our governing equation now reads

$$\frac{\partial \rho}{\partial t} + c(\rho) \frac{\partial \rho}{\partial x} = \nu \frac{\partial^2 \rho}{\partial x^2} \tag{2.240}$$

where $c(\rho)$ is defined as before. When comparing (2.204) and (2.240), one easily realizes that the difference relies in the addition in the latter of a diffusive term (the term in ρ_{xx}), also called a viscous term (which comes from fluid dynamics models, then ν is called the viscosity). A small amount of diffusion or viscosity usually makes the mathematical model more realistic. Note that if ν is small, the diffusion term $\nu\rho_{xx}$ becomes relevant only when ρ_{xx} is large, that is in a region where ρ_x changes rapidly, i.e. in the region where the shock occurs. The effect of this viscous/diffusion term is most easily explained with the Burgers equation; this is what we do next.

2.6 Viscous effects and Burger's equation

Several times throughout this Chapter, we have looked into what we called the **inviscid Burgers equation**. It is interesting to note that this equation is a special of the more general **Burger's equation** (or **viscous Burgers equation**). The viscous Burgers equation is one of the most studied example of nonlinear diffusion equation. In 1948, Burgers derived this celebrated equation as a simplified form of the Navier-Stokes equation with the goal of studying some aspects of turbulence. The viscous Burgers equation also appears in gas dynamics, in the theory of sound waves and in more involved models in traffic flow modelling. This second-order equation is generally written in the following form

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2} \tag{2.241}$$

The difference between this equation and the inviscid case is the addition of a second-order derivative of u in space; we have seen in Section 2.1.3 that this is a diffusion term and that diffusion tends to spread and smooth profiles. As a consequence, this equation constitutes a basic example of competition between **dissipation** (due to the linear diffusion) and **steepening** (like during shock formation due to the nonlinear transport term).

2.6.1 Travelling wave solution

Consider the viscous Burgers equation (2.241). It is clear that if $\nu = 0$, one recovers the inviscid Burger's equation. We have shown in the previous sections that this equation may lead to shock formation and propagation. What happens when $\nu \neq 0$?

First, let's consider solutions satisfying the following boundary conditions

$$u \to u_1, \quad \text{as } x \to -\infty \tag{2.242}$$
$$u \to u_2, \quad \text{as } x \to +\infty \tag{2.243}$$

with $u_1 > u_2$. Let us see if this boundary value problem admits nonlinear travelling waves.

For that suppose that a travelling wave solution exists, i.e. a wave with a permanent form

$$u(x,t) \equiv u(x - ct) \tag{2.244}$$

We can transform the independent variables $(x,t) \to \eta = x - ct$. In this case, the partial derivatives read

$$\frac{\partial}{\partial t} = \frac{\partial \eta}{\partial t}\frac{d}{d\eta} = -c\frac{d}{d\eta} \quad \text{and} \quad \frac{\partial}{\partial x} = \frac{\partial \eta}{\partial x}\frac{d}{d\eta} = \frac{d}{d\eta} \tag{2.245}$$

As a consequence, the viscous Burgers equation can be rewritten

$$-cu' + uu' = \nu u'' \tag{2.246}$$

where the prime denotes a derivative with respect to η. If we integrate once, we obtain

$$-cu + \frac{1}{2}u^2 = \nu u' + K \tag{2.247}$$

where K is an integration constant. We can evaluate this constant using the boundary conditions. Indeed, we know that in the limit where $\eta \to \pm\infty$ (i.e. in the limit of $x \to \pm\infty$), we have $u'(\eta) \to 0$. So we can write

$$-cu_1 + \frac{1}{2}u_1^2 = K = -cu_2 + \frac{1}{2}u_2^2 \tag{2.248}$$

In particular, this allows us to write that the wave speed is given by

$$c = \frac{1}{2}\frac{u_1^2 - u_2^2}{u_1 - u_2} = \frac{\left[\frac{1}{2}u^2\right]_{+\infty}^{-\infty}}{[u]_{+\infty}^{-\infty}} \tag{2.249}$$

which should look familiar! This simplifies to

$$c = \frac{1}{2}(u_1 + u_2) \tag{2.250}$$

Provided the expression for c, we can now obtain that

$$K = -\frac{1}{2}(u_1 + u_2)u_1 + \frac{1}{2}u_1^2 = -\frac{1}{2}u_1 u_2 \tag{2.251}$$

We finally obtain the following ODE for u

$$\nu\frac{du}{d\eta} = \frac{1}{2}u^2 - \frac{1}{2}(u_1 + u_2)u + \frac{1}{2}u_1 u_2 \tag{2.252}$$

In Problem Sheet #3, you will solve this separable ODE to finally show that

$$u(x,t) = \frac{u_2 + u_1 e^{-\frac{(u_1-u_2)(x-ct)}{2\nu}}}{1 + e^{-\frac{(u_1-u_2)(x-ct)}{2\nu}}} \tag{2.253}$$

We have exhibited here a travelling wave solution to the viscous Burgers equation, whose profiles for various times are shown in Fig. 2.41.

Remark. *Fig. 2.41 shows that as $\nu \to 0$, the solution profile seems to be converging to the shock solution we have previously obtained in the case of the inviscid Burgers equation. This observation forms the basis of the so-called vanishing viscosity method.*

Here, we have looked for a particular type of solution to a particular problem. Can we do better than that? The answer is yes!

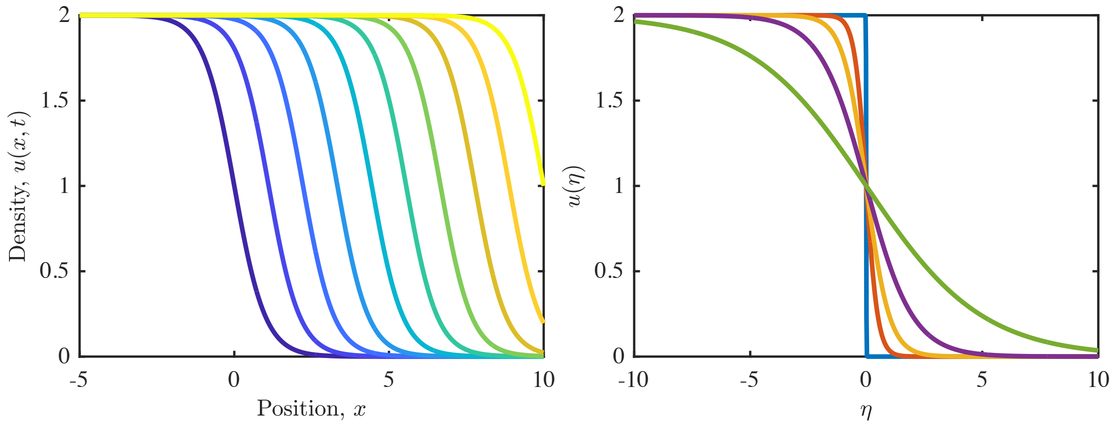


Figure 2.41 (*Left*) Solution profiles for various times (time increasing from blue to yellow) for $u_1 = 2$, $u_2 = 0$ and $\nu = 1$ leading to a wave travelling to the right with speed $c = 1$. (*Right*) Profiles $u(\eta)$ for various values of the viscosity $\nu = 0.01$ (blue), 0.5 (orange), 1 (yellow), 2 (purple), 5 (green). Decreasing the viscosity sharpens the solution profile.

2.6.2 Cole-Hopf transformation

The success of Burgers equation is in large part due to the fact that the **initial value problem can be solved analytically**. In fact, using the so-called **Cole-Hopf transformation**, one can reduce this problem to a diffusion equation. Let us see how this can be done. First, rewrite the viscous Burgers equation as

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left[\frac{1}{2}u^2 - \nu\frac{\partial u}{\partial x}\right] = 0 \tag{2.254}$$

Now in the plane (x, t), we define the planar vector field

$$\mathbf{F} = \left(-u, \frac{1}{2}u^2 - \nu\frac{\partial u}{\partial x}\right) \tag{2.255}$$

The curl of this planar field is given by

$$\frac{\partial F_x}{\partial t} - \frac{\partial F_t}{\partial x} = -\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left[\frac{1}{2}u^2 - \nu\frac{\partial u}{\partial x}\right] = 0 \tag{2.256}$$

This vector field is therefore curl-free and we know that there exists a potential $\psi(x, t)$ such that

$$\frac{\partial \psi}{\partial x} = -u \quad \text{and} \quad \frac{\partial \psi}{\partial t} = \frac{1}{2}u^2 - \nu\frac{\partial u}{\partial x} \tag{2.257}$$

Combining these two definitions, we find that ψ is solution to the equation

$$\frac{\partial \psi}{\partial t} = \frac{1}{2}\left(\frac{\partial \psi}{\partial x}\right)^2 + \nu\frac{\partial^2 \psi}{\partial x^2} \tag{2.258}$$

Now this is still a nonlinear equation and it may appear like we have not really simplified our problem; it would make our lives much easier if we could get rid of the nonlinear term! So let us write $\psi(x, t) = g(\varphi(x, t))$; by the chain rule, we have

$$\frac{\partial \psi}{\partial t} = g'(\varphi)\frac{\partial \varphi}{\partial t}, \quad \frac{\partial \psi}{\partial x} = g'(\varphi)\frac{\partial \varphi}{\partial x}, \quad \text{and} \quad \frac{\partial^2 \psi}{\partial x^2} = g''(\varphi)\left(\frac{\partial \varphi}{\partial x}\right)^2 + g'(\varphi)\frac{\partial^2 \varphi}{\partial x^2} \tag{2.259}$$

Substituting these expressions in (2.258), we obtain the following expression

$$g'(\varphi)\left[\frac{\partial \varphi}{\partial t} - \nu\frac{\partial^2 \varphi}{\partial x^2}\right] = \left[\frac{1}{2}(g'(\varphi))^2 + \nu g''(\varphi)\right]\left(\frac{\partial \varphi}{\partial x}\right)^2 \tag{2.260}$$

In particular, we can see that if we take $g(\varphi) = 2\nu \ln \varphi$, we have

$$\frac{1}{2}(g'(\varphi))^2 + \nu g''(\varphi) = 0 \tag{2.261}$$

and the RHS of the previous equation cancels out, ridding us of the nonlinear term. In this case, our equation reduces to

$$\frac{\partial \varphi}{\partial t} - \nu \frac{\partial^2 \varphi}{\partial x^2} = 0 \tag{2.262}$$

which is a linear diffusion equation for the auxiliary function $\varphi(x,t)$. As ψ was defined such that $u = -\partial\psi/\partial x$, we define the **Cole-Hopf transformation** as the introduction of the auxiliary function $\varphi(x,t)$ such that

$$u = -\frac{2\nu}{\varphi}\frac{\partial \varphi}{\partial x} \tag{2.263}$$

or equivalently

$$\varphi(x,t) = \exp\left[-\frac{1}{2\nu}\int u(x,t)dx\right] \tag{2.264}$$

Via the Cole-Hopf transformation, solving the viscous Burgers equation for the original function $u(x,t)$ is converted to solving a linear diffusion equation for the auxiliary function φ. Given an initial condition $u(x,0) = u_0(x)$, one also needs to transform it into initial conditions for the auxiliary function $\varphi(x,t)$.

In conclusion, we can see that if we knew how to solve a simple linear diffusion equation, we would be able to analytically solve an important nonlinear problem. **It sounds like it may be the perfect time to introduce the Chapter on second-order PDEs...**

*[This version was compiled on **February 15, 2024**.]*