The method of characteristics

(Black provides ambience; blue is background; red is righteous (i.e. the good stuff - the examples); green is go (i.e. try it); purple refers to additional examples solved in the lectures (hopefully).)

Consider the PDE,

$$u_t + cu_x = s, \qquad u(x,0) = f(x). \tag{1}$$

To begin, we won't supply any boundary conditions and assume that we're solving (1) on the infinite line with $u \to 0$ as $x \to \pm\infty$. The coefficients c and s (which is not to be confused with the Laplace transform variable of previous lectures) are taken to be constants for now.

Now, consider the new variables, (x_0, τ), with $(x_0, \tau) = (x - ct, t)$, or, equivalently, $(x,t) = (x_0 + c\tau, \tau)$. According to the chain rule we have

$$\frac{\partial}{\partial \tau} = \frac{\partial t}{\partial \tau}\frac{\partial}{\partial t} + \frac{\partial x}{\partial \tau}\frac{\partial}{\partial x} = \frac{\partial}{\partial t} + c\frac{\partial}{\partial x} \qquad \left(\frac{\partial t}{\partial \tau} = 1 \;\&\; \frac{\partial x}{\partial \tau} = c\right). \tag{2}$$

If we therefore change variables from (x,t) to (x_0,τ) and put $U(x_0,\tau) = u(x,t)$, the PDE simply becomes

$$\frac{\partial U}{\partial \tau} = s \qquad \Longrightarrow \qquad U(x_0,\tau) = U(x_0,0) + s\tau.$$

But $U(x_0, 0) = f(x_0)$ since $x = x_0$ at $t = \tau = 0$, and so, in terms of the original variables,

$$u(x,t) = f(x - ct) + st.$$

In other words, by making a clever change of variables, we may turn the PDE into what is effectively an ODE (because derivatives with respect to the variable x_0 do not appear), integrate that equation, and finally change variables back to the original ones to find the required solution.

The trick works far more generally: if $c = c(x,t)$ and $s = s(x,t)$, then we may still (implicitly) define the new variables,

$$t = \tau \qquad \& \qquad x = x_0 + \int_0^\tau c(x(x_0,\tau),\tau)d\tau,$$

which again imply (2). In other words, solving the PDE in (1) is equivalent to solving the two ODEs,

$$\frac{\partial x}{\partial \tau} = c(x,\tau) \qquad \& \qquad \frac{\partial U}{\partial \tau} = s(x,\tau),$$

subject to the initial conditions $x = x_0$ and $U = f(x_0)$ at $\tau = 0$.

At this stage we may as well overhaul the notation: the variable x_0 appears purely as a parameter in the new system, and there is really no need of the full machinery of multi-variable calculus (partial derivatives *etc.*). Moreover, persevering with the replacement of the original pair, $u(x,t)$ and t, with the new pair, $U(x_0,\tau)$ and τ, is hardly worthwhile. So, for simplicity we will simply recycle the former notation instead of the latter. Thus, rather than solving (1) as a PDE, we will instead solve the ODE system,

$$\frac{dx}{dt} = c(x,t), \qquad \frac{du}{dt} = s(x,t), \qquad \text{with} \quad x = x_0 \quad \& \quad u = f(x_0) \quad \text{at} \quad t = 0.. \tag{3}$$

We need only eliminate x_0 at the end of the calculation to complete the solution to the problem.

e.g. Solve

$$u_t + (x - t)u_x = 0, \qquad u(x,0) = e^x.$$

We read off the coefficients: $c \equiv x - t$ and $s = 0$. Therefore our ODE system in (3) becomes

$$\frac{dx}{dt} = x - t \qquad \& \qquad \frac{du}{dt} = 0.$$

Hence, $x = t + 1 + (x_0 - 1)e^t$ and $u = f = e^{x_0}$. Eliminating $x_0 = 1 + (x - t - 1)e^{-t}$ furnishes $u = \exp[1 + (x - t - 1)e^{-t}]$. Since $u_t = (t - x)e^{-t}u$ and $u_x = e^{-t}u$ for this solution, we observe that it certainly solves the original PDE.

e.g. Solve

$$u_t + (x - t)u_x = 3x + 1, \qquad u(x, 0) = e^x.$$

The relevant ODEs are

$$\frac{dx}{dt} = x - t \qquad \& \qquad \frac{du}{dt} = 3x + 1.$$

Again we find $x = t + 1 + (x_0 - 1)e^t$, which is now needed to integrate the second ODE:

$$\frac{du}{dt} = 3t + 4 + 3(x_0 - 1)e^t$$

giving

$$u = e^{x_0} + \frac{3t^2}{2} + 4t + 3(x_0 - 1)(e^t - 1) = \exp[1 + (x - t - 1)e^{-t}] + \frac{3t^2}{2} + 4t + 3(x - t - 1)(1 - e^{-t}).$$

Try it out for yourself on the problem,

$$u_t - \frac{t^2}{x^2}u_x = 4\cos t, \qquad u(x, 0) = x^3.$$

The solution is $u(x, t) = 4\sin t + x^3 + t^3$.

$$e.g. \qquad u_t + xu_x = x^2, \qquad u(x, 0) = f(x),$$

The solution is $u(x, t) = \frac{1}{2}x^2(1 - e^{-2t}) + f(xe^{-t})$.

Quasi-linear PDEs

Our generalization of the PDE (1) to a problem that can be attacked by the method of characteristics does not stop with making the coefficients c and s depend on x and t. In fact, we can just as well apply the methodology if these coefficient also depend on u, but not its partial derivatives. *i.e.* we can solve the PDE,

$$u_t + c(x, t, u)u_x = s(x, t, u), \qquad u(x, 0) = f(x). \tag{4}$$

The corresponding ODE system is

$$\frac{dx}{dt} = c(x, t, u), \qquad \frac{du}{dt} = s(x, t, u), \qquad \text{with} \quad x = x_0 \quad \& \quad u = f(x_0) \quad \text{at} \quad t = 0.. \tag{5}$$

The only complication is (5) is now a pair of fully coupled ODEs (previously, the $x-$ equation could be solved independently of the $u-$equation).

Note that the PDE in (4) is now potentially nonlinear, as we could have terms like $cu_x = uu_x$ and $s = u^2$. The restriction to coefficients c and s that do not depend on the partial derivatives of u lead us to call this a "quasi-linear" PDE (rather than a fully nonlinear one).

At this stage, we introduce some further terminology: the ODEs in (5) are called the "characteristic equations." The solutions, which one can write formally as $x = x(t; x_0)$ and $u = u(t; x_0)$, corresponds to curves in the three-dimensional space (t, x, u) that intersect $(0, x_0, f(x_0))$; these are the "characteristic curves". Often we concern ourselves simply with the projection of these curves onto the $(t, x)-$plane, which furnishes what people usually call a "characteristics diagram." Drawing such a diagram can be quite insightful, as we shall see later.

$$e.g. \qquad u_t + xu_x = u^{-1}, \qquad u(x, 0) = \sqrt{x^2 + 1}.$$

The characteristic equations are

$$\frac{dx}{dt} = x \qquad \& \qquad \frac{du}{dt} = u^{-1}, \qquad \Longrightarrow \qquad x = x_0 e^t \qquad \& \qquad u = \sqrt{x_0^2 + 1 + 2t} \tag{6}$$

in view of the initial condition. Hence,

$$u(x, t) = \sqrt{x^2 e^{-2t} + 1 + 2t}.$$

Try

$$u_t + x^3 u_x = -u, \qquad u(x, 0) = \frac{x^2}{1 + x^2}.$$

You should arrive at $u(x, t) = x^2 e^{-t}/[1 + x^2(1 + 2t)]$.

$$e.g. \qquad u_t + xu_x = -\tfrac{1}{2}u^3, \qquad u(x, 0) = f(x),$$

The solution is $u(x, t) = f(xe^{-t})/\sqrt{1 + t[f(xe^{-t})]^2}$.

The characteristics diagram for (6) is shown in figure 1(a). The projections of the characteristic curves (or "characteristics" for short) fan out from their initial positions on the $x-$axis. Evidently, the characteristics cover the entire space-time diagram, and so our procedure provides a solution of the PDE everywhere. Adding a boundary somewhere, with some boundary condition imposed there, clearly requires some revision of the solution strategy.

Boundary-value problems

Consider the PDE,

$$u_t + u_x = xt, \qquad u(x, 0) = e^{-x}. \tag{7}$$

Following the method of characteristics (*i.e.* replacing the PDE with the system (5) and then solving it), gives $u(x, t) = e^{t-x} + \frac{1}{2}xt^2 - \frac{1}{6}t^3$. But this is the PDE solved with Laplace transforms earlier, although there we used the domain, $x > 0$ and $t > 0$, and added a boundary condition at $x = 0$: $u(0, t) = F(t)$. Our specific example took $F(t) = 0$, but we may as well leave $F(t)$ in there to see how things generalize. How can we modify the method of characteristics to deal with this situation?

Looking at figure 1(b), we see the characteristics diagram for this PDE. It is particularly simple, consisting of the family of straight lines, $x = x_0 + t$ (or $t = x - x_0$). The way we are thinking of the problem is that we initiate all the characteristics by pinning them at points along the entire $x-$axis. But when the domain is $x > 0$ and there is a boundary condition at $x = 0$, this is no longer correct. As illustrated in the figure, our domain lies on the right of the (red) $t-$axis and so there is now a second family of characteristics that are initiated at points there, rather than from positions on the negative $x-$axis. In other words, we should take $x = x_0 + t$ for the characteristics starting at $(x, t) = (x_0, 0)$ with $x_0 > 0$, and $x = t - t_0$ for those begining at $(x, t) = (0, t_0)$ with $t_0 > 0$. For the

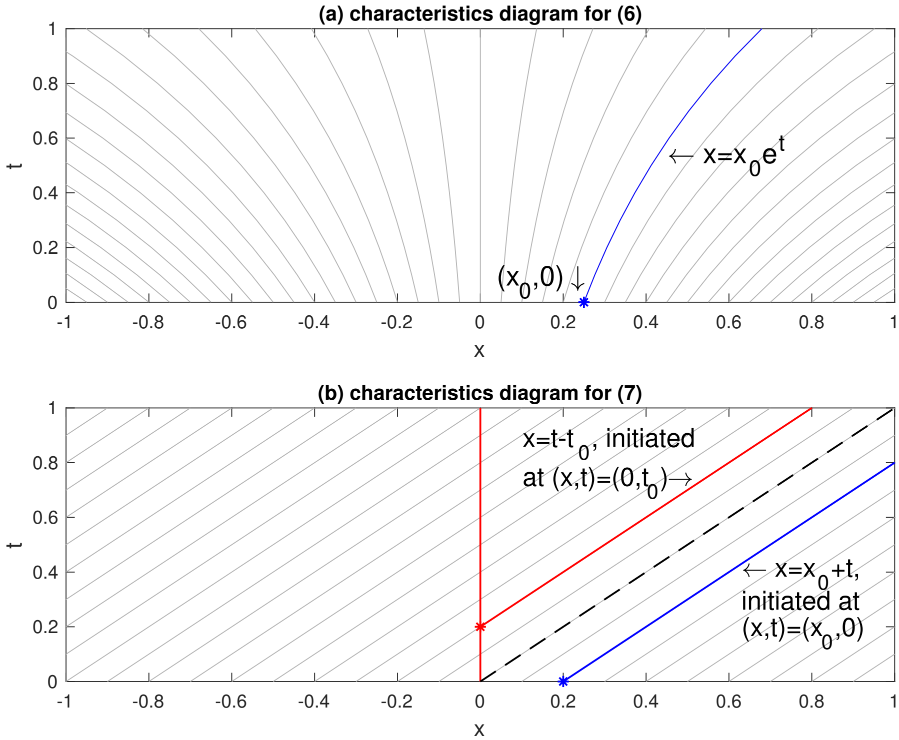

Figure 1: Characteristics diagram for (a) PDE (6) and (b) PDE (7).

former, the starting value for u is specified by the initial condition: $u = f(x_0) = e^{-x_0}$ at $t = 0$. For the latter, the boundary condition provides the starting value: $u = F(t_0)$ at $t = t_0$.

Armed with this insight, we may now solve (7) for $x > 0$ along with the boundary condition $u(0,t) = F(t)$. The characteristics starting along the positive $x-$axis have the solution $x = x_0 + t$ and $u = e^{t-x} + \frac{1}{2}xt - \frac{1}{6}t^3$, as before. This family of characteristics lies to the right of the special characteristic $x = t$ that emerges from the origin; *i.e.* in $x > t$. Those which begin on the $t-$axis, however, have the solution

$$x = t - t_0 \qquad \& \qquad u = \tfrac{1}{3}t^3 - \tfrac{1}{2}t_0t^2 + \tfrac{1}{6}t_0^3 + F(t_0) \equiv \tfrac{1}{2}xt^2 - \tfrac{1}{6}t^3 + \tfrac{1}{6}(t-x)^3 + F(t-x).$$

(the latter coming from the solution of $du/dt = xt = (t - t_0)t$ with the relevant starting value, $u = F(t_0)$ at $t = t_0$). The second family lies above the special characteristic $x = t$; *i.e.* in $t > x$. Finally, we combine the two results using step functions, to write

$$u(x,t) = \left[\tfrac{1}{2}xt^2 - \tfrac{1}{6}t^3 + \tfrac{1}{6}(t-x)^3 + F(t-x)\right] H(t-x) + \left[e^{t-x} + \tfrac{1}{2}xt - \tfrac{1}{6}t^3\right] H(x-t)$$

$$= \tfrac{1}{2}xt^2 - \tfrac{1}{6}t^3 + \tfrac{1}{6}(t-x)^3 H(t-x) + e^{t-x} - e^{t-x}H(t-x) + F(t-x)H(t-x),$$

as found previously with the Laplace transform when $F = 0$ (and using $H(x-t) = 1 - H(t-x)$).

$$e.g. \qquad u_t + u_x = u, \qquad u(x,0) = 0, \qquad u(0,t) = te^{-t}.$$

The solution is $u(x,t) = (t-x)e^{2x-t}H(t-x)$.

Implicit solutions

Although we can apply the method of characteristics to quasi-linear PDEs, there is an awkward detail hidden within the coupled ODEs in (5). Generally, the solution is *implicit*, as illustrated by the following example.

$$e.g. \qquad u_t + uu_x = 0, \qquad u(x,0) = f(x), \qquad -\infty < x < \infty,$$

We have the characteristic equations,

$$\frac{dx}{dt} = u, \qquad \frac{du}{dt} = 0, \qquad \text{with} \quad x = x_0 \quad \& \quad u = f(x_0) \quad \text{at} \quad t = 0..$$

The second ODE tells us that u does not change along each characteristic curve, and is therefore given by the initial value: $u = f(x_0)$. Moreover, the characteristics are the straight lines $x = x_0 + ut$. Thus, we may immediately integrate the first ODE to obtain the *implicit* solution $u = f(x - ut)$. For some forms of the initial condition function, $f(x)$, we may be able to algebraically solve this relation for $u(x,t)$ (as done below). But, in general, an explicit solution is not possible.

For example, consider the initial condition $u(x,0) = f(x) = 1 - x^2$ for $-1 \le x \le 1$ and $u(x,0) = f(x) = 0$ for $|x| > 1$. Thus, $u(x,t) = 1 - x_0^2$ for $-1 \le x_0 \le 1$ and $u(x,t) = 0$ for $|x_0| > 1$. The solution for $-1 \le x_0 \le 1$ can be be developped further: we have

$$u = 1 - x_0^2 = 1 - (x - ut)^2 \qquad \Longrightarrow \qquad u = \frac{1}{t^2}\left(\sqrt{\tfrac{1}{4} - xt + t^2} - \tfrac{1}{2} + xt\right); \tag{8}$$

the choice of sign when solving the quadratic for u is dictated by making sure one recovers the correct initial condition in the limit $t \to 0$. As the characteristic curves can move away from their initial points, we should be careful about where the interval $-1 \le x_0 \le 1$ translates to at later times. Fortunately, $f(\pm 1) = 0$ and so the characteristics bordering this interval are *vertical* on the characteristics diagram (presented shortly); *i.e.* $x = x_0$ if $x_0 \pm 1$. Thus, the solution in (8) applies for $-1 \le x \le 1$. Outside of this interval for x, the solution is, of course, zero for all time.

Although we have successfully found an explicit solution in (8) for the previous problem, there are some very thorny details lurking underneath its surface. We can bring these out by drawing the characteristics diagram, as done in figure 2.

In looking at this diagram it is helpful to recall that $u = 0$ for $x \ge 1$ and $x \le -1$. Hence $x = x_0$ over this region, and the characteristics are all vertical. Within $-1 \le x \le 1$, the characteristics are tilted, however, according to the value of $f(x_0)$ (the equation of these straight lines is $x = x_0 + f(x_0)t$, or $t = (x - x_0)/f(x_0)$). The greatest tilt corresponds to the highest value of u, which is $u = 1$ and achieved for $x_0 = 0$ or $x = t$. The construction of the characteristics diagram therefore unavoidably leads to the crossing of the characteristics, as seen in figure 2.

The crossing of characteristics is a complete disaster! To see this, we note that the characteristics all possess their own value of u, which is fixed at the initial value for the problem at hand. Thus, any point along a particular characteristic has that value for the solution. If two characteristics cross, there are therefore two possible values for u at the crossing point. In other words the solution is multi-valued and there can be no unique solution, which is the prime goal of most mathematical problems.

In fact, we can immediately observe the multi-valued solution simply by plotting the implicit result in (8), which is also done in figure 2: one selects a grid of points for x_0 and computes $f(x_0)$; we then plot those values of $u = f(x_0)$ against $x = x_0 + tf(x_0)$. The procedure is effortlessly encoded in MATLAB. As indicated by the figure, the multi-valuedness of the solution arises because the slanted characteristics demand that the lumpy initial condition continually tilt over to the right as time

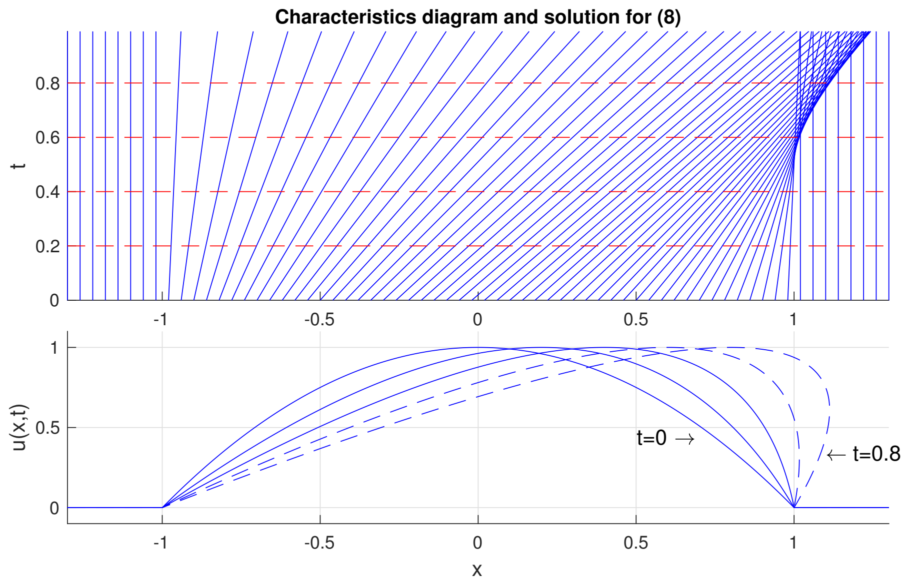

Figure 2: Characteristics diagram for (8); the snapshots of the solution shown at the bottom correspond to the times shown by the red dashed lines in the characteristics diagram. The solutions for $t = 0.6$ and 0.8 (blue dashed lines) are multi-valued and do not make sense.

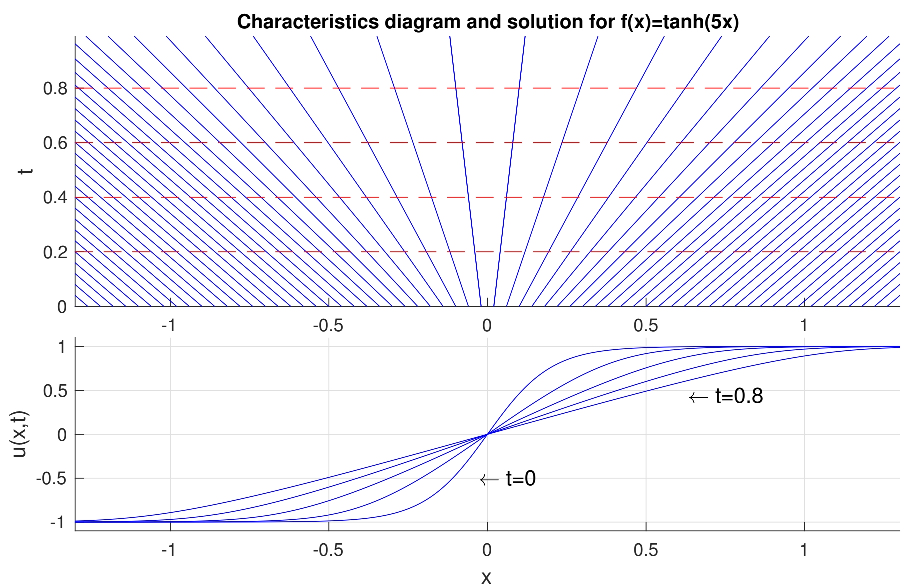

Figure 3: Characteristics diagram for $u_t + uu_x = 0$ with $u(x, 0) = \tanh(5x)$. Again, the time of the snapshots of the solution shown at the bottom correspond to the red dashed lines in the characteristics diagram.

progresses. At some instant, u becomes vertical, and then leans over itself; the solution takes three possible values over the region where there is an over-hang.

This is not an issue with the method of characteristics: the issue is with the nonlinear PDE itself. The solution only remains unique and is mathematically sensible upto a certain time. Nonlinear equations can be like that. We can easily compute the time at which things go wrong by noting that the disaster of the crossing characteristics begins when the solution first becomes vertical. That is, when u_x diverges. But our implicit solution of the PDE is

$$u(x,t) = f(x_0) = f(x - u(x,t)t) \qquad \Longrightarrow \qquad u_x = (1 - tu_x)f'(x_0) = 2x_0(tu_x - 1)$$

(using the chain rule), if $-1 \le x_0 \le 1$. We may re-arrange this to find:

$$u_x = \frac{f'(x_0)}{1 + tf'(x_0)} \longrightarrow \frac{2x_0}{2x_0 t - 1}. \tag{9}$$

Thus, u_x diverges when $2x_0 t = 1$, which happens first for the largest value of x_0 over the range $[-1, 1]$. In other words, the solution fails and becomes+ multi-valued for $t = \frac{1}{2}$ at the position $x_0 = x = 1$ (as seen in the figure). At the moment, we cannot proceed into $t > \frac{1}{2}$.

Equation (9) also highlights how the issue with crossing characteristics does depend on the initial condition. For example, with $u(x,0) = f(x) = \tanh(ax)$ (a is a parameter), the characteristics diagram and solution are shown in figure 3. In this example, the straight lines of the characterstics never cross and our solution (which remains implicit and cannot be solved for explicitly) is good for all time. Indeed, u_x never diverges for this initial condition because $f'(x_0) = a\ \mathrm{sech}^2(ax_0) > 0$ implying the denominator in (9) never vanishes.

The initial slope of the solution is therefore critical to whether the solution to $u_t + uu_x = 0$ becomes multi-valued or remains sensible. For this PDE, if $f'(x_0)$ is negative somewhere, the solution will steepen up as the characteristics from that region tilt over towards one another, eventually prompting the solution to bend over itself and become multi-valued. This is referred to a "shock steepening". For the opposite situation, with $f'(x_0) > 0$ everywhere, the characteristics are slanted away from one another, leading to the solution continually flattening out (as in figure 3). This is sometimes called a "rarefaction."

Consider $u_t + uu_x = 0$ with the initial condition $u(x,0) = -\tanh(ax)$. Construct the characteristics diagram and show that the solution becomes multi-valued for $t = a^{-1}$.

Conservation Laws

Consider a density field $\rho(x,t)$ that describes the amount of some sort of "stuff" per unit length. Over an interval $a \le x \le b$, the total amount of "stuff" is $\int_a^b \rho(x,t)\,dx$. This amount changes because "stuff" can be transported into and out of the interval, and if "stuff" is produced or destroyed within $[a,b]$. We express this mathematically as

$$\frac{d}{dt}\int_a^b \rho(x,t)\,dx = J(a,t) - J(b,t) + \int_a^b s(x,t)\,dx = -\left[J(x,t)\right]_{x=a}^{x=b} + \int_a^b s(x,t)\,dx. \tag{10}$$

Here, $J(x,t)$ denotes the transport, or *flux*, of "stuff" along the $x-$axis, so $J(a,t)$ denotes the influx to the interval through $x = a$, and $J(b,t)$ represents the outflux at $x = b$. The total production or generation of "stuff" is described by $\int_a^b s\,dx$, where $s(x,t)$ denotes the density of sources and sinks at each position and time. The specification of both $J(x,t)$ and $s(x,t)$ requires some sort of physical model relevant to whatever the "stuff" is.

Because (10) expresses how much the amount of stuff is conserved as a result of transport and production/destruction, we refer to it as a "conservation law". In particular, owing to the fact that it involves integrals over the interval $[a, b]$, we call this the "integral form" of a conservation law. An alternative, "differential form" is derived by first expressing (10) as a single integral: since

$$\frac{d}{dt}\int_a^b \rho\, dx \equiv \int_a^b \rho_t\, dx \qquad \& \qquad [J(x,t)]_{x=a}^{x=b} \equiv \int_a^b J_x\, dx,$$

we observe that

$$\int_a^b (\rho_t + J_x - s)\, dx = 0. \tag{11}$$

But the limits of the interval, a and b, are arbitrary positions, and the only way to ensure that the integral on the left of (11) vanish for every possible choice of a and b is to demand that the integrand itself vanishes. Thus

$$\frac{\partial \rho}{\partial t} + \frac{\partial J}{\partial x} = s, \tag{12}$$

which is the differential form of the conservation law.

Traffic flow

A popular model in the application of the method of characteristics to conservation laws is traffic flow. Here, $\rho(x, t)$ refers to the density of cars following a road. We ignore any overtaking and the cars coming in the opposite direction. In the absence of any side roads, breakdowns or parking, there is no source or sink of cars and so conservation of cars demands that $\rho_t + J_x = 0$. Last, we need a model for the flux $J(x, t)$. For this, we observe that $J = \rho V$ where V is the car speed. The speed V should be largest when there are no other cars on the road, and then decrease steadily until a traffic jam occurs at the maximum possible car density. Suitably normalizing so that the maximum density is reached for $\rho = 1$, we take $V = 1 - \rho$. Thus, $J = \rho(1 - \rho)$, and so

$$\rho_t + (1 - 2\rho)\rho_x = 0, \qquad \rho(x, 0) = f(x), \tag{13}$$

if $f(x)$ is the initial car distribution.

We may easily solve (13) by the method of characteristics: the characteristics equations are

$$\frac{dx}{dt} = 1 - 2\rho, \qquad \frac{d\rho}{dt} = 0, \qquad \text{with} \quad x = x_0 \quad \& \quad \rho = f(x_0) \quad \text{at} \quad t = 0..$$

Thus,

$$x = x_0 + (1 - 2\rho)t \qquad \& \qquad \rho = f(x_0) = f(x - t + 2t\rho).$$

Once more, the solution ρ is constant along each of the characteristics, and again these are all straight lines. The problem is much the same as we considered above, with another implicit solution.

A pile-up:

For the ramp-like initial condition,

$$\rho(x, 0) = \begin{cases} 0 & x < -\epsilon \\ \frac{1}{2}\rho_0(1 + x/\epsilon) & -\epsilon \le x \le \epsilon \\ \rho_0 & x > \epsilon, \end{cases} \tag{14}$$

the implicit solution can be re-arranged to give the explicit solution,

$$\rho = \tfrac{1}{2}\rho_0 \frac{\epsilon + x - t}{\epsilon - \rho_0 t}, \qquad \text{for} \; -\epsilon \le x_0 = x - (1 - 2\rho)t \le \epsilon.$$

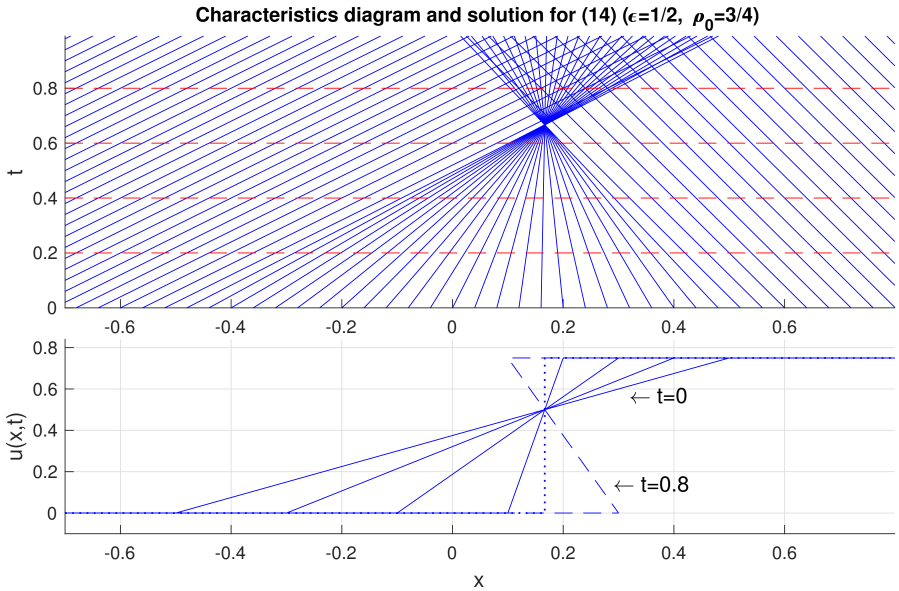

Figure 4: Characteristics diagram for (14) with $\epsilon = \frac{1}{2}$ and $\rho_0 = \frac{3}{4}$. Red dashed lines show the times of the snapshots of the solution; the solution for $t = 0.8$ (blue dashed lines) is multi-valued, and that at $t = \epsilon\rho_0^{-1}$ (dotted line) is discontinuous.

Given that $\rho(-\epsilon) = 0$ and $\rho(\epsilon) = \rho_0$, the borders of this interval corresponds to $x = -\epsilon + t$ and $x = \epsilon + (1 - 2\rho_0)t$, respectively. For $x < -\epsilon + t$ ($x_0 < -\epsilon$), we have $\rho = 0$, and for $x > \epsilon + (1 - 2\rho_0)t$ ($x_0 > \epsilon$) we have $\rho = \rho_0$.

The characteristics diagram for the solution is shown in figure 4. The ramp in the initial condition steepens up as time progresses, with the cars at the back catching up to those at the front. The solution becomes vertical for $t = \epsilon\rho_0^{-1}$, at the point $x = \epsilon(\rho_0^{-1} - 1)$. For $t > \epsilon\rho_0^{-1}$ the solution no longer makes sense as it becomes multi-valued. Evidently, the cars at the back pile up. All this happens simply because widely separated cars travel faster down the road than closely spaced ones. The lower-density tail of the ramp must therefore catch up to the higher-density file of cars ahead of the ramp.

Discontinuous solutions to the conservation law:

The key to avoiding problematic solutions where the characteristics cross is to allow discontinuities to appear in place of the multi-valued sections. This is hard to implement for the differential form of the conservation law in (12), as the space derivatives no longer make sense. However, there is nothing wrong with inserting a discontinuous solution into the integral form (10).

To explore this in more detail, let us imagine that there is discontinuity at $x = x_*(t)$. We allow this position to be a function of time as we have no reason currently to see why it must always stay in the same place. The integral conservation law can then be broken up to take explicit account of the discontinuity:

$$\frac{d}{dt}\int_a^{x_*} \rho(x,t)\, dx + \frac{d}{dt}\int_{x_*}^b \rho(x,t)\, dx = J(a,t) - J(b,t) + \int_a^{x_*} s(x,t)\, dx + + \int_{x_*}^b s(x,t)\, dx. \quad (15)$$

Each of the integrals now has a continuous integrand. Moreover, by Leibnitz's rule

$$\frac{d}{dt}\int_a^{x_*} \rho\, dx = \int_a^{x_*} \rho_t\, dx + \frac{dx_*}{dt}\rho(x_*^-, t) \quad \& \quad \frac{d}{dt}\int_{x_*}^b \rho\, dx = \int_{x_*}^b \rho_t\, dx - \frac{dx_*}{dt}\rho(x_*^+, t).$$

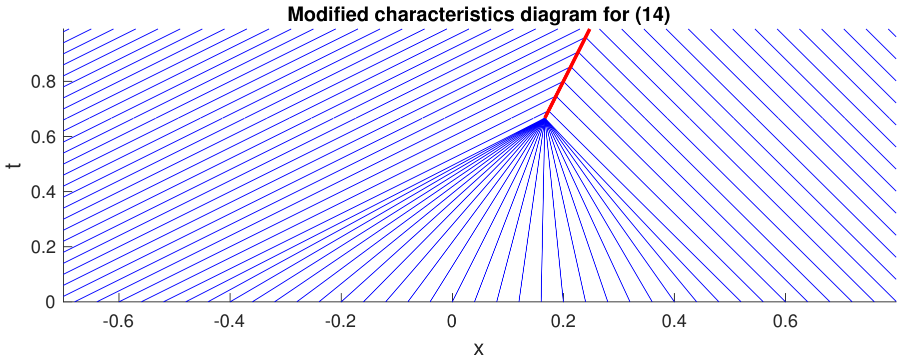

Figure 5: The modified characteristics diagram for (14), demonstrating how the insertion of a discontinuity (shock) at the position given by (17) (thick red line) prevents the characteristics from crossing.

Here, the notation $\rho(x_*^-, t)$ and $\rho(x_*^+, t)$ refer to the limits from either the left or right, respectively. Moreover, since the integrands are all now continuous, if we take the limit $a \to x_*^-$ and $b \to x_*^+$ in (15), all the integrals converge to zero, leaving

$$\frac{dx_*}{dt}[\rho(x_*^-, t) - \rho(x_*^+, t)] = J(x_*^-) - J(x_*^+), \qquad \text{or} \qquad \frac{dx_*}{dt} = \frac{J(x_*^+) - J(x_*^-)}{\rho(x_*^+, t) - \rho(x_*^-, t)}. \tag{16}$$

Thus, unless the fluxes into and out of the discontinuity are in balance ($J(x_*^-) = J(x_*^+)$), this structure, or "shock", must move.

For the traffic-flow problem, we have $J = \rho(1-\rho)$ and so

$$\frac{dx_*}{dt} = \frac{\rho^+(1-\rho^+) - \rho^-(1-\rho^-)}{\rho^+ - \rho^-} = 1 - \rho^+ - \rho^-.$$

where we have introduced the notation $\rho^\pm = \rho(x_*^\pm, t)$. In the pile-up problem considered previously, the solution becomes vertical at $x = \epsilon(\rho_0^{-1} - 1)$ and $t = \epsilon\rho_0^{-1}$, at which instant all the values of ρ that originally lay along the initial ramp now line up to form a step. That is, a discontinuity, or shock, forms with $\rho^- = 0$ and $\rho^+ = \rho_0$. Our consideration of the integral form of the conservation law now tells us that this shock will subsequently move with speed $\dot{x}_* = 1 - \rho_0$. In other words, for $t \geq \epsilon\rho_0^{-1}$, the solution becomes discontinuous, taking the form a step from $\rho = 0$ upto $\rho = \rho_0$. The position of the shock is given by

$$x_*(t) = (1-\rho_0)(t - \epsilon\rho_0^{-1}) + \epsilon(\rho_0^{-1} - 1) = (1-\rho_0)t \tag{17}$$

(bearing in mind where and when it first appears).

The insertion of the shock allows us, therefore, to avoid a multi-valued solution. Simultaneously, as illustrated in figure 5, it also prevents the characteristics from crossing by consuming the offenders.

Traffic lights

Green light to red light

Consider a line of cars with density $\rho = \rho_0$ (speed $V = 1 - \rho_0$) that is passing a traffic light at $x = 0$ which is green. At $t = 0$, however, the light turns red. No cars can now pass through $x = 0$; we must treat that location as a solid barrier and solve separate problems in $x > 0$ and $x < 0$.

First, for $x > 0$ and $t > 0$, we must deal with the boundary condition $\rho(0, t) = 0$, along with the initial condition $\rho(x, 0) = \rho_0$. The conflict between the two means that for $t = 0^+$, the solution

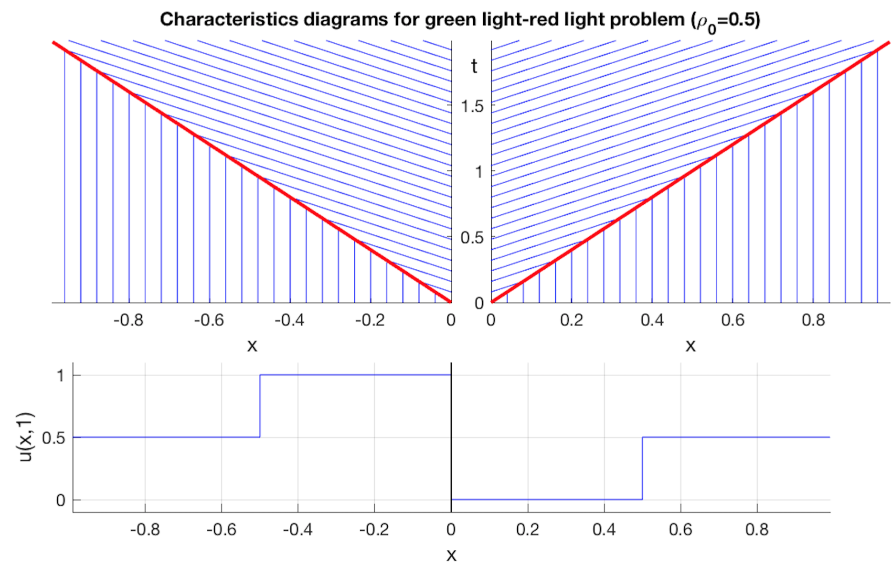

Figure 6: One the right, we have the characteristic diagram ahead of the red light ($x > 0$); on the left, we have the the diagram behind the light ($x < 0$). The shocks are shown by the thicker red lines.

jumps from $\rho = 0$ at $x = 0$ to $\rho = \rho_0$ at $x = 0^+$. In other words, a shock forms at the back of the line of traffic as in the pile-up problem considered above. This shock moves to the right with speed $\dot{x}_* = (1 - \rho_0)$, and so the shock position is $x_* = (1 - \rho_0)t$. The characteristics diagram consists of two regions: there is a family of characteristics that begins on the $x-$axis, given by $x = x_0 + (1 - 2\rho_0)t$, and a second family starting along the $t-$axis with $x = t - t_0$. Both become consumed when they reach the shock.

Second, in $x < 0$, as no cars can pass through $x = 0$ from the left, the traffic immediately jams up and becomes stationary, implying the boundary condition $\rho(0, t) = 1$. Again, this is in conflict with the inital condition $\rho(x, 0) = \rho_0$, and so another shock forms at $t = 0^+$ and $x = 0^-$, jumping from $\rho^- = \rho_0$ upto $\rho^+ = 1$. This shock travels with the speed $\dot{x}_* = -\rho_0$, and so $x_* = -\rho_0 t$. Once more there are two families of characteristics, both of which again disappear on colliding with the shock.

Red light to green light

Next we consider how a lane of stationary cars in $x < 0$ becomes set into motion by the change of a red light to green. In this case, the initial condition is the step function $\rho(0, t) = H(-x)$. One might think that the discontinuity here sets the stage for a shock solution. However, this is not the case: the cars further back from the light cannot move initially, whereas the cars at the front are free to accelerate away. Thus, the step should actually topple backwards and flatten out. To show this, we solve this problem in a more round-about fashion by reverting to the initial condition,

$$\rho(x, 0) = \begin{cases} 1 & x < -\epsilon \\ \frac{1}{2}(1 - x/\epsilon) & -\epsilon \le x \le \epsilon \\ 0 & x > \epsilon, \end{cases} \tag{18}$$

which gives

$$\rho = \frac{\epsilon - x + t}{2(t + \epsilon)}, \qquad \text{for } -\epsilon - t \le x \le \epsilon + t.$$

For $x < -\epsilon - t$, ρ remains at 1, whereas $\rho = 0$ for $x > \epsilon + t$. This solution describes a downwards slanting ramp over the interval $[-\epsilon, \epsilon]$ that expands outwards and flattens with time.

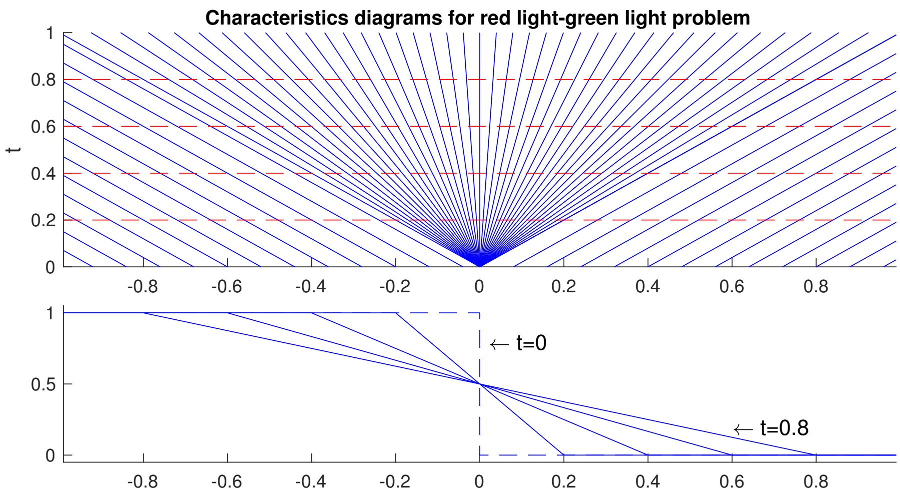

Figure 7: Characteristics diagram and sample snapshots of the solution for the red light to green light problem.

In fact, we may now take the limit $\epsilon \to 0$ to obtain the solution to the original traffic light problem:

$$\rho = \rho(x,0) = \begin{cases} 1 & x < -t \\ \frac{1}{2}(1 - x/t) & -t \le x \le t \\ 0 & x > t, \end{cases} \tag{19}$$

The characteristics diagram for this solution takes the form plotted in figure 7: one family of characteristics emanates from the positive $x-$axis with $x = x_0 + t$, and one from the negative $x-$axis with $x = x_0 - t$. Inbetween, a third family is launched from the origin, fanning out to fill the intervening wedge. This "expansion fan" is the remnant of the ramp in the continuous initial condition (18). Note that one may recover the solution in (19) without removing the initial step in the manner of (18): one simply observes $x_0 = 0$ for all the characteristics in the fan, implying that $0 = x - (1-2\rho)t$, which immediately gives the solution for ρ within $-t \le x \le t$.

Red light to green light for a finite lane

What happens when the stopped lane of traffic ends at $x = -L$? The initial condition is now $\rho(x,0) = 1$ for $-L \le x \le 0$ and $\rho(x,0) = 0$ otherwise. The step at $x = 0$ is expected to topple over backwards in the manner of the expansion fan outlined above. The step at $x = -L$, on the other hand, is much like the tail end of the advancing traffic lane in $x > 0$ for green light to red light problem. *i.e.* the discontinuity at $x = -L$ should remain shocking. In fact, since $\rho^- = 0$ and $\rho^+ = 1$ here, the shock remains stationary (as it should, since all the cars are still jammed up at the back). Thus, to begin with, an expansion fan appears at the front, with a stationary shock at the back.

At some instant, however, the rear of the fan reaches the stationary shock (or more poetically, the "shock hits the fan"). The fan is given by (19), and so the fan reaches the shock at $t = L$. From then on, $\rho^+ < 1$ and must be given by the solution for the fan. *i.e.*

$$\rho^- = 0 \qquad \& \qquad \rho^+ = \frac{1}{2}\left(1 - \frac{x_*}{t}\right) \qquad \Longrightarrow \qquad \frac{dx_*}{dt} = \frac{1}{2}\left(1 + \frac{x_*}{t}\right) \qquad \text{or} \qquad \frac{d}{dt}(x_* t^{-1/2}) = \tfrac{1}{2}t^{-1/2}.$$

Hence,

$$x_*(t) = t - 2\sqrt{tL}.$$

The characteristics diagram for this problem is shown in figure 8. Note that $\dot{x}_*(L) = 0$, so the shock moves left without a break in speed as it begins to move, and that $x_*(t) \to t$ for large times, thereby

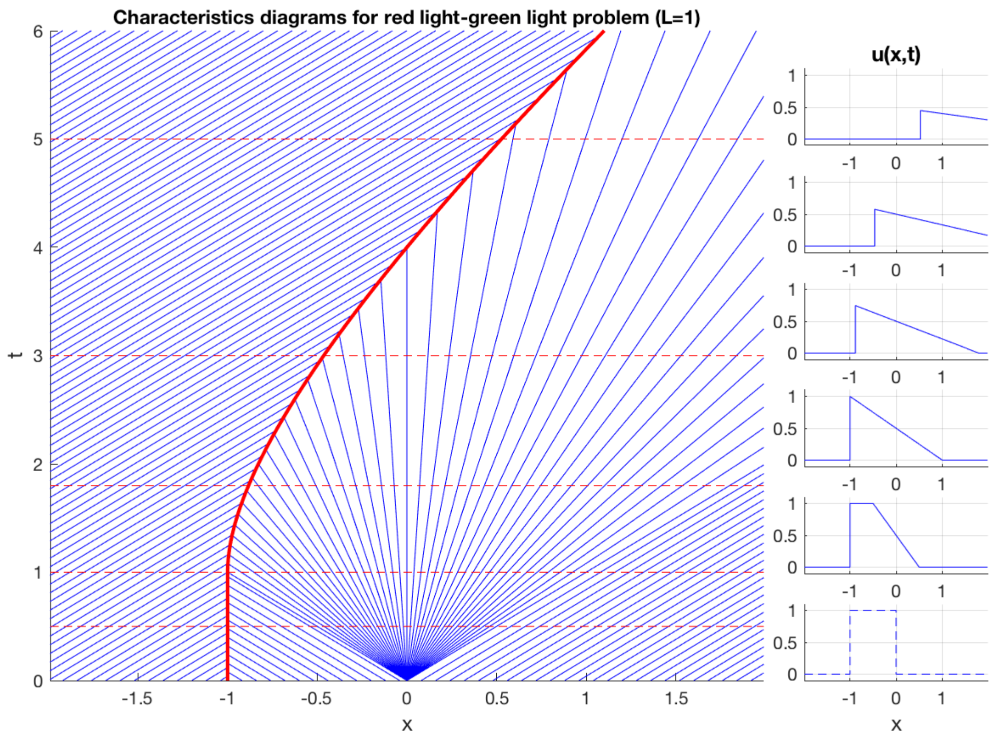

Figure 8: Characteristics diagram and sample snapshots of the solution for the red light to green light problem for a finite lane of traffic occupying $-1 < x < 0$. The shock is shown by the thicker red line.

asymptoting to a straight line with the same slope as the characteristics from $x_0 < -1$ and $x_0 > 0$ (which arises because $\rho^+ \to 0$ for $t \gg 1$).

The equal-areas rule

In most of the preceding examples, the path taken by the shock is straightforward to determine as the limiting values of the car density to either side, $\rho^\pm$, are fixed by the form of the initial condition or given analytically by an expansion fan. In more general settings, there is no such guarantee, and one does not have access to convenient expressions for $\rho^\pm$. The shock path is therefore hard to predict. However, there is a simple geometrical construction that allows us to qualitatively place the shock.

First, we go back to the integral form of the conservation law in (10). Without any sources or sinks, we may consider $a \to -\infty$ and $b \to \infty$ to observe that

$$\frac{d}{dt}\int_{-\infty}^{\infty} u(x,t)\, dx = -\left[J(x,t)\right]_{-\infty}^{\infty}.$$

Thus, without any net influx or outflux ($[J(x,t)]_{-\infty}^{\infty} = 0$), the total amount of "stuff" is conserved and given by the initial condition:

$$\int_{-\infty}^{\infty} u(x,t)\, dx = \int_{-\infty}^{\infty} u(x,0)\, dx.$$

This places an important constraint on the shock position: when we add the discontinuity to render the solution single-valued, it must be placed in such a way that "stuff" is conserved, or equivalently, that the area under the curve of $u(x,t)$ as a function of x at any given t is the same. But, there

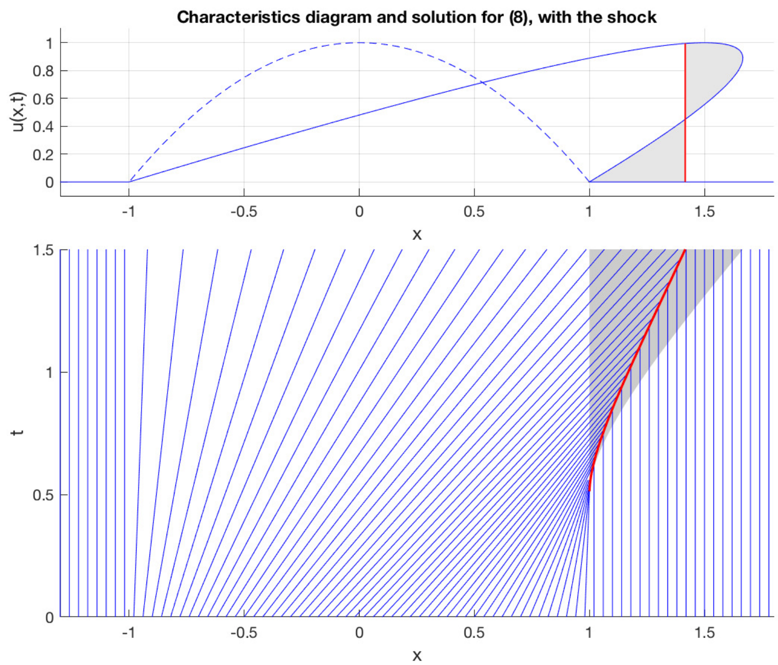

Figure 9: Characteristics diagram and solution for (8), but now including the shock (thick red line). The equal areas construction is shown by the shaded regions at the top; the region of the characteristics diagram over which the characteristics would have crossed without the shock is shaded below.

is a clever way of enforcing this constraint: one takes the multi-valued solution provided by the characteristics solution, then adds a vertical line connecting the top branch with the bottom branch over the region(s) where $u(x,t)$ is multi-valued. More specifically, one adds a shock this way in such a manner that the two lobes that are created to either side of the discontinuity *have the same area*. This construction eliminates the multi-valued parts of the solution, and leaves the same *signed area* as that for the multi-valued curve. *i.e.* conserves the amount of "stuff".

To illustrate, we return to problem (8): $u_t + uu_x = 0$, with the piece-wise parabolic initial condition. The solution was plotted in figure 2. We take an even later time than shown there for illustration: see figure 9. The key to place the shock is to add the vertical red line as shown, so that the two shaded regions have equal area. The construction simultaneously deletes the offensive region of crossings in the characteristics diagram, as also plotted in figure 9.

Note that, although it is not needed to find the shock path, we can also verify the conservation law explicitly in the red light to green light problem discussed above: as illustrated by figure 8, after the shock hits the fan, the solution takes the form of a ramp, declining from $\rho^+ = \frac{1}{2}(1 - x_*/t)$ at $x = x_*$, down to $\rho = 0$ at the front of the fan ($x = t$). The area is therefore

$$\tfrac{1}{2}(t - x_*)\rho^+ = \frac{1}{4t}(t - x_*)^2 = L,$$

which is, indeed, the area of the initial condition.

$$e.g. \qquad u_t + uu_x = 0, \qquad u(x,0) = f(x) = \frac{x}{1+x^2} \qquad \left(J = \tfrac{1}{2}u^2\right).$$

We have

$$u = f(x - ut) = \frac{x - ut}{1 + (x - ut)^2} \qquad \text{or} \qquad x = ut + \frac{1 - \sqrt{1 - 4u^2}}{2u}.$$

(the latter can be useful for plots). Since $u_x = f'(x_0)/[1 + tf'(x_0)]$, the shock time is given smallest positive value of $-1/f'(x_0)$, or $t = t_s = 8$. At this time, two shocks appear at the positions $x_0 \pm \sqrt{3}$, or $x = \pm 3\sqrt{3}$.

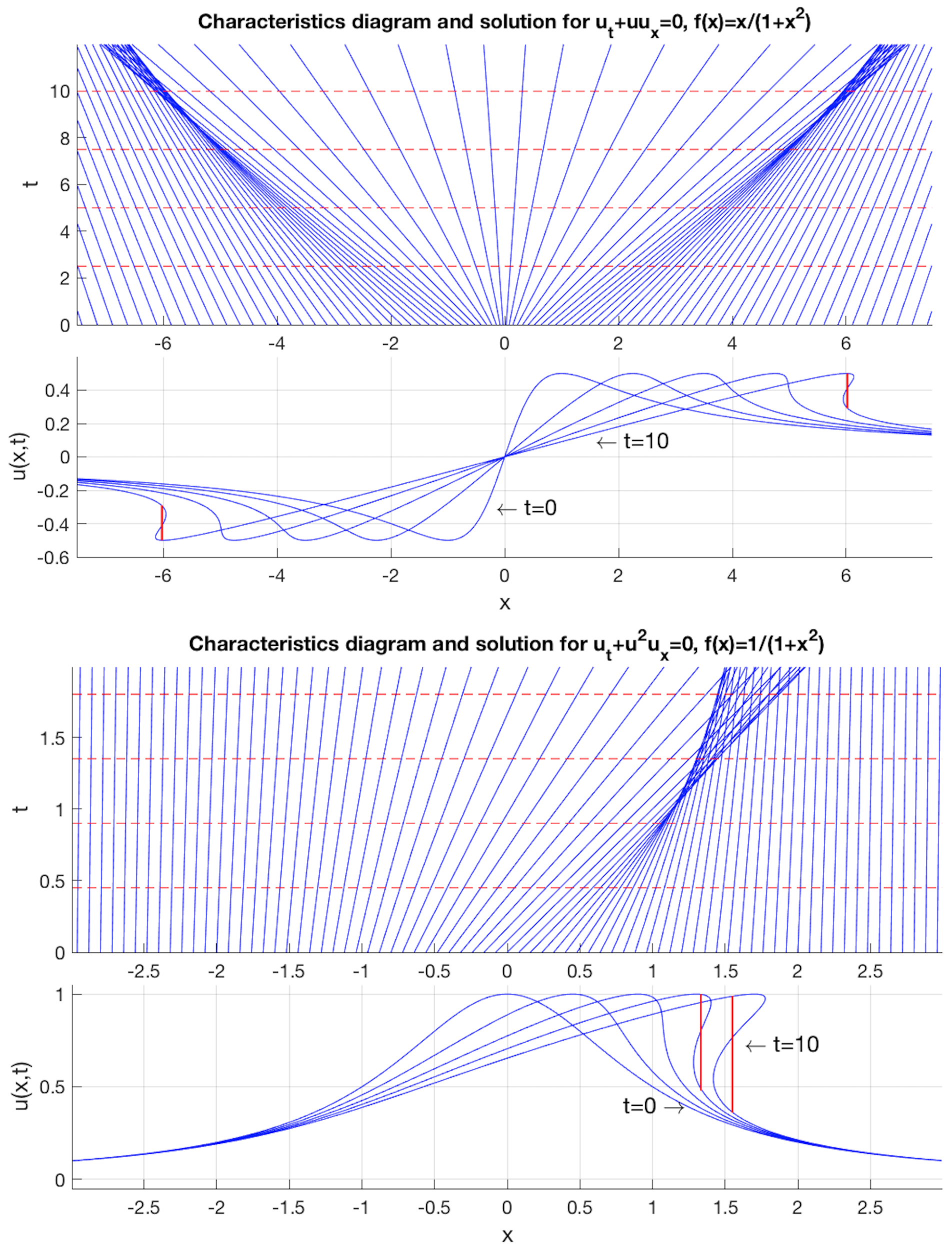

Figure 10: Rough equal-area constructions.

$$e.g. \qquad u_t + u^2 u_x = 0, \qquad f(x) = \frac{1}{1+x^2} \qquad (J = \tfrac{1}{3}u^3)$$

The implicit solution is

$$u = f(x - u^2 t) = \frac{1}{1 + (x - u^2 t)^2} \qquad \text{or} \qquad x = u^2 t \pm \sqrt{u^{-1} - 1}.$$

The shock time is computed by considering $u_x = f'(x_0)/[1 + 2tuf'(x_0)]$, which diverges when

$$t = -\left[2f(x_0)f'(x_0)\right]^{-1} = \frac{(1 + x_0^2)^3}{4x_0}.$$

This function of x_0 has a minimum positive value of $t = t_s = 54/25\sqrt{5}$ at $x_0 = 1/\sqrt{5}$, or $x = \frac{1}{2}\sqrt{5}$.