MAT 1060
More on First Order PDEs and the Method of Characteristics

Adam Morgan
adam.morgan@mail.utoronto.ca

November 29, 2022 and December 2, 2022

1 Introduction

In these lecture notes, I take the reader deeper into the theory and practice of first-order PDEs. The method of characteristics (reviewed in the first few pages) is applied to discuss conditions under which general first-order PDEs have *local* solutions, and show how these solutions can sometimes fail to exist *globally*. A better title might therefore be "how the method of characteristics can screw up", but as we'll see this is a bit pessimistic: at least for an interesting class of quasilinear equations there is a way to weaken the notion of "solution" to get global existence. The latter ideas lead us into the fascinating world of discontinuity, or **shock**, formation in evolution equations.

For the reader's benefit I have included many references, related both to the core material and to tangential but still interesting topics. I mention in particular the excellent textbook of Logan [8], which provided my model for section 5 on quasilinear conservation laws. Additionally, I have included a review of the method of characteristics for linear equations, taken directly from my own lecture notes for MAT 351.

2 Review: Basics of the Method of Characteristics

At the start of this course, you learned how to solve the first-order linear PDE

$$a(t,x)u_t + b(t,x)u_x = 0. \tag{2.1}$$

by realizing that (2.1) is equivalent to the vanishing of the directional derivative of $u(t,x)$ in the direction

$$\begin{pmatrix} a(t,x) \\ b(t,x) \end{pmatrix} \in \mathbb{R}_t \times \mathbb{R}_x.$$

In other words, $u(t,x)$ is constant along implicitly defined curves $x(t)$ whose tangent vector field is given by $(a(t,x),b(t,x))$. Such curves obey the ODE

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{b(t,x(t))}{a(t,x(t))}. \tag{2.2}$$

Why? We know the tangent line to the curve at $(t,x(t))$ lies along $(a(t,x(t)),b(t,x(t)))$, hence the slope of this line (which is precisely $\mathrm{d}x/\mathrm{d}t$) is b/a. Solutions to (2.2) are called the **characteristic curves** (or simply **characteristics**) of (2.1). As you have seen already, we can reduce finding the general solution of (2.1) to the simpler problem of integrating the ODE (2.2) and finding the characteristics. This solution procedure is called the **method of characteristics (MOC)**, and it is one of the few systematic analytical techniques you can deploy to exactly solve PDEs.

But, there is another approach to the MOC that is sometimes more user-friendly than integrating (2.2): instead of looking at characteristics as implicit curves in the xt plane, we can consider them as *parametric* curves $(t(s),x(s))$ (where here s is the parameter in question). You have already seen the basics of this point of view in lectures. Since we still want the vector field (a,b) to be tangent to the characteristics, we must have

$$\frac{\mathrm{d}t}{\mathrm{d}s} = a(t(s),x(s)), \tag{2.3a}$$

$$\tag{2.3b}$$

$$\frac{\mathrm{d}x}{\mathrm{d}s} = b(t(s),x(s)). \tag{2.3c}$$

The system of ODEs above is sometimes called the **characteristic system** of (2.1). Note that the characteristic system implies (at least formally) the implicit definition of a characteristic curve given by (2.2). Additionally, by the chain rule, (2.1), and (2.3) we find

$$\frac{\mathrm{d}}{\mathrm{d}s}u\left(t(s),x(s)\right) = 0,$$

so indeed the solution is still constant along characteristics.

We also need to label each distinct characteristic with a parameter τ, so our notation should really look like

$$(t(s;\tau),x(s;\tau)).$$

In the implicit version of MOC, the integration constant that arises from solving (2.2) plays the same role as τ. Recognize that τ picks out what characteristic we're on, and s gives (roughly) our displacement along that characteristic: we then have a way of re-parameterizing (t,x) space so that characteristics become coordinate lines

$$\tau = \text{constant}.$$

In practice, there is some freedom in determining the characteristic label τ. To keep life as simple as possible, τ should always be chosen based on where the solution to the PDE is already known. For example, if we are given an initial condition

$$u(0,x) = u_0(x)$$

then the τ-axis should be precisely $\{t = 0\}$. Incidentally, I recommend you use the parametric approach to MOC *only when you are given initial data* (or boundary data, etc): when you are interested in finding the general solution to a PDE, the implicit approach works a lot better. That being said, there are initial-value problems where the implicit approach is a lot cleaner, so generally speaking you should be comfortable with both methods.

All of this discussion is much easier to understand after a few examples:

Example 2.1 (Transport Equation Again)**.** Given $c \in \mathbb{R}$, consider the initial value problem for the transport equation,

$$\begin{cases} u_t + cu_x = 0 \\ u(0,x) = u_0(x) \end{cases}. \tag{2.4}$$

You already know the solution to this problem is

$$u(t,x) = u_0(x - ct),$$

which just means the graph of the initial state is transported without modification towards $\mathrm{sgn}(c)\infty$ a constant speed c. We now re-derive this solution using the parametric form of MOC. The characteristic equations are

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1,$$

$$\frac{\mathrm{d}x}{\mathrm{d}s} = c.$$

These equations are easily integrated in terms of unknown functions of τ:

$$\begin{pmatrix} t(s;\tau) \\ x(s;\tau) \end{pmatrix} = \begin{pmatrix} s + t_0(\tau) \\ cs + x_0(\tau) \end{pmatrix}$$

To fix the t_0, x_0, we need to make a clever choice of τ. We want to treat $\{s = 0\}$ as the set where we already know the solution, but this is precisely $\{t = 0\}$. Therefore, we set the τ-axis equal to the x-axis: we label characteristics by their x-intercepts (see Figure 1 for a sketch). This means

$$t_0(\tau) = 0, \quad x_0(\tau) = \tau.$$

We conclude that $t = s$ and

$$\tau = x - ct. \tag{2.5}$$

Next, since

$$\frac{\mathrm{d}}{\mathrm{d}s} u\left(t(s;\tau), x(s;\tau)\right) = 0,$$

we necessarily have

$$u(t(s;\tau), x(s;\tau)) = f(\tau)$$

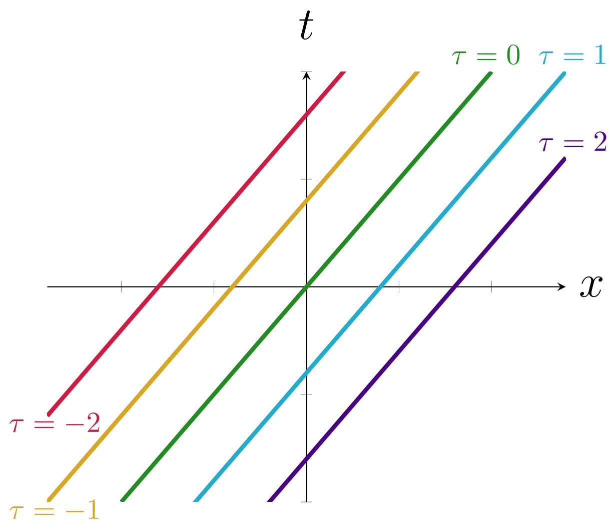

Figure 1: Various characteristics of the transport equation in the xt plane. The characteristic lines are all labelled by their x-intercept τ.

for an unknown function $f(\tau)$. Setting $s = 0$ gives

$$f(\tau) = u(0, \tau) = u_0(\tau).$$

We conclude that, in (s, τ) variables, the solution to the transport equation is

$$u(t(s;\tau), x(s;\tau)) = u_0(\tau).$$

We then use (2.5) to rewrite this solution in terms of the natural variables (t, x):

$$u(t, x) = u_0(x - ct).$$

Let's review our work so far. First, we wrote down and solved ODEs for the characteristic curves, lines in the xt plane where our solution $u(t, x)$ is constant. Then, since we are given data at $t = 0$, we chose to label characteristics by their intersections with the $\{t = 0\}$ axis, which we called τ. This choice amounts to parameterizing characteristics by $s = t$. Said differently, characteristics give us a new coordinate system in the xt plane: characteristics are precisely the grid lines $\tau =$ constant, and the lines $t =$ constant become the grid lines $s =$ constant. See Figure 2 for an illustration. In particular, this picture tells us that the given initial data is "flowed out" from the $t = 0$ axis along the characteristics: to compute $u(t, x)$, all we need to do is find out what (unique!) characteristic passes through (t, x), then we can use the initial condition and the constancy of u along characteristics to finish. We see then that MOC works *because it tells us how to construct a coordinate system where solving our PDE becomes trivial*.

Remark. *In the language of differential geometry, we might say that the method of characteristics gives us a clean way of building a useful* ***foliation*** *of the space of independent variables (in the above example, the xt plane). For more on first-order PDEs and characteristics from a geometric viewpoint, see* [5, Chapters 9 and 19].

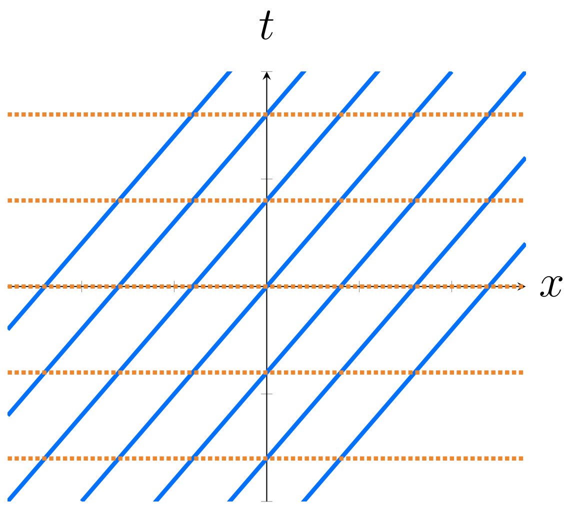

Figure 2: Sketch of the re-parameterization of the xt plane in terms of (s, τ) as suggested by the method of characteristics for the transport equation. The solid diagonal grid lines are characteristics $\tau =$ constant and the dotted horizontal grid lines are $s = t =$ constant.

Example 2.2 (Variable Coefficients)**.** To hammer in the discussion of our previous example, let's solve the more complicated initial value problem

$$\begin{cases} u_t + txu_x = 0 \\ \quad u(0, x) = u_0(x) \end{cases}. \tag{2.6}$$

We start by writing out the characteristic ODEs:

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1,$$

$$\frac{\mathrm{d}x}{\mathrm{d}s} = t(s)x(s),$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = 0.$$

To integrate the second equation, we first need to solve for $t(s)$. As in the previous example, since we are given the solution's values at $t = 0$, we should label the characteristics by their x-intercepts. That is, the τ-axis is precisely the x-axis. This means

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1 \Rightarrow t(s; \tau) = s.$$

Then, the ODE for $x(s; \tau)$ becomes

$$\frac{\mathrm{d}x}{\mathrm{d}s} = sx.$$

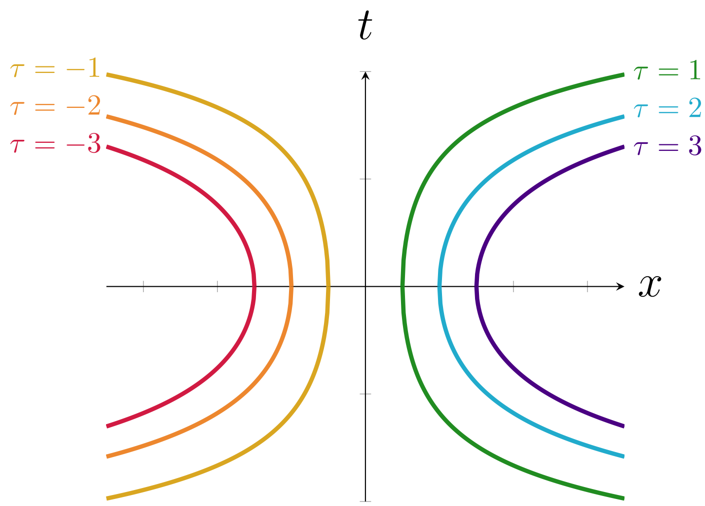

Figure 3: Various characteristics of $u_t + txu_x = 0$. The t axis is the $\tau = 0$ characteristic.

This separable ODE (together with the condition $x(0;\tau) = \tau$) is immediately integrated to yield

$$\tau = xe^{-\frac{1}{2}s^2}. \tag{2.7}$$

The characteristic curves for this PDE are then given by

$$\begin{pmatrix} t(s;\tau) \\ x(s;\tau) \end{pmatrix} = \begin{pmatrix} s \\ \tau e^{\frac{1}{2}s^2} \end{pmatrix}.$$

See Figure 3 for a sketch of a few of these curves. Since $\frac{\mathrm{d}u}{\mathrm{d}s} = 0$, we find that

$$u(t(s;\tau), x(s;\tau)) = f(\tau)$$

for an unknown function $f(\tau)$. Plugging in $s = 0$ and using (2.7), we find that the solution to our initial-value problem is

$$u(t,x) = u_0\left(xe^{-\frac{1}{2}t^2}\right).$$

3 Existence and Uniqueness for Scalar First-Order PDE

In this section I briefly discuss some general results related to well-posedness of the Cauchy problem for first order PDEs. First, we'll see an easy example to introduce a general condition under which well-posedness definitely fails. Once we have a systematic understanding of this general condition, I will state (but not prove) some general results on existence and uniqueness for nonlinear PDEs. Additionally, I make sure to emphasize the different behaviour of *quasilinear* versus *fully nonlinear* equations when it comes to general well-posedness results.

3.1 Characteristic Initial Conditions

In this subsection, we discuss an important notion critical to understanding a necessary condition for well-posedness of a first-order Cauchy problem. We open with a simple example:

Example 3.1. For a fixed nonzero $c \in \mathbb{R}$, consider the following Cauchy problem for the transport equation:

$$\begin{cases} u_t + cu_x = 0 & \forall\ (t,x) \in \mathbb{R} \times \mathbb{R} \\ u|_{x=ct} = u_0(x) & \forall\ x \in \mathbb{R} \end{cases} \tag{3.1}$$

where $u(t,x)$ is real-valued and the initial state u_0 is a nice function. We want to show this problem is not well-posed. To begin with, recall that the characteristics of the PDE in question are straight lines in the xt plane:

$$\begin{pmatrix} t(s,\tau) \\ x(s;\tau) \end{pmatrix} = \begin{pmatrix} s + t_0(\tau) \\ cs + x_0(\tau) \end{pmatrix}$$

where the τ-dependent intercepts are determined from the Cauchy data. This suggests changing variables so our characteristics become grid lines:

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto \begin{pmatrix} \eta \\ \xi \end{pmatrix} = \begin{pmatrix} t \\ x - ct \end{pmatrix}.$$

In the new coordinates, (3.1) becomes

$$\begin{cases} u_\eta = 0, \\ u|_{\xi=0} = u_0(\eta). \end{cases} \tag{3.2}$$

Solving the PDE gives

$$u = U(\xi)$$

for an arbitrary function $U(\xi)$. However, we get into trouble when we plug in the Cauchy data:

$$u|_{\xi=0} = U(0) = u_0(\eta).$$

Therefore, we cannot find a solution unless u_0 is identically constant! Even if u_0 is constant, we find that (3.2) is solved by any function of the form

$$u(\eta,\xi) = u_0 + U(\xi)$$

with $U(\xi)$ any function obeying $U(0) = 0$ (for example, $U(\xi) = \xi$, $U(\xi) = \sin\xi$, $U(\xi) =$ any homogeneous polynomial, $U(\xi) =$ any odd function, ...). So, even if we are lucky enough to have initial data that guarantees *existence*, we find *uniqueness* fails!

What has gone wrong here? Let's stick with our dependent variables being $(t,x) \in \mathbb{R}^2$ for the sake of exposition. Recall how the MOC game is played: given (t,x), we can compute our solution $u(t,x)$ by finding what characteristic intersects (t,x), then tracing backwards along this characteristic until we hit the Cauchy curve. Of course, if the method works, characteristics can never intersect. Thus, for the example problem (3.1), if you pick $(t,x) \notin \{(t,ct)\}$, you can move along the characteristic cutting (t,x) until kingdom come and never cut the Cauchy curve. Moving this logic backwards, the Cauchy data cannot be flowed out along characteristics to foliate the entire xt plane, or even a tiny open subset of the xt plane, since the only characteristic we can flow along is the Cauchy curve itself. We cannot then take existence for granted and certainly must throw out any hope of uniqueness, for in this case any function vanishing on the Cauchy curve can be added to a solution to generate another solution! Summarily, the reason (3.1) is not well-posed is that the Cauchy data *is given on a characteristic.*

To drive this concept home, let's look at another Cauchy problem that's a little more involved than our motivating problem for the transport equation.

Example 3.2. In this example, we show that the Cauchy problem

$$\begin{cases} u_t + 3t^2 u_x = tu & \forall\ (t,x) \in \mathbb{R}\times\mathbb{R} \\ u|_{x=t^3} = \cos(t) & \forall\ t \in \mathbb{R} \end{cases} \tag{3.3}$$

has no solution. Since the PDE in question has order one, we should definitely start by finding the characteristics. The characteristic equations in this case are

$$\frac{\mathrm{d}t}{\mathrm{d}s} = 1,$$

$$\frac{\mathrm{d}x}{\mathrm{d}s} = 3t^2,$$

$$\frac{\mathrm{d}u}{\mathrm{d}s} = tu.$$

Using the first two ODEs in the above system, we can express the characteristics implicitly via

$$\frac{\mathrm{d}x}{\mathrm{d}t} = 3t^2$$

which means each characteristic has the form

$$x = t^3 + \xi \tag{3.4}$$

for a constant ξ, see figure 4 for a sketch. The Cauchy curve Γ is the special case $\xi = 0$! Therefore, our Cauchy data is given on a characteristic, and we immediately know we cannot expect well-posedness. But we want something stronger, namely *non-existence* of any solution, so we have a little more work to do.

As with the transport equation, we change variables using the characteristics for inspiration:

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto \begin{pmatrix} \eta \\ \xi \end{pmatrix} = \begin{pmatrix} t \\ x - t^3 \end{pmatrix}.$$

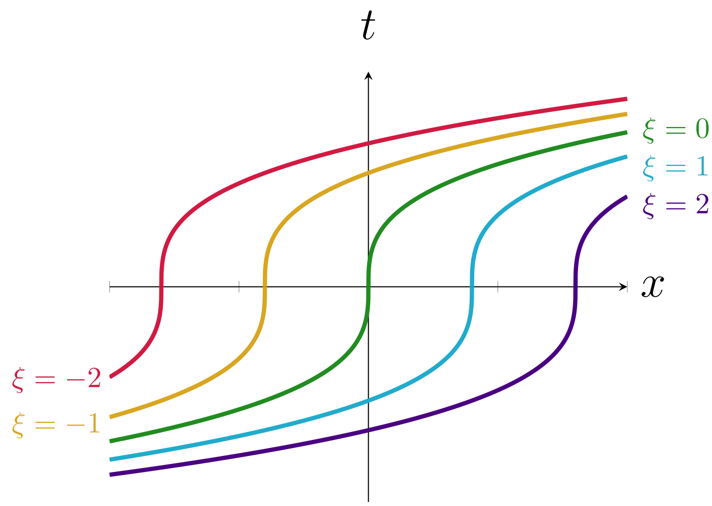

Figure 4: Various characteristics of the PDE from (3.3) in the xt plane. The characteristic lines are all labelled by ξ as defined in (3.4).

After using the chain rule, we find that (3.3) may be written as

$$\begin{cases} u_\eta = \eta u, \\ u|_{\xi=0} = \cos(\eta). \end{cases} \tag{3.5}$$

Integrating the ODE gives

$$u(\eta, \xi) = U(\xi) e^{\frac{1}{2}\eta^2}$$

for some function $U(\xi)$. Plugging in the Cauchy data, we get

$$u(\eta, 0) = U(0) e^{\frac{1}{2}\eta^2} = \cos \eta.$$

Clearly, no real number $U(0)$ can make the above statement hold for all $\eta \in \mathbb{R}$. We conclude that (3.1) has no solution.

In conclusion, we have shown that, in order to have any chance at well-posedness for a first-order Cauchy problem, the Cauchy curve must not be a characteristic. A bit of thought shows that even allowing the Cauchy curve to be tangent to a characteristic at a point x^0 may cause trouble. If our Cauchy curve is never tangent to a characteristic, we say that our Cauchy data is **noncharacteristic**.

3.2 A Survey of Local Existence and Uniqueness Results for First-Order Cauchy Problems

In this subsection, I introduce some basic local existence and uniqueness results for first-order Cauchy problems subject to the condition of noncharacteristic initial data. We work in a great deal of generality here, but the emphasis is still on ideas and examples rather than complete rigorous proofs (which the interested reader may find in the references cited throughout the discussion below).

First, we must quantitatively formulate what it means for Cauchy data to be noncharacteristic. Let's start by specializing to quasilinear equations. Assume we are given an open set $U \subseteq \mathbb{R}^n$, a smooth orientable hypersurface Γ with $U \cap \Gamma \neq \emptyset$, and three smooth functions

$$c = c(z,x)\colon \mathbb{R} \times U \to \mathbb{R}^n,$$

$$f = f(z,x)\colon \mathbb{R} \times U \to \mathbb{R}, \quad \text{and}$$

$$u_0 = u_0(x)\colon \Gamma \to \mathbb{R}.$$

The general quasilinear Cauchy problem reads

$$\begin{cases} c\left(u(x),x\right) \cdot \nabla u + f(u(x),x) = 0 \quad \forall\ x \in U, \\ \hfill u|_{U\cap\Gamma} = u_0. \end{cases} \tag{3.6}$$

If $c(z,x)$ is independent of z and $f(z,x)$ is affine in z, then the PDE is actually *linear*; this is the case we've considered so far in the examples. The characteristic ODEs for (3.6) are, if we define $z(s;\tau) = u(x(s;\tau))$,

$$\frac{\mathrm{d}x}{\mathrm{d}s}(s;\tau) = c(z(s;\tau), x(s;\tau)),$$

$$\frac{\mathrm{d}z}{\mathrm{d}s}(s;\tau) = f(z(s;\tau), x(s;\tau)).$$

In other words, the slope of the characteristic labelled τ evaluated at displacement s is $c(z(s;\tau), x(s;\tau))$. If we want out Cauchy data to be noncharacteristic, then, $c(z(s;\tau), x(s;\tau))$ must never be tangent to Γ: if it is, flowing along characteristics won't take us off Γ. Equivalently, the projection of $c(z(s;\tau), x(s;\tau))$ onto the unit outward normal of Γ, denoted $\mathbf{n}$, must never vanish:

$$c\left(u_0(x),x\right) \cdot \mathbf{n}\left(x\right) \neq 0 \quad \forall\ x \in U \cap \Gamma. \tag{3.7}$$

Under this condition, quasilinear Cauchy problems with smooth coefficient functions always admit unique local solutions.

Theorem 3.3 (Local Existence and Uniqueness for First-Order Quasilinear Cauchy Problems)**.** *Let $U \subseteq \mathbb{R}^n$ be an open set and let $\Gamma \subseteq \mathbb{R}^n$ be a smooth orientable hypersurface with $U \cap \Gamma \neq \emptyset$. Assume we are given three smooth functions*

$$c = c(z,x)\colon \mathbb{R} \times U \to \mathbb{R}^n,$$

$$f = f(z,x)\colon \mathbb{R} \times U \to \mathbb{R}, \quad \textit{and}$$

$$u_0 = u_0(x)\colon \Gamma \to \mathbb{R}.$$

Let $\mathbf{n}(x)$ *denote the unit outward normal to* Γ *evaluated at* $x \in \Gamma$*. Assume there is some* $x^0 \in U \cap \Gamma$ *such that the noncharacteristic condition*

$$c\left(u_0\left(x^0\right), x^0\right) \cdot \mathbf{n}\left(x^0\right) \neq 0$$

holds. Then, there exists an open neighbourhood V *of* x^0 *with* $V \cap \Gamma \neq 0$ *and a* C^1 *function* $u\colon V \to \mathbb{R}$ *solving the Cauchy problem*

$$\begin{cases} c\left(u(x), x\right)\right) \cdot \nabla u + f(u(x), x) = 0 \quad \forall\, x \in V, \\ \hfill u|_{V\cap\Gamma} = u_0. \end{cases} \tag{3.8}$$

For a proof of the above theorem in the case $n = 2$, see [4, §1.4, 1.5] (really, most of the ingredients are in [2, §3.2] as well). I emphasize the *local* nature of the conclusion here: the second half of these notes deals with the widely-encountered phenomenon of finite-time shock formation (spatial derivative blowup) for quasilinear Cauchy problems, and even for linear equations global existence is a bit sticky:

Example 3.4. Consider the Cauchy problem

$$\begin{cases} yu_x - xu_y = 0 \\ \hfill u|_{y=0} = x^3 \end{cases}. \tag{3.9}$$

We show that this problem does not admit a global solution $u(x, y)$. By "global", we mean defined for all $(x, y) \in \mathbb{R}^2$. This occurs despite the initial data being noncharacteristic, as the reader can easily verify for themselves (indeed, the characteristic cutting the Cauchy curve is orthogonal to the Cauchy curve at the intersection points).

To start with, let's write the characteristics implicitly:

$$\frac{\mathrm{d}x}{\mathrm{d}y} = -\frac{y}{x}. \tag{3.10}$$

This is a separable ODE, so solving it is trivial. In terms of an arbitrary constant R, the solution to this ODE is given implicitly by

$$x^2 + y^2 = R^2. \tag{3.11}$$

Therefore, the characteristics of the PDE in (3.9) are circles. In this case, the move to a new coordinate system where characteristics become grid lines is precisely the move to *polar coordinates*. In fact, the polar coordinate form of the PDE is

$$\partial_\theta u = 0,$$

which simply says that u is a radial function.

Now, let us focus on the $R = 1$ characteristic. Call this curve γ_1. The solution must be constant along γ_1, so to find $u|_{\gamma_1}$ all we need to do is evaluate u at one point along γ_1.

Obviously, we know the solution at $(x, y) = (1, 0) \in \gamma_1$ since our data is given on the x-axis. Therefore,

$$u|_{\gamma_1} = x^3|_{(x,y)=(1,0)} = 1.$$

However, we also have $(-1, 0) \in \gamma_1$, which would imply

$$u|_{\gamma_1} = x^3|_{(x,y)=(-1,0)} = -1,$$

a contradiction. So, MOC gives a multiply-valued solution, which of course is not a solution at all! We recognize that *the given data is not adapted to the characteristics of the governing PDE*. That is, the cause of ill-posedness of (3.9) is that the Cauchy data given on $\{y = 0\}$ is ambiguously defined on characteristics. I remark that we could simply ask for new initial conditions: as long as the solution prescribed on $\{y = 0\}$ is *even* (for example, $u(x, 0) = x^2$), the solution given by MOC is unique and exists globally.

How do we extend the notion of noncharacteristic data, and therefore theorem 3.3, to fully nonlinear equations? We cannot undertake a complete discussion for the sake of time, but it turns out that uniqueness has to be given up (we'll see a specific example of a simple noncharacteristic Cauchy problem with nonunique solutions below). We still have local existence, however:

Theorem 3.5 (Local Existence for First-Order Cauchy Problems). *Let $U \subseteq \mathbb{R}^n$ be an open set and let $\Gamma \subseteq \mathbb{R}^n$ be a smooth orientable hypersurface with $U \cap \Gamma \neq \emptyset$. Assume we are given two smooth functions*

$$\begin{aligned} F = F(p, z, x) &\colon \mathbb{R}^n \times \mathbb{R} \times U \to \mathbb{R}, \quad \textit{and} \\ u_0 = u_0(x) &\colon \Gamma \to \mathbb{R}. \end{aligned}$$

Let $\mathbf{n}(x)$ denote the unit outward normal to Γ evaluated at $x \in \Gamma$. Assume there is some $x^0 \in U \cap \Gamma$ such that

- *there exists $p^0 \in \mathbb{R}^n$ satisfying the compatibility conditions*

$$\begin{cases} p^0 - \left(p^0 \cdot \mathbf{n}\left(x^0\right)\right)\mathbf{n} = \nabla_\Gamma u_0\left(x^0\right) \\ \\ F\left(p^0, u_0\left(x^0\right), x^0\right) = 0, \end{cases} \tag{3.12}$$

  where ∇_Γ denotes the surface gradient on Γ, and

- *the noncharacteristic condition*

$$\nabla_p F\left(p^0, u_0\left(x^0\right), x^0\right) \cdot \mathbf{n}\left(x^0\right) \neq 0 \tag{3.13}$$

  also holds.

Then, there exists an open neighbourhood V of x^0 with $V \cap \Gamma \neq 0$ and a C^1 function $u\colon V \to \mathbb{R}$ solving the Cauchy problem

$$\begin{cases} F(\nabla u, u, x) = 0 \quad \forall\ x \in V \\ \qquad u|_{V\cap\Gamma} = u_0. \end{cases} \tag{3.14}$$

For a proof of the above result, see [2, §3.2] or [4, §1.7, 1.8]. Technically speaking, such a proof actually shows that our definition of characteristics as curves in the x-domain U is not sufficient to handle fully nonlinear equations. Indeed, the characteristics we have been dealing with are actually **projected characteristics** arising from projecting the "real" characteristics down from the (p, z, x)-domain. However, as we've already seen, for quasilinear equations (including many of the equations of fluid dynamics) one can obtain a perfectly satisfying theory using only projected characteristics, so we shouldn't give ourselves a hard time about glossing over this detail.

Example 3.6 ([4, §1.8])**.** Consider the following fully nonlinear Cauchy problem involving the **eikonal equation**:

$$\begin{cases} |\nabla u(x,y)| = 1 \quad \forall\ (x,y) \in \mathbb{R}^2 \\ u|_{x^2+y^2=1} = 0. \end{cases} \tag{3.15}$$

One can show that the solution $u(x, y)$ to (3.15) gives the signed distance from (x, y) to the unit circle (for more on signed distance functions and their use in computer graphics and other practical applications, see [9]). However, notice that solutions to (3.15) are not unique, for if u solves this problem then so does $-u$. This reflects the obvious statement that the sign in our definition of a signed distance function is just a matter of convention! This nonuniqueness happens *despite* the Cauchy data satisfying both the compatibility conditions (3.12) and the noncharacteristic condition (3.13), as we now show.

- If we parameterize the Cauchy curve by the angle $\theta \in [0, 2\pi)$, then the compatibility conditions read

$$\begin{cases} \begin{pmatrix} p^0 \\ q^0 \end{pmatrix} \cdot \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} = 0, \\ \\ \left(p^0\right)^2 + \left(q^0\right)^2 = 1 \end{cases} \tag{3.16}$$

  since $\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$ is tangent to the Cauchy curve. This system has *two* solutions

$$\begin{pmatrix} p^0 \\ q^0 \end{pmatrix} = \pm \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$

- For any point θ along our Cauchy curve, the noncharacteristic condition reads

$$\begin{pmatrix} p^0 \\ q^0 \end{pmatrix} \cdot \begin{pmatrix} -\cos\theta \\ -\sin\theta \end{pmatrix} \neq 0.$$

  Of course, from solving the compatibility conditions, we know the left-hand side of the above is $\pm 1 \neq 0$.

Thus, at every $\theta \in [0, 2\pi)$, we can find two $\begin{pmatrix} p^0 \\ q^0 \end{pmatrix} \in \mathbb{R}^2$ so that the hypotheses of theorem 3.5 are satisfied. Therefore, lack of uniqueness of solutions to the compatibility conditions (3.12) can easily give rise to nonuniqueness of solutions to the corresponding PDE.

Example 3.7. The theorems presented in this section all work for *real, scalar* first-order PDEs. When working with systems of first-order PDEs, however, the theory becomes much more subtle. For instance, in 1957 Lewy [7] showed that there exist smooth functions

$$\begin{aligned} f &= f(x,y,t)\colon \mathbb{R}^3 \to \mathbb{R} \quad \text{and} \\ g &= g(x,y,t)\colon \mathbb{R}^3 \to \mathbb{R} \end{aligned}$$

such that the 2×2 linear system

$$\begin{cases} v_x + w_y + yv_t + xw_t = f(x,y,t), \\ \\ w_x - v_y + yw_t - xv_t = g(x,y,t). \end{cases} \tag{3.17}$$

admits no solutions! This is even worse than the worst-case failures we saw when handling Cauchy data prescribed on a characteristic: at least in that situation, non-existence only occurred for the Cauchy problem and not the PDE itself. We can also write this system as a scalar PDE for a complex-valued function $u = v + iw$ defined on $(z,t) \in \mathbb{C} \times \mathbb{R}$. Letting $z = x + iy$ and $F = f + ig$, (3.17) becomes

$$\frac{\partial u}{\partial \overline{z}} - iz\frac{\partial u}{\partial t} = F(z,t).$$

For further details on Lewy's example see [4, Ch. 8].

4 An Introduction to Shock Formation for Quasilinear First-Order PDE

We have seen so far that quasilinear Cauchy problems always admit local solutions, provided the noncharacteristic condition holds. However, we have also seen that we can't say much *a priori* about global solutions, even when the PDE in question is linear. Now, we explore another way that global existence can eventually break down, namely the formation of **shocks** or discontinuities in our solution. We'll also see how to continue solutions past the time of shock formation by introducing an appropriate notion of "weak solution". Finally, we'll explore through examples some of the key ideas related to computing "weak solutions" and how to prove such solutions are unique.

4.1 Finite-Time Derivative Blowup for Conservation Laws

We consider here the initial-value problem for one-dimensional scalar **conservation laws**. These involve choosing two smooth functions

$$f = f(z)\colon \mathbb{R} \to \mathbb{R} \quad (\text{the } \textbf{flux function}) \quad \text{and}$$

$$u_0 = u_0(x)\colon \mathbb{R} \to \mathbb{R},$$

and $T > 0$, then trying to find

$$u = u(t, x) \colon [0, T] \times \mathbb{R} \to \mathbb{R}$$

solving

$$\begin{cases} u_t + \partial_x (f(u)) = 0 & \forall\ (t, x) \in (0, T] \times \mathbb{R} \\ \qquad\quad u|_{t=0} = u_0(x) & \qquad\qquad \forall\ x \in \mathbb{R} \end{cases}. \tag{4.1}$$

Notice that such a PDE is simply the infinitesimal expression of the following integral conservation law: for all $a, b \in \mathbb{R}$ with $a < b$,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_a^b u \ \mathrm{d}x = - \left[f(u(t, x)) \right]_a^b .$$

Pretending that $u(t, x)$ is a mass density function, the above says that the mass inside $[a, b]$ can only change if there is a net flux into or out of $[a, b]$ (in this sense, calling f a "flux function" is sane). Since conservation laws express the painfully obvious fact that conserved things are conserved, they are ubiquitous in mathematical physics, most notably in fluid dynamics (really, I should say their higher-dimensional analogues are ubiquitous).

For concreteness, most of our discussion will centre around the special case

$$f(u) = \frac{1}{2} u^2,$$

giving rise to

$$\begin{cases} u_t + u u_x = 0 & \forall\ (t, x) \in (0, T] \times \mathbb{R} \\ \quad u|_{t=0} = u_0(x) & \qquad\qquad \forall\ x \in \mathbb{R} \end{cases}, \tag{4.2}$$

called the initial value problem for the **(inviscid) Burgers equation**. In one dimension, we can get a good picture of what can go wrong for a general conservation law by examining Burgers' equation in detail.

In practice, you should think of Burgers' equation as a nonlinear variant of the transport equation: the constant transport speed is replaced with the value of the solution $u(t, x)$ itself. Assuming $u_0 > 0$, by analogy with the transport equation we can guess $u > 0$ for all times (we'll prove this in detail later). Thus, points higher on the graph of $u(t, x)$ move to the right faster than points lower on the graph. Eventually, we may expect that the high points will outpace the low points and the solution will become multiply-valued! See Figure 5 for an illustration of this wave-steepening process. So, using simple intuition, we guess that global existence generically fails for (4.2).

We use MOC to prove that this intuition is perfectly accurate. We start by writing out

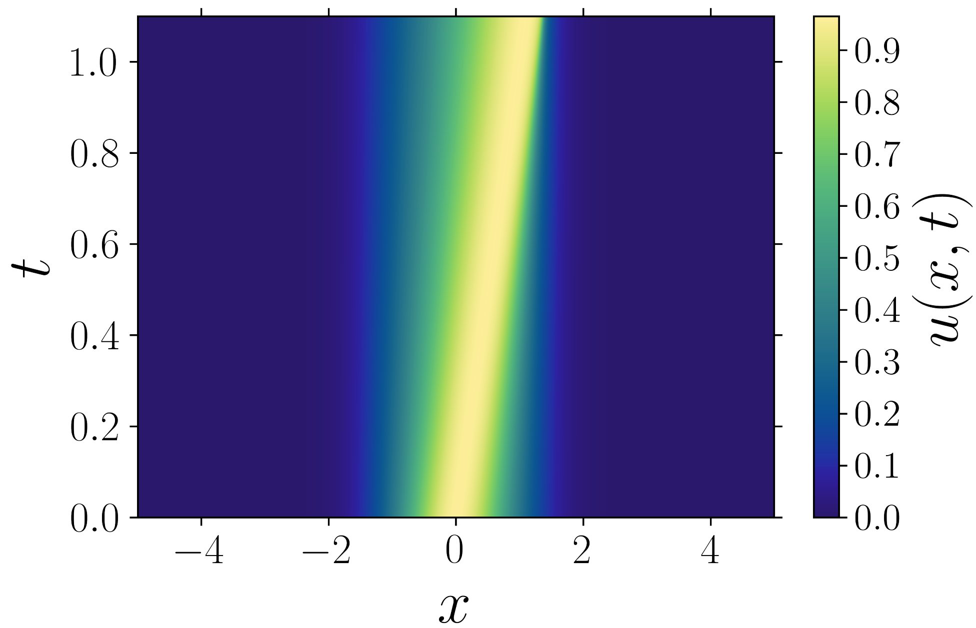

Figure 5: Filled spacetime contour plot of nice initial data evolving according to Burgers' equation. Note how the solution steepens throughout the motion as points of large u move faster.

the characteristic ODEs for (4.2):

$$\frac{\mathrm{d}t}{\mathrm{d}s}(s;\tau) = 1,$$

$$\frac{\mathrm{d}x}{\mathrm{d}s}(s;\tau) = u(t(s;\tau), x(s;\tau)),$$

$$\frac{\mathrm{d}u}{\mathrm{d}s}(s;\tau) = 0.$$

We find that

$$\begin{pmatrix} 1 \\ u_0(x) \end{pmatrix} \cdot (1,0) = 1 \neq 0$$

so the initial data is noncharacteristic and a unique local-in-time solution is guaranteed to exist by theorem 3.3. Parameterizing the initial data curve by $\tau \mapsto (0,\tau)$ and solving the characteristic equations tells us that

$$t(s;\tau) = s,$$

$$u(t(s;\tau), x(s;\tau)) = u_0(\tau),$$

$$x(s;\tau) = \tau + s u_0(\tau).$$

I emphasize that the characteristics (t, x) depend strongly on the initial data u_0; this is a hallmark of nonlinearity. Re-arranging the above formulas gives the following implicit expression for $u(t, x)$:

$$u(t,x) = u_0 \left(x - u(t,x)t\right). \tag{4.3}$$

The reader should compare this with the general solution of the transport equation. Additionally, (4.3) implies the following:

Lemma 4.1 (Weak Maximum Principle for Burgers' Equation)**.** *Suppose* $u_0 \colon \mathbb{R} \to \mathbb{R}$ *is smooth and nonnegative. Then, for all* t *such that a solution* $u(t, x)$ *to* (4.2) *exists, that solution obeys*

$$u(t,x) \geq 0 \quad \forall\, x \in \mathbb{R}.$$

□

Now, I cautioned you at the beginning of this section that eventually the local solution to a quasilinear Cauchy problem is liable to break down in finite time. In other words, we may not be able to take $T = +\infty$ in (4.2). To see why, let's sketch the characteristics for (4.3) in the special case of a Gaussian initial state

$$u_0(x) = e^{-x^2}.$$

See figure 6. From this figure, we see a huge problem: the characteristics eventually intersect at some $t = t_b$! Since the whole point of characteristics is to give a coordinate system where solving our PDE becomes trivial, a situation where the characteristics *fail to provide a coordinate system at all* definitely represents a breakdown of the method. The "solution" would have to be multivalued at the spacetime intersection point. We must then impose $T < t_b$.

How do we calculate t_b explicitly? To start, observe that figure 6 substantiates our intuition that the solution becomes multivalued because points high on the graph of $u(t, x)$ eventually overtake points lower on the graph: the characteristics with $\tau \approx 0$ have a lower slope in the xt plane than the characteristics with $\tau \gg 1$. Now, to any reasonable person, the model of a multivalued function is a vertical line, or a line with infinite slope. Consequently, the first time t_b where the solution becomes multivalued is the first time where

$$u_x(t_b, x) = +\infty.$$

Using (4.3) and the chain rule gives, for any $t \in (0, t_b)$,

$$u_x(t,x) = \frac{u_0'(x-ut)}{1+tu_0'(x-ut)}. \tag{4.4}$$

Therefore, if $u_0' < 0$ somewhere (as it is for $u_0 = e^{-x^2}$), we can expect the derivative to blow up at time

$$t_b = \inf_{\tau \in \mathbb{R}} \left(-\frac{1}{u_0'(\tau)}\right), \tag{4.5}$$

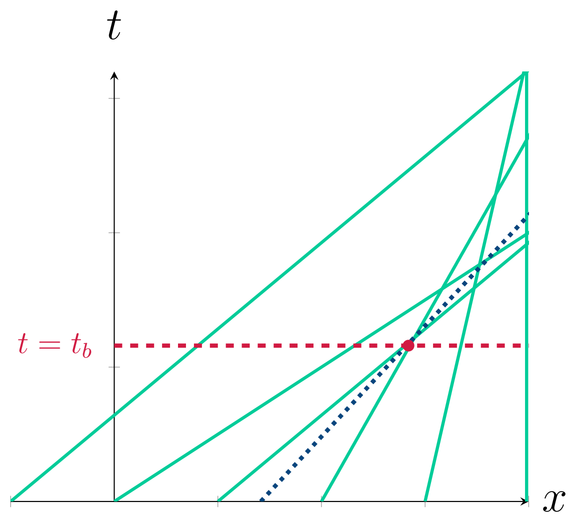

Figure 6: Sketch of the characteristics for Burgers' equation with $u_0 = e^{-x^2}$. The dotted line indicates the characteristic along which a discontinuity forms at the first intersection time: we show in the main text that this characteristic has label $\tau = 2^{-\frac{1}{2}}$.

where we have recalled that $\tau = x - ut$. If u_0 is monotonically increasing, then blowup never occurs, but if u_0 is nonconstant, smooth, and rapidly decaying it eventually needs to have a local extremizer, meaning that the risk of blowup is present. For example, with our Gaussian initial state, a first-year calculus exercise shows that the minimum of $u_0'(x)$ is

$$\min_{\tau\in\mathbb{R}} u_0'(\tau) = -\sqrt{\frac{2}{e}}$$

and it occurs at

$$\tau_b = 2^{-\frac{1}{2}}.$$

Consequently, the blowup time is

$$t_b = \sqrt{\frac{e}{2}} \approx 1.16.$$

For a general conservation law, one can show that (4.5) generalizes to

$$t_b = \inf_{\tau\in\mathbb{R}} \left(-\frac{1}{\frac{\mathrm{d}}{\mathrm{d}\tau}\left(f'(u_0(\tau))\right)} \right). \tag{4.6}$$

For Burgers' equation, $\frac{\mathrm{d}}{\mathrm{d}\tau}\left(f'(u_0(\tau))\right) = f''(u_0)u_0' = u_0'$, so this formula is consistent. I remark that this expression implies that derivative blowup is related to *convexity* of the flux function.

In summary, we have shown that the Cauchy problem for a quasilinear conservation law typically does not admit global-in-time solutions. In particular, the method of characteristics may be used to show that nice initial data give rise to solutions that become multivalued in finite time. Really, since we want to deal with *functions*, we should say "discontinuous" instead of "multivalued", and we do so in the sequel.

4.2 Integral Solutions of Conservation Laws or: How I Learned to Stop Worrying and Love Propagating Bad Initial Data

So far, we have seen that solutions of quasilinear equations can become nasty in finite time. But, such nastiness often has physical significance. For instance, the breaking of waves against a beach should reasonably be modelled by a nice, localized initial state eventually overturning on itself like a solution to Burgers' equation. A broken wave on the beach does not pop out of existence, so there should be some mathematically consistent way of describing the propagation of the wave *after* it breaks. In other words, the Universe is OK with discontinuity. Accordingly, we might ask if there is a way to continue solutions to Burgers' equation past the blowup time t_b. Of course, this would require weakening our notion of "solution" to include discontinuous functions. By now, we all know the right way to do this is to hide all the discontinuity under an integral sign, leading to...

Definition 4.2. *We say* $v \in [0,\infty)\times\mathbb{R} \to \mathbb{R}$ *is a* ***test function*** *if it is smooth and has compact support.*

Note that the support of v does not have to be strictly contained inside $[0,\infty)\times\mathbb{R}$. That is, for an arbitrary test function v, $x \mapsto v(0,x)$ is not necessarily identically 0.

Definition 4.3. *We say* $u \in L^{\infty}([0,T]\times\mathbb{R})$ *is an* ***integral solution*** *to* (4.1) *if, for all test functions* $v(t,x)$*, we have*

$$0 = \int_0^{\infty}\int_{-\infty}^{\infty} uv_t + f(u)v_x \ \mathrm{d}x\ \mathrm{d}t + \int_{-\infty}^{\infty} u_0 v|_{t=0}\ \mathrm{d}x. \tag{4.7}$$

Integration by parts shows that, if $u(t,x)$ is a C^1 solution to (4.1), then it is also an integral solution (do the computation!). Notice also that, if we're in the business of looking for integral solutions, we can easily take initial data $u_0 \in L^{\infty}(\mathbb{R})$ instead of asking for a smooth u_0.

Hopefully, then, we should be able to show that after a discontinuity forms in a classical C^1 solution to some conservation law, it can be continued for a longer time as an integral solution. After the C^1 solution breaks, the resulting discontinuity should propagate through spacetime along a smooth curve $\gamma = \{(t,x) = (t,\gamma(t))\}$; for now, we wantonly assume the initial data has been prepared so that only a single discontinuity is allowed to form. On either side of the curve, then, the solution should remain smooth. Thus, to get further insight into how discontinuities propagate, we should try to derive a formula for this $\gamma(t)$.

From now on, we specialize to integral solutions $u(t,x)$ to (4.1) that are smooth on the complement of the smooth curve $\gamma = \{(t,x) = (t,\gamma(t))\}$. The curve γ cuts $(0,\infty)\times\mathbb{R}$ into two open pieces V_ℓ, V_r such that $u|_{V_\ell}$ and $u|_{V_r}$ are both smooth. Let $\mathbf{n}$ denote the unit normal to γ pointing out of V_ℓ and into V_r. Explicitly, a second-year calculus exercise shows

$$\mathbf{n} = \frac{(-\gamma'(t),1)^{\mathrm{T}}}{\left|(-\gamma'(t),1)^{\mathrm{T}}\right|}. \tag{4.8}$$

Using the definition of an integral solution, we have for all test functions v that

$$\begin{aligned}0 = &\int_{V_\ell} uv_t + f(u)v_x \ \mathrm{d}x \ \mathrm{d}t + \int_{\mathbb{R}_x \cap V_\ell} u_0 v|_{t=0} \ \mathrm{d}x \\ &+ \int_{V_r} uv_t + f(u)v_x \ \mathrm{d}x \ \mathrm{d}t + \int_{\mathbb{R}_x \cap V_r} u_0 v|_{t=0} \ \mathrm{d}x.\end{aligned}$$

Define

$$u_\ell = (u|_{V_\ell})\,|_\gamma, \quad u_r = (u|_{V_r})\,|_\gamma.$$

Applying the divergence theorem in V_ℓ and V_r and using the compact support of the test function v gives

$$\begin{aligned}0 &= -\int_{V_\ell} v\left(u_t + \partial_x\left(f(u)\right)\right) \ \mathrm{d}x \ \mathrm{d}t - \int_{V_r} v\left(u_t + \partial_x\left(f(u)\right)\right) \ \mathrm{d}x \ \mathrm{d}t \\ &\quad + \int_\gamma \left[(u_\ell, f(u_\ell))^{\mathrm{T}} \cdot \mathbf{n} - (u_r, f(u_r))^{\mathrm{T}} \cdot \mathbf{n}\right] v \ \mathrm{d}s \\ &\quad - \int_{\mathbb{R}_x \cap V_\ell} v \ \ (u, f(u))^{\mathrm{T}} \cdot (-1, 0)^{\mathrm{T}} \ \mathrm{d}x - \int_{\mathbb{R}_x \cap V_r} v \ \ (u, f(u))^{\mathrm{T}} \cdot (-1, 0)^{\mathrm{T}} \ \mathrm{d}x \\ &\quad + \int_{\mathbb{R}_x \cap V_\ell} v \ \ (u, f(u))^{\mathrm{T}} \cdot (-1, 0)^{\mathrm{T}} \ \mathrm{d}x + \int_{\mathbb{R}_x \cap V_r} v \ \ (u, f(u))^{\mathrm{T}} \cdot (-1, 0)^{\mathrm{T}} \ \mathrm{d}x \\ &= -\int_{V_\ell} v\left(u_t + \partial_x\left(f(u)\right)\right) \ \mathrm{d}x \ \mathrm{d}t - \int_{V_r} v\left(u_t + \partial_x\left(f(u)\right)\right) \ \mathrm{d}x \ \mathrm{d}t \\ &\quad + \int_\gamma \left[(u_\ell, f(u_\ell))^{\mathrm{T}} \cdot \mathbf{n} - (u_r, f(u_r))^{\mathrm{T}} \cdot \mathbf{n}\right] v \ \mathrm{d}s,\end{aligned}$$

where $\mathrm{d}s$ is the length element along γ. Remember that u is a classical solution to the PDE in V_ℓ and V_r, so the above simplifies to

$$0 = \int_\gamma \left[(u_\ell, f(u_\ell))^{\mathrm{T}} \cdot \mathbf{n} - (u_r, f(u_r))^{\mathrm{T}} \cdot \mathbf{n}\right] v \ \mathrm{d}s.$$

Since v is arbitrary, we conclude

$$(u_\ell, f(u_\ell))^{\mathrm{T}} \cdot \mathbf{n} = (u_r, f(u_r))^{\mathrm{T}} \cdot \mathbf{n}.$$

Using (4.8) and introducing the jump notation

$$[[\phi]] \doteq \phi_\ell - \phi_r,$$

we obtain the **Rankine-Hugoniot jump condition (RH)**

$$\gamma'(t) = \frac{[[f(u)]]}{[[u]]}. \tag{4.9}$$

RH tells us that we can find the speed of a propagating discontinuity provided we can compute

1. the jump in our solution across the discontinuity, and
2. the jump in our flux across the discontinuity.

This allows us to reverse-engineer the discontinuity curve itself (the integration constant will be obvious in most examples). Of course, for a hard problem like continuing our Gaussian initial state through the blowup time t_b, RH might be a little tough to apply in practice. So, for the remainder of these notes, we only use simple model problems with piecewise constant initial data u_0. These are called **Riemann problems**. Knowing how to solve Riemann problems actually allows one to construct efficient numerical methods for solving more complex problems with propagating discontinuities, such as the breaking Gaussian. For details on such numerical methods, see [6].

Example 4.4 (Compressive Riemann Problem for Burgers)**.** Consider solving Burgers' equation (4.2) with the initial state

$$u_0(x) = \begin{cases} 1 & x \leq 0 \\ 0 & x > 0 \end{cases}. \tag{4.10}$$

We already know the characteristics for this problem from our earlier work:

$$\begin{pmatrix} t(s;\tau) \\ x(s;\tau) \end{pmatrix} = \begin{pmatrix} s \\ \tau + s u_0(\tau) \end{pmatrix}.$$

Figure 7 shows these characteristics, and we see they are essentially useless since they are guaranteed to intersect immediately after $t = 0$. We can resolve this by allowing the initial discontinuity at $x = 0$ to propagate with speed given by the Rankine-Hugoniot condition. Taking inspiration from the (admittedly flawed) characteristics, we want to have a discontinuity moving from $-\infty$ to $+\infty$: the $x \leq 0$ part of the initial solution wants to move to the right, and the $x > 0$ part wants to remain stationary (intuitively, this is why such initial data with a downward jump is called "compressive"). Owing to the simplicity of our initial data, we know that for all times t we shall have

$$u_\ell = 1, \quad u_r = 0.$$

Since Burgers' equation has flux $f(u) = \frac{1}{2}u^2$, the Rankine-Hugoniot condition (4.9) gives

$$\gamma'(t) = \frac{1}{2}.$$

Using that the discontinuity is located at $x = 0$ when $t = 0$, our curve of discontinuity must be

$$\gamma = \left\{ (t, x) = \left(t, \frac{1}{2}t \right) \right\}.$$

The strategy is then to form a coherent picture of the solution by drawing this line γ, then drawing the characteristics, with the caveat that we stop drawing when the characteristics touch γ. This is done in figure 8. Intuitively, such a procedure makes sense: we should be able

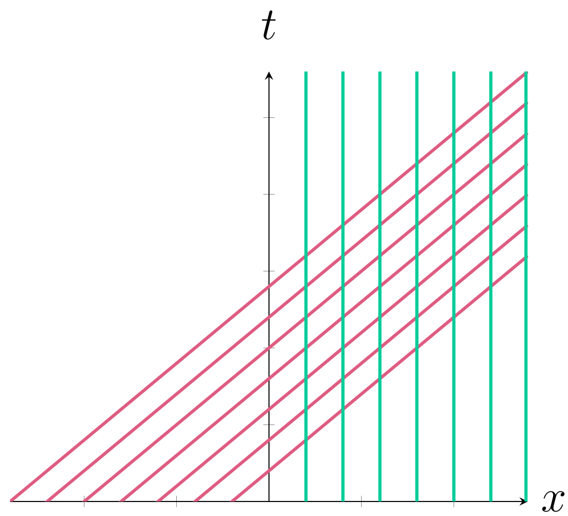

Figure 7: Sketch of the characteristics for Burgers' equation with compressive data given by (4.10). We can introduce a shock to fix the ambiguity arising from the intersections.

to flow along a characteristic until we hit the discontinuity, at which point the characteristic gets "eaten" by the discontinuity front and travels with it. This procedure shows that

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

is an RH-satisfying integral solution of Burgers' equation with initial state (4.10).

Example 4.5 (Expansive Riemann Problem for Burgers)**.** Now, we consider the opposite problem to the previous example, namely solving (4.2) with the initial state

$$u_0(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}. \tag{4.11}$$

Since the opposite of compression is expansion, we call this kind of upward-jumping initial data "expansive". Figure 9 shows the characteristics for this initial state. Right away, something is off: while characteristics never intersect, there is a wedge in the $x > 0$ sector that no characteristic touches! So, we have to somehow fill this sector in with information.

Following the developments of the previous example, we may want to introduce a moving discontinuity. If we reasonably assume that we always have

$$u_\ell = 0, \quad u_r = 1,$$

then RH gives that the line of discontinuity must again be

$$\gamma = \left\{ (t,x) = \left(t, \frac{1}{2}t \right) \right\}.$$

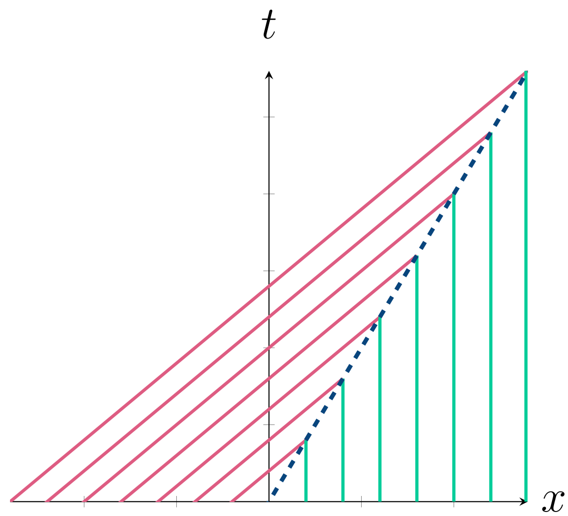

Figure 8: Sketch of the characteristics for Burgers equation with compressive initial data (4.10), but this time a discontinuity has been introduced to resolve the intersections from figure 7. The dashed line gives the trajectory of the discontinuity, with slope 2.

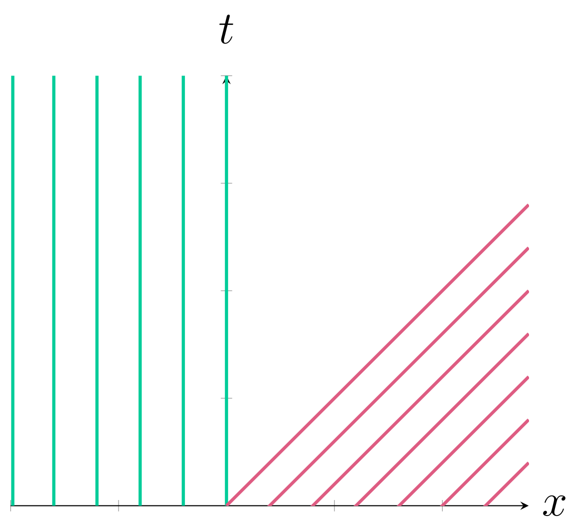

Figure 9: Sketch of the characteristics for Burgers' equation with expansive initial data (4.11). To construct a globally-defined integral solution, we must fill in the blank region!

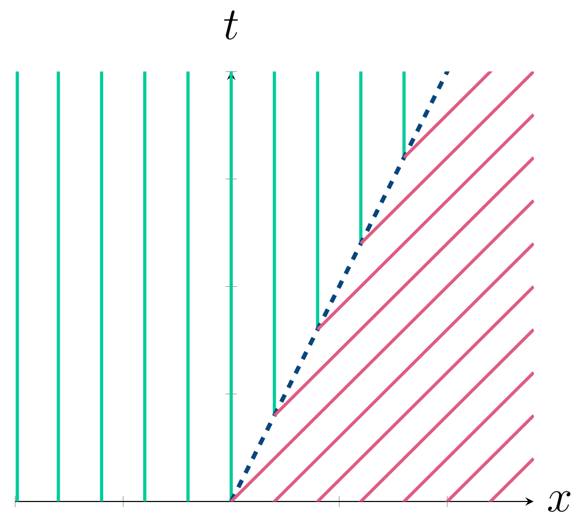

Figure 10: Sketch of the characteristics for Burgers equation with initial data (4.11), where a line of discontinuity (dashed) has been added to fix the issues with figure 9.

So, we want to add this line to our characteristic sketch in figure 9. To make this addition self-consistent, we must have $u \equiv 0$ above the discontinuity and $u \equiv 1$ below the discontinuity. As demonstrated in figure 10, this is enough to fill up the blank wedge and give us an integral solution:

$$u(t,x) = \begin{cases} 0 & x \leq \frac{t}{2} \\ 1 & x > \frac{t}{2} \end{cases} . \tag{4.12}$$

However, in the world of nonlinear PDEs, nothing is easy. The integral solution we have just constructed is not unique. To see this, notice that the "raw" characteristics in figure 9 do not cross, so introducing a discontinuity is not strictly necessary. Instead, we could fill up the blank wedge by saying that the solution continuously, but not smoothly, interpolates between the left state $u \equiv 0$ and the right state $u \equiv 1$. Said differently, the information contained in the initial discontinuity expands out to nicely fill the empty region, rather than propagating through space as a line of discontinuity. This would mean setting

$$u(t,x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{t} & 0 < x < t \\ 1 & x \geq t \end{cases} . \tag{4.13}$$

The reader should verify that this function is indeed an integral solution to Burgers' equation. See figure 11 for a sketch of this solution. The $0 < x < t$ lines in this figure make up what is called a **rarefaction fan** (rarefaction is a synonym for expansion).

We conclude with a bit of deeper discussion about the expansive Riemann problem. We have shown solutions are not unique, but which of the two solutions should we prefer, if indeed one is preferable at all? We can demand (after the discussion in [2, §3.4]) that following a characteristic backward in time should never lead us to a discontinuity, except

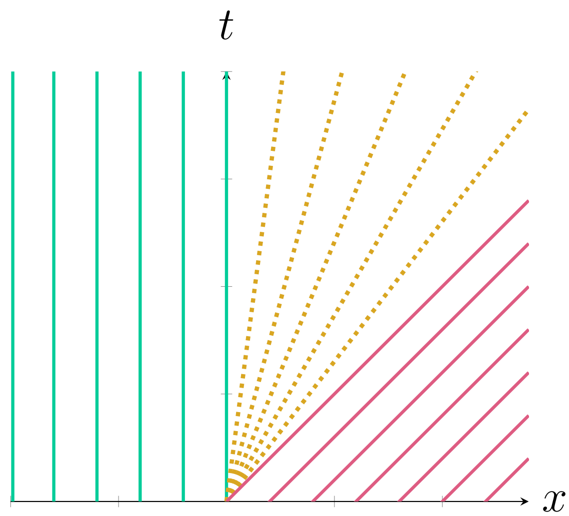

Figure 11: Sketch of the characteristics for Burgers equation with the expansive initial data (4.11), where a rarefaction fan (represented here by dotted lines) has been added to fix the issues in figure 9.

for a discontinuity at $t = 0$. This is a qualitative way of prescribing an **entropy condition** on our integral solutions. Satisfactorily explaining the name "entropy" without going through a computational engineering graduate program is impossible, so for now I'll simply state that the entropy condition quantitatively thus: if we have an integral solution $u(t, x)$ to a conservation law, then for each curve of discontinuity $\gamma(t)$ in our solution, we must have

$$f'(u_\ell) > \gamma'(t) > f'(u_r) . \tag{4.14}$$

The reader should think about why this inequality agrees with the qualitative statement above. Additionally, the reader can quickly verify that the discontinuous solution to the compressive Riemann problem indeed satisfies the entropy condition, while the discontinuous solution to the expansive problem does not! Therefore, for Burgers equation, we allow discontinuities only for compressive initial data, and for expansive initial data we instead introduce rarefaction fans. All this leads to a fun new word:

Definition 4.6. *Suppose $u(t, x)$ is an integral solution to* (4.1) *with a moving discontinuity given by the space-time curve*

$$\gamma = \{(t, x) = (t, \gamma(t))\} .$$

We say that the line of discontinuity constitutes a ***shock*** *if*

- *the Rankine-Hugoniot condition* (4.9) *is satisfied, and*
- *the entropy condition* (4.14) *is satisfied.*

In particular, the solution to the compressive Riemann problem for Burgers' equation includes a shock.

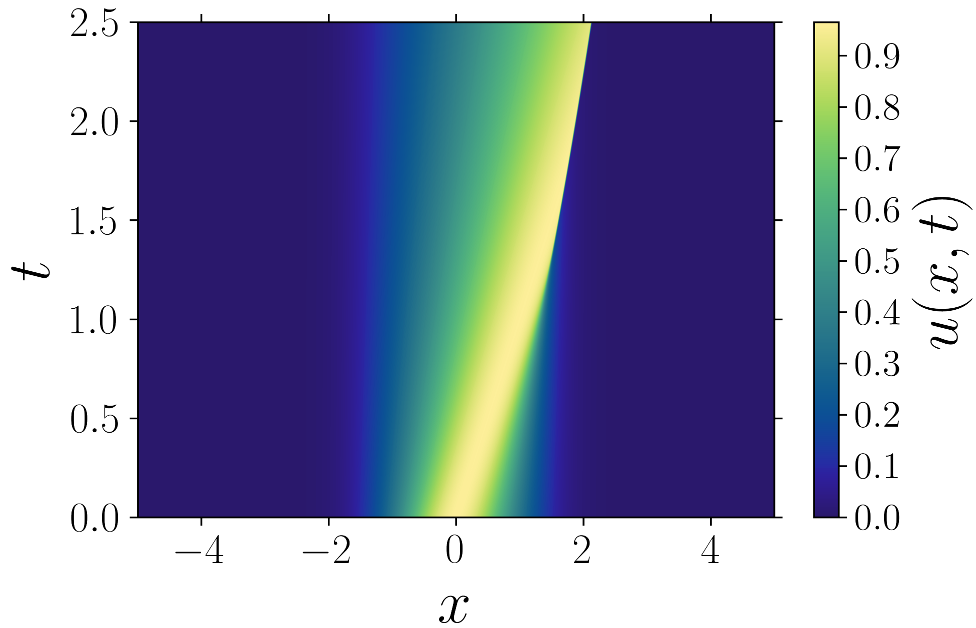

Figure 12: Filled spacetime contour plot of a Gaussian evolving according to Burgers' equation. Note the shock formation around $t \approx 1.16$, as we predicted in the main text.

Remark. *The notion of an entropy condition is, as we have already seen, critical to refining our class of integral solutions and therefore critical to developing a useful global well-posedness theory for conservation laws. The reader looking to learn this theory rigorously can find introductions in various books and lecture notes, including* [1]*,* [3]*, or* [10]*. On a related note, entropy conditions also allow us to design* ***conservative*** *numerical methods that can sensibly reproduce shocks and other interesting integral solutions: see Figure 12 for a plot of such a conservative numerical solution to Burgers' equation. Note how this numerical solution is valid even past the shock time!*

In summary, we have shown the following:

1. solutions of the Cauchy problem for quasilinear conservation laws generically become discontinuous in finite time t_b;
2. these discontinuous solutions can be continued past t_b as weakened "integral solutions", for example, shocks or rarefaction fans;
3. the Rankine-Hugoniot condition and the entropy condition allow one to determine the correct spacetime path traced out by a moving discontinuity.

References

[1] Gregory A. Chechkin and Andrey Yu. Goritsky, *S. N. Kruzhkov's lectures on first-order quasilinear PDEs*, Analytical and numerical aspects of partial differential equations, 2009, pp. 1–67. Translated from the Russian by Boris P. Andreianov.

[2] Lawrence C. Evans, *Partial differential equations*, second edition, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010.

[3] Lars Hörmander, *Lectures on nonlinear hyperbolic differential equations*, Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 26, Springer-Verlag, Berlin, 1997.

[4] Fritz John, *Partial differential equations*, fourth edition, Applied Mathematical Sciences, vol. 1, Springer-Verlag, New York, 1982.

[5] John M. Lee, *Introduction to smooth manifolds*, second edition, Graduate Texts in Mathematics, vol. 218, Springer, New York, 2013.

[6] Randall J. LeVeque, *Finite volume methods for hyperbolic problems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2002.

[7] Hans Lewy, *An example of a smooth linear partial differential equation without solution*, Ann. of Math. (2) **66** (1957), 155–158.

[8] J. David Logan, *An introduction to nonlinear partial differential equations*, second edition, Pure and Applied Mathematics (Hoboken), Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2008.

[9] Stanley Osher and Ronald Fedkiw, *Level set methods and dynamic implicit surfaces*, Applied Mathematical Sciences, vol. 153, Springer-Verlag, New York, 2003.

[10] Joel Smoller, *Shock waves and reaction-diffusion equations*, second edition, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 258, Springer-Verlag, New York, 1994.