AST560 2025: Hyperbolic Partial Differential Equations

Ammar H. Hakim

March 20, 2025

Contents

1 On Propagation of Small Perturbations. Definition of Hyperbolic PDEs.

Hyperbolic partial differential equation (PDEs) appear everywhere in physics. A simple reason for this is that time exists (as far as we know), and that there is a maximum speed at which causal signals can propagate. In this section we first work through an informal definition of hyperbolic PDEs, giving the rigorous definition later.

Definition 1 (Hyperbolic PDE (Informal))**.** *A hyperbolic PDE is in one in which small perturbations around a uniform equilibrium propagate at a finite speed without any damping or dispersion.*

To understand how this definition leads to a mathematical condition, let us look at a generic system of first-order partial differential equations in 1D

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0 \tag{1}$$

where $\mathbf{Q}(x,t)$ is a column vector of m *conserved variables* and $\mathbf{F} = \mathbf{F}(\mathbf{Q})$ is the corresponding column vector of *fluxes*. To understand how a small perturbation propagates in this system we will linearize around a uniform initial condition $\mathbf{Q}(x,0) = \mathbf{Q}_0$ by writing

$$\mathbf{Q}(x,t) = \mathbf{Q}_0 + \mathbf{Q}_1(x,t) \tag{2}$$

where $\mathbf{Q}_1(x,t)$ is a small perturbation. Substituting this in Eq. (1) we get

$$\frac{\partial \mathbf{Q}_1}{\partial t} + \frac{\partial}{\partial x}\mathbf{F}(\mathbf{Q}_0 + \mathbf{Q}_1) = 0. \tag{3}$$

Expanding

$$\mathbf{F}(\mathbf{Q}_0 + \mathbf{Q}_1) = \mathbf{F}(\mathbf{Q}_0) + \frac{\partial \mathbf{F}}{\partial \mathbf{Q}}(\mathbf{Q}_0)\mathbf{Q}_1 + \dots \tag{4}$$

we get

$$\frac{\partial \mathbf{Q}_1}{\partial t} + \mathbf{A}(\mathbf{Q}_0)\frac{\partial \mathbf{Q}_1}{\partial x} = 0. \tag{5}$$

Here we have dropped all quadratic and higher nonlinearities and defined

$$\mathbf{A} \equiv \frac{\partial \mathbf{F}}{\partial \mathbf{Q}} \tag{6}$$

as the *flux Jacobian*. Note that the flux Jacobian is a $m \times m$ matrix and, in general, will depend on the state $\mathbf{Q}$. When $\mathbf{A}$ does not depend on $\mathbf{Q}$ the system is called *linear*. The system Eq. (5) is the linearized equation for the propagation of the perturbation $\mathbf{Q}_1$. To understand the structure of this equation we can look at a single mode $\mathbf{Q}_1(x,t) = e^{ikx}e^{-i\omega t}$ to get

$$[k\mathbf{A}(\mathbf{Q}_0) - \omega\mathbf{I}]\mathbf{Q}_1 = 0 \tag{7}$$

where $\mathbf{I}$ is a $m \times m$ unit matrix. This is an eigenvalue problem and we see that the propagation of the mode will depend on the eigenvalues of the flux Jacobian $\mathbf{A}(\mathbf{Q}_0)$. Let λ^p be the eigenvalues of $\mathbf{A}(\mathbf{Q}_0)$. Then the *dispersion relation* is

$$\omega^p = k\lambda^p. \tag{8}$$

Clearly, the mode will propagate without any damping if all the λ^p are *real*. Also, there is no dispersion in this system as both the phase- and group-velocities are constant. Hence, this linearization shows that the PDE Eq. (1) will be hyperbolic if the eigenvalues of the flux Jacobian are *all real*.

Though we arrived at the condition for hyperbolicity via a linearization, this condition is, in fact, the defining property of hyperbolic PDEs even in the nonlinear case. This leads to the rigorous definition:

Definition 2 (Hyperbolic PDE)**.** *A PDE*

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0$$

is called hyperbolic *if the flux Jacobian*

$$\mathbf{A} \equiv \frac{\partial \mathbf{F}}{\partial \mathbf{Q}}$$

has all real eigenvalues and a complete set of right eigenvectors (or, in other words, is diagonalizable) for all valid solution states $\mathbf{Q}$*. The system is called* strictly hyperbolic *if all eigenvalues are distinct and non-zero.*

Remark 1. *Note that in this definition we have added an additional condition: that* $\mathbf{A}$ *must be diagonalizable. We will see why this condition is needed below. When this condition is relaxed strange things can happen, for example, the formation of delta-function solutions. Note we also added the phrase "for all valid solution states". In general, not all solution states are valid. For example, for the Euler equations, which are hyperbolic, valid states are those in which the* density and pressure are positive*. Often the valid states of a PDE are also referred to as their* invariant domains*.*

Remark 2. *Most of the hyperbolic PDEs we encounter in physics are* not *strictly hyperbolic. There are often repeated eigenvalues and even eigenvalues that are zero.*

We will denote λ^p as the eigenvalues of the flux Jacobian, and $\mathbf{r}^p$, $p = 1, \ldots, m$ as the complete set of right eigenvectors (column vectors). Corresponding to these right eigenvectors we will denote the left eigenvectors (row vectors) as l^p. Note we must have $\mathrm{l}^p \mathbf{r}_q = \delta^p{}_q$.

We can extend this definition to a PDE in multiple dimensions as follows.

Definition 3 (Multi-Dimensional Hyperbolic PDE)**.** *A PDE*

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial y} + \frac{\partial \mathbf{H}}{\partial z} = 0$$

is called hyperbolic *if, for all unit vectors* $\mathbf{n} = n_x \boldsymbol{\sigma}_x + n_y \boldsymbol{\sigma}_y + n_z \boldsymbol{\sigma}_z$*, the flux Jacobian*

$$\mathbf{A} \equiv n_x \frac{\partial \mathbf{F}}{\partial \mathbf{Q}} + n_y \frac{\partial \mathbf{G}}{\partial \mathbf{Q}} + n_z \frac{\partial \mathbf{H}}{\partial \mathbf{Q}}$$

has all real eigenvalues and a complete set of right eigenvectors (or, in other words, is diagonalizable) for all valid solution states $\mathbf{Q}$*. We call the hyperbolic PDE* isotropic *if the eigenvalues of* $\mathbf{A}$ *do not depend on the unit vector* $\mathbf{n}$*.*

Physically, if a hyperbolic PDE is isotropic, then the speed of propagation does not depend on the direction of the wave. Most hyperbolic PDEs one encounters are isotropic.

Though we have written the PDE (see Eq. (1)) in conservation form, a more general form is the *quasilinear* form

$$\frac{\partial \mathbf{V}}{\partial t} + \mathbf{B} \frac{\partial \mathbf{V}}{\partial x} = 0. \tag{9}$$

Here $\mathbf{V}$ is a column vector of m quantities and $\mathbf{B} = \mathbf{B}(\mathbf{V})$ is a $m \times m$ matrix that, in general, depends on $\mathbf{V}$. We call a quasilinear PDE hyperbolic if the eigenvalues of $\mathbf{B}$ are all real and $\mathbf{B}$ is diagonalizable.

We can transform a equation written in conservation law form to the quasilinear form, but it is not always possible to do the opposite. Consider the 1D conservation law

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0 \tag{10}$$

and let $\mathbf{V}$ be related by the invertible transform.

$$\mathbf{Q} = \varphi(\mathbf{V}). \tag{11}$$

Then, substituting this in the conservation law we get the quasilinear form

$$\frac{\partial \mathbf{V}}{\partial t} + \left(\varphi'\right)^{-1}\mathbf{A}\varphi'\frac{\partial \mathbf{V}}{\partial x} = 0. \tag{12}$$

Here φ' is the Jacobian of the transform and $\mathbf{A}$ is the flux Jacobian of the conservation law. Define

$$\mathbf{B} \equiv \left(\varphi'\right)^{-1}\mathbf{A}\varphi'. \tag{13}$$

Let λ^p, $\mathbf{r}^p$ and l^p be the eigenvalues, right- and left-eigenvectors of $\mathbf{B}$ respectively. Then, we have

$$\mathbf{B}\mathbf{r}^p = \lambda^p\mathbf{r}^p = \left(\varphi'\right)^{-1}\mathbf{A}\varphi'\mathbf{r}^p. \tag{14}$$

Left multiply by φ' to see that λ^p are also the eigenvalues of $\mathbf{A}$, and $\varphi'\mathbf{r}^p$ are the right eigenvectors of $\mathbf{A}$. In the same way we see that $\mathrm{l}^p\left(\varphi'\right)^{-1}$ are the left eigenvectors of $\mathbf{A}$. Hence, knowing the eigensytem of the quasilinear form of the equations we can easily compute the eigensytem of the conservation law form. Often, for complex nonlinear equations it is, in fact, much easier to work with the quasilinear form and so the above procedure is very useful.

2 Examples of Hyperbolic PDEs

We will now look at a sequence of examples of hyperbolic PDEs that arize in fluid mechanics and plasma physics. One can multiple the examples given below to many more systems: hyperbolic PDEs are essentially universal, and arise

2.1 Linear Advection Equation

One example we have already seen and studied earlier is the advection equation

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial x}(uf) = 0. \tag{15}$$

This is the simplest *linear* scalar hyperbolic PDE. It has a single eigenvalue $\lambda^1 = u$. Despite its extreme simplicity, the linear advection equation is the fundamental equation that provides insight into solving more complex systems of equations, including nonlinear equations, specially numerically.

2.2 Burgers Equation

The simplest nonlinear scalar hyperbolic PDE is the *Burgers equation*

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) = 0. \tag{16}$$

For this equation, $Q = u$ and $F(Q) = u^2/2$. Hence, the flux Jacobian is just a scalar $A = u$, and the single eigenvalue is $\lambda^1 = u$. Notice something very important: the eigenvalue depends on the solution itself. Hence, the propagation speed, in general, will vary from point to point, and, in fact, may change signs.

2.3 Maxwell Equations

Maxwell equations are the simplest *system* of linear hyperbolic PDEs that occur in physics. In free-space, we can write them in curl form as

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0 \tag{17}$$

$$\epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} = 0 \tag{18}$$

Here, $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic flux density, and ϵ_0, μ_0 are permittivity and permeability of free space. The speed of light is determined from $c = 1/(\mu_0\epsilon_0)^{1/2}$.

These are linear equations and hence the eigensytem is independent of the value of the electromagnetic fields. In 1D Maxwell equations can be written as

$$\frac{\partial}{\partial t}\begin{bmatrix} E_x \\ E_y \\ E_z \\ B_x \\ B_y \\ B_z \end{bmatrix} + \frac{\partial}{\partial x}\begin{bmatrix} 0 \\ c^2 B_z \\ -c^2 B_y \\ 0 \\ -E_z \\ E_y \end{bmatrix} = 0. \tag{19}$$

The eigenvalues of this system are $-c, -c, c, c, 0, 0$. Note the repeated eigenvalues and also the zero eigenvalues. Hence, Maxwell equations are *not* strictly hyperbolic. Despite this, they posses a complete set of right eigenvectors. A simple calculation (easily done in a computer algebra system) shows that the right eigenvectors are the columns of the

following matrix

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & \frac{1}{c} & 0 & -\frac{1}{c} & 0 & 0 \\ -\frac{1}{c} & 0 & \frac{1}{c} & 0 & 0 & 0 \end{bmatrix} \tag{20}$$

and the left eigenvectors are the rows of the matrix

$$\mathbf{L} = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & -\frac{c}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{c}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{c}{2} \\ 0 & 0 & \frac{1}{2} & 0 & -\frac{c}{2} & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{21}$$

2.4 The Isothermal Euler Equations

The simplest system of *nonlinear* hyperbolic PDEs that appears in fluid mechanics is the isothermal Euler equations. These equations describe the motion of an ideal fluid (that is, one with no viscosity or heat-conduction) with a *constant* temperature. In 1D these equations can be written as

$$\frac{\partial}{\partial t}\begin{bmatrix} \rho \\ \rho u \end{bmatrix} + \frac{\partial}{\partial x}\begin{bmatrix} \rho u \\ \rho u^2 + \rho a^2 \end{bmatrix} = 0 \tag{22}$$

where ρ is the fluid mass density, u is the fluid velocity and a is a constant with units of speed. (This is, in fact, the sound speed). The first of these equations is the *continuity equation* and the second is the *momentum equation*.

To compute the eigenvalues of the isothermal Euler equation (and other nonlinear systems) it is actually much easier to work with the quasilinear form. The momentum equation can be written as

$$\rho\frac{\partial u}{\partial t} + \underbrace{u\frac{\partial \rho}{\partial t} + u\frac{\partial}{\partial x}(\rho u)}_{0} + \rho u\frac{\partial u}{\partial x} + a^2\frac{\partial \rho}{\partial x} = 0. \tag{23}$$

where we used the continuity equation to eliminate the second and third terms. Hence, we get that the isothermal Euler equation is equivalent to the quasilinear system

$$\frac{\partial}{\partial t}\begin{bmatrix} \rho \\ u \end{bmatrix} + \begin{bmatrix} u & \rho \\ a^2/\rho & u \end{bmatrix}\frac{\partial}{\partial x}\begin{bmatrix} \rho \\ u \end{bmatrix} = 0. \tag{24}$$

Thus, we have

$$\mathbf{B} = \begin{bmatrix} u & \rho \\ a^2/\rho & u \end{bmatrix}. \tag{25}$$

Computing the eigenvalues of this matrix we get that the eigenvalues of the isothermal Euler equations are $\lambda^{1,2} = u \pm a$. We can also show that as long as $a > 0$, we have a complete set of eigenvectors, hence showing that the isothermal Euler equations are hyperbolic.

2.5 The Euler Equations

Probably the most important hyperbolic PDEs are the Euler equations for ideal fluids. These equations are extremely useful, both by themselves, and also as building blocks of more complex equations in fluid dynamics and plasma physics. For example, the compressible Navier-Stokes equations can be built from these by adding non-ideal terms (viscosity and heat-conduction). For plasma physics, Euler equations can describe *two-fluid* plasmas, in which each plasma species is treated as a separate charged fluid, coupled by electromagnetic fields and collisions. In fact, these two-fluid systems (with additional physics to include non-ideal terms) contain physics beyond MHD and are required to properly study, for example, magnetic reconnection, instabilities in various fusion systems, and space and astrophysical plasmas.

In 1D the Euler equations can be written in conservative form as

$$\frac{\partial}{\partial t}\begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix} + \frac{\partial}{\partial x}\begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uw \\ (E+p)u \end{bmatrix} = 0 \tag{26}$$

where

$$E = \rho\varepsilon + \frac{1}{2}\rho(u^2 + v^2 + w^2) \tag{27}$$

is the total energy and ε is the internal energy of the fluid. The pressure is given by an equation of state (EOS) $p = p(\varepsilon, \rho)$. For an ideal gas the EOS is

$$p = (\gamma - 1)\rho\varepsilon. \tag{28}$$

For a neutral fluid, usually we set $\gamma = 1.4$, while for a plasma $\gamma = 5/3$.

Computing the eigensystem of the Euler equations is not trivial. However, the task is made much more tractable by transforming the system to quasilinear form and using a computer algebra system. We find is that the eigenvalues are

$$\lambda^1, \lambda^2, \lambda^3, \lambda^4, \lambda^5 = u - c, u, u, u, u + c, \tag{29}$$

where c is the sound-speed given by

$$c = \sqrt{\frac{\partial p}{\partial \rho} + \frac{p}{\rho^2}\frac{\partial p}{\partial \varepsilon}}. \tag{30}$$

In general, to compute the sound speed we need the equation of state (EOS). Sometimes, the EOS is given as an analytical formula, but more often is in tabulated form. The National Ignition Facility (NIF), for example, is an experiment to collect the equation of state for elements (and mixtures) used in nuclear weapons. Less aggressively (but more energetically), for example, neutron star crust simulations also require a semi-analytical EOS to properly simulate their dynamics. The eigenvalues are all real if the sound-speed is real. This, of course, depends on the EOS. For the ideal gas EOS, we see that

$$c = \sqrt{\frac{\gamma p}{\rho}}. \tag{31}$$

This remains real as long as $\rho > 0$ and $p > 0$. These conditions on density and pressure (the invariant domain) restrict the valid states a fluid can take.

The right eigenvectors are given by the columns of the matrix

$$R = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ u - c & 0 & 0 & u & u + c \\ v & 1 & 0 & v & v \\ w & 0 & 1 & w & w \\ h - uc & v & w & h - c^2/b & h + uc \end{bmatrix}. \tag{32}$$

Here, h is the *enthalpy* of the fluid given by

$$h = (E + p)/\rho. \tag{33}$$

Also

$$b = \frac{1}{\rho}\frac{\partial p}{\partial \varepsilon}. \tag{34}$$

From this discussion it becomes clear that the Euler equations are an extremely complex, nonlinear system of equations. Things become even more complicated for general EOS. Of course, further extensions to these fundamental systems are possible, in particular, to (special and general) relativistic problems. For high-energy astrophysics problems, in fact, one must use these relativistic extensions to the Euler equations as the energies and speeds encountered in extreme plasma environments (for example, around black holes and neutron stars) can be ultra-relativistic.

2.6 The Ideal MHD Equations

In plasma physics the single most important set of equations are the ideal MHD equations. These are the foundation of all of plasma physics, including the theory of equilibrium and stability of tokamaks and stellarators, and the proper understanding of everything from space to astrophysical plasmas.

Ideal MHD is a pre-Maxwell theory in the sense that the displacement currents are ignored (hence, no electromagnetic waves) and the plasma is treated as a conducting fluid. The equations consist of the *continuity* equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{35}$$

and the *momentum equation*

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{\nabla p}{\rho} = \frac{1}{\mu_0 \rho} (\nabla \times \mathbf{B}) \times \mathbf{B} \tag{36}$$

where $\mathbf{B}$ is the magnetic field. The evolution of the field is determined by the *induction equation*

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0 \tag{37}$$

where $\mathbf{E} = -\mathbf{u} \times \mathbf{B}$ (ideal Ohm's Law). Finally, for an ideal plasma, the pressure evolves according to

$$\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p = -\gamma p \nabla \cdot \mathbf{u}. \tag{38}$$

Using some algebraic manipulations we can show that the momentum equation can be written in the conservative form as

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) + \nabla \cdot \mathbf{T} = 0 \tag{39}$$

where

$$\mathbf{T} = \rho \mathbf{u} \otimes \mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \mathbf{g} - \frac{1}{\mu_0} \mathbf{B} \otimes \mathbf{B}. \tag{40}$$

is a *symmetric* second-order tensor describing the momentum flux, and $\mathbf{g}$ is the metric tensor.

Some rather hairy manipulations allow us to derive evolution equations for the kinetic-energy

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho \mathbf{u}^2 \right) + \nabla \cdot \left[\frac{1}{2} \rho \mathbf{u}^2 \mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \mathbf{u} \right] = \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \nabla \cdot \mathbf{u} + \frac{1}{\mu_0} \mathbf{u} \cdot \left[(\mathbf{B} \cdot \nabla) \mathbf{B} \right], \tag{41}$$

the internal energy

$$\frac{\partial}{\partial t} \left(\frac{p}{\gamma - 1} \right) + \nabla \cdot \left(\mathbf{u} \frac{p}{\gamma - 1} \right) = -p \nabla \cdot \mathbf{u}, \tag{42}$$

and the magnetic field energy

$$\frac{\partial}{\partial t} \left(\frac{\mathbf{B}^2}{2\mu_0} \right) - \nabla \cdot \left(\frac{1}{\mu_0} (\mathbf{u} \cdot \mathbf{B}) \mathbf{B} - \frac{\mathbf{B}^2}{2\mu_0} \mathbf{u} \right) = -\frac{\mathbf{B}^2}{2\mu_0} \nabla \cdot \mathbf{u} - \frac{1}{\mu_0} \mathbf{u} \cdot [(\mathbf{B} \cdot \nabla) \mathbf{B}]. \tag{43}$$

These evolution equations for the various components of the energy show how energy can be exchanged between various components of the fluid and the field: internal energy and kinetic energies are exchanged via the fluid compressibility, and kinetic and field energies through a complicated set of terms involving compressibility and currents and electric field (effectively, an $\mathbf{E} \cdot \mathbf{J}$ term).

Rather neatly, if we add all the components of the energy up the right-hand sides all cancel, and we get that the *total energy*

$$\mathcal{E} \equiv \frac{1}{2} \rho \mathbf{u}^2 + \frac{p}{\gamma - 1} + \frac{\mathbf{B}^2}{2\mu_0} \tag{44}$$

evolves as

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \left[(\mathcal{E} + p^*) \mathbf{u} - \frac{1}{\mu_0} (\mathbf{u} \cdot \mathbf{B}) \mathbf{B} \right] = 0 \tag{45}$$

where the total (fluid + magnetic-field) pressure is defined as

$$p^* \equiv p + \frac{\mathbf{B}^2}{2\mu_0}. \tag{46}$$

Of course, the evolution of the total energy equation is very important. However, I should point out that equally important are the evolution of the components of the energy, as these fundamentally determine energizaton processes, and also how turbulence manifests itself in MHD. In fact, it is *extremely difficult* (perhaps impossible) to develop generic numerical schemes that account for both proper transfer between various components of the energy, and also at the same time account for shocks and other features of compressibility. As nature would have it, these are the cutting-edge regimes of interest for modern, high-energy plasma physics.

It is clear that computing the eigensystem of this formidable set of equation is highly non-trivial. However, again, it is easiest to work in the quasilinear form of the equations and use a computer algebra system. One finds that indeed the ideal MHD equation is hyperbolic.

3 Discontinous Solutions and Entropy Conditions

One of the defining features of the hyperbolic PDEs is the formation of *shocks*, *rarefaction fans* and *contact discontinuities*. We will now see how these arise by looking first at the simplest nonlinear scalar hyperbolic PDEs, the Burgers equation. However, first consider the scalar quasilinear PDE

$$\frac{\partial V}{\partial t} + a(V)\frac{\partial V}{\partial x} = 0 \tag{47}$$

for $-\infty < x < \infty$. For this system we simply have a single eigenvalue, $\lambda^1 = a(V)$. This equation also indicates that if we define *characteristics* in the (x, t) plane:

$$x = \lambda^1 t + C = a(V)t + C \tag{48}$$

where C is some constant, then the solution will remain constant along those characteristics. (Of course, different characteristics will, in general, have different values of V).

We can use this *method of characteristics* to solve the Burgers equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) = 0 \tag{49}$$

with initial conditions $u(x, 0) = u_0(x)$. Here the characteristic speed is just u, and hence the characteristics are $x = ut + C$. Now consider some point x_0 at $t = 0$. Then we must have $C = x_0$. As the characteristic velocity at that point is $u_0(x_0)$, the characteristic must be

$$x = u_0(x_0)t + x_0. \tag{50}$$

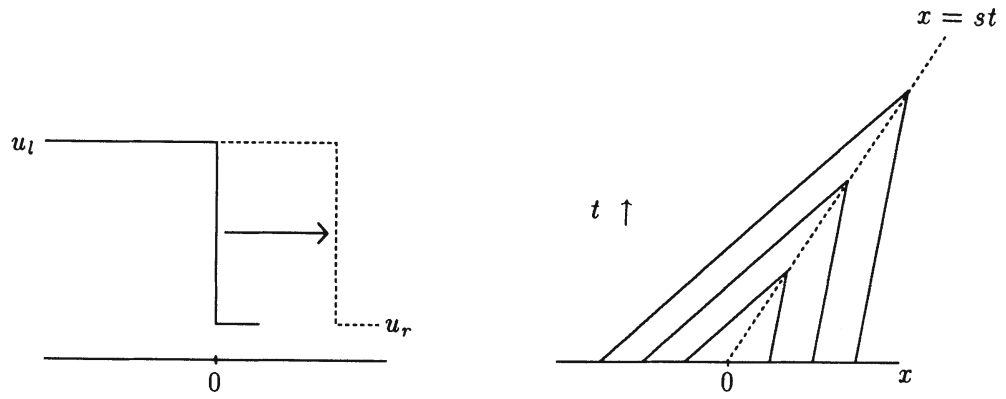

Figure 1: To avoid multi-valued solutions due to converging characteristics (right) we must allow *discontinuous* solutions to Burgers equation. These *shock solutions* (left) are only defined *weakly*, in the sense that they satisfy the weak-form of the equation and not the strong form.

As the solution remains constant along characteristics, the exact solution to Burgers equation can be easily written as

$$u(x,t) = u_0(x_0) \tag{51}$$

where x_0 is determined from Eq. (50). At this point we may think we have found a general solution to the 1D Burgers equation. However, it is not so. Consider the case shown in the Fig. 1. Here we see that characteristics *intersect* after some finite-time. As the solutions are constant along characteristics this would mean that the solution at the point of intersection are multi-valued! To avoid this undesirable situation we must introduce *shock solutions*, that is, solutions that are *discontinuous*.

Discontinuous solutions, of course, do not have well-defined deriviatives (their derivatives are *distributions*, or delta functions). Hence, we are led to consider the notion of *weak-solutions* of hyperbolic PDEs. These are "derived" by multiplying Eq. (1) by a compact, smooth function $\phi(x,t)$ (that, a function that has smooth derivatives and vanishes outside some bounded region) and integrating over all time and space:

$$\int_0^\infty \int_{-\infty}^\infty \phi(x,t)\left[\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x}\right] dx\, dt = 0. \tag{52}$$

Integrating by parts we get

$$\int_0^\infty \int_{-\infty}^\infty \left[\frac{\partial \phi}{\partial t}\mathbf{Q} + \frac{\partial \phi}{\partial x}\mathbf{F}\right] dx\, dt = -\int_{-\infty}^\infty \phi(x,0)\mathbf{Q}(x,0)dx. \tag{53}$$

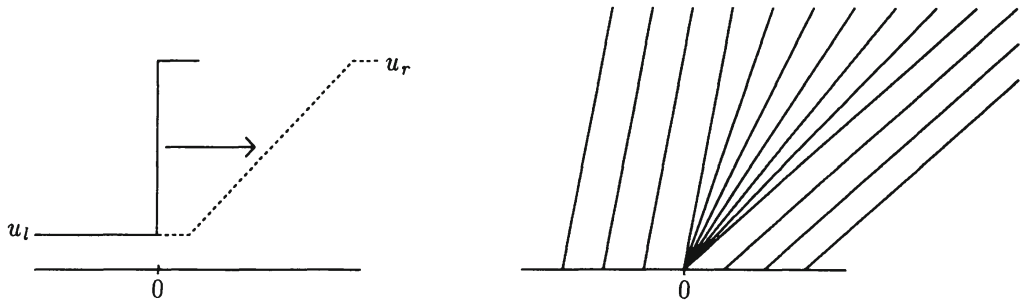

Figure 2: Diverging charactertics (right) leave behind a "void" in the (x, t) plain, allowing us to fill it with many possible ways. This *non-uniqueness* of solutions is solved by introducing *entropy*, a scalar quantity that must be *non-increasing* for all valid solutions of the PDE.

Note that constructing this weak-form from the PDE is not fully rigorous. In fact, we must *turn the process around* and forget the PDE and instead use the weak-form to *define* valid the solutions to hyperbolic PDEs.

Definition 4 (Weak-solution)**.** *A function* $\mathbf{Q}(x, t)$ *is said to be a weak-solution if it satisfies the weak-form, Eq. (53), for all compact, smooth* $\phi(x, t)$*.*

A shock solution to a hyperbolic PDE is one that satisfies the weak-form. Let s the *shock-speed* and the discontinuity exist at some point x and time t. Taking a small space-time interval around (x, t) we can show that the jump in $\mathbf{Q} = \mathbf{Q}_R - \mathbf{Q}_L$ and jump in flux $\mathbf{F} = \mathbf{F}_R - \mathbf{F}_L$ must be related by the Rankine-Hugoniot jump condition

$$s(\mathbf{Q}_R - \mathbf{Q}_L) = \mathbf{F}_R - \mathbf{F}_L. \tag{54}$$

Plugging in the specific case for Burgers equation we get

$$s(u_R - u_L) = \frac{1}{2}(u_R^2 - u_L^2) \tag{55}$$

or that

$$s = \frac{1}{2}(u_R + u_L). \tag{56}$$

Hence, a shock front in Burgers equation will move with the *mean* of the characteristic speeds computed from the left/right states.

Rather unfortunately, the weak-form of the solution does not uniquely determine all solutions to hyperbolic PDEs. This is again seen in the Burgers equations already, in the

case when the characteristics *diverge*. See Fig. 2. This shows that when the characteristics diverge they leave behind a "void" in the (x, t) plain. We could fill this in many ways yet satisfy the weak-form of the PDE. Hence, diverging characteristics lead to non-uniqueness of solutions!

One way to fix this non-uniqueness is to introduce a *entropy* function. This is a scalar function, $S = S(\mathbf{Q})$ that satisfies the PDE

$$\frac{\partial S}{\partial t} + \frac{\partial F_s}{\partial x} \leq 0 \tag{57}$$

where $F_s = F_s\big(s(\mathbf{Q})\big)$ is an *entropy flux*. That is, the entropy is a *non-increasing* function of time. Imposing that entropy be *decreasing* across any solution we construct across diverging characteristics, picks out a unique solution. This solution is called a *rarefaction* wave, and for Burgers equation is show on the left of Fig. 2.

4 A Gallery of Riemann Problems

The *Riemann problem* is a fundamental problem in the study of hyperbolic PDEs. It is an initial-value problem that looks at the decomposition of an initial discontinuity at $x = 0$. For nonlinear hyperbolic PDEs the structure of the solutions can be surprisingly complicated. For several nonlinear hyperbolic PDEs one can construct the exact solution to the Riemann problem, though this is a highly non-trivial task, and often involves complex root finding and nonlinear algebraic solves.

Formally, the Riemann problem can be stated as follows. For the hyperbolic PDE

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0 \tag{58}$$

find $\mathbf{Q}(x, t)$ for $t > 0$ on $-\infty < x < \infty$ given the initial conditions

$$\mathbf{Q}(x, 0) = \mathbf{Q}_L \quad x < 0 \tag{59}$$

$$\mathbf{Q}(x, 0) = \mathbf{Q}_R \quad x > 0. \tag{60}$$

Below, I present the solution to Riemann problems for various example equation systems presented above.

5 Exact Solution to 1D Linear Hyperbolic System

(I am dropping the bold fonts for Q and F below. In most modern applied math papers on hyperbolic PDEs vectors are no longer distinguished by bold fonts, but should be apparent by context.)

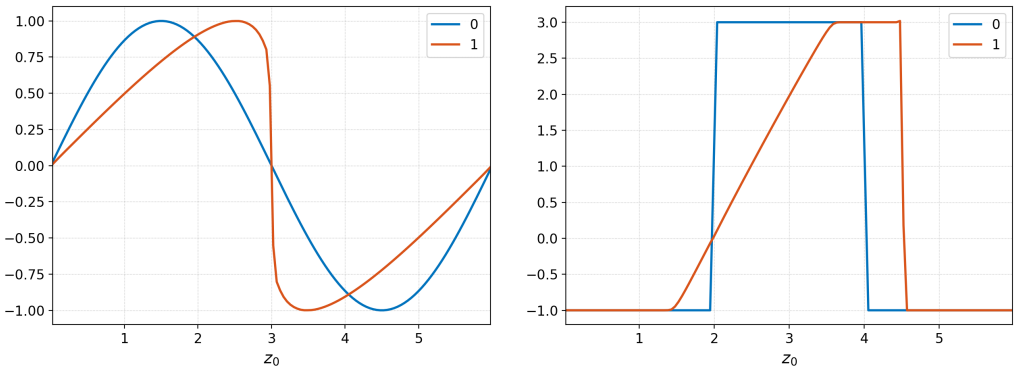

Figure 3: (Left) Steepening (orange) of an initial mode (blue) due to the nonlinear characteristic velocity of the Burgers equation. (Right) Rarefaction fan and shock in the Burgers Riemann problem.

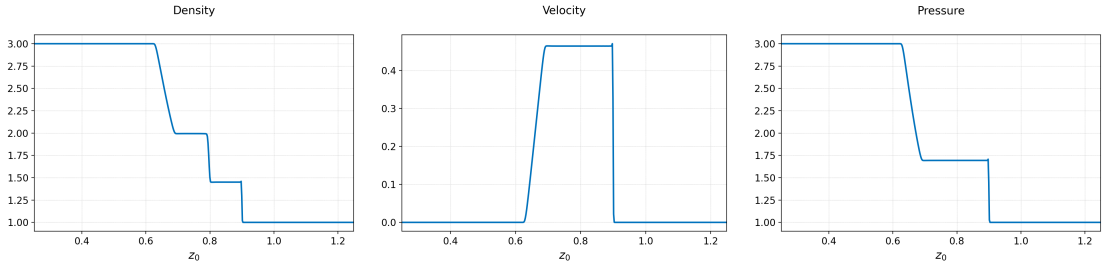

Figure 4: Density (left), velocity (middle) and pressure (right) from the solution of the *Sod-Shock* Euler Riemann problem for Euler equations. These figures show three features: a rarefaction fan, a contact discontinuity (across which only density is discontinous), and a shock (across which all quantities have a jump).

Consider the one-dimensional system of equations

$$\frac{\partial Q}{\partial t} + \frac{\partial F}{\partial x} = 0 \tag{61}$$

Where $Q(x,t)$ is a vector of m conserved quantities and $F(Q)$ is the flux function. At first, let us assume the system is linear and write the flux as

$$F = AQ \tag{62}$$

where A is a constant $m \times m$ matrix. We will assume that this system is hyperbolic, i.e. the eigenvalues of A are real and the eigenvectors are complete. Let λ_p be the eigenvalues and

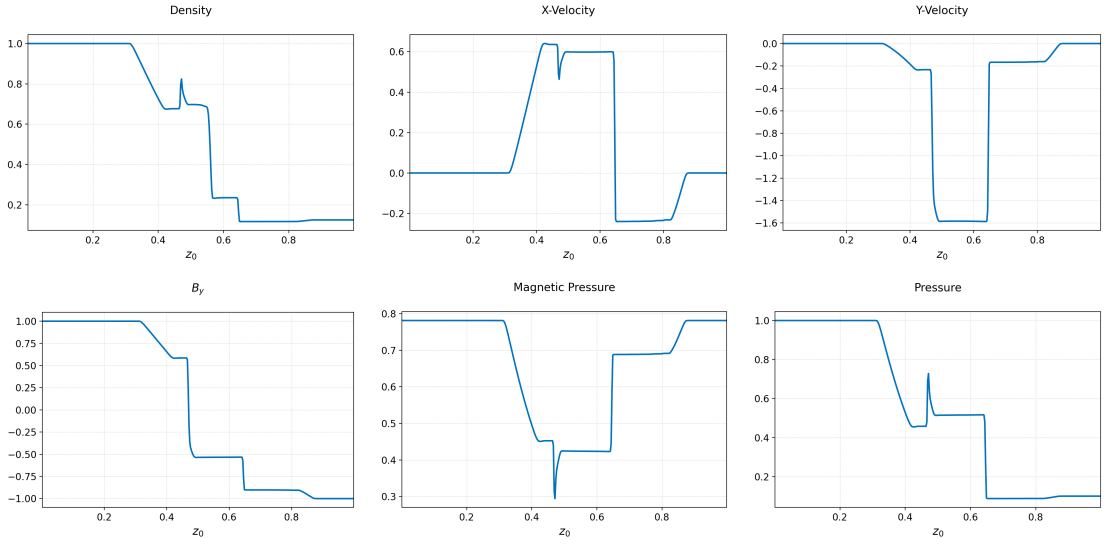

Figure 5: Density, X- and Y-velocities (top, left to right), B_y, magnetic pressure and fluid pressure (bottom, left to right) from the solution to the *Brio-Wu* MHD Riemann problem. Besides the features seen in the Euler Riemann problem, also seen is the *compound wave* (spike-like structure). Also notice that as the magnetic field is frozen into the fluid, complex structures form in the B_y component of the field also.

r^p and l^p, $p = 1, \ldots, m$ be the right and left eigenvectors respectively. We will represent right eigenvectors as column vectors, and left eigenvectors as row vectors.

To solve Eq. (61) we first convert it to a system of uncoupled advection equations by multiplying it from the left by l^p. This gives

$$\frac{\partial w^p}{\partial t} + \lambda_p \frac{\partial w^p}{\partial x} = 0 \tag{63}$$

where we have defined the *Riemann variables* $w^p = l^p Q$. Note that given w^p we can recompute Q from

$$Q = \sum_p w^p r^p. \tag{64}$$

A useful identity is

$$\sum_p w^p \lambda_p r^p = \sum_p w^p A r^p = A \sum_p w^p r^p = AQ = F(Q). \tag{65}$$

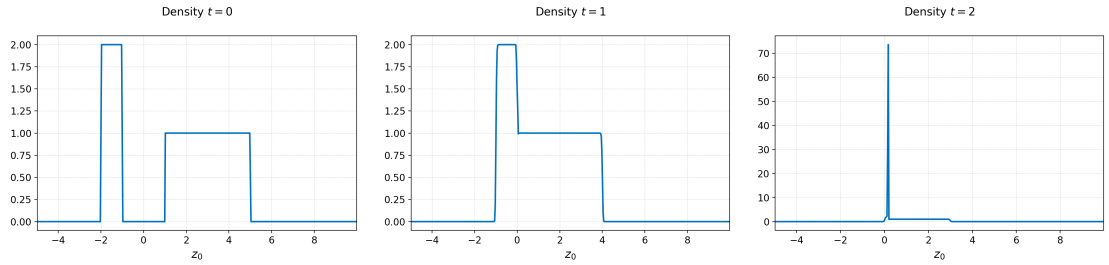

Figure 6: Riemann problem solution to the *cold* fluid equations. For this system there is only a single eigenvalue and the eigensystem is not diagonalizable. Hence, this system is *not* hyperbolic. This leads to the formation of δ-function shocks. The left plot shows two "clouds" moving towards each other, the middle plot shows the clouds starting to merge, and the right plot shows the δ-shock.

Now consider a domain $-\infty < x < \infty$ and the initial conditions $w_0^p(x) = l^p Q_0(x)$. Each advection equation for the Riemann variables can be solved exactly as

$$w^p(x,t) = w_0^p(x - \lambda_p t) \tag{66}$$

From this, the exact solution to the linear system can be obtained as

$$Q(x,t) = \sum_p w_0^p(x - \lambda_p t) r^p = \sum_p l^p Q_0(x - \lambda_p t) r^p. \tag{67}$$

6 Basic formulation of finite-volume schemes

The finite-volume (FV) scheme for the system Eq. (61) is a method to update the *cell-average* of the solution in time. Consider a cell $I_j \equiv [x_{j-1/2}, x_{j+1/2}]$ with uniform cell-spacing $\Delta x \equiv x_{j+1/2} - x_{j-1/2}$. Integrate Eq. (61) to get

$$\frac{\partial Q_j}{\partial t} + \frac{F_{j+1/2} - F_{j-1/2}}{\Delta x} = 0 \tag{68}$$

where Q_j are cell-average quantities

$$Q_j(t) = \frac{1}{\Delta x} \int_{I_j} Q(x,t)\, dx \tag{69}$$

and $F_{j\pm1/2} = F(Q_{j\pm1/2})$ are the fluxes at the *cell-edges* $x_{j\pm1/2}$. Notice that in effect finite-volume scheme uses the *mean flux gradient* in the cell

$$\frac{1}{\Delta x} \int_{I_j} \frac{\partial F}{\partial x}\, dx = \frac{F_{j+1/2} - F_{j-1/2}}{\Delta x}. \tag{70}$$

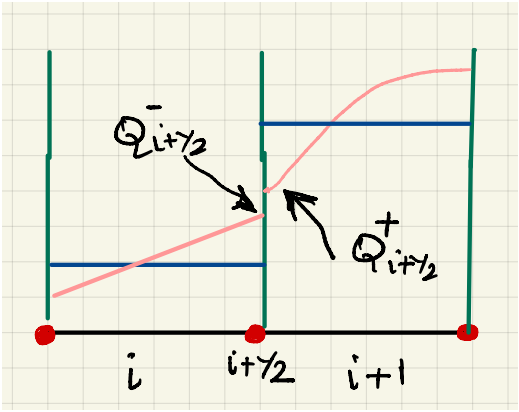

Figure 7: Recovery is used to compute left/right values $Q^{\pm}_{i+1/2}$ from a set of cell-averages and are then fed into a numerical flux function to update the solution.

to update the *cell average* in time. This is a subtle point and often the cause of misinterpreting the accuracy or order of a FV scheme.

At this point the discrete expression is exact, but only formally: given the cell-averages we can't uniquely determine the edge values $Q_{j\pm1/2}$ to insert into the edge fluxes. The FV scheme is an approximation in which these edge values are *recovered* (approximately) from the cell-averages and then used in a *numerical-flux* to update the cell-average approximation. We can write this as

$$\frac{\partial Q_j}{\partial t} + \frac{G(Q^+_{j+1/2}, Q^-_{j+1/2}) - G(Q^+_{j-1/2}, Q^-_{j-1/2})}{\Delta x} = 0 \tag{71}$$

where $G(Q_L, Q_R)$ the *numerical-flux* function and $Q^+_{j\pm1/2}$ and $Q^-_{j\pm1/2}$ are recovered values just the right and left of the interface $j \pm 1/2$ respectively (see Fig. 7). For consistency with the exact form Eq. (68) we must ensure that the numeric flux is Lipschitz continuous[1] and *consistent*: when $Q_L = Q_R$ then $G(Q_L, Q_R)$ must reduce to the physical flux function, i.e.

$$\lim_{Q_{L,R}\to Q} G(Q_L, Q_R) = F(Q). \tag{72}$$

Hence, to completely specify a finite-volume scheme we must design algorithms for each of the following three steps:

- **Step 1**: A recovery scheme (possibly with limiters) to compute the left/right interface values $Q_{L,R}$ at each interface using a set of cell-average values around that interface,

[1]This is somewhat of a technical restriction which ensures that the derivative of the numerical-flux with each of its independent variables is *bounded*.

- **Step 2**: A numerical flux function that takes the left/right values and returns a consistent approximation to the physical flux, and
- **Step 3**: A time-stepping scheme to advance the solution in time and compute the cell-averages at the next time-step.

7 First-order upwind scheme for one-dimension linear hyperbolic systems

Let us first consider a linear system in which we use the cell-averages directly as the left/right edge values, skipping the recovery step completely. This *first-order* scheme for linear equations will form the basis for higher-order schemes for nonlinear systems. We will also use a simple forward-Euler time-stepper to simplify the third-step.

Instead of directly discretizing the coupled system of equations (in which upwind direction may not not be clear), we can instead solve the linear advection equations for the Riemann variables using an upwind method and the convert the solution for w^p to Q using Eq. (64). We can write a first-order upwind method for the Riemann variables as

$$w_j^{p,n+1} = w_j^{p,n} - \frac{\Delta t}{\Delta x}\left(G^p(w_{j+1}^{p,n}, w_j^{p,n}) - G^p(w_j^{p,n}, w_{j-1}^{p,n})\right) \tag{73}$$

where the numerical flux function for the Riemann variables, $G^p(w_R, w_L)$, is defined as

$$G^p(w_R^p, w_L^p) = \frac{\lambda_p}{2}(w_R^p + w_L^p) - \frac{|\lambda_p|}{2}(w_R^p - w_L^p). \tag{74}$$

Note that this choice of flux ensures that if $\lambda_p > 0$ then $G^p(w_R, w_L) = \lambda_p w_L$ and if $\lambda_p < 0$ then $G^p(w_R, w_L) = \lambda_p w_R$, ensuring proper upwinding of the values at cell interfaces.

To convert this to a scheme for Q instead, multiply by r^p and sum over p to get

$$Q_j^{n+1} = Q_j^n - \frac{\Delta t}{\Delta x}\left(G(Q_{j+1}^n, Q_j^n) - G(Q_j^n, Q_{j-1}^n)\right). \tag{75}$$

The numerical flux $G(Q_R, Q_L)$ is computed from

$$G(Q_R, Q_L) = \sum_p r^p G^p(w_R^p, w_L^p) = \frac{1}{2}\sum_p \lambda_p r^p (w_R^p + w_L^p) - \frac{1}{2}\sum_p |\lambda_p| r^p (w_R^p - w_L^p). \tag{76}$$

We can write the first term, using identity Eq. (65) as $(F(Q_R) + F(Q_L))/2$. To rewrite the second term introduce

$$\lambda_p^+ = \max(\lambda_p, 0) \tag{77}$$

$$\lambda_p^- = \min(\lambda_p, 0). \tag{78}$$

Note that in terms of these we can write $\lambda_p = \lambda_p^+ + \lambda_p^-$ and $|\lambda_p| = \lambda_p^+ - \lambda_p^-$. Using the latter identity the numerical flux can be written as

$$G(Q_R, Q_L) = \frac{1}{2}\big(F(Q_R) + F(Q_L)\big) - \frac{1}{2}(A^+\Delta Q_{R,L} - A^-\Delta Q_{R,L}) \tag{79}$$

where the *fluctuations* $A^\pm \Delta Q$ are defined as

$$A^\pm \Delta Q_{R,L} \equiv \sum_p r^p \lambda_p^\pm (w_R^p - w_L^p) = \sum_p r^p \lambda_p^\pm l^p (Q_R - Q_L). \tag{80}$$

The fluctuations satisfy the *flux-difference* or *flux-jump* identity

$$A^+\Delta Q_{R,L} + A^-\Delta Q_{R,L} = \sum_p r^p \underbrace{(\lambda_p^+ + \lambda_p^-)}_{\lambda^p}(w_R^p - w_L^p) = F(Q_R) - F(Q_L). \tag{81}$$

Using this identity, the numerical flux can also be written as

$$G(Q_R, Q_L) = F(Q_L) + A^-\Delta Q_{R,L} = F(Q_R) - A^+\Delta Q_{R,L}. \tag{82}$$

Using this final identity, the complete first-order update for the linear system can be written entirely in terms of fluctuations as

$$Q_j^{n+1} = Q_j^n - \frac{\Delta t}{\Delta x}\left(A^-\Delta Q_{j+1/2} + A^+\Delta Q_{j-1/2}\right). \tag{83}$$

Note that instead, dividing by Δt and taking limits as $\Delta t \to 0$, this can be written in the *semi-discrete* or method-of-lines form

$$\frac{\partial Q_j}{\partial t} = -\frac{1}{\Delta x}\left(A^-\Delta Q_{j+1/2} + A^+\Delta Q_{j-1/2}\right). \tag{84}$$

This system of ODEs can be solved using, for example, a SSP-RK stepper we studied before. Further if the left/right edge values $Q_{L,R}$ used in the fluctuations are recovered (and not merely the cell-average values) then we will get a spatially high-order scheme.