AST242 LECTURE NOTES PART 4

Contents

1. Numerical Approximations to Partial Differential Equations 1
1.1. Elliptic, hyperbolic and parabolic partial differential equations 1
1.2. Boundary conditions 2
1.3. General classification for linear systems 2
1.4. Hyperbolic equations in conservational law form 4
1.5. Lagrangian vs Eulerian integration approaches 5
1.6. Finite differences 6
1.7. Truncation Error, Accuracy and Consistency 7
1.8. Big Oh notation 8
1.9. Finding coefficients for a scheme – an example 8
1.10. Stability (von Neumann analysis) 9
1.11. Upwind differencing 12
1.12. Upwind vs Downwind 13
1.13. Upwind method for linear systems with positive and negative characteristics 14
1.14. The Modified equation – Numerically generated diffusion and dispersion 15
1.15. General Issues for Finite Differencing Schemes 16
1.16. Some Simple Finite Differencing Schemes 17
1.17. Grids in different coordinate systems 18
1.18. Boundary Conditions 20
2. Conservative methods and Riemann Solvers 20
2.1. Conservation Laws and shock speeds 20
2.2. Conservative schemes 21
2.3. The Riemann Problem 23
2.4. Riemann Problem, the example of linearized gas dynamics 25
2.5. Riemann Problem and the Hugoniot locus 27
2.6. Shocks and Rarefraction Waves in Burger's equation 28
2.7. Riemann Problem and Hugoniot locus for a Non-Linear System 30
2.8. Godunov's Method 32
2.9. Roe's approximate Riemann solver 33
3. Acknowledgments 34

1. Numerical Approximations to Partial Differential Equations

1.1. **Elliptic, hyperbolic and parabolic partial differential equations.** An example of an **elliptic** differential equation is the Poisson equation for the gravitational potential $\Phi(x, y, z)$

$$\nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = 4\pi G\rho(\mathbf{x}) \tag{1}$$

Elliptic equations are often associated with boundary value problems in which at every point inside a domain of interest the solution depends on the data provided at the boundary of the domain.

Hyperbolic partial differential equations are often associated with specifying initial conditions and finding solutions at later times. A simple example is the wave equation

$$\frac{\partial^2\rho}{\partial t^2} - c^2\frac{\partial^2\rho}{\partial x^2} = 0 \tag{2}$$

or the linear advection equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0 \tag{3}$$

Parabolic partial different equations require more than just an initial condition to be specified for a solution. For example the conditions on the boundary could be specified at all times as well as the initial conditions. An example is the one-dimensional diffusion equation

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial x}\left(K\frac{\partial\rho}{\partial x}\right) \tag{4}$$

with diffusion coefficient $K > 0$.

Fluid equations can be a mixture of hyperbolic and elliptical presenting an additional difficulty for solving them robustly.

For example the advection diffusion equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = K\frac{\partial^2 u}{\partial x^2} \tag{5}$$

is considered parabolic even though the left hand side is advective.

1.2. **Boundary conditions.** Hyperbolic equations such as the wave equation have have two derivatives in time and two in space. An initial condition (both function and its derivative) is required and boundary conditions on both side of the domain. Contrast this with Laplace's equation for the gravitational potential ($\nabla \cdot \Phi = 0$) which is an elliptic partial differential equation. In this case conditions on the entire boundary are needed and specify the solution everywhere. Either the function or its

derivative must be specified on each of the boundaries and changing the conditions at one point will change the solution everywhere. What accounts for the difference in these two? Consider the elliptic operator

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{6}$$

A solution that oscillates in x like $e^{\pm ikx}$ will typically behave exponential in y as $e^{\pm ky}$. A solution specified at one boundary (along x) may grow exponentially in the y direction. Solutions are ill specified unless they are constrained on all boundaries. The hyperbolic operator in comparison

$$\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2} \tag{7}$$

has oscillatory solutions in both x and t and so solutions remain bounded.

1.3. **General classification for linear systems.** Consider two general linear equations

$$a_1 \frac{\partial u}{\partial x} + b_1 \frac{\partial u}{\partial y} + c_1 \frac{\partial v}{\partial x} + d_1 \frac{\partial v}{\partial y} = f_1 \tag{8}$$

$$a_2 \frac{\partial u}{\partial x} + b_2 \frac{\partial u}{\partial y} + c_2 \frac{\partial v}{\partial x} + d_2 \frac{\partial v}{\partial y} = f_2 \tag{9}$$

with real coefficients and real functions f_1, f_2. We can write this as

$$\begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \frac{\partial \mathbf{W}}{\partial x} + \begin{pmatrix} b_1 & d_1 \\ b_2 & d_2 \end{pmatrix} \frac{\partial \mathbf{W}}{\partial y} = \mathbf{A} \frac{\partial \mathbf{W}}{\partial x} + \mathbf{B} \frac{\partial \mathbf{W}}{\partial y} = \mathbf{f} \tag{10}$$

with

$$\mathbf{W} = \begin{pmatrix} u \\ v \end{pmatrix}, \qquad \mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \tag{11}$$

and $\mathbf{A}, \mathbf{B}$ are the matrices

$$\mathbf{A} = \begin{pmatrix} a_1 & c_1 \\ a_2 & c_2 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} b_1 & d_1 \\ b_2 & d_2 \end{pmatrix} \tag{12}$$

If we can invert $\mathbf{A}$, we can write

$$\frac{\partial \mathbf{W}}{\partial x} + \mathbf{A^{-1}B} \frac{\partial \mathbf{W}}{\partial y} = \mathbf{A}^{-1}\mathbf{f} \tag{13}$$

If the matrix $\mathbf{A^{-1}B}$ has two real eigenvalues then the equations are hyperbolic. If one eigenvalue is zero then equations are parabolic and if the eigenvalues are complex the equations are elliptic. We can think of elliptic equations as having no characteristics, those of hyperbolic equations as having two characteristics, and those that are parabolic with one real eigenvalue as having one characteristic.

The wave equation, Laplace's equation, and the advection equation are easily put into the above form. The heat equation requires addition of a term dependent on u. If the matrix $\mathbf{A}$ is not invertible one can switch the roles of x and t and can invert the matrix $\mathbf{B}$ instead.

For a second order equation put in the form

$$au_{tt} + bu_{xt} + cu_{xx} + du_x + eu_t + f = 0 \tag{14}$$

The equation is hyperbolic if $ac - b^2 < 0$, parabolic if $ac - b^2 = 0$ and elliptic if $ac - b^2 > 0$.

1.3.1. *Examples.* The wave equation can be written as

$$\begin{pmatrix} u \\ v \end{pmatrix}_{,t} + \begin{pmatrix} 0 & -c \\ -c & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_{,x} = 0 \tag{15}$$

The eigenvalues are $\pm c$.

Laplaces's equation can be written as

$$\begin{pmatrix} u \\ v \end{pmatrix}_{,t} + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_{,x} = 0 \tag{16}$$

There are no real eigenvalues.

The diffusion equation ($u_t = Ku_{xx}$) can be written as

$$\begin{aligned} u_t - Kv_x &= 0 \\ u_x &= v \end{aligned} \tag{17}$$

or as

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_{,t} + \begin{pmatrix} 0 & -K \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_{,x} = \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{18}$$

The matrix on the left is not invertible. There are not two non-zero eigenvalues. One eigenvalue is real. The system is considered parabolic.

A note on the Shroedinger equation. The time independent Shroedinger equation contains a Laplacian operator. Finding eigenfunctions for time independent problems is an elliptic problem. The time dependent Shroedinger equation has $ih\partial_t$ in it. This is complex and gives wavelike solutions. Above we have discussed real coefficients. With complex coefficients it is possible to compute how wave functions evolve with time using a hyperbolic scheme.

1.4. **Hyperbolic equations in conservational law form.** Often our fluid equations can be put in conservation law form. Consider the conservation law

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0 \tag{19}$$

where ∂_t is short for $\frac{\partial}{\partial t}$. Here $\mathbf{F}$ is a vector of fluxes for conserved quantities. For example, the isothermal gas in 1d

$$\mathbf{U} = (\rho, j) \qquad \mathbf{F} = (j, c^2\rho + j^2/\rho) \tag{20}$$

with $j = \rho u$ and c the sound speed so our differential equation looks like

$$\partial_t \begin{pmatrix} \rho \\ j \end{pmatrix} + \partial_x \begin{pmatrix} j \\ c^2\rho + j^2/\rho \end{pmatrix} = 0 \tag{21}$$

For hydrodynamics in one dimension $j = \rho u$ is the momentum flux. The pressure is a force per unit are or a change in momentum per unit area, so we can think of pressure as a momentum flux. And in more than 1 dimension it contributes to the diagonal terms in the stress tensor. For the isothermal gas $\rho c^2 = p$ and the momentum flux is $\rho c^2 + j^2/\rho = p + \rho u^2$. Here ρu^2 we can recognize as ram pressure and is also a momentum flux.

Using the Jacobian $\mathbf{J}(\mathbf{U}) = \frac{\partial \mathbf{F}(\mathbf{U})}{\partial \mathbf{U}}$, we can write the differential equation in quasi-linear (but not conservation law) form

$$\partial_t \mathbf{U} + \mathbf{J}(\mathbf{U})\partial_x \mathbf{U} = 0 \tag{22}$$

If the Jacobian has real eigenvalues and a complete set of linearly independent eigenvectors. Then the system is hyperbolic.

For example our isothermal 1d gas has

$$\mathbf{J}(\mathbf{U}) = \begin{pmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{pmatrix} \tag{23}$$

This Jacobian has two eigenvalues $u \pm c$ and the Riemann invariants are eigenvectors.

Consider small perturbations with

$$\rho(x,t) = \rho_0 + \rho_1(x,t)$$
$$j(x,t) = j_1(x,t)$$

where $\rho_1 \ll \rho_0$ and j_1 are small. Our state vectors

$$\mathbf{U}_0 = (\rho_0, 0)$$
$$\mathbf{U}_1 = (\rho_1, j_1)$$

To first order in the small quantities our partial differential equation (equation 21) looks like

$$\partial_t \mathbf{U}_1 + \mathbf{J}(\mathbf{U}_0)\partial_x \mathbf{U}_1 = 0 \tag{24}$$

Inserting $\mathbf{U}_0$ into our Jacobian matrix we find

$$\mathbf{J}(\mathbf{U}_0) = \begin{pmatrix} 0 & 1 \\ c^2 & 0 \end{pmatrix} \tag{25}$$

Because our Jacobian matrix is now full of constants, the problem is linear.

$$\rho_{1t} + j_{1x} = 0$$
$$j_{1t} + c^2\rho_{1x} = 0$$

Taking the time derivative of the first equation and x derivative of the second equation

$$\rho_{1tt} + j_{1xt} = 0$$
$$j_{1xt} + c^2\rho_{1xx} = 0$$

We find

$$\rho_{1tt} - c^2\rho_{1xx} = 0$$

which is the wave equation. So small perturbations travel at the speed of sound. Here we used an isothermal gas but we would have arrived at the same result using an isobaric equation of state $p(\rho)$ with $c^2 = dp/d\rho$.

We can write $\mathbf{J} = \mathbf{R\Lambda L}$ where Λ is a diagonal matrix containing the eigenvalues. The transformations R and L are composed of left and right eigenvectors. We can define characteristic variables as $v = R^{-1}U$ or $v = LU$ so that v is in terms of the eigenvectors.

1.5. **Lagrangian vs Eulerian integration approaches.** Lagrangians methods are particle based. An example is integrating N bodies interaction with gravitational forces. Another example is modeling traffic flow by updating positions and velocities of cars. This is in the viewpoint of individual car drivers. Eulerian techniques instead work in a fixed background coordinate system. Traffic flow can be modeled by computing the mean number density and mean car velocity as a function of position on the road, from the viewpoint of an outside viewer rather car drivers.

The two approaches give different types of codes, for example SPH vs Grid based simulations for hydrodynamics. In the limit of large numbers of particles and ultra-fine grids the two should give consistent results. The two approaches tend to be used in different settings. SPH is often used when phenomena over a large dynamic range is simulated and high accuracy is not as important. Grid based methods are more popular when the result must be accurate and when shocks must be resolved.

1.6. **Finite differences.** There are a variety of approximations we could use for a derivative $\frac{\partial f}{\partial x}$. We could use

$$\frac{f(x+h)-f(x-h)}{2h} \quad \text{or} \quad \frac{f(x+h)-f(x)}{h} \quad \text{or} \quad \frac{f(x)-f(x-h)}{h} \tag{26}$$

where h represents our grid spacing. Each of these is accurate to at least first order in h (the one on the left is accurate to second order). Let us consider higher order approximation using a Taylor series.

$$f(x+jh) = f(x) + f'(x)(jh) + f''(x)\frac{(jh)^2}{2!} + \cdots + f^n(x)\frac{(jh)^n}{n!} + O(h^{n+1}) \tag{27}$$

Using the Taylor expansion

$$\begin{aligned} f(x+h) &= f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + O(h^4) \\ f(x-h) &= f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + O(h^4) \end{aligned} \tag{28}$$

We see that

$$\frac{f(x+h)-f(x)}{h} \sim f'(x) + O(h) \tag{29}$$

is a first order approximation. Whereas,

$$\frac{f(x+h)-f(x-h)}{2h} \sim f'(x) + O(h^2) \tag{30}$$

is second order. Adding the expressions for $f(x+h)$ and $f(x-h)$ and manipulating we find a second order expression for the second derivative

$$f''(x) = h^{-2}\left[f(x+h) + f(x-h) - 2f(x)\right] + O(h^4) \tag{31}$$

We can now use the above to construct an example of a second order in space and first order in time scheme for advancing the diffusion equation

$$\frac{\partial u}{\partial t} = K\frac{\partial^2 u}{\partial x^2} \tag{32}$$

We denote u_j^n as the value of $u(x,t)$ at $x = j\Delta x$ or at the j-th grid point and the n-th timestep or time $t = n\Delta t$.

Consider the following scheme approximating the diffusion equation

$$u_j^{n+1} = u_j^n + K\frac{\Delta t}{(\Delta x)^2}(u_{j+1}^n + u_{j-1}^n - 2u_j^n) \tag{33}$$

accurate to first order in Δt and second order in Δx. This is an *explicit* scheme as the term on the left hand side that is the next timestep only depends on terms (on the right hand side) that are at the current time step.

1.7. **Truncation Error, Accuracy and Consistency.** Given a scheme and a differential equation how do we define its accuracy? We can estimate a truncation error between that given by the scheme and that by the equation and use Taylor series to measure this difference. Remember our differential equation

$$\frac{\partial u}{\partial t} = K\frac{\partial^2 u}{\partial x^2} \tag{34}$$

Let us write the above scheme (equation 33) as

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - K\frac{(u_{j+1}^n + u_{j-1}^n - 2u_j^n)}{(\Delta x)^2} = 0 \tag{35}$$

The Taylor expansion of each term is

$$u_j^{n+1} = u_j^n + \Delta t\frac{\partial u}{\partial t} + \frac{(\Delta t)^2}{2}\frac{\partial^2 u}{\partial t^2} + O((\Delta t)^3) \tag{36}$$

$$u_{j+1}^n = u_j^n + \Delta x\frac{\partial u}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \frac{(\Delta x)^3}{6}\frac{\partial^3 u}{\partial x^3} + \frac{(\Delta x)^4}{4!}\frac{\partial^4 u}{\partial x^4} + O((\Delta x)^5) \tag{37}$$

$$u_{j-1}^n = u_j^n - \Delta x\frac{\partial u}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 u}{\partial x^2} - \frac{(\Delta x)^3}{6}\frac{\partial^3 u}{\partial x^3} + \frac{(\Delta x)^4}{4!}\frac{\partial^4 u}{\partial x^4} + O((\Delta x)^5) \tag{38}$$

We insert these into the above finite difference equation finding

$$\epsilon_j^n = \frac{\partial u}{\partial t} - K\frac{\partial^2 u}{\partial x^2} + \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2} - K\frac{(\Delta x)^2}{12}\frac{\partial^4 u}{\partial x^4} + O((\Delta t)^2, (\Delta x)^4) \tag{39}$$

Using our differential equation we estimate the truncation error as

$$\epsilon_j^n = O(\Delta t, \Delta x^2) \tag{40}$$

The scheme is first order **accurate** in time and second-order accurate in space.

A scheme is said to be **consistent** if the truncation error goes to zero as $\Delta t \to 0$ and $\Delta x \to 0$ independently. If the truncation error is of the form $\epsilon_i^n = O((\Delta t)^p, (\Delta x)^q)$ we say the scheme is order p in t and order q in x. What is meant by independently? Both Δt and Δx must go to zero for the error to go to zero. However it doesn't matter what order we take the limit or if Δx depends on Δt.

1.8. **Big Oh notation.** For two functions $f(h)$ and $g(h)$ we say that

$$f(h) \quad \text{is} \quad O(g(h)) \quad \text{as} \quad h \to 0 \tag{41}$$

If there is a constant C such that

$$\left|\frac{f(h)}{g(h)}\right| < C \qquad \text{for all } h \text{ sufficiently small} \tag{42}$$

1.9. **Finding coefficients for a scheme – an example.** Supposing you want a second order approximation to $\frac{\partial u}{\partial x}$ that involves u_j, u_{j-1} and u_{j-2}. This situation you might find at a boundary. We expand u_j, u_{j-1}, u_{j-2} using a Taylor series about the j grid point. We add the sums together, but each term is weighted by an unknown coefficient. We group the terms by the derivatives evaluated at the j grid point. We set the sum equal to the desired derivative, $\frac{\partial u}{\partial x}$. This gives us 3 equations in 3 coefficients, one for each derivative. We solve for the coefficients.

We describe our scheme as

$$au_j + bu_{j-1} + cu_{j-2} \tag{43}$$

Here a, b, c are the unknown coefficients. Using a Taylor series we expand out the difference between our scheme and our desired derivative

$$\begin{aligned} \epsilon &= (au_j + bu_{j-1} + cu_{j-2}) - \frac{\partial u}{\partial x} \\ &= (a+b+c)u_j - (b+2c)h\frac{\partial u}{\partial x} + (b+4c)\frac{h^2}{2}\frac{\partial^2 u}{\partial x^2} + O(h^3) - \frac{\partial u}{\partial x} \end{aligned} \tag{44}$$

where $h = \Delta x$. We want ϵ to be small so that our scheme is accurate. To make ϵ small we set each group of terms to zero. This reduces to the following equations for our coefficients

$$\begin{aligned} a+b+c &= 0 \\ (b+2c)h+1 &= 0 \\ b+4c &= 0 \end{aligned} \tag{45}$$

We solve this set of equations. The solution is $a = \frac{3}{2h}, b = -\frac{2}{h}, c = \frac{1}{2h}$ and so our scheme is

$$\frac{\partial u}{\partial x} \approx \frac{1}{h}\left(\frac{3}{2}u_j - 2u_{j-1} + \frac{1}{2}u_{j-2}\right). \tag{46}$$

The neglected terms in the expansion are proportional to $h^3\frac{\partial^3 u}{\partial x^3}$. Because these terms are proportional to the coefficients a, b, c and these are proportional to h^{-1} the remaining terms would be proportional to h^2. The scheme is therefor a second order in x approximation to $\frac{\partial u}{\partial x}$. We expect 2 grid points are needed for a first order approximation to the derivative, whereas three are needed for a second derivative. Three grid points are needed for a second order approximation to the first derivative.

1.10. **Stability (von Neumann analysis).** Above we considered the following scheme approximating the diffusion equation

$$u_j^{n+1} = u_j^n + K\frac{\Delta t}{(\Delta x)^2}(u_{j+1}^n + u_{j-1}^n - 2u_j^n) \tag{47}$$

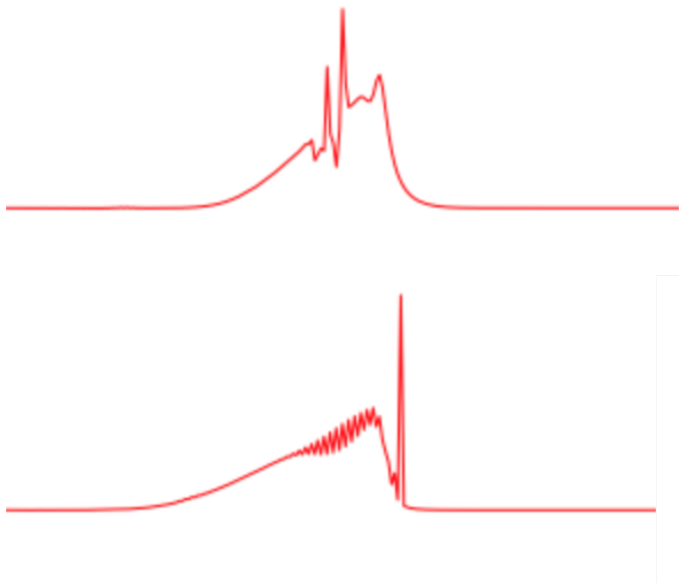

FIGURE 1. Some examples of unstable hydrodynamics integrations. Here I increased the timestep to just above that given by the CFL condition.

Suppose we try and *improve* the above scheme so that it is second order in *time* with the following

$$u_j^{n+1} = u_j^{n-1} + 2K\frac{\Delta t}{(\Delta x)^2}(u_{j+1}^n + u_{j-1}^n - 2u_j^n) \tag{48}$$

What we are going to find is that even though the the scheme in equation 48 looks better (is higher order in time) it is actually unconditionally unstable (for any choice of time step) and so completely unusable. In contrast the scheme in equation 47, even though it is lower order in time, is stable, as long as we choose the time step appropriately.

Let's assume our function is restricted to a domain of size X with the number of grid points an integer, L,

$$L = \frac{X}{\Delta x} \tag{49}$$

and has periodic boundary conditions. Our scheme is an L dimensional map, $\mathbf{u}^n \to \mathbf{u}^{n+1}$. Using a discrete Fourier transform for our vector $\mathbf{u}$ we can make a vector of Fourier coefficients $\hat{\mathbf{u}}$. Our scheme also gives a map in Fourier space $\hat{\mathbf{u}}^n \to \hat{\mathbf{u}}^{n+1}$. We can look at the map for each wave vector k. If the map is unstable for one of the

Fourier coefficients, then that Fourier coefficient will grow and we will see a wave-like structure grow during the integration (see Figure 1).

Represent u with a Fourier series in space

$$u_j = \frac{1}{L}\sum_{k=1}^{L} \hat{u}_k e^{-2\pi ikj/L} \tag{50}$$

and inverse transform

$$\hat{u}_k = \sum_{j=1}^{L} u_j e^{2\pi ikj/L} \tag{51}$$

Fourier amplitudes, $\hat{u}_k$, evolve with time. At time $t = n\Delta t$ and $x = x_j = j\Delta x$

$$u_j^n = \frac{1}{L}\sum_{k=1}^{L} \hat{u}_k^n e^{-2\pi ikj\Delta x/X} = \frac{1}{L}\sum_{k=1}^{L} \hat{u}_k^n e^{-ij\phi_k} \tag{52}$$

with phase angle

$$\phi_k \equiv \frac{2\pi k\Delta x}{X} \tag{53}$$

At the same time but at position $j+1$ or $x = x_{j+1} = (j+1)\Delta x$

$$u_{j+1}^n = \frac{1}{L}\sum_{k=1}^{L} \hat{u}_k^n e^{-i(j+1)\phi_k} = \frac{1}{L}\sum_{k=1}^{L} \hat{u}_k^n e^{-ij\phi_k} \times e^{-i\phi_k}, \tag{54}$$

likewise,

$$u_{j-1}^n = \frac{1}{L}\sum_{k=1}^{L} \hat{u}_k^n e^{-ij\phi_k} \times e^{i\phi_k}. \tag{55}$$

Now insert our Fourier expansion into the scheme of equation (48)

$$\begin{aligned} \frac{1}{L}\sum_{j=1}^{L} \hat{u}_k^{n+1} e^{-ij\phi_k} &= \frac{1}{L}\sum_{j=1}^{L} \hat{u}_j^{n-1} e^{-ij\phi_k} \\ &+ \frac{2K\Delta t}{(\Delta x)^2 L}\sum_{j=1}^{L} \hat{u}_j^n \left[e^{-ij\phi_k} \times e^{-i\phi_k} + e^{-ij\phi_k} \times e^{i\phi_k} - 2e^{-ij\phi_k}\right] \end{aligned} \tag{56}$$

Dropping the sums and the factor $e^{-ij\phi_k}$,

$$\begin{aligned} \hat{u}_k^{n+1} &= \hat{u}_k^{n-1} + d\left(\hat{u}_k^n e^{-i\phi_k} + \hat{u}_k^n e^{-i\phi_k} e^{-i\phi_k} - 2\hat{u}_k^n\right) \\ &= \hat{u}_k^{n-1} + d\hat{u}_k^n\left(e^{-i\phi_k} + e^{-i\phi_k} - 2\right) \\ &= \hat{u}_k^{n-1} + 2d\hat{u}_k^n(\cos\phi_k - 1) \end{aligned} \tag{57}$$

with

$$d \equiv 2K\frac{\Delta t}{(\Delta x)^2} \tag{58}$$

We can write equation 57 as a matrix

$$\begin{bmatrix} \hat{u}^{n+1} \\ \hat{u}^{n} \end{bmatrix} = \mathbf{G} \begin{bmatrix} \hat{u}^{n} \\ \hat{u}^{n-1} \end{bmatrix} \tag{59}$$

with amplification matrix $\mathbf{G}$,

$$\mathbf{G} = \begin{bmatrix} 2d(\cos\phi_k - 1) & 1 \\ 1 & 0 \end{bmatrix} \tag{60}$$

Each timestep we reapply the matrix $\mathbf{G}$. If the solution grows exponentially for some value of ϕ then we say the scheme is unstable.

Our phase angle ϕ_k depended on a wavevector k. If there is a value of ϕ_k that leads to exponential growth then any perturbation that has wavevector k (or wavelength $2\pi/k$ will grow exponentially. We would rather not have numerically solutions that falsely grew small initial perturbations from something small to something large very quickly. Note if you by mistake adopt an unstable numerical scheme you will see a particular wavelength swamp your numerically generated solution in a few dozen timesteps.

A stability requirement is that the eigenvalues of the amplification matrix should not lead to growth which means that their magnitudes should be less than 1 (note they could be complex). The eigenvalues of $\mathbf{G}$ (equation 60) are

$$\lambda_{\pm} = da \pm \sqrt{d^2a^2 + 1} \tag{61}$$

where $a = (\cos\phi_k - 1)$. However if $\phi_k = \pi/2$ then $|\lambda_-| = |-d - \sqrt{d^2+1}| > 1$. Repeated applications of $\mathbf{G}$ will lead to amplification so the scheme is *unstable*. Since the scheme is unstable no matter what size timestep is used we say the scheme (equation 48) is *unconditionally* unstable.

A small change in the scheme can make it stable. Our previous scheme that I write again here

$$u_j^{n+1} = u_j^n + K\frac{\Delta t}{(\Delta x)^2}(u_{j+1}^n + u_{j-1}^n - 2u_j^n) \tag{62}$$

has amplification matrix

$$\mathbf{G} = 1 - 2\frac{K\Delta t}{(\Delta x)^2}(1 - \cos\phi_k) \tag{63}$$

We make sure that $|\mathbf{G}| < 1$ for all possible ϕ_k. When $2\frac{K\Delta t}{(\Delta x)^2}(1-\cos\phi_k) < -2$ then $|\mathbf{G}| > 1$. We find stability only when

$$\Delta t < \frac{(\Delta x)^2}{2K}. \tag{64}$$

We say that the scheme (equation 62) is *conditionally* stable as stability depends on the timestep. The above condition sets the choice of timestep with explicit schemes for the diffusion equation. The above Fourier analysis of stability is known as von Neumann stability analysis.

The conditionally stable scheme (equation 62) is vastly superior (as it can be used) to the unstable scheme (equation 48), even though the unstable scheme is second order in Δt and so a higher order approximation to the diffusion equation than the stable one.

1.10.1. *Physical meaning of stability.* Consider again the differential equation that we are approximating

$$\frac{\partial u}{\partial t} = K\frac{\partial^2 u}{\partial x^2} \tag{65}$$

where the diffusion coefficient K has units of cm^2/s. Information propagates over a distance Δx in a timescale $t_{diffusion} = (\Delta x)^2/K$. The above condition for stability (equation 64) is equivalent to demanding that the timestep be shorter than the timescale for information to propagate between grid points.

1.11. **Upwind differencing.** Consider the advection equation

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{66}$$

and asymmetric first order scheme

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + O(\Delta t) + a\frac{u_j^n - u_{j-1}^n}{\Delta x} + O(\Delta x) = 0 \tag{67}$$

which can be written explicitly as

$$u_j^{n+1} = u_j^n + \sigma(u_{j-1}^n - u_j^n) \tag{68}$$

with

$$\sigma \equiv \frac{a\Delta t}{\Delta x}. \tag{69}$$

We can write this as

$$\mathbf{u}^{n+1} = L\mathbf{u}^n = [I + \sigma(S^{-1} - I)]\mathbf{u}^n \tag{70}$$

where S is a shift operator and I an identity operator. In Fourier space this becomes $\hat{u}_k^{n+1} = G\hat{u}_k^n$ and our operator G is the function

$$\lambda = 1 + \sigma(e^{i\phi_k} - 1) \tag{71}$$

where we have directly replaced G with λ so that it is clear we are thinking about it as an eigenvalue with amplitude that determines the stability of the map in Fourier space. For stability we require that $|\lambda| < 1$ for all possible ϕ_k with eigenvalue λ. Take the complex conjugate of λ

$$\bar{\lambda} = 1 + \sigma(e^{-i\phi_k} - 1) \tag{72}$$

$$|\lambda|^2 = \lambda\bar{\lambda} = 1 - 4\sigma(1-\sigma)\sin^2\frac{\phi_k}{2} \tag{73}$$

and we have stability only if $\sigma < 1$ (assuming that our velocity $a > 0$). This corresponds to the condition

$$\frac{a\Delta t}{\Delta x} < 1 \tag{74}$$

which is known as the CFL or *Courant-Friedrichs-Lewy* condition. Physically information is propagated along a grid cell of size Δx in a time $t \sim \Delta x/a$. Our CFL condition

$$\Delta t < \frac{\Delta x}{a} \tag{75}$$

is consistent with this timescale.

1.12. **Upwind vs Downwind.** Supposing we had not used an upwind scheme but instead had used a downwind one or

$$u_j^{n+1} = u_j^n + \sigma(u_j^n - u_{j+1}^n) \tag{76}$$

This gives

$$\begin{aligned} \lambda &= 1 + \sigma(1 - e^{i\phi_k}) \\ \bar{\lambda} &= 1 + \sigma(1 - e^{-i\phi_k}) \end{aligned} \tag{77}$$

But in this case

$$|\lambda|^2 = 1 + 4\sigma(1+\sigma)\sin^2\frac{\phi_k}{2} \tag{78}$$

and $|\lambda|$ is always greater than one (for $a > 0$). Note that for $a > 0$ the upwind scheme is stable as long as the CFL condition is satisfied but the downwind scheme is unconditionally unstable.

Consider flipping the sign of a in the advection equation so that the upwind direction becomes our previously downwind direction. It makes sense that the above

equation for $|\lambda|^2$ (equation 78) is the same as equation 73 but with the sign of σ and a flipped.

1.13. **Upwind method for linear systems with positive and negative characteristics.** We consider the system

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = 0 \tag{79}$$

where $\mathbf{A} = R\Lambda R^{-1}$ and Λ a diagonal matrix with nonzero and real (but not necessarily all positive) eigenvalues, λ_i.

How does one modify the upwind method for a linear system with characteristic velocities λ_i with some positive and some negative?

Let

$$\begin{aligned} \lambda_p^+ &= \max(\lambda_p, 0) \\ \lambda_p^- &= \min(\lambda_p, 0) \end{aligned} \tag{80}$$

and matrices where we replace negative or positive eigenvalues with zeros

$$\begin{aligned} \Lambda^+ &= \mathrm{diag}(\lambda_1^+, \lambda_2^+, ..., \lambda_m^+) \\ \Lambda^- &= \mathrm{diag}(\lambda_1^-, \lambda_2^-, ..., \lambda_m^-) \end{aligned} \tag{81}$$

Note $\Lambda = \Lambda^+ + \Lambda_-$. We can define

$$\mathbf{A}^+ = R\Lambda^+ R^{-1} \qquad \mathbf{A}^- = R\Lambda^- R^{-1} \tag{82}$$

In the basis of our eigenvectors the upwind scheme can written as

$$\mathbf{V}_j^{n+1} = \mathbf{V}_j^n - \frac{\Delta t}{\Delta x}\Lambda^+(\mathbf{V}_j^n - \mathbf{V}_{j-1}^n) - \frac{\Delta t}{\Delta x}\Lambda^-(\mathbf{V}_{j+1}^n - \mathbf{V}_j^n) \tag{83}$$

Back in our original basis

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \frac{\Delta t}{\Delta x}\mathbf{A}^+(\mathbf{U}_j^n - \mathbf{U}_{j-1}^n) - \frac{\Delta t}{\Delta x}\mathbf{A}^-(\mathbf{U}_{j+1}^n - \mathbf{U}_j^n) \tag{84}$$

giving us a compact notation for an upwind scheme when we have both positive and negative characteristic velocities.

1.14. **The Modified equation – Numerically generated diffusion and dispersion.** Let's look at the form of our error for our upwind scheme for the linear advection equation $u_{,t} + au_{,x} = 0$. We recall the scheme is

$$u_j^{n+1} = u_j^n + \sigma(u_{j-1}^n - u_j^n) \tag{85}$$

with $\sigma = a\Delta t/\Delta x$. The Taylor expansion of each term is

$$u_j^{n+1} = u_j^n + \Delta t\frac{\partial u}{\partial t} + \frac{(\Delta t)^2}{2}\frac{\partial^2 u}{\partial t^2} + O((\Delta t)^3) \tag{86}$$

$$u_{j+1}^n = u_j^n + \Delta x\frac{\partial u}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 u}{\partial x^2} + \frac{(\Delta x)^3}{6}\frac{\partial^3 u}{\partial x^3} + \frac{(\Delta x)^4}{4!}\frac{\partial^4 u}{\partial x^4} + O((\Delta x)^5) \tag{87}$$

$$u_{j-1}^n = u_j^n - \Delta x\frac{\partial u}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 u}{\partial x^2} - \frac{(\Delta x)^3}{6}\frac{\partial^3 u}{\partial x^3} + \frac{(\Delta x)^4}{4!}\frac{\partial^4 u}{\partial x^4} + O((\Delta x)^5) \tag{88}$$

Now substitute these back into our difference equation we find

$$\Delta t\frac{\partial u}{\partial t} = \sigma\left[-\Delta x\frac{\partial u}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 u}{\partial x^2} - \frac{(\Delta x)^3}{6}\frac{\partial^3 u}{\partial x^3}\right] + O(\Delta t^2, \Delta x^3) \tag{89}$$

becoming

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = \frac{a\Delta x}{2}\frac{\partial^2 u}{\partial x^2} + O(\Delta t, \Delta x^2) \tag{90}$$

While our finite difference technique is $O(\Delta x)$ approximating the linear advection equation, it is $O((\Delta x)^2)$ approximating the above differential equation.

Our finite difference technique is integrating to a higher precision an equation that is not the same as one desired and has additional terms. In particular the term proportional to $\frac{\partial^2 u}{\partial x^2}$ is **diffusive** so we have an advection diffusion equation. The amount of diffusion depends on $a\Delta x$. Our original equation had solutions in the form $f(x - at)$. Because of the diffusive term a function with sharp edge will smooth as it propagates. The amplitude of a traveling wave will decrease with time. Consider inserting a wave $u \propto e^{i(\omega t - kx)}$ into the above equation. We find a dispersion relation

$$\omega = ak + ik^2 a\Delta x/2 \tag{91}$$

Leading to amplitude of the wave decaying with $e^{-k^2 a\Delta x/2}$.

Another example is the second order Lax-Wendroff method which is a second order accurate finite difference method for the linear advection equation. The Lax-Wendroff scheme approximates the linear advection equation

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0 \tag{92}$$

with

$$u_j^{n+1} = u_j - \frac{a\Delta t}{2\Delta x}(u_{j+1}^n - u_{j-1}^n) + \frac{a^2(\Delta t)^2}{2(\Delta x)^2}(u_{j+1}^n + u_{j-1}^n - 2u_j^n). \tag{93}$$

The scheme is based on the Taylor expansion to second order

$$u_j^{n+1} = u_j^n + \Delta t\frac{\partial u_j^n}{\partial t} + \frac{(\Delta t)^2}{2}\frac{\partial^2 u_j^n}{\partial t^2}. \tag{94}$$

The Lax-Wendroff method is equivalent to a modified equation to order $O((\Delta x)^3)$ that is in the form

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = \mu\frac{\partial^3 u}{\partial x^3} \tag{95}$$

The above is a **dispersive** equation. If we insert $u \propto e^{i(\omega t - kx)}$ we find a dispersion relation

$$\omega = ak + \mu k^3 \tag{96}$$

Each frequency propagates at a slightly different velocity. For the Lax-Wendroff method $\mu = (\Delta x)^2 a(\nu^2 - 1)/6$ where $\nu = a\Delta t/\Delta x$ is the Courant number. For stability $\nu < 1$ so we expect $\mu < 0$ and all wave numbers travel more slowly than they should be. The smaller wavenumbers have larger k values and travel the fastest. A discontinuity will not remain sharp. High frequency waves will arrive first giving oscillations in front of a traveling discontinuity.

1.15. **General Issues for Finite Differencing Schemes.** Numerical integration by finite differencing introduces errors. Numerical instability is intolerable and so only schemes that are stable can be used. Implicit schemes can achieve stability and accuracy but are more difficult to solve than explicit schemes. Errors introduced can take different forms such as numerical or artificial viscosity, diffusion and dispersion. Typically first order schemes are more diffusive than second order schemes and so don't well resolve discontinuities or shocks and energy is lost during propagation of waves or discontinuities. Higher order schemes can be dispersive and cause unphysical oscillations near discontinuities. Shocks may not propagate at the desired sound speed. More complex methods such as those involving Riemann solvers may solve some of these problems but at the expense of additional computational complexity and associated inflexibility in the code making it more difficult to add additional physical processes into the code.

For most hyperbolic problems one expects a stability condition based on the CFL condition setting the timestep $\Delta t \lesssim \Delta x/a$ where a is the speed of characteristics. For diffusive systems a condition on the timestep $\Delta t \lesssim (\Delta x)^2/K$ for diffusion coefficient K. For mixed advection and parabolic systems we would expect that both conditions must be satisfied.

1.16. **Some Simple Finite Differencing Schemes.** If you expect a positive velocity at all positions in your grid then you can use a one-sided first order scheme. Remember that it must be upwind to be stable. For the linear advection equation $u_t + au_x = 0$, and $a > 0$ choose

$$u_j^{n+1} = u_j^n - a\frac{\Delta t}{\Delta x}(u_j^n - u_{j-1}^n) \tag{97}$$

If $a < 0$ choose

$$u_j^{n+1} = u_j^n - a\frac{\Delta t}{\Delta x}(u_{j+1}^n - u_j^n) \tag{98}$$

When I modify these schemes to cover a non-linear setting I have found that a fixed boundary can present a problem causing instability. We have had trouble on the boundary when trying to model a 1-dimensional wind or accretion flow.

A first order scheme that is stable for both positive and negative velocities is the Lax-Friedrichs scheme or

$$u_j^{n+1} = \frac{u_{j+1}^n + u_{j-1}^n}{2} - a\frac{\Delta t}{\Delta x}\frac{(u_{j+1}^n - u_{j-1}^n)}{2} \tag{99}$$

For a nonlinear system in conservation law form

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left[f(u)\right] = 0 \tag{100}$$

and using a modified flux the Lax-Friedrichs method becomes

$$u_j^{n+1} = \frac{u_{j+1}^n + u_{j-1}^n}{2} - \left(\frac{\Delta t}{2\Delta x}\right)\left[f(u_{j+1}^n) - f(u_{j-1}^n)\right] \tag{101}$$

This stably covers a non-linear setting like Euler's equation and conservation of mass in 1dimension and can be stable near a fixed boundary.

The Lax-Wendroff scheme is a second order scheme with

$$u_j^{n+1} = u_j^n - a\left(\frac{\Delta t}{\Delta x}\right)\frac{(u_{j+1}^n - u_{j-1}^n)}{2} + a^2\left(\frac{\Delta t}{\Delta x}\right)^2\frac{(u_{j+1}^n + u_{j-1}^n - 2u_j^n)}{2} \tag{102}$$

This scheme can be modified to cover an advection-diffusion equation.

These are examples of schemes than can be quickly employed to explore solutions a simple set of hyperbolic differential equations. However, more sophisticated schemes would improve upon numerical viscosity, diffusion and dispersion, and ability to accurately portray discontinuities.

1.17. **Grids in different coordinate systems.** Some examples to be added here!

Consider a grid with spacing that is set with a function, for example, by $x = \log y$ and a differential equation that depends on y, t. It is straightforward to determine dx for the grid spacing $dx = dy/y$ and so modify a 1d finite differencing scheme.

Changes in 1d coordinate system are straightforward but varying the 2 or 3 coordinate system is less so because differential operators become more complex (for example going from cartesian coordinates to cylindrical coordinates).

The timestep for the entire grid is usually based on the minimal value of the CFL (or related) condition.

Name	Difference Equations	Stencil
Backward Euler	$U_j^{n+1} = U_j^n - \frac{k}{2h}A(U_{j+1}^n - U_{j-1}^n)$	
One-sided	$U_j^{n+1} = U_j^n - \frac{k}{h}A(U_j^n - U_{j-1}^n)$	
One-sided	$U_j^{n+1} = U_j^n - \frac{k}{h}A(U_{j+1}^n - U_j^n)$	
Lax-Friedrichs	$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{k}{2h}A(U_{j+1}^n - U_{j-1}^n)$	
Leapfrog	$U_j^{n+1} = U_j^{n-1} - \frac{k}{2h}A(U_{j+1}^n - U_{j-1}^n)$	
Lax-Wendroff	$U_j^{n+1} = U_j^n - \frac{k}{2h}A(U_{j+1}^n - U_{j-1}^n)$ $+\frac{k^2}{2h^2}A^2(U_{j+1}^n - 2U_j^n + U_{j-1}^n)$	
Beam-Warming	$U_j^{n+1} = U_j^n - \frac{k}{2h}A(3U_j^n - 4U_{j-1}^n + U_{j-2}^n)$ $+\frac{k^2}{2h^2}A^2(U_j^n - 2U_{j-1}^n + U_{j-2}^n)$	

TABLE 10.1

Finite difference methods for the linear problem $u_t + Au_x = 0$.

FIGURE 2. From Randall LeVeque's book Numerical Methods for Conservation Laws. The stencil for the backward Euler equation doesn't make any sense to me, it should be the same as the Lax Wendroff stencil.

1.18. **Boundary Conditions.** The easiest boundary condition to use is a periodic periodic boundary. Excepting in the case of periodic boundaries, a finite difference scheme must be modified specifically for each boundary. When the integrated quantities are more than 1 dimensional (density and velocity) then you must choose which variable is fixed on a fixed boundary and this determines the sign that pulses get when they reflect off of the boundary. I have found that a perfectly stable scheme

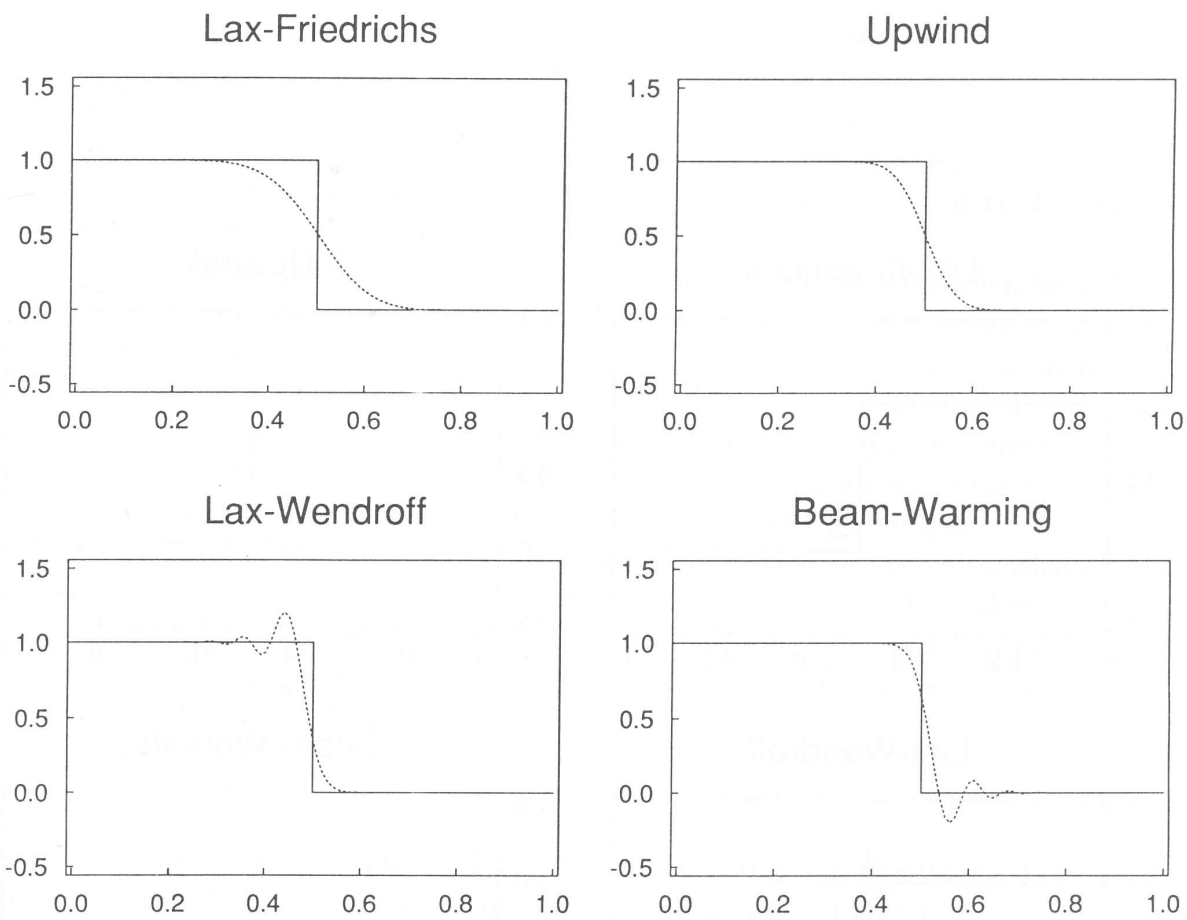

Figure 11.1. Numerical and exact solution to (11.1) with $h = 0.01$ and the following methods: (a) Lax-Friedrichs, (b) Upwind, (c) Lax-Wendroff, (d) Beam-Warming.

FIGURE 3. From Randall LeVeque's book Numerical Methods for Conservation Laws

can be unstable at a boundary, (presumably that means that the boundary condition has been badly chosen).

When waves are generated, it is challenging to keep them from bouncing off the boundaries and interfering with the region of interest. Rotating systems present a particular challenge.

I should give some examples of boundary conditions (first order ones) here! As for example in the problem set 8 for PHY256.

2. Conservative methods and Riemann Solvers

2.1. **Conservation Laws and shock speeds.** Conservation laws can be used to estimate shock speeds. Consider equation

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0 \tag{103}$$

We can consider the conditions at a shock front that is moving with the speed s. The Rankine-Hugoniot conditions become

$$s(U_2 - U_1) = F(U_2) - F(U_1) \tag{104}$$

If the system is linear then $\mathbf{F}(\mathbf{U}) = \mathbf{A}\mathbf{U}$ for matrix $\mathbf{A}$.

$$s\Delta U = \mathbf{A}\Delta U \tag{105}$$

The above implies that a *single* discontinuity propagating at speed s can only do so if s is an eigenvalue of $\mathbf{A}$. Thus the eigenvalues of $\mathbf{A}$ give the characteristic velocities and these are the same thing as shock velocities or velocities of discontinuity in the problem. When the system is not linear we can still estimate shock speeds with equation (104). This gives us a feeling why it is useful to consider our fluid equations in conservation law form. A finite difference technique based on conservation law form will more truthfully match shock speeds.

For example, we can write Burger's equation in the quasi-linear form

$$\partial_t u + u\partial_x u = 0 \tag{106}$$

instead of in conservation law form $\partial_t u + \partial_x \left(\frac{u^2}{2}\right) = 0$.

Consider the following upwind method based on the non-conservation law form of Burger's equation

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} u_j^n (u_j^n - u_{j-1}^n) \tag{107}$$

Let's look at a boundary with

$$u_j^0 = \begin{cases} 1 & \text{for} \quad j < 0 \\ 0 & \phantom{\text{for}} \quad j \geq 0 \end{cases} \tag{108}$$

Unfortunately it is easy to verify that our finite difference technique gives $u_j^1 = u_j^0$ (no change) for all j. This is obviously wrong as the discontinuity should propagate with the velocity 1. This is an example of a case where the velocity of the discontinuity or shock is very badly estimated by the finite difference scheme.

2.2. **Conservative schemes.** Consider again the scalar conservation law

$$u_{,t} + f(u)_{,x} = 0 \tag{109}$$

One way around the problem posed above caused by a non-conservative scheme is to adopt a finite difference scheme in the form

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}\left[G(u^n; j) - G(u^n; j-1)\right] \tag{110}$$

where the flux function $G(u)$ could depend on a number of spatial positions (j values). Here $G(u^n; j-1)$ is the same as $G(u^n; j)$ but with all spatial positions shifted by 1.

If we choose our flux functions G correctly then the discontinuity will propagate at the appropriate speed. When characteristics near a discontinuity converge we know the discontinuity will propagate at a speed s given by the Rankine-Hugoniot condition

$$s(u_2 - u_1) = f(u_2) - f(u_1) \tag{111}$$

If we choose G to be consistent with the flux $f(u)$ then the discontinuity should propagate at the right speed.

A **conservative method or scheme** for a conservation law is a numerical method in the form

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{\Delta x}\left[f_{j-\frac{1}{2}} - f_{j+\frac{1}{2}}\right] \tag{112}$$

where the numerical flux

$$f_{j+\frac{1}{2}} = G(u_{j-l_L}, ..., u_{j+l_R}) \tag{113}$$

is an approximation to the physical flux $f(u)$. Here l_L, l_R are two non-negative numbers. The indexes involve $1/2$ arise from considering volumes over regions of space. For example if we consider two boxes, each centered at the position given by $j, j+1$ then we want the flux through the interface between them or at $j+1/2$. The simplest example would be if G only involved two positions or

$$u_j^{n+1} = u_n^n - \frac{\Delta t}{\Delta x}\left[G(u_j^n, u_{j+1}^n) - G(u_{j-1}^n, u_j^n)\right] \tag{114}$$

We need to specify a function $G(u_l, u_r)$ that is a function of left and right positions.

Let's explore a different scheme to Burger's equation that is in conservation law form. Using

$$G(u_l, u_r) = f(u_l) = \frac{1}{2}u_l^2 \tag{115}$$

we find

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{2\Delta x}\left[(u_j^n)^2 - (u_{j-1}^n)^2\right] \tag{116}$$

This is an "upwind" method, and the choice of whether to use the right or left position for G would depend on the sign of the characteristic velocities, specifying the upwind direction.

Or we could chose for our flux function an average,

$$G(u_l, u_r) = \frac{1}{2}\left[f(u_l) + f(u_r)\right] \tag{117}$$

and giving

$$u_j^{n+1} = u_n^n - \frac{\Delta t}{2\Delta x}\left[f(u_{j+1}^n) - f(u_{j-1}^n)\right] \tag{118}$$

This unfortunately is unconditional **unstable**.

The modified flux

$$G(u_l, u_r) = \frac{1}{2}\left[f(u_l) + f(u_r)\right] - \frac{\Delta t}{2\Delta x}(u_r - u_l) \tag{119}$$

gives

$$u_j^{n+1} = \frac{1}{2}\left[u_{j-1}^n + u_{j+1}^n\right] - \frac{\Delta t}{2\Delta x}\left[f(u_{j+1}^n) - f(u_{j-1}^n)\right] \tag{120}$$

which is known as the **Lax-Friedrichs method** (and using fluxes here), and this is conditionally stable. The extra term in the flux is based on an approximation to $(\Delta t)^2 u_x/(2\Delta x)$ and so adds a diffusive flux term to G. The modification amounts to adding some artificial viscosity to the centered flux function G.

As we have seen above a non-conservative method can lead to solution with the shock propagating at the wrong speed. This motivates using conservation law forms of the differential equations and considering characteristic velocities. An important development in numerical approximations to hydrodynamics was Godunov's method. Godunov proposed a way to make use of the characteristics information within the framework of a conservative method. Godunov suggested solving local Riemann problems forward in time at each grid interface. Solutions to the Riemann problem give substantial information about the characteristic structure and are exact solutions to conservation laws.

Before we go on to talk about Godunov's method and Riemann solvers we first will explore what is known as the Riemann problem.

2.3. **The Riemann Problem.** Consider a linear hyperbolic system

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = 0 \tag{121}$$

with a discontinuous initial condition

$$\mathbf{U}(x, t=0) = \begin{cases} \mathbf{U_L} & \text{for} \quad x < 0 \\ \mathbf{U_R} & \phantom{\text{for}} \quad x > 0 \end{cases} \tag{122}$$

Previously we discussed how to estimate a shock velocity or the velocity of a discontinuity from a conservation law but we did it in only 1 dimension. The Riemann problem is similar except we can have more than 1 characteristic velocity. What this means is that a discontinuity may break up into a series of discontinuities each traveling with its own speed.

For this to be hyperbolic, the matrix $\mathbf{A}$ must have m distinct and real eigenvalues $\lambda_1 < \lambda_2 < ... < \lambda_m$ where m is the dimension of the vector $\mathbf{U}$. We write expand the initial $\mathbf{U_L}, \mathbf{U_R}$ in terms of eigenvectors of A which we call $\mathbf{K_i}$.

$$\mathbf{U_L} = \sum_{i=1}^{m} \alpha_i \mathbf{K}_i \qquad \mathbf{U_R} = \sum_{i=1}^{m} \beta_i \mathbf{K}_i \tag{123}$$

Each eigenvector $\mathbf{K}_i$ has its own characteristic velocity λ_i. We can figure out the discontinuity propagation speed for each eigenvector. In eigenvector space we have m scalar problems

$$\frac{\partial w_i}{\partial t} + \lambda_i \frac{\partial w_i}{\partial x} = 0 \tag{124}$$

and m initial conditions

$$w_i(x, t = 0) = \begin{cases} \alpha_i & \text{for} \quad x < 0 \\ \beta_i & \phantom{\text{for}} \quad x > 0 \end{cases} \tag{125}$$

The solutions to each of these m initial value problems is

$$w_i(x, t) = w_i(x, 0)(x - \lambda_i t) = \begin{cases} \alpha_i & \text{for} \quad x - \lambda_i t < 0 \\ \beta_i & \phantom{\text{for}} \quad x - \lambda_i t > 0 \end{cases} \tag{126}$$

Our solution for the full problem can be described as

$$\mathbf{U}(x, t) = \sum_{i=I+1}^{m} \alpha_i \mathbf{K}_i + \sum_{i=1}^{I} \beta_i \mathbf{K}_i \tag{127}$$

where I is the maximum value of i for which $x - \lambda_i t > 0$ (remember we put all our eigenvalues in increasing order).

2.3.1. *2d Riemann problem.* A 2 dimensional problem is a good one to start with as it is not trivial but illustrates what happens when there are two characteristic velocities.

We expect two discontinuities one propagating at the characteristic of the first eigenvalue, the second one propagating at the second eigenvalue. For $x - \lambda_1 t < 0$ the solution is equal to the left state $\mathbf{U_L}$. For $x - \lambda_2 t > 0$ the solution is equal to the right sate $\mathbf{U_R}$. It is only for $x - \lambda_1 t > 0$ and $x - \lambda_2 t < 0$ or $\lambda_1 t < x < \lambda_2 t$ that the solution is interesting. In this region, often called the star region, the new solution involves a sum of the initial condition in the left and right states.

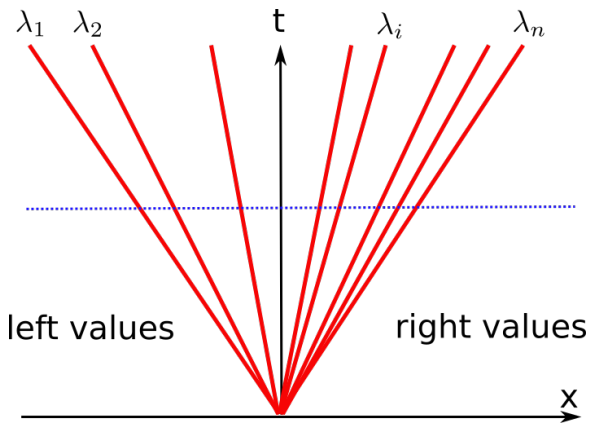

FIGURE 4. Structure of the solution of the Riemann problem for a general m dimensional linear hyperbolic system with constant coefficients. λ_i are the characteristic velocities of the system in order of increasing size. This figure shows characteristics so the horizontal axis is x and the vertical axis is t. At a particular time draw a horizontal line across the plot. The solution should have a jump crossing each characteristic. The x positions of each jump is read off from the horizontal line.

We look at the left and right eigenvalue decomposition for $m = 2$

$$\begin{aligned} \mathbf{U}_L &= \alpha_1 \mathbf{K}_1 + \alpha_2 \mathbf{K}_2 \\ \mathbf{U}_R &= \beta_1 \mathbf{K}_1 + \beta_2 \mathbf{K}_2 \end{aligned} \tag{128}$$

In the star region we get the coefficient from the slower characteristic from the right and the coefficient from the faster coefficient from the left,

$$\mathbf{U}_* = \beta_1 \mathbf{K}_1 + \alpha_2 \mathbf{K}_2 \tag{129}$$

The full solution is

$$\mathbf{U}(x,t) = \begin{cases} \mathbf{U}_L & & x < \lambda_1 t \\ \mathbf{U}_* & \text{for} & \lambda_1 t < x < \lambda_2 t \\ \mathbf{U}_R & & x > \lambda_2 t \end{cases} \tag{130}$$

2.4. **Riemann Problem, the example of linearized gas dynamics.** In one dimension recall that our continuity equation and Euler's equation can be written

$$\begin{aligned} \rho_t + \rho u_x + u \rho_x &= 0 \\ u_t + u u_x + c_s^2 \rho_x / \rho &= 0 \end{aligned} \tag{131}$$

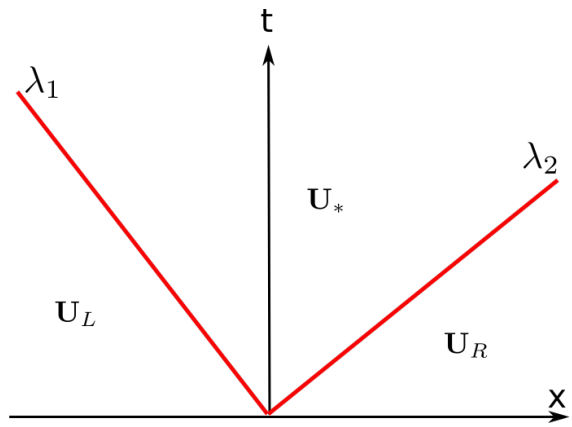

FIGURE 5. Structure of the solution of the Riemann problem for a 2 dimensional linear hyperbolic system with constant coefficients.

where c_s is the sound speed, and we have neglected gravity. Consider perturbations around a steady state

$$\begin{aligned} \rho(x,t) &= \rho_0 + \rho_1(x,t) \\ u(x,t) &= 0 + u_1(x,t) \end{aligned} \tag{132}$$

and u_1 is small compared to c_s and ρ_1 is small compared to ρ_0.

Taking only first order terms our differential equations become

$$\begin{aligned} \rho_{1,t} + \rho_0 u_{1,x} &= 0 \\ u_{1,t} + \frac{c_s^2}{\rho_0}\rho_{1,x} &= 0 \end{aligned} \tag{133}$$

The above equations are know as the linearized equations of gas dynamics. The equations can be written in matrix form as

$$\begin{pmatrix} \rho_1 \\ u_1 \end{pmatrix}_t + \begin{pmatrix} 0 & \rho_0 \\ c_s^2/\rho_0 & 0 \end{pmatrix} \begin{pmatrix} \rho_1 \\ u_1 \end{pmatrix}_x = 0 \tag{134}$$

The eigenvalues are

$$\lambda_1 = -c_s \qquad \lambda_2 = c_s \tag{135}$$

These eigenvalues correspond to right eigenvectors

$$\mathbf{K}_1 = \begin{bmatrix} \rho_0 \\ -c_s \end{bmatrix} \qquad \mathbf{K}_2 = \begin{bmatrix} \rho_0 \\ c_s \end{bmatrix} \tag{136}$$

Now we consider our initial conditions

$$\mathbf{U}(x, t=0) = \begin{cases} \mathbf{U_L} & \text{for} \quad x < 0 \\ \mathbf{U_R} & \phantom{\text{for}} \quad x > 0 \end{cases} \tag{137}$$

First let us decompose the left state

$$\mathbf{U_L} = \begin{bmatrix} \rho_L \\ u_L \end{bmatrix} = \alpha_1 \begin{bmatrix} \rho_0 \\ -c_s \end{bmatrix} + \alpha_2 \begin{bmatrix} \rho_0 \\ c_s \end{bmatrix} \tag{138}$$

We solve for the coefficients finding

$$\alpha_1 = \frac{c_s\rho_L - \rho_0 u_L}{2c_s\rho_0}, \qquad \alpha_2 = \frac{c_s\rho_L + \rho_0 u_L}{2c_s\rho_0} \tag{139}$$

Likewise we can solve for the coefficients of the eigenvectors for the right hand data

$$\mathbf{U_R} = \begin{bmatrix} \rho_R \\ u_R \end{bmatrix} = \beta_1 \begin{bmatrix} \rho_0 \\ -c_s \end{bmatrix} + \beta_2 \begin{bmatrix} \rho_0 \\ c_s \end{bmatrix} \tag{140}$$

finding for the coefficients

$$\beta_1 = \frac{c_s\rho_R - \rho_0 u_R}{2c_s\rho_0}, \qquad \beta_2 = \frac{c_s\rho_R + \rho_0 u_R}{2c_s\rho_0} \tag{141}$$

Recall the solution in the star region (see previous subsection) is

$$\mathbf{U}_* = \beta_1\mathbf{K}_1 + \alpha_2\mathbf{K}_2 \tag{142}$$

Using our eigenvectors

$$\mathbf{U}_* = \begin{bmatrix} \rho_* \\ u_* \end{bmatrix} = \beta_1 \begin{bmatrix} \rho_0 \\ -c_s \end{bmatrix} + \alpha_2 \begin{bmatrix} \rho_0 \\ c_s \end{bmatrix} \tag{143}$$

Inserting our relations for β_1 and α_2 and simplifying we find that

$$\begin{aligned} \rho_* &= \frac{1}{2}(\rho_L + \rho_R) - \frac{1}{2}(u_R - u_L)\rho_0/c_s \\ u_* &= \frac{1}{2}(u_L + u_R) - \frac{1}{2}(\rho_R - \rho_L)c_s/\rho_0 \end{aligned} \tag{144}$$

The full solution is

$$(\rho, u)(x,t) = \begin{cases} (\rho_L, u_L) & & x < -c_s t \\ (\rho_*, u_*) & \text{for} & -c_s t < x < c_s t \\ (\rho_R, u_R) & & x > c_s t \end{cases} \tag{145}$$

2.5. **Riemann Problem and the Hugoniot locus.** For a two dimensional linear system, there are two characteristic velocities (eigenvalues, λ_1, λ_2), and two corresponding eigenvectors. A single discontinuity propagating with speed, s, must satisfied the Rankine-Hugoniot condition. This implies that $U_L - U_R$ must be proportional to one of the eigenvectors.

For arbitrary left and right values, U_L, U_R, the difference between the two vectors likely is not aligned with one of the eigenvectors. In this case two discontinuities arise, one propagating at λ_1 and the other at λ_2. The state in between the two discontinuities we called U_*. Each discontinuity must satisfy the Rankine-Hugoniot

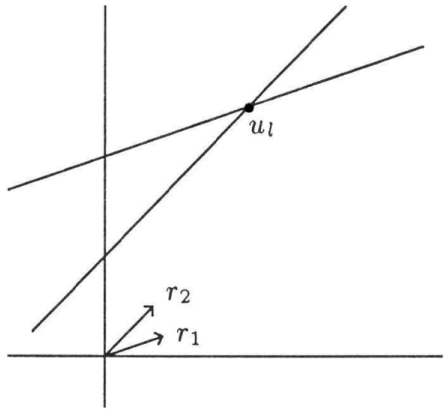

FIGURE 6. The Hugoniot locus for the state U_L consists of all states that differ from U_L by a scalar multiple of one of the eigenvectors, r_1 or r_2, shown on the lower left. This plot is for a two-dimensional linear system, $U_t + AU_x = 0$. Axes would be ρ and j for the two dimensional linearized isothermal gas dynamics case. This Figure has been take from Leveque's book (his figure 6.3).

condition. For each discontinuity we can consider the Rankine-Hugoniot locus (two lines, for the linear system). We connect U_L and U_L by two Rankine-Hugoniot loci (one for each discontinuity) that intersect at the value U_*.

See Figures 6 and 7.

2.6. **Shocks and Rarefraction Waves in Burger's equation.** For the linear systems with constant coefficients the values of the characteristics don't depend on the initial conditions. Let's consider a one dimensional non-linear example before we generalize the Riemann problem to the non-linear case. Recall Burger's equation

$$u_{,t} + uu_{,x} = 0 \qquad\qquad u_{,t} + \left(\frac{u^2}{2}\right)_{,x} = 0 \tag{146}$$

The form on the left shows that characteristic velocities are given by u itself. The conservation law form on the right shows a flux of $u^2/2$.

Consider the following two initial conditions

$$u(x, t = 0)_A = \begin{cases} 1 & \text{for} \quad x < 0 \\ \frac{1}{2} & \phantom{\text{for}} \quad x \geq 0 \end{cases} \tag{147}$$

$$u(x, t = 0)_B = \begin{cases} \frac{1}{2} & \text{for} \quad x < 0 \\ 1 & \phantom{\text{for}} \quad x \geq 0 \end{cases} \tag{148}$$

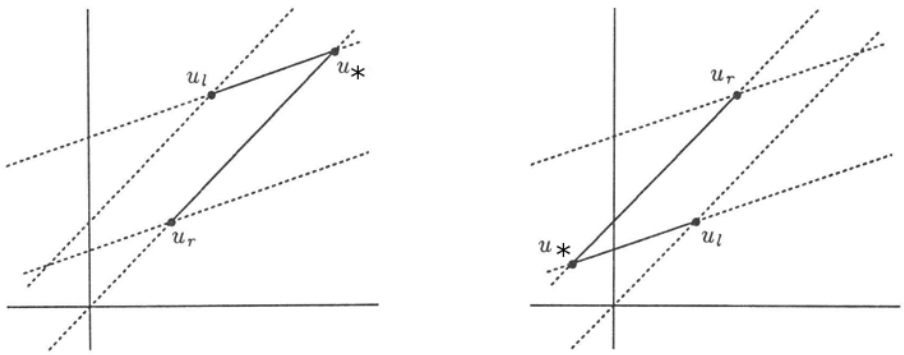

FIGURE 7. When U_L and U_R do not lie along a single eigenvector, two discontinuities propagate, one at each characteristic velocity (or eigenvalue). The intermediate state, U_* lies at the intersection of the Hugoniot loci (each parallel to an eigenvector direction). To get from U_L to U_R we must first travel along the r_1 direction, where r_1 corresponds to the eigenvector with the slower characteristic velocity, and then along the r_2 direction, corresponding to the faster characteristic velocity. This Figure has been take from Leveque's book (his figure 6.4).

For u_A the characteristic velocity on the left side ($x < 0$) is 1 and the characteristic velocity on the right side ($x > 0$) is 1/2. The characteristic velocity is larger on the left than the right so the characteristics are converging. We will get a shock. Using the Ranking Hugoniot relation we can find the shock velocity from the conservation law form of the equation. Recall the jump condition for a conservation law $U_{,t} + F(U)_{,x} = 0$.

$$s(U_2 - U_1) = F(U_2) - F(U_1) \tag{149}$$

for shock velocity s. For our problem with the first set of initial conditions $u_2 = 1/2$ and $u_1 = 1$, the fluxes are $f_2 = u_2^2/2 = 1/8$, and $f_1 = u_1^2/2 = 1/2$

$$s = \frac{1/8 - 1/2}{1/2 - 1} = \frac{3/8}{1/2} = 3/4 \tag{150}$$

What does the solution for u_A look like?

$$u(x,t) = \begin{cases} 1 & \text{for} \quad x < st \\ \frac{1}{2} & \phantom{\text{for}} \quad x \geq st \end{cases} \tag{151}$$

Now let us consider the second initial condition

$$u(x,t=0)_B = \begin{cases} \frac{1}{2} & \text{for} \quad x<0 \\ 1 & \quad\quad\quad x \geq 0 \end{cases} \tag{152}$$

In this case the characteristic velocity is lower on the left than on the right and the characteristics are diverging. We will get a rarefraction wave. If we assume a fan of characteristics coming from the discontinuity at $t=0$ the solution is

$$u(x,t) = \begin{cases} 1/2 & \text{for} \quad x<t/2 \\ x/t & \quad\quad\quad t/2<x\leq t \\ 1 & \quad\quad\quad x \geq t \end{cases} \tag{153}$$

For a multidimensional non-linear system there are three cases to consider

- Shock wave. Characteristics converge. The Rankine-Hugoniot condition can be applied to the discontinuity to estimate its speed of propagation.
- Rarefaction wave. Characteristics diverge. There is a smooth transition. The initial discontinuity becomes a smooth transition region.
- Contact Wave. Two characteristics are the same velocity. The Rankine-Hugoniot condition applies to the discontinuity.

When we have three hydrodynamic conservation law equations, one for mass, one for momentum and one for energy, we find three eigenvalues for the Jacobian matrix instead of two eigenvalues. The three eigenvalues are $u \pm c_s, u$ where u is the velocity and c_s the sound speed. The third eigenvalue can give you contact discontinuities.

When we have a linear system, then characteristics don't converge or diverge. When the system is non-linear, the Rankine-Hugoniot loci are curves rather than lines as shown on Figure 6 and 7. Solutions that are discontinuous are not necessarily unique, however not all of them are physically relevant, for example some of them may be unstable. We can consider our linear equation as the vanishing viscosity limit of a second order differential equation and try to chose discontinuous solutions that are vanishing viscosity limits of a second order differential equation. This choice is often referred to as an *entropy condition*. The rarefaction fan, with solution that is proportional to x/t, is such a limit and so is considered a physically relevant solution.

2.7. **Riemann Problem and Hugoniot locus for a Non-Linear System.** Supposing one one side of a jump we have $\mathbf{U}_1$ and flux $\mathbf{F}(\mathbf{U}_1)$. We can ask what values of $\mathbf{U}_2$ and velocity s are allowed. The Rankine-Hugoniot jump condition relates s and $\mathbf{U}_2$ for a specific $\mathbf{U}_1$. The jump condition gives curves for $\mathbf{U}_2$, where each value corresponds to a particular velocity, s. The set of points on these curves is often called the **Hugoniot locus**. There may be more than one curve. If $\mathbf{U}_2$ lies along the p-th Hugoniot curve then we say that $\mathbf{U}_2$ and $\mathbf{U}_1$ are connected by a p-shock. We can parametrize each curve with a variable ξ where $s_p(\xi)$. At $\xi=0$, we assert that

$\mathbf{U}_{2,p}(\xi = 0) = \mathbf{U}_1$, and $s_p(\xi = 0) = 0$, corresponding to a shock with zero velocity and no jump.

The jump condition gives for each curve

$$\mathbf{F}(\mathbf{U}_{2,p}(\xi)) - \mathbf{F}(\mathbf{U}_1) = s_p(\xi)(\mathbf{U}_{2,p}(\xi) - \mathbf{U}_1). \tag{154}$$

Differentiating this expression with respect to ξ and setting $\xi = 0$ gives

$$\mathbf{F}'(\mathbf{U}'_{2,p}(0))\mathbf{U}'_{2,p}(0) = s'_p(0)(\mathbf{U}_{2,p}(0) - \mathbf{U}_1) + s_p(0)\mathbf{U}'_{2,p}(0), \tag{155}$$

and using the condition for $\xi = 0$,

$$\mathbf{F}'(\mathbf{U}_1)\mathbf{U}'_{2,p}(0) = s_p(0)\mathbf{U}'_{2,p}(0). \tag{156}$$

The above relation implies that $\mathbf{U}'_{2,p}(0)$ is a right eigenvector of $\mathbf{F}'(\mathbf{U}_1)$ and that $s_p(0)$ is an eigenvalue of this matrix.

For example consider the one dimensional gas dynamic equations for an isothermal fluid.

$$\rho_t + j_x = 0 \tag{157}$$

$$j_t + \left(\frac{j^2}{\rho} + a^2\rho\right)_x = 0 \tag{158}$$

where j is the mass flux. This can be written

$$u_t + f(u)_x = 0 \tag{159}$$

where

$$u = \begin{pmatrix} \rho \\ j \end{pmatrix} \tag{160}$$

and

$$f(u) = \begin{pmatrix} j \\ \frac{j^2}{\rho} + a^2\rho \end{pmatrix} \tag{161}$$

The Jacobian of the matrix is

$$f'(u) = \begin{bmatrix} 0 & 1 \\ a^2 - \frac{j^2}{\rho^2} & 2j/\rho \end{bmatrix} \tag{162}$$

and eigenvalues are

$$\lambda_\pm = \frac{j}{\rho} \pm a \tag{163}$$

and eigenvectors

$$r_\pm = \begin{pmatrix} 1 \\ j/\rho \pm a \end{pmatrix}. \tag{164}$$

The Rankine-Hugoniot condition becomes

$$j_2 - j_1 = s(\rho_2 - \rho_1) \tag{165}$$

$$\left(\frac{j_2^2}{\rho_2} + a^2\rho_2\right) - \left(\frac{j_1^2}{\rho_1} + a^2\rho_1\right) = s(j_2 - j_1). \tag{166}$$

Solving for j_2 and s in terms of ρ_2

$$j_2 = \frac{\rho_2 j_1}{\rho_1} \pm a\sqrt{\frac{\rho_2}{\rho_1}}(\rho_2 - \rho_1) \tag{167}$$

$$s = \frac{j_1}{\rho_1} \pm a\sqrt{\frac{\rho_2}{\rho_1}}. \tag{168}$$

We can parametrize the curves with ξ using

$$\rho_{2,p} = \rho_1(1+\xi) \tag{169}$$

Rewriting our solutions

$$u_{2,-} = u_1 + \xi \begin{pmatrix} \rho_1 \\ j_1 - a\rho_1\sqrt{1+\xi} \end{pmatrix}, \qquad s_- = \frac{j_1}{\rho_1} - a\sqrt{1+\xi} \tag{170}$$

$$u_{2,+} = u_1 + \xi \begin{pmatrix} \rho_1 \\ j_1 + a\rho_1\sqrt{1+\xi} \end{pmatrix}, \qquad s_+ = \frac{j_1}{\rho_1} + a\sqrt{1+\xi}. \tag{171}$$

Note that equation 156 related the eigenvalues and eigenvectors of the Jacobian matrix at u_1 to the Hugoniot locus. We can verify that the derivative $\lim_{\xi\to 0}\frac{\partial u_{2,+}}{\partial \xi}(\xi)$ is proportional to the positive right eigenvector and that $\lim_{\xi\to 0}\frac{\partial u_{2,-}}{\partial \xi}(\xi)$ is proportional to the left eigenvector. Likewise the velocities approach the eigenvalues, $\lim_{\xi\to 0} s_\pm(\xi) = \lambda_\pm$, as expected.

Note that not all solutions of the Rankine-Hugoniot condition may be physically relevant (this problem is related to entropy conditions and limits of equations with finite viscosity). Also there may not be a solution to the Riemann problem (loci may not intersect).

As is true for the linear case conditions on either side of a discontinuity may not lie along a locus. For the non-linear case, the locus is not a line but a curve. We must consider a series of discontinuities that connect conditions on either side u_l, u_r. For a two dimensional system, two discontinuities are required. Instead of connecting u_l, u_r with intersecting parallel lines (where lines are parallel to eigenvectors) we connect them with two intersecting curves. The slopes of these curves at u_l, u_r are eigenvectors of the Jacobian. As was true for the linear case, the path must first move along the locus with the slowest eigenvalue and then on the locus with the faster eigenvalue.

2.8. **Godunov's Method.** To update a cell value u_j^n to a new value u_{j+1}^n Godunov proposed solving two local Riemann problems, that for u_{j-1}^n, u_j^n and that for u_j^n, u_{j+1}^n. He proposed taking an average of the combined solutions of these two Riemann problems and constructing u_j^{n+1} from it.

What is meant by a *local* Riemann problem? We first consider a scalar conservation law

$$u_{,t} + f(u)_{,x} = 0 \tag{172}$$

The local Riemann problem for u_j^n and u_{j+1}^n is solving the equation for future times with an initial condition

$$u(x, t = 0) = \begin{cases} u_j^n & \text{for} \quad x > 0 \\ u_{j+1}^n & \quad\quad\quad x \geq 0 \end{cases} \tag{173}$$

The local Riemann problem is a Riemann problem at the intercell boundary.

In conservation law form

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n + \frac{\Delta t}{\Delta x}\left[\mathbf{F}_{j-\frac{1}{2}} - \mathbf{F}_{j+\frac{1}{2}}\right] \tag{174}$$

where $\mathbf{F}_{j+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{j+\frac{1}{2}})$ and $\mathbf{U}_{j+\frac{1}{2}}$ is the solution to the Riemann problem the intercell boundary between j and $j+1$, or $RP(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n)$.

Two situations $u_* < 0$, (negative speed in star region), $u_* > 0$ (positive speed in star region). For each of these two, there are 4 cases to consider. Left solution, right solution, star/shock solution, and fan/rarefraction wave solution.

2.9. **Roe's approximate Riemann solver.** One popular method is to approximate the non-linear system with a linear one. In this case the Riemann problems solved are those for the linear system. For the non-linear system

$$u_t + f(u)_x = 0 \tag{175}$$

at a particular grid location with neighboring values u_l, u_r, we solve the Reimann problem for the linear system

$$u_t + Au_x = 0 \tag{176}$$

where the matrix A is an approximation based on $f(u)$ and depends on u_l, u_r.

The matrix A is called the Roe matrix. Requirements for the matrix, A,

(1) **Hyperbolicity**. The matrix A has complete set of real eigenvalues and eigenvectors.
(2) **Consistency**.

$$\lim_{u_l \to u, u_r \to u} A(u_l, u_r) = f'(u) \tag{177}$$

(3) **Conservation**. Obeys the conservation law so that discontinuities will propagate at the appropriate velocity.

$$A(u_r - u_l) = f(u_l) - f(u_r). \tag{178}$$

A good choice of A would be one that is the Jacobian, $f'(\bar{u})$, but evaluated at a weighted average of u_l, u_r. For gas dynamics the third condition seems to be satisfied when $\bar{u}$ depends on the average of u_l, u_r where both are weighted by the square root of the density. Because a weighted average is used, the second condition is satisfied. Both Toro and LeVeque in their books show that the third condition is satisfied with this choice of A, but it not obvious how one chooses A for a general system. I think there is an extended theory by Roe and collaborators exploring how to choose A.

An additional problem is posed by rarefraction waves as linear systems don't exhibit them and this approximate solve approximates using a linear system. One approach to this issue is to modify the scheme to obtain entropy satisfying solutions.

More here xxx

2.9.1. *Notes.* For a non-linear problem all possible combinations of rarefraction and shock waves for each of the two local Riemann problems must be considered in constructing the Godunov fluxes, and averaging over them for the new timestep. In approximate solvers the local Riemann problems are approximated by linear systems and the procedure for constructing the Godunov fluxes is simpler.

Godunov's method matches shock speeds. Godunov originally used a first order upwind scheme which since it is only first order has quite a bit of numerical dissipation. This is why in many numerical textbooks the displayed Godunov solution appears very smooth and does not show a sharp jump across a discontinuity. There is a large body of literature developing higher order versions but with the same idea of relating the numerical scheme to local Riemann problems and using conservative methods.

3. Acknowledgments

Following *Numerical Methods in Astrophysics* by Bodenheimer et al. 2007, *Computational Aerodynamics and Fluid Dynamics* by J.J. Chattot, *Numerical Methods for Conservation Laws* by Randall J. LeVeque, and *Riemann Solvers and Numerical Methods for Fluid Dynamics, A practical introduction* by Eleuterio F. Toro.

Two Figures taken From Toro's very clear book and four other figures taken from LeVeque's book.