Chapter 4. Nonlinear Hyperbolic Problems

1. Introduction

Reading: Durran sections 3.5-3.6, Mesinger and Arakawa (1976) Chapter 3 sections 6-7. Supplementary reading: Tannehill et al sections 4.4 and 4.5 – Inviscid and viscous Burgers equations.

Nonlinear problems creates two important problems in CFD:

1. They generate nonlinearity instability.
2. New waves can be generated in nonlinear problems via nonlinear wave interaction.

The stability analysis we discussed in the previous Chapter refers to linear stability and linear instability – because they do not require nonlinearity in the equation.

The above two issues are specific to nonlinear equations.

Many processes in the atmosphere can be nonlinear – many physical processes, such as phase changes are nonlinear. In the Navier-Stokes equations, the most significant nonlinear term is the advection term.

The simplest equation including nonlinear advection is the Burgers Equation:

\[
\text{Inviscid: } \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0 \quad \text{(Hyperbolic)} \tag{1a}
\]

\[
\text{Viscous: } \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2} \quad \text{(Parabolic)} \tag{1b}
\]

We can rewrite the advection term \( u \frac{\partial u}{\partial x} = \frac{\partial}{\partial x} \left( \frac{u^2}{2} \right) \) - the nonlinearity is often called quadratic nonlinearity.

There is a fundamental difference between the inviscid Burger's equation (1a) and the linear advection equation, \( \frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \), we discussed in last chapter, where \( c \) is a constant.
1. In the linear problem, all points on the wave move at the same speed, \( c \), the shape of the wave remain unchanged:

For a nonlinear equation (1b), the wave \textit{advects itself} such the local speed depends on the wave amplitude and the shape of the wave change in time:

The process is called \textit{nonlinear steepening}, and eventually results in \textit{shock waves} and \textit{overturning} if the flow is inviscid. In this case, the characteristics coalesce into a group where multiple values of \( u \) exist for a given \( x \).

2. Nonlinear problems \textit{create new waves modes}. This was evident in the previous problem where we \textit{start with a single sine wave} and ended up with a \textit{step-like function}. Clearly the step function can be not represented by a single wave \( \rightarrow \) new waves have been generated! For nonlinear problems, the \textit{principle of superposition does not apply}!
To illustrate this, consider \( u = \sin(kx) \). Plug it into the advection term

\[
\frac{\partial u}{\partial x} = k \sin(kx) \cos(kx) = k \sin(2kx) / 2.
\]

Now the system contains a new wave – \( \sin(2kx) \), whose wave number is \( 2k \), and wavelength is \( L = \pi/k \), half of the wave length of the original \( 2\pi/k \).

The new wave can interact with itself and the original one, the process goes on and on and a entire spectrum of waves will result! This process is the source of aliasing error, to be discussed soon.

Despite of its nonlinear, Burger's equation has analytical solutions.

For the inviscid case, one of the examples is:

if \( u(x, t=0) = -U \tanh(kx) \)

then \( u(x, t) = -U \tanh[k(x-ut)] \).

Note that the solution is an implicit function of \( u \), and it has to be solved iteratively for the value of \( u \).

For the viscous case, an example is:

if \( u(x, t=0) = -U \tanh(kx) \)

then the steady state solution is

\[
u(x, t) = -U \tanh(ux/2\nu) .
\]

Here, dissipation of energy within the shock is exactly balanced by the conversion of kinetic energy from infinity.

References for exact solutions:


Solution techniques

Many solution techniques discussed earlier for linear advection equation can be used for Burgers equation. We will not discuss them in details here, but we will look the behavior of the solutions:
Sample Solutions to the Inviscid Burgers' Equation

Figure 4-27 Numerical solution of Burgers' equation using Lax method.

Figure 4-26 Solution for right-moving discontinuity time-centered implicit method, delta form.

Figure 4-25 Solution of Burgers' equation using Beam-Warming (trapezoidal) method.
Because of the nonlinear steepening, the solution contains sharp gradient near the step – numerical schemes tend to perform poorly near sharp gradient, and most schemes, especially high order ones, generates small scale oscillations near the sharp gradient – monotonic schemes are particularly good at dealing with sharp gradient, because they are designed to prevent overshoot and undershoot from being generated.

With conventional schemes, there is a tendency for the small-scale noises to grow quickly and eventually destroy the solution or cause instability. Such instability occurs only in nonlinear problems, and was first discussed by the developer of NGM, Norman Phillips (1959), and the instability is called Nonlinear Instability.

2. Nonlinear Instability

Linear instability occurs when the linear stability criteria is violated, usually when $\Delta t$ is too large.

Nonlinear instability occurs when waves shorter than $2\Delta x$ are generated and feed energy spuriously into the wavelengths near but larger than $2\Delta x$. The energy buildup becomes catastrophic.

The generation of waves with wavelength $< 2\Delta x$ is a consequence of aliasing (c.f., p.35-42. Mesinger and Arakawa 1976. Read it).

**Aliasing:**

Consider a function $u = \sin(kx)$.

We know that the shortest wave that can be represented by a grid has a wavelength of $2\Delta x$, and a wave number of $k_{\text{max}} = 2\pi/(2\Delta x) = \pi/\Delta x \rightarrow$ the largest wave number is $k_{\text{max}} = \pi/\Delta x$.

We saw earlier for the nonlinear advection term $u \frac{\partial u}{\partial x}$

$$
\frac{du}{dx} = k \sin(kx) \cos(kx) = k \sin(2kx) / 2 .
$$

If $k = k_{\text{max}}$, then the new wave has a wave number of $2k_{\text{max}}$, corresponding to a wavelength of $(2\Delta x)/2 = \Delta x$ - too short to be represented on the grid!

Therefore, nonlinear interaction between waves can generated waves that are unresolvable by the original grid!

Then what happens to these unresolvable waves? They are spuriously presented as, or aliased as, resolvable waves!
Consider a wave with W.L. = 4/3Δx (<2Δx). With only three grid points to represent one wavelength, it cannot tell it apart from the 4Δx wave. In fact, the grid mis-represents it as 4Δx wave!

![Wave diagram](image)

**Figure 6.1** A wave of wave length 4Δx/3, misrepresented by the finite difference grid as a wave of wave length 4Δx.

Consider now a general case of a function $u$ that contains harmonic components:

$$u = \sum_n u_n$$

nonlinear term will be of the form

$$\sin(k_1x) \sin(k_2x) = \frac{\cos(k_1-k_2)x - \cos(k_1+k_2)x}{2}$$

→ two new waves, $k_1 \pm k_2$, are created!

Even if the calculation is started with all wavelengths $\geq 2\Delta x$, waves $< 2\Delta x$ will be generated, through nonlinear interaction.
To generalize, let’s write
\[
\cos(kx) = \cos[2k_{\text{max}} - (2k_{\text{max}} - k)]x_i \\
= \cos \left( \frac{2\pi x_i}{\Delta x} \right) \cos \left( \frac{2\pi}{\Delta x} - k \right) x_i + \sin \left( \frac{2\pi x_i}{\Delta x} \right) \sin \left( \frac{2\pi}{\Delta x} - k \right) x_i.
\]

Since \( x_i = i \Delta x \), and \( i \) is integer,
\[
\sin \left( \frac{2\pi x_i}{\Delta x} \right) = \sin \left( \frac{2\pi i \Delta x}{\Delta x} \right) = 0, \quad \cos \left( \frac{2\pi x_i}{\Delta x} \right) = 1 \Rightarrow
\]
\[
\cos(kx) = \cos[2k_{\text{max}} - k]x \Rightarrow
\]

Knowing only those values at the grid points, we cannot distinguish between wavenumber \( k \) and \( 2k_{\text{max}} - k \), thus, if \( k > k_{\text{max}} \) (W.L. < \( 2\Delta x \)), then \( k \) is really misrepresented as (or aliased as)
\[
k^* = 2k_{\text{max}} - k.
\]

Thus, the aliased wave \( k^* \) is less than \( k_{\text{max}} \) by an amount equal to that by which \( k \) was greater than \( k_{\text{max}} \):

```
0   k*   k_{\text{max}}   k   2k_{\text{max}}
```

Figure 6.2 Misrepresentation of a wave number \( k > k_{\text{max}} \) in accordance with (6.4).

Back to our example, let W.L. = \( 4/3 \Delta x \), this is aliased as
\[
k^* = 2\pi/(2\Delta x) - 2\pi/(4/3\Delta x) = 2\pi/(4\Delta x) \Rightarrow 4\Delta x \text{ wave}
\]

– the same as we saw earlier by the graphic means.

Note that the waves generated by aliasing are always near \( 2\Delta x \) – energy start to pile up in the form of short wave noises. In the next section, we will look at ways to control such pileup.
3. Controlling Nonlinear Instability

3.1. Consequences of N.L. Instability

If a flow contains many modes, it is useful to examine the distribution of energy (a measure of the amplitude of the modes) as a function of wavenumber:

\[ E = \frac{\sum \mu_k^2}{2} \]

In a numerical simulation, aliasing occurs near \( 2\Delta x \rightarrow \) energy is shifted to small scales and the short waves grow with time \( \rightarrow \) nonlinear instability.

3.2. Filter Method

Phillips (1959) showed that catastrophic growth of wave disturbances can be prevented in a 2-level geostrophic model, by periodically applying a spectral filter, which eliminates waves shorter than or equal to \( 4\Delta x \).

The method decomposes the solution into Fourier modes (waves / harmonics), and recomposes them without hence eliminating the shortest waves.

![Graph](image)

**Fig. 1.** Disturbance kinetic energy as a function of time. The solid curve was obtained without smoothing, the computations breaking down at about 56 days. The dashed curve was obtained by periodically introducing a filtering procedure.
Orszag (1971) later showed that it is sufficient to eliminate only waves equal to or shorter than $3\Delta x$ (see hand out).

The use of spectral filter is very expensive in grid point model. Doing it in spectral models is straightforward, however, since the solution is already in the spectral form.

### 3.3. Spatial Smoothing or Damping

In this case, we apply, at chosen intervals (often every time step), a spatial smoother similar in form to the term in our parabolic diffusion equation.

We want the smoothing to be **selective**, so that only the short (aliased) waves get damped.

Filter types:

- **Low-pass:** allows low-frequency or long wavelength waves to pass through
- **High-pass:** allows high-frequency or short wavelength waves to pass through
- **Band-pass:** allows intermediate waves to pass through

What is desired here:
There exists many types of filters. Let look at one that this commonly used, the 1-2-1 or Shapiro filter:

$$\bar{u}_j = u_j + 0.5v(u_{j+1} - 2u_j + u_{j-1})$$

(2)

where the $\bar{u}_j$ is the value after smoothing.

To see what the smoother does, we need to look at the response function $\sigma$ defined by

$$\bar{u} = \sigma u .$$

- all a filter does is changing the wave amplitude (a well-designed filter should not change the phase). Here $\sigma$ might be a function of $k, \Delta x, v$ etc., much like $|\lambda|$ earlier.

The method for obtaining $\sigma$ is very similar to the method for von Neumann stability analysis.

Let $u = A \exp(ikx_j)$. Plug it into (2) $\Rightarrow$

$$\bar{u} = [1 - v(1 - \cos(k\Delta x))]A \exp(ikx_j) \rightarrow$$

$$\sigma = [1 - v(1 - \cos(k\Delta x))] \quad -- \text{response function of filter (2).}$$

See Figure.

One can create multi-dimensional smoothers by successive applications of 1-D smoothers, one can also design fully MD ones.
3.4. Smoothing via numerical diffusion

This method damps the aliased waves by adding a smoothing or diffusion term to the prognostic equations (called computational mixing term in the ARPS – which also helps to suppress small scale noises created by dispersion and physical processes. Actually, ARPS uses advective formulations that conserve the advected quantities and their variances – therefore nonlinear instability due to aliasing is reasonably controlled even without the smoothing).

Consider for example of the CTCS case for linear advection:

\[
\begin{pmatrix}
-u \\
-\delta_{xx}u \\
-\delta_{xxxx}u \\
\delta_{xxxxx}u
\end{pmatrix} = \alpha
\begin{pmatrix}
-u \\
-\delta_{xx}u \\
-\delta_{xxxx}u \\
\delta_{xxxxx}u
\end{pmatrix}^{n-1}
\]

(3)

the right RHS terms are called zero, 2nd, 4th and 6th order numerical diffusion / smoothing, respectively. Note that the diffusion term is evaluated at time level n-1 – this makes the time integration forward in time relative to this term – remember that forward-in-time is (conditionally) stable for diffusion term but centered-in-time scheme is absolutely unstable.

We can find the response function to be

\[
|\lambda|^2 = 1 - 2\alpha\Delta t + \frac{2 - 2\cos(k\Delta x)}{\Delta x^2} + \frac{6 - 8\cos(k\Delta x) + 2\cos(2k\Delta x)}{\Delta x^4} + \frac{20 - 30\cos(k\Delta x) + 12\cos(2k\Delta x) - 2\cos(3k\Delta x)}{\Delta x^6}
\]

(4)

and they are plotted in the following figure. We can see that this term selectively damps shorter waves, and the higher order schemes are more selective, which is desirable. Given \(|\lambda|\), you can estimate the amplitude change due the diffusion for different wavelength after given number of time steps.
\[ (1 - \text{parenthesized terms in ( )}) \]

normalized for complete damping at \( RAX = \pi \)
3.5. Lagrangian or Semi-Lagrangian Formulation

The cause of nonlinear instability is the nonlinear advection term in momentum equations. If we can get rid of this term, we can eliminate the instability!

This can be achieved by solving the advection problem in a Lagrangian or Semi-Lagrangian framework.

With Lagrangian methods, the pure advection problem is

\[
\frac{du}{dt} = 0
\]

i.e., \( u \) is conserved along the trajectory, which is also the characteristic curve \( (dx/dt=c) \) in this case.

In the purely Lagrangian method, the grid points move with the flow, and the grid can become severely deformed.

Semi-Lagrangian method is based on a regular grid – it finds the solution at grid points by finding the values of \( u \) at the *departure points* – the location where the parcels come from. Spatial interpolation is usually needed to find the value at the departure point. We will cover this topic in more details later.
3.6. Use of conservation to control nonlinear instability

Recall that aliasing acts to feed energy into small-scale components. It is possible to control (not prevent) aliasing by forcing the total energy or other physical properties (e.g., enstrophy – squared voroticity) to be conserved – just as the continuous system does.

If such constraints are satisfied, the energy spectrum cannot grow without bound!

Consider 2-D advection in a non-divergent flow:

\[ \frac{\partial A}{\partial t} + u \frac{\partial A}{\partial x} + v \frac{\partial A}{\partial y} = 0 \]  
\[ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \]

We can write the advection-form equation (5a) in a flux-divergence form:

\[ \frac{\partial A}{\partial t} + \frac{\partial (uA)}{\partial x} + \frac{\partial (vA)}{\partial y} = 0 \]

What is conserved for this system of equations?

We first the domain integration of the first moment of A to be:

\[ \frac{\partial}{\partial t} \left[ \int \int A \, dx \, dy \right] = -\int \int \frac{\partial (uA)}{\partial x} \, dx \, dy - \int \int \frac{\partial (vA)}{\partial y} \, dx \, dy \]

\[ -\int [(uA)_L - (uA)_R] \, dy - \int [(vA)_T - (vA)_B] \, dx = 0 \]

for a periodic domain. For non-periodic domain, we can see that the change in the domain integration of A equals to the net flux through the lateral boundaries – there is no interior source or sink in A.

We say the domain integral of the first moment of A is conserved by this system of equations.

Let’s now look at the conservation of the second moment of A, i.e., $A^2$:

Multiply (6) by $A$ →
\[
A \frac{\partial A}{\partial t} + A \frac{\partial (uA)}{\partial x} + A \frac{\partial (vA)}{\partial y} = 0 \to \\
\frac{\partial}{\partial t} \left( \frac{A^2}{2} \right) + \frac{\partial (uA^2)}{\partial x} + \frac{\partial (vA^2)}{\partial y} - uA \frac{\partial A}{\partial x} - vA \frac{\partial A}{\partial y} = 0
\] (8)

Multiply (5b) by \( A \) \( \to \)

\[
\frac{\partial}{\partial t} \left( \frac{A^2}{2} \right) + uA \frac{\partial A}{\partial x} + vA \frac{\partial A}{\partial y} = 0
\] (9)

(8) + (9) \( \to \)

\[
\frac{\partial A^2}{\partial t} + \frac{\partial (uA^2)}{\partial x} + \frac{\partial (vA^2)}{\partial y} = 0
\] (10)

we have a conservation equation for \( A^2 \) in the flux divergence form too!

For a periodic domain, we have

\[
\frac{\partial}{\partial t} \left[ \iint A^2 dx dy \right] = 0
\]

therefore the second-moment of \( A \) is also conserved by the continuous system.

What about the discrete equations? Do they also conserve these quantities? Not all discrete forms do. We will show one conservative example in the following.

Conservation for the Discrete System

Consider the case of staggered Arakawa C-grid:
and the following second-order FD formulation:

\[ \delta_x A + \delta_y (u\vec{A}^x) + \delta_y (v\vec{A}^y) = 0 \]  \hspace{1cm} (11a) \]

\[ \delta_x u + \delta_y v = 0 \]  \hspace{1cm} (11b) \]

Q: Does this system conserve \( A \) and \( A^2 \)?

Note that \( \vec{A}^x \) is defined at \( u \) point and \( \vec{A}^y \) at \( v \) point, we denote \( \vec{A}^x = B \) and \( \vec{A}^y = C \) →

\[ \sum_{i,j} \delta_x (u\vec{A}^x) = (u_i B_1 - u_o B_0) / \Delta x + (u_2 B_2 - u_1 B_1) / \Delta x + \ldots \]
\[ + (u_{N-1} B_{N-1} - u_{N-2} B_{N-2}) / \Delta x + (u_N B_N - u_{N-1} B_{N-1}) / \Delta x = 0 \]

with periodic B.C.

The same is true to the flux in \( y \) direction – therefore \( A \) is conserved.

Conservation of \( A^2 \) is a little more complicated to show. We will make use of two identities (you can check them out for yourself):

\[ \delta_x (\vec{P}^x Q) = P \delta_x Q + (\delta_x P) \vec{Q}^x \]  \hspace{1cm} (12a) \]

\[ \vec{P}^x \delta_x P = \delta_x (P^2 / 2) \]  \hspace{1cm} (12b) \]

Multiply (11a) by \( A \) →

\[ A \delta_x A = -A \delta_x (u\vec{A}^x) - A \delta_y (v\vec{A}^y) \]  \hspace{1cm} (13) \]

Look at only the 1st term on RHS of (13):

\[ -A \delta_x (u\vec{A}^x) \]

Let \( P = A, \) \( Q = u\vec{A}^x, \) (12a) becomes

\[ \delta_x (\vec{A}^x u\vec{A}^x) = A \delta_x (u\vec{A}^x) + (\delta_x A)(u\vec{A}^x) \rightarrow \]
\[ A \delta_x (u\vec{A}^x) = \delta_x (\vec{A}^x u\vec{A}^x) - u \delta_x (A^2 / 2) \]  \hspace{1cm} (14) \]
The 2nd term on RHS of (14) is not in the flux form. Using (12a) again, let $P = \frac{A^2}{2}, Q = u$

\[ \delta_x (\frac{A^2}{2} u) = (\frac{A^2}{2}) \delta_x u + \delta_x (\frac{A^2}{2} u) \]

therefore

\[ u \delta_x (\frac{A^2}{2} u) = \delta_x (\frac{A^2}{2} u) - (\frac{A^2}{2}) \delta_x u \]  

(15)

Now (14) becomes

\[ A \delta_x (u \overline{A}^x) = \delta_x (\overline{A}^x u \overline{A}^x) - \delta_x (\frac{A^2}{2} u) + (\frac{A^2}{2}) \delta_x u \]  

(16)

- only the last term is not in the flux form.

For the y direction, we can also get

\[ A \delta_y (v \overline{A}^y) = \delta_y (\overline{A}^y v \overline{A}^y) - \delta_y (\frac{A^2}{2} v) + (\frac{A^2}{2}) \delta_y v \]  

(17)

\[(16) + (17) \Rightarrow \]

\[ A \delta_x (u \overline{A}^x) + A \delta_y (v \overline{A}^y) = .... + (\frac{A^2}{2}) (\delta_x u + \delta_y v) \]

the last term is zero because of (11b)!

Therefore

\[ \sum_{ij} A \delta_{2i} A = 0 \Rightarrow \sum_{ij} (A^n A^{n+1} - A^{n-1} A) = 0 \Rightarrow \]

\[ \sum_{ij} (A^n A^{n+1}) = \sum_{ij} (A^{n-1} A^n) \]

$A^n A^{n+1}$ is not exactly $A^2$ due to the temporal discretization – we say $A^2$ is quasi-conserved!

Comments on Conservations:

• Conservation is generally a good thing to have in a model – can be used to check the correctness of code – if you know your scheme conserves, check if the domain integral changes in time.
• Don't want to use schemes that are known to conserve poorly.
• It is not always possible to conserve all conservative quantities of the continuous system, however.
Nonlinear advection schemes that conserve more quantities

Arakawa derived and compared several methods for dealing with the nonlinear advection of a barotropic vorticity equation

\[ \frac{\partial \zeta}{\partial t} + u \frac{\partial \zeta}{\partial x} + v \frac{\partial \zeta}{\partial y} = 0 \]  

(18)

with \( \zeta = \nabla^2 \psi \); \( u = \frac{\partial \psi}{\partial y} \); \( v = \frac{\partial \psi}{\partial x} \)

where \( \zeta \) is vorticity and \( \psi \) is the streamfunction.

(18) can be rewritten as

\[ \frac{\partial \zeta}{\partial t} + \frac{\partial \psi}{\partial y} \frac{\partial \zeta}{\partial x} + \frac{\partial \psi}{\partial x} \frac{\partial \zeta}{\partial y} = 0 \]  

(19)

where the advection can be written as a Jacobian:

\[ J(\psi, \zeta) = \frac{\partial \psi}{\partial x} \frac{\partial \zeta}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial \zeta}{\partial x} \]  

(20)

Arakawa came up with seven different forms of discretization for the Jacobian (called Arakawa Jacobians), some conserve total energy and enstrophy (\( \zeta^2 \)).

The following figures shows the kinetic energy spectrum, and the total kinetic energy and enstrophy as function of time, using different Jacobians:
Fig. 12. Comparison of the time variation of the mean square vorticity (units arbitrary) during a numerical integration with the seven finite-difference Jacobians under consideration. (Arakawa, 1970). Reprinted with permission of the publisher American Mathematical Society from SIAM-AMS Proceedings. Copyright © 1970, Vol. 2, Fig. 5, p. 35.
FIG. 13. Comparison of the time variation of the kinetic energy during a numerical integration with the seven finite-difference Jacobians under consideration (Arakawa, 1970). Reprinted with permission of the publisher American Mathematical Society from SIAM-AMS Proceedings. Copyright © 1970, Vol. 2, Fig. 6, p. 36.
Read Mesinger and Arakawa (1976) GARP Report, section 7 (handout).
Review before second exam (March 20)

Chapter 2, Part II

- Stability analysis – several methods
- von Neumann stability analysis
- Explicit and Implicit methods for heat transfer equation and their stability properties
- Multi-dimensional methods for solving heat transfer equation - Direct extension, directional splitting, fully MD

Chapter 3. Hyperbolic equations

- CFL stability condition
- Computational domain of dependence and its relevance to computational stability
- Phase and amplitude errors for advection schemes
  - Modified equation
  - Definition of errors
  - Derivation of errors
- Computational modes of multi-time level schemes
- Methods for suppressing computational modes
- Comparison of phase and amplitude accuracy of several common schemes
- Practical measure of dissipation and dispersion errors
- Concept of monotonicity
- Methods for multi-dimensional advection and their stability properties

Chapter 4. Nonlinear Hyperbolic equations

- Aliasing, nonlinear instability – their origin
- Methods for controlling nonlinear instability

Ref: Chapter IV of Mesinger and Arakawa (1976) GARP Report

4.1. Introduction

It is assumed that you are familiar with the shallow water equations and associated theories. If not, consult Holton or Haltiner and Williams.

The following is a set of linear 1D shallow water equations:

\[
\frac{\partial u'}{\partial t} + \overline{u} \frac{\partial u'}{\partial x} + \frac{\partial \phi'}{\partial x} = 0 \quad (21a)
\]

\[
\frac{\partial \phi'}{\partial t} + \overline{u} \frac{\partial \phi'}{\partial x} + \Phi \frac{\partial u'}{\partial x} = 0 \quad (21b)
\]

- \( \overline{u} \) = constant base state flow
- \( \Phi = gH \) = g * mean depth of the water = constant
- \( u \to u' \) = perturbation velocity
- \( \phi = gh' \) = perturbation geopotential height

Issues to consider with respect to numerical solution

1) More than 1 variable
2) Equations coupled
3) Can support multiple physical modes
4) There are more possibilities of grid layout (see figure below)

![Diagram](image)

Figure 2.1 Three types of lattice considered for the finite difference solution of 2.1.
4.2. The differential solution

Performing standard analysis by assuming

\[ \psi = \Psi \exp[i(kx - \omega t)] \]  \hfill (22)

gives

\[ \omega = k(\bar{u} \pm \sqrt{\Phi}) \]  \hfill (23)

which is called the dispersion relation.

From (23) \( \rightarrow \)

\[ c = \frac{\omega}{k} = \bar{u} + \sqrt{\Phi}. \]

In the phase speed, there are slow mode represented by \( \bar{u} \) (advection) and fast mode given by \( \sqrt{\Phi} \) (surface gravity waves). Since \( c \) is constant, the waves are non-dispersive.

Group velocity \( c = \frac{\partial \omega}{\partial k} = \bar{u} + \sqrt{\Phi} \)

represents the speed of wave energy propagation.

What about the characteristics (we have seen this before – see example problem given at the end of Chapter 1). Make use of the auxiliary equations, we have the following equations in matrix form:

\[
\begin{bmatrix}
1 & \bar{u} & 0 & 1 \\
0 & \Phi & 1 & \bar{u} \\
dt & dx & 0 & 0 \\
0 & 0 & dt & dx
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
u_t \\
u_x \\
\phi_t \\
\phi_x
\end{bmatrix}
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
\frac{du}{dt} \\
\frac{d\phi}{dt}
\end{bmatrix}
\] \hfill (24)

Setting the determinant of the coefficient matrix to zero gives

\[
\left( \frac{dx}{dt} \right)^2 - 2\bar{u} \frac{dx}{dt} + (\bar{u}^2 - \Phi) = 0 \rightarrow
\]

\[ \frac{dx}{dt} = \bar{u} \pm \sqrt{\Phi} \]

which is the characteristics equations.
The compatibility equations can be found to be

\[ u \pm \phi = \text{constant along } \frac{dx}{dt} = \bar{u} \pm \sqrt{\Phi}. \]  

(25)

(25) can be rewritten as

\[
\begin{align*}
\left[ \frac{\partial}{\partial t} + (\bar{u} + \sqrt{\Phi}) \frac{\partial}{\partial x} \right] \left( \bar{u} + \phi \frac{\partial}{\partial \sqrt{\Phi}} \right) &= 0 \\
\left[ \frac{\partial}{\partial t} + (\bar{u} - \sqrt{\Phi}) \frac{\partial}{\partial x} \right] \left( \bar{u} - \phi \frac{\partial}{\partial \sqrt{\Phi}} \right) &= 0
\end{align*}
\]

(26a, 26b)

which are two decoupled equations describing wave disturbances 'adveected' by the respective propagation speeds. \( \bar{u} \pm \phi / \sqrt{\Phi} \) are known as the Riemann invariants, as said before.

Equations (26) can also be obtained using matrix method (see example problem solution given at the end of Chapter 1).

4.3. Discretization for the Shallow Water Equations

4.3.1. Forward-backward scheme

We know that FTCS is unstable for pure advection equations, and this is also true to the shallow water equations.

But, we can obtain a stable scheme if we use backward scheme for the second equation. Let’s look at the simple case of \( \bar{u} = 0 \), i.e., there is not mean flow:

\[
\begin{align*}
\delta_{x,t} u + \delta_{2,x} \phi^n &= 0 \\
\delta_{x,t} \phi + \Phi \delta_{2,x} u^{n+1} &= 0
\end{align*}
\]

Since forward scheme is used for the first eq. and backward scheme used for the second, the overall scheme is called forward-backward scheme. We can show that it is conditionally stable.
Stability Analysis

Assume that

\[ u^n_j = A_\lambda^n \exp(i k x_j) \]
\[ \phi^n_j = B_\lambda^n \exp(i k x_j) \]  

(28)

Note here A and B could be complex so as to account for possible phase difference between u and \( \phi \).

Plug (28) into (27) \( \rightarrow \)

\[ (\lambda^{n+1} - \lambda^n) A + B_\lambda^n \frac{\Delta t}{2 \Delta x} (e^{i \Delta x} - e^{-i \Delta x}) = 0 \]
\[ (\lambda^{n+1} - \lambda^n) B + \Phi A_\lambda^n \frac{\Delta t}{2 \Delta x} (e^{i \Delta x} - e^{-i \Delta x}) = 0 \]

(29)

or

\[ (\lambda - 1) A + i B_\lambda \frac{\Delta t}{\Delta x} \sin(k \Delta x) = 0 \]
\[ (\lambda - 1) B + i \Phi A_\lambda \frac{\Delta t}{\Delta x} \sin(k \Delta x) = 0 \]

(30)

or

\[ \begin{pmatrix} \lambda - 1 & i \frac{\Delta t}{\Delta x} \sin(k \Delta x) \\ i \Phi \frac{\Delta t}{\Delta x} \lambda \sin(k \Delta x) & \lambda - 1 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \]  

(30')

(30') is a simultaneous linear system of equations for A and B. It has non-trivial solutions if and only if the determinant of the coefficient matrix equals to zero. \( \rightarrow \)

\[ \lambda^2 - \lambda[2 - \Phi a^2] + 1 = 0 \quad [\text{where } a = \Delta t / \Delta x \sin(k \Delta x)] \]

\[ \lambda_{\pm} = \frac{\Phi a^2 - 2 \pm \sqrt{(2 - \Phi a^2)^2 - 4}}{2} \]  

(31)

If the radical is negative, then \( |\lambda_{\pm}| = 1 \). I.e., if

\[ (2 - \Phi a^2)^2 \leq 4 \]
\[ |2 - \Phi a^2| \leq 2 \]
\[ \Phi a^2 \leq 4 \rightarrow \]

\[ \Delta t \leq \frac{2\Delta x}{\sqrt[2]{\Phi \sin(k\Delta x)}} \rightarrow \]

\[ \Delta t \leq \frac{2\Delta x}{\sqrt[2]{\Phi}} \]  \hspace{1cm} (32)

which is the stability condition! Here \( \sqrt[2]{\Phi} \) is the disturbance propagation speed in the absence of base-state advective flow. When the mean flow is non-zero, the condition is

\[ \Delta t \leq \frac{2\Delta x}{|\bar{u} \pm \sqrt[2]{\Phi}|}. \]

Note the factor of 2 in the condition – the use of forward-backward scheme actually allows a Courant number of 2 to be used! This is due to the fact the backward scheme is actually a kind of 'implicit' scheme.

**4.3.2. CTCS scheme**

\[ \delta_{2x} u + \bar{u} \delta_{2x} u + \delta_{2x} \phi = 0 \]
\[ \delta_{2x} \phi + \bar{u} \delta_{2x} \phi + \Phi \delta_{2x} u = 0 \]

(here we assume a non-staggered grid)

Similar stability analysis leads to:

\[ \Delta t \leq \frac{\Delta x}{|\bar{u} \pm \sqrt{\Phi}|} \]  \hspace{1cm} (33)

which is twice as restrictive as that for forward-backward scheme. Also it contains a computational mode.

**Grid Splitting**

When using non-staggered grid for the above equations, we can also run into the grid-splitting problem. We discussed this issue in the past.
One way of avoiding grid splitting is to use staggered grid – in which different variables are located as different points of a grid mesh.

Let's stagger u and φ (h in the figure) in the following way:

![Staggered grid](image)

**Figure 1.2** A grid with two dependent variables that are carried at alternate grid points.

Our FD equation using CTCS scheme is then

\[
\delta_{x}u + \bar{u}\delta_{x}u + \delta_{x}\phi = 0 \quad \text{at u point}
\]

\[
\delta_{x}\phi + \bar{u}\delta_{x}\phi + \Phi\delta_{x}u = 0 \quad \text{at φ point}
\]

(34)

Note the key difference in the third term of each equation from the previous non-staggered CTCS scheme. Also the equations are solved at different grid point.

Stability analysis will show the stability condition is

\[
\Delta t \leq \frac{\Delta x}{|\bar{u} \pm 2\sqrt{\Phi}|}
\]

which, for zero mean flow case, is twice as restrictive as the non-staggered version.

However, since the pressure gradient force and velocity divergence terms are differenced over one Δx interval, and these are terms responsible for the gravity wave propagation, the solution should be more accurate, since the effective grid spacing is half as large.

**4.3.3. Treatment of insignificant fast modes**

(Reading: Durran Chapter 7 – Physically insignificant fast waves)

We obtained earlier the phase speed of shallow water waves:
\[ c = \bar{u} + \sqrt{gH} \]

it contains two modes. The slower advective mode and the faster gravity wave mode:

\[ \bar{u} \sim 10 \text{m/s} \]

\[ \sqrt{gH} \sim \sqrt{10 \times 10000} \sim 200 \text{m/s} \text{ for external gravity waves} \]

\[ |\bar{u}| \ll \sqrt{gH} \text{ for many problems.} \]

Gravity waves are not important in global models in which the resolutions are usually too coarse to resolve them adequately anyway.

GW are often important for mesoscale flows. For mesoscale models, often compressible equations are used which support fast sound waves – so sound wave play a similar role as the gravity waves in large scale model in limiting the time step size (when using explicit schemes).

When the fast mode is not important, we don’t want it to be the one that limits the time step size.

There are in general two ways to deal with this problem – one is to treat the terms responsible for the fast modes implicitly, and the other uses different time step sizes for fast and slow modes and the method is called mode splitting method. ARPS uses the latter to deal with fast sound waves (hence the large and small time steps).

4.4.4. Semi-implicit method

We will look at the first one here. Since the PGF term in u equation and the velocity divergence term in \( \phi \) equation are responsible for gravity waves, we treat them implicitly, using time average.

Again we look at the non-staggered case:

\[ \delta_{2t} u + \bar{u} \delta_{s\tau} u + \delta_{s\tau} \phi = 0 \]
\[ \delta_{2t} \phi + \bar{u} \delta_{s\tau} \phi + \Phi \delta_{s\tau} u = 0 \] (35)

The time averages makes the scheme implicit. Since only some of the terms are treatly implicitly, the scheme is called semi-implicit.

Stability of the system – only the advective velocity \( \bar{u} \) appears in the stability condition therefore much larger time step can be used (see Durran 7.2.3).
Analysis shows that the fast mode in the numerical solution is actually slowed down – i.e., there is a lagging phase error with this mode – it is okay if this mode is considered unimportant, like the sounding waves in atmospheric flows.

Solution procedure for (35)

1) Computer $\phi_{n+1}^j$ for all $j$ by eliminating $u_{n+1}^j$ from the 2nd equation using the first:

$$\phi_{j}^{n+1} - \frac{\Phi \Delta t^2}{4\Delta x^2} \left[ \phi_{j-2}^{n+1} - 2\phi_{j}^{n+1} + \phi_{j+2}^{n+1} \right] = f,$$

the right hand side is known.

2) Two effectively decoupled tridiagonal system of equations have to be solved, one for even $j$ and one for odd $j$ (can lead to grid splitting).

3) Once $\phi_{n+1}^j$ is known, we can plug it into $u$ equation to obtain $u_{n+1}^j$.

4) If a staggered grid is used, then only one tridiagonal system of equations has to be solved. The total amount of calculation is about the same as the non-staggered case since there the number of equations is halved.

5) For 2D or 3D problems, the semi-implicit scheme results in a Helmholtz equation that can't be as easily solved as the 1D tridiagonal equation.

Tapp and White is one of the first to use semi-implicit method in the compressible UK Met Office mesoscale model (Tapp and White 1976 QJRMS).

4.4.5. Mode-splitting Method

For info on mode-splitting method for compressible model, see Klemp and Wilhemson (1978) and Durran Section 7.3.2.

Skamarock and Klemp (1982) discusses that stability issues associated with the mode-splitting methods as applied to compressible system of equations.


4.4. The Arakawa Grids

(p.47 in Mesinger and Arakawa 1976)

Arakawa (Arakawa and Lamb 1977) introduced a variety of staggered grids in trying to find an accurate method for handling geostrophic adjustment process, which we know relies on inertia gravity waves. Inertia gravity waves are dispersive, they disperse ageostrophic energy.

To describe inertia GW, we need to include rotational effect into the shallow water equations:

\[
\frac{\partial u}{\partial t} + g \frac{\partial h}{\partial x} - f v = 0 \\
\frac{\partial v}{\partial t} + g \frac{\partial h}{\partial y} + f u = 0 \\
\frac{\partial h}{\partial t} + H \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0
\]

For 1-D version of this problem, i.e. for \( \frac{\partial}{\partial y} = 0 \) case, the dispersion equation for the exact solution is

\[
\omega = \left( f^2 + k^2 g H \right)^{1/2}.
\]

Arakawa defined 5 different grids, all of which has the same number of dependent variables per unit area – so that the computational time is about the same.
Figure 3.1 Five types of lattice considered for the finite difference solution of (3.1).
For each of the above grid, the FD equation can be written as

\[ \frac{\partial u}{\partial t} = -g \frac{\partial v}{\partial x} + f v, \quad \frac{\partial v}{\partial t} = -g \frac{\partial u}{\partial y} - f u, \]  
\[ \frac{\partial h}{\partial t} = -H \left( \frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y} \right), \]  
\[ (3.2)_A \]

\[ \frac{\partial u}{\partial t} = -g \frac{\partial \bar{u}}{\partial x} + f v, \quad \frac{\partial v}{\partial t} = -g \frac{\partial \bar{v}}{\partial y} - f u, \]
\[ \frac{\partial h}{\partial t} = -H \left( \frac{\partial \bar{u}}{\partial x} + \frac{\partial \bar{v}}{\partial y} \right), \]  
\[ (3.2)_B \]

\[ \frac{\partial u}{\partial t} = -g \frac{\partial \bar{u}}{\partial x} + f \bar{v}^{xy}, \quad \frac{\partial v}{\partial t} = -g \frac{\partial \bar{v}}{\partial y} - f \bar{u}^{xy}, \]
\[ \frac{\partial h}{\partial t} = -H \left( \delta_x \bar{u} + \delta_y \bar{v} \right), \]  
\[ (3.2)_C \]

\[ \frac{\partial u}{\partial t} = -g \frac{\partial \bar{u}}{\partial x} + f \bar{v}^{xy}, \quad \frac{\partial v}{\partial t} = -g \frac{\partial \bar{v}}{\partial y} - f \bar{u}^{xy}, \]
\[ \frac{\partial h}{\partial t} = -H \left( \delta_x \bar{u}^{xy} + \delta_y \bar{v}^{xy} \right), \]  
\[ (3.2)_D \]

\[ \frac{\partial u}{\partial t} = -g \delta_x h + f v, \quad \frac{\partial v}{\partial t} = -g \delta_y h - f u, \]
\[ \frac{\partial h}{\partial t} = -H (\delta_x u + \delta_y v). \]  
\[ (3.2)_E \]
We want to find the numerical dispersion relations and compare them with the exact solution. For 1-D problem, the dispersion relations are (note $v$ is our $\omega$, $d = \Delta x$)

\[
\left(\frac{v}{f}\right)^2 = 1 + \left(\frac{\lambda}{d}\right)^2 \sin^2 kd, \quad (3.6)_A
\]

\[
\left(\frac{v}{f}\right)^2 = 1 + 4 \left(\frac{\lambda}{d}\right)^2 \sin^2 \frac{kd}{2}, \quad (3.6)_B
\]

\[
\left(\frac{v}{f}\right)^2 = \cos^2 \frac{kd}{2} + 4 \left(\frac{\lambda}{d}\right)^2 \sin^2 \frac{kd}{2}, \quad (3.6)_C
\]

\[
\left(\frac{v}{f}\right)^2 = \cos^2 \frac{kd}{2} + \left(\frac{\lambda}{d}\right)^2 \sin^2 kd, \quad (3.6)_D
\]

\[
\left(\frac{v}{f}\right)^2 = 1 + 2 \left(\frac{\lambda}{d}\right)^2 \sin^2 \frac{kd}{\sqrt{2}}, \quad (3.6)_E
\]

They are plotted in the following figure:
The phase speed and group velocities for each of these grid can be plotted together with the exact solution:

**Figure 3.2** The functions $|v|/f$ given by (3.4) and (3.6), with $\lambda/d = 2$.

The phase speed and group velocities for each of these grid can be plotted together with the exact solution:

**Figure 7.5** The phase velocity $c = \dot{\omega}/\mu$ and the group velocity $d\dot{\omega}/d\mu$ from Schoenstadt (1978) as functions of $\mu d/\pi$ for the four grids as indicated. The differential solution is also included. These results use the values: $gH = 10^4$ m$^2$ sec$^{-2}$, $f = 10^{-4}$ sec$^{-1}$, $d = 500$ km.
We can see that for the 1-D problem, B and C grid perform the best.

A and D are not good at all. Energy of waves shorter than $4\Delta x$ propagates in the wrong direction.

E is reasonable good.

For 2-D problem, the $\omega/f$ is plotted in the following:

We can see C grid is closest to the exact solution given in (A), and B grid is not as good in 2-D, especially along the diagonal direction in the plot.
5. Boundary Conditions for Hyperbolic Equations

(ref. Chapter 8, Durran)

5.1. Introduction

In numerical models, we have to deal with two types of boundary conditions:

a) Physical

- e.g., ground (terrain), coast lines, the surface of a car when modeling flow around a moving car.
- internal boundaries / discontinuities

b) Artificial / Numerical

- must impose them to limited integration domain, but they should act as if they don’t exist
- the boundary should be transparent to "disturbances" moving across the boundary
- there can be different kinds of forcing at the boundaries, e.g., lifting by mountain slope and heating at the surface
- these boundaries should be well-posed mathematically
- often we have to over-specify the boundary condition, e.g., when a grid is one-way nested inside the coarser grid
- it has been shown that no well-posed lateral b.c. exists for the shallow water equations and also for the Navier-Stokes equations
- still a lot of debate in this area. B.C. are often critical because they can exert enormous control over the interior solution

As you might suspect, B.C. for hyperbolic problems are closely related to the theory of characteristics – information propagates along characteristic paths.

Consider the well-posed problem in a 1-D domain \( x \geq 0 \) (only one b.c. at \( x=0 \), but to solve the problem numerically, we have to place a computational boundary somewhere at \( x>0 \)).

\[
\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0
\]

(36)

I.C.: \( u(x,0) = f(x) \)
B.C.: \( u(0,t) = g(x) \)
And \( f(0) = g(0) \) for consistency between the I.C. and B.C.
From our earlier discussion, we know that

\[ du = 0 \text{ along } \frac{dx}{dt} = c, \text{ i.e., } x = ct + \beta \]

where \( \beta \) is a constant to be determined by I.C.

Look at an x-t diagram:

Consider the characteristic passing through \((x_1, t_1)\):

\[ u(x_1, t_1) = u(x_2, 0) = f(x_2) = f(x_1 - ct_1) \]

For any \((x,t)\) such that \(x - ct \geq 0\), that solution can be related to the I.C. \(f(x)\), i.e.,

\[ u(x,t) = f(x-ct) \text{ for } x \geq ct. \]

Consider now point \((x_3, t_3)\). In this case, using the MOC, we see that

\[ u(x_3, t_3) = u(0, t^*) = g(t_3 - x_3/c) \rightarrow \]

solution has dependency on the B.C. \(g(t)\) and not the I.C. Thus in general,

\[ u(x,t) = g(t-x/c) \text{ if } x<ct. \]

Now, if we have to impose a boundary condition at \(x = L\), the problem becomes ill-posed because we've over-specified the solution at \(x = L\), i.e., no condition is required there!
It is unlikely that the solution given by \( f(x_1 - ct_1) \), for example, will match whatever condition we impose at \( x = L \). The problem is that, in the general case, the B.C. depends on the solution, which isn’t known at \( x = L \) a priori! What happens if we have a whole spectrum of waves that propagate at difference speeds? We can’t supply a B.C. for each one!

**5.2. Number of Boundary Conditions**

For the previous 1-D advection problem, we need only one B.C. Now let’s look at the 1-D shallow water equations in the absence of mean flow:

\[
\begin{align*}
\frac{\partial u'}{\partial t} + \frac{\partial \phi'}{\partial x} &= 0 \\
\frac{\partial \phi'}{\partial t} + \Phi \frac{\partial u'}{\partial x} &= 0
\end{align*}
\]

\((37a) \quad (37b)\)

\[0 \leq x \leq L.\]

Recall the characteristic form of the system for \( \Phi = \text{constant} \)

\[
\begin{align*}
\frac{\partial A}{\partial t} + \sqrt{\Phi} \frac{\partial A}{\partial x} &= 0 \\
\frac{\partial B}{\partial t} - \sqrt{\Phi} \frac{\partial B}{\partial x} &= 0
\end{align*}
\]

\((38a) \quad (38b)\)

where \( A = u + \phi / \sqrt{\Phi} \) and \( B = u - \phi / \sqrt{\Phi} \).

Clearly, there are 2 pure advection equations in the Riemann invariants \( A \) and \( B \), with wave speeds of \( \pm \sqrt{\Phi} \). We have separated the waves, or eigenvalues, and in general, the number of boundary conditions equals to the number of eigenvalues. This doesn’t tell what the B.C. should, however. Just how many.

In practice, the number of B.C. also depends on the particular grid structure used.

In our case,

\[\lambda_1 = +\sqrt{\Phi} \rightarrow \text{right moving wave} \rightarrow \text{must specify L.H. boundary condition}\]

\[\lambda_2 = -\sqrt{\Phi} \rightarrow \text{left moving wave} \rightarrow \text{must specify R.H. boundary condition}\]

Note that we specify the B.C. from where the wave originates, but not to where it’s going!
5.3. Sample B.C. and Wave Reflection

For limited area models that contain artificial lateral boundaries, we desire to let incident waves pass through without reflection, i.e., as if the boundary wasn’t there at all. This is the behavior for the exact or differential solution.

\[ \frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad c > 0 \text{ and constant, } \quad -\infty \leq x \leq \infty, \quad t \geq 0. \quad (39) \]

To look at reflection at the boundary, we need consider only the spatical derivative.

\[ \frac{\partial u}{\partial t} + c \partial_x^u = 0 \quad -\infty \leq x \leq L \quad (40) \]

This equation describes a right-moving wave. If there is reflection at \( x=L \), the reflected wave must be computational in origin, because the physical equation doesn’t support left-moving wave.

Our center-in-space discretization cannot be applied at \( x=L \) (since it needs \( u \) at \( L+1 \)), so something else must be done. Note that no B.C. should be specified at \( x=L \), except due to the fact that computer has limited memory and computing power so you can’t make the computational domain infinitely large.

Approximations to the PDE at \( x=L \):

1. \( u_L = 0 \) \quad Fixed or rigid boundary
2. \( u_L = u_{L-1} \) \quad Zero gradient (about \( u_{L-\Delta x/2} \))
3. \( u_L = 2u_{L-1}-u_{L-2} \) \quad Linear extrapolation

One can also use special forms of the PDE, e.g., upstream at the boundary:
Question: What happens to a wave when these B.C. are applied to the semi-discrete equation (40)?

Assume solutions of the form

\[ u = A \exp[i(kx_j - \omega t)] \rightarrow \]
\[ -i \omega A e^{ikx_j - \omega \Delta x} + A e^{ikx_j - \omega \Delta x} (e^{ik\Delta x} - e^{-ik\Delta x})/(2\Delta x) \]

\[ \omega_d = \frac{c \sin(k\Delta x)}{\Delta x} = \frac{c \sin(p_d)}{\Delta x} \quad (p_d = k\Delta x) \]

Now, for the PDE we have

\[ \omega_e = kc = \frac{k \Delta x}{\Delta x} c = \frac{p_e c}{\Delta x} \]

where \( p_e \) = exact value of \( k\Delta x \). We now want to compare these two frequencies.
Note that, in the F.D. solution, two values of kΔx correspond to a single ωd, whereas in the exact solution, ωe is linear in kΔx.

We therefore say that the F.D. solution is non-monochromatic, i.e., there is more than one wavelength per frequency.

We can gain insight into the behavior of the solution by creating a monochromatic form, like we did in the leapfrog scheme when we wrote β1 = f(β2) = β. In that spirit, we write:

\[ u_j^n = [A e^{ipj} + B e^{i(\pi - p)j}] e^{-i\omega t} \quad 0 \leq p \leq \pi / 2 \]

where the 2nd term takes care of 'reflection'.

The second term is the computational mode in space. Note that the slope of the curve for p> π/2 is opposite to that for p≤ π/2. Slope = \( \partial \omega / \partial k \) = group velocity – we will return to this shortly.

Phase speed:

\[ c_d = \frac{\omega_d}{k} = \frac{-\omega_d}{p_d / \Delta x} = c \frac{\sin(p_d)}{p_d} > 0 \]

which is the same for both modes (0 ≤ p ≤ π / 2).

Note that, for small p, \( c_d \sim c \).

We now start to see that phase speed isn't a good indicator of wave reflection, because it does not represent the propagation of wave energy. In the above case, the phase speed is always positive – so it has no way of indicating reflection.

Group velocity:

\[ c_g = \frac{\partial \omega_d}{\partial (p_d / \Delta x)} = c \cos(p_d) \]

\[ c_g > 0 \quad \text{for } p_d = p \quad (0 \leq p \leq \pi / 2) \]

\[ c_g < 0 \quad \text{for } p_d = \pi - p. \]

If \( c_g < 0 \), this means that energy is propagating in a direction opposite to the \( c_g \) of the exact solution (which is positive), an indication of reflection!

We can thus interpret reflection in terms of group velocity - the reflected waves (or more accurately wave packets) transport energy in the opposite direction upon reflection.

It is possible to determine the amplitude, \( r \), of the reflected wave. See Matsuno, JMSJ, 44, 145-157 (1966).
Let \( A = \) amplitude of incident wave \( = 1.0 \)
\( B(=r) = \) amplitude of reflected (computational mode) wave

(notation is from the monochromatic solution)

\[
  u_j = e^{-i\omega t} \left[ e^{i\omega j} + re^{i(\pi-\omega)j} \right]
\]

Example: Zero gradient lateral boundary condition

\[
  u_L = u_{L-1}
\]

without loss of generality, we can let \( L=0 \) \( \rightarrow \)

\[
  u_0 = u_{-1}
\]

With \( j=0 \) and \( j=-1 \) in our solution, we have

\[
  1 + r = e^{-ip} + re^{-i(\pi-p)}
\]

where ‘d’ on \( p \) has been dropped, and the \( e^{-i\omega t} \) ‘canels’ via linear independence for a single frequency \( \omega \).

\[
  r(1+e^{ip}) = e^{-ip} - 1 \rightarrow \\
  r = \frac{e^{-ip} - 1}{1+e^{ip}} = \frac{e^{-ip/2}(e^{-ip/2} - e^{ip/2})}{e^{ip/2}(e^{-ip/2} + e^{ip/2})} = e^{-ip} \frac{-i\sin(p/2)}{\cos(p/2)} \\
  = -[\cos(p) - i\sin(p)]i\tan(p/2) \rightarrow \\
  | r | = \tan(p/2) \quad \text{where } p = k \Delta x = p_d \ (0 \leq p \leq \pi/2)
\]

Note: The reflected amplitude depends upon both \( k \) of the incident wave and \( \Delta x \).

4\( \Delta x \) wave (\( p = \pi/4 \)) \( \rightarrow \) \( | r | = 1.0 \) \( \rightarrow \) complete reflection without change in amplitude.

32 \( \Delta x \) wave (\( p=\pi/16 \)) \( \rightarrow \) \( | r | = 0.0982 \) \( \rightarrow \) relatively weak reflection

\( \rightarrow \) \( | r | \sim 1/ (\text{incident wavelength}) \).

Short waves are reflected much more than long waves – it is understandable – infinitely long waves, i.e., constant field, should pass through a zero gradient boundary freely.
### 5.4. Reflection for the Shallow Water Equations – Radiation Boundary Condition

Let’s consider a semi-discrete shallow water equations system:

\[
\begin{align*}
\frac{\partial u}{\partial t} + \bar{u}\delta_{2,4}u + \sqrt{\Phi}\delta_{2,4}\hat{\phi} &= 0 \\
\frac{\partial \hat{\phi}}{\partial t} + \bar{u}\delta_{2,4}\hat{\phi} + \sqrt{\Phi}\delta_{2,4}u &= 0
\end{align*}
\]

Here, we let \( \hat{\phi} = \phi / \sqrt{\Phi} \) so that the two equations are completely symmetric. This is to make the following discussions easier.

The system describes waves moving in the + or – directions depending upon the relationship of \( \bar{u} \) and \( \Phi^{1/2} \). If \( \bar{u} < 0 \) and \( |\bar{u}| > \Phi^{1/2} \), then the wave will go the left. This is the foundation for the Klemp and Wilhelmson (1978) radiation boundary condition.

Often we want to allow such waves to pass through an artificial lateral boundary while minimizing the reflection.

Because we have 2 waves in this system, there will be 2 physical modes and 2 computational modes – one for each wave. Using the same analysis as before, we have

**Physical Modes:**

\[(u, \phi)_1 = (\tilde{u}, \tilde{\phi})_1 e^{i(k_1x - \omega_1t)} \quad \text{with} \quad \omega_1 = \frac{\bar{u} + \sqrt{\Phi}}{\Delta x} \sin(k_1\Delta x)\]

\[(u, \phi)_2 = (\tilde{u}, \tilde{\phi})_2 e^{i(k_2x - \omega_2t)} \quad \text{with} \quad \omega_2 = \frac{\bar{u} - \sqrt{\Phi}}{\Delta x} \sin(k_2\Delta x)\]

where \( \tilde{u} \) and \( \tilde{\phi} \) are the amplitude of \( u \) and \( \hat{\phi} \), respectively.

**Computational Modes** (the "image" about \( p=\pi/2 \))

\[(u, \phi)_3 = (\tilde{u}, \tilde{\phi})_3 e^{i(k_3x - \omega_3t)} \quad \text{with} \quad \omega_3 = \frac{\bar{u} + \sqrt{\Phi}}{\Delta x} \sin(k_3\Delta x) \quad \text{and} \quad k_3 = \frac{\pi}{\Delta x} - k_1\]

\[(u, \phi)_4 = (\tilde{u}, \tilde{\phi})_4 e^{i(k_4x - \omega_4t)} \quad \text{with} \quad \omega_4 = \frac{\bar{u} - \sqrt{\Phi}}{\Delta x} \sin(k_4\Delta x) \quad \text{and} \quad k_4 = \frac{\pi}{\Delta x} - k_2\]

Note: \( \omega_3 = \omega_1 \) and \( \omega_4 = \omega_2 \), i.e., 2 horizontal wavenumbers corresponding to the same value of frequency.
Let’s now look at the directions of phase speed and group velocity of these modes.

**Phase Speed**

**Physical modes:**
\[ c_p > 0 \text{ for both modes if } \bar{u} > \Phi^{1/2} \]
\[ c_p > 0 \text{ for one and } < 0 \text{ for the other if } 0 < \bar{u} < \Phi^{1/2} \]

**Computational Modes:**
Same as above

**Group Velocity**

\[ c_{g1} = -c_{g3} = (\bar{u} + \Phi^{1/2}) \cos(k_1 \Delta x) \]
\[ c_{g2} = -c_{g4} = (\bar{u} - \Phi^{1/2}) \cos(k_2 \Delta x) \]

where \( 0 \leq k_{1,2} \Delta x \leq \pi/2 \) for a monochromatic solution. Thus there are always 2 group velocities > 0 and two < 0.

**Example:** Consider what happens at an outflow boundary when \( \Phi^{1/2} > \bar{u} > 0 \). Here, we want to specify a boundary condition at \( x = L \).

Now, we know that
\[ u = u_1 + u_2 + u_3 + u_4 \]
\[ \hat{\phi} = \hat{\phi}_1 + \hat{\phi}_2 + \hat{\phi}_3 + \hat{\phi}_4 \]

From the analytical solution, we know that
\[ \omega = k (\bar{u} + \Phi^{1/2}) \Rightarrow \]
\[ c_p = c_g, \]

i.e., the energy and phase always propagate in the same direction.

Going back to our 4 solutions, we find that for \( \Phi^{1/2} > \bar{u} > 0 \) (this case only), i.e., the that is dominated by the wave propagation rather than advection.

Mode 1: Phase > 0, group > 0 \( \Rightarrow \) physical mode
Mode 2: Phase < 0, group < 0 \( \Rightarrow \) physical mode
Mode 3: Phase > 0, group < 0 \( \Rightarrow \) computational mode
Mode 4: Phase < 0, group > 0 \( \Rightarrow \) computational mode

Now, mode 1 is the physical incident wave, and if it gets reflected, its energy will come back in modes 2 and 3 because their group velocities are negative. Mode 4 won’t be involved in the reflection.

As a result, we can analyze the reflection in terms of 1, 2 and 3 only:
Why the minus sign on the last term in the $\phi$ equation?

Recall that the Riemann invariant of the right-moving wave is $u_{1,3} + \phi_{1,3}/\Phi^{1/2} = u_{1,3} + \hat{\phi}_{1,3}$.

$$u_{1,3} + \hat{\phi}_{1,3} = 2(e^{ik_1 x} + re^{ik_3 x})e^{-i\omega t},$$

which is the physical mode 1 with its computational counterpart 3. The minus sign is needed to make sure that the above 2 equations when added together form a Riemann invariant that does not involve left-ward propagation waves - those that are supported by the other characteristic equation, i.e.,

$$\partial u + \partial (u - \phi) = 0.$$

Now, let’s examine the reflection properties for various boundary conditions applied to this particular semi-discrete form of the shallow water equations.

First, it’s important to recognize that the solutions at the boundary must be continuous → frequency of the incident and reflected waves must be identical at the boundary. So, if $\omega_1 = \omega_2$ ($\omega_1 = \omega_3$ already), we have,

$$k_2 \approx k_1 \frac{\vec{u} + \Phi^{1/2}}{\vec{u} - \Phi^{1/2}}.$$

[Here we assumed that $k \Delta x$ is small (good for high resolution) so that $\sin(k \Delta x) \approx k \Delta x$.]

Therefore, $k_2 > k_1$, or the reflected wave is always smaller in scale than the incident wave. (Note that at a critical layer where $\Phi^{1/2} = \vec{u}$, the analysis is not valid and a different approach has to be taken.)
Now, let’s apply the above analysis to different types of boundary conditions.

Based on earlier analysis, we rewrite (41) as

\[ u = [e^{ik_1 x} + re^{(\frac{\pi}{\Delta x} - k_1) x} + R e^{ik_2 x}] e^{-i\omega x} \]

\[ = [e^{ik_1 x} + r(-1)^j e^{ik_1 x} + R e^{ik_2 x}] e^{-i\omega x} \]

\( x = j\Delta x \)

Similarly we have

\[ \phi = [e^{ik_1 x} + r(-1)^j e^{ik_1 x} - R e^{ik_2 x}] e^{-i\omega x} \]

**Case I**: \( u_L = 0 \), i.e., we have the rigid wall.

\[ u_L = 0 \Rightarrow \frac{\partial \phi}{\partial x} \bigg|_L = 0 \] (if \( \bar{u} = 0 \), i.e., can’t have \( \bar{u} \neq 0 \) if \( u = 0 \) at the boundary)

Let \( L = 0 \) without loss of generality:

\[
\begin{array}{c c c c c c c}
-1 & \quad & 0 & \quad & \quad & \quad & \\
\end{array}
\]

therefore \( u_0 = 0, \phi_0 = \phi_{-1} \).

For the first equation, we have

\[ 1 + r + R = 0 \]

For the second, we have

\[ 1 + r - R = [e^{-ik_2 \Delta x} - re^{-ik_1 \Delta x} - R e^{-ik_2 \Delta x}] \]

Assuming that \( e^x \approx 1 + x \), we have

\[ 1 + r - R = 1 - ik_1 \Delta x - r + rik_1 \Delta x - R + Rik_2 \Delta x \]

\[ r = i\Delta x(-k_1 + rk_1 + Rk_2) / 2 \]

Equating the real and imaginary parts, we have
\[ r = 0 \]
\[ R = -1 \]

\[ \rightarrow \quad k_1 = -k_2. \]

We showed earlier that

\[ k_2 = k_1 \frac{u + \Phi^{1/2}}{u - \Phi^{1/2}} \]

which agrees with current result when \( u = 0 \).

In summary, we found total reflection in the physical mode, which is the expected condition for a rigid lateral boundary. This is also undesirable if the true problem does not actually have a boundary here. When there is indeed a wall, we often use the so-called mirror boundary condition (as in ARPS for B.C. option one), which finds its basis in our analysis.

Case II: Wave radiating lateral boundary condition

At the boundary \( L \), we use

\[ \frac{\partial u}{\partial t} + (\bar{u} + c^*) \frac{u_L - u_{L-1}}{\Delta x} = 0 \]

for \( \phi \), use the governing equation itself with one-sided difference:

\[ \frac{\partial \phi}{\partial t} + \bar{u} \frac{\phi_L - \phi_{L-1}}{\Delta x} + \Phi \frac{u_L - u_{L-1}}{\Delta x} = 0. \]

In the above, \( c^* \) is an estimate of the dominant wave speed at the lateral boundary (can be set constant, or computed as was done by Orlanski 1976).

Doing a reflection analysis, we find (show for yourselves) that

\[ R = \frac{(\bar{u} - \Phi^{1/2})(\Phi^{1/2} - c^*)}{(\bar{u} + \Phi^{1/2})(\Phi^{1/2} + c^*)} \quad \text{(for reflected physical mode)} \]
\[ r = 0 \quad \text{(for computational mode)} \]

[See Klemp and Lilly, 1978, JAS, p78]

If we do a good job of estimating \( c^* \), then we can make \( R \sim 0 \).
Again, the reflection is in the physical mode, and that is not good. Keep in mind that this
analysis assumes that \( k \Delta x \ll 1 \).

In complicated models, one typically replaces the normal momentum equation at the
lateral boundary with

\[
\frac{\partial u}{\partial t} + (\bar{u} + c^*) \frac{\partial u}{\partial x} = 0
\]  
(42)

It neglects pressure gradient force responsible for the wave propagation, the effect can be
thought as being accounted for in \( c^* \).

In (42), the sign of \( u + c^* \) determines inflow or outflow boundary. Other equations are
then solved using their original governing equations, with one-sided difference when
necessary. This type of condition is called Sommerfeld radiation condition. Orlanski
(1976) applies equation like (42) to all variables, but in practice, it does not work very
well – it leads to over-specification of conditions and enhancement of the computational
mode.

ARPS has options for four variations for radiation lateral boundary conditions – all
follow Sommerfeld condition, the difference being the way \( c^* \) is determined and whether
(42) is applied in large or small time steps.

Klemp and Lilly (1978) gives more details on the analysis of radiation lateral boundary
conditions.

It should be noted that, in terms of characteristics, our shallow water system is

\[
\frac{\partial A}{\partial t} + (\bar{u} + \Phi^{1/2}) \frac{\partial A}{\partial x} = 0
\]

\[
\frac{\partial B}{\partial t} + (\bar{u} - \Phi^{1/2}) \frac{\partial B}{\partial x} = 0
\]

with \( A = u + \phi/\Phi^{1/2} \) and \( B = u - \phi/\Phi^{1/2} \), the Riemann invariants.

The radiation condition \( \frac{\partial u}{\partial t} + (\bar{u} + c^*) \frac{\partial u}{\partial x} = 0 \) is equivalent to setting to zero the
amplitude of \( B \). So, the radiation B.C. can be regarded as a condition written in terms of
the characteristics, and that's why a PGF does not occur. In fact, if \( B=0 \), at the lateral
boundary, then

\[
A + B = A + 0 = 2u
\]

/        /

4-49
\[
\frac{\partial (2u)}{\partial t} + (\bar{u} + \Phi^{1/2}) \frac{\partial (2u)}{\partial x} = 0
\]

which is the Radiation Condition we used.

References:


5.4. Other Boundary Conditions

Rayleigh Damping (Sponge)

Here we include a regional (zone) next to the lateral boundary in which zero-order (Rayleigh) damping is applied to the prognostic variables, to ‘absorb’ incident wave energy and damp possibly reflected energy.

\[
\delta_2 u = -r(x)(u - \bar{u})^{n-1} = -(u - \bar{u})^{n-1} / \tau(x)
\]

where \( r \) is the Rayleigh damping coefficient and \( \tau \) the corresponding e-folding time of damping. The smaller \( r \) is, the longer it takes to damp.

Rayleigh damping is usually needed when the lateral boundary conditions are over-specified, such as the case of externally forced boundary (e.g., when a grid is forced by the solution of another model, the coarse grid solution of the same model or by analysis – the case of one-way nesting). ARPS uses Rayleigh damping with externally forced boundary option.
Viscous sponge

It takes the form of second-order diffusion

$$\delta_{2,i} u = K \delta_{x,i} u^{n-1}$$

In this case, short waves are selectively damped. It does not damp long reflected waves effectively, however. It is often used in combination with the Rayleigh damping, as in the ARPS.

Top boundary condition

In atmospheric models, the top boundary of the computational domain often has to be placed at a finite height – creating an artificial top boundary. Vertically propagating, e.g., internal gravity, waves can reflect off the boundary and interact with flow below – creating problems. One example of radiation top boundary condition is that of Klemp and Durran (1983). See also Durran section (8.3.2).

Reference:

5.5. Additional References on Boundary Conditions (most significant ones with *). Most recent references not included


Elvius, T., and A. Sundstrom, 1973: Computationally efficient schemes and boundary conditions for a fine-mesh barotropic model based on the shallow-water equations. Tellus, 25, 132-156.


