Numerical Weather Modeling

Fourth’s Year Lectures

Lecturer Dr. Thaer Obaid Roomi
Announcements

Subject: Numerical Weather Modeling

Text: Class notes

Course grade: 1/4 on homework, 1/4 on each of two midterms, and 1/4 on final. The final will emphasize the latter part of the course, and will be held during final weeks.

Access to instructor: As you may know, I have posted office hours, but students in this class are welcome to come to me with questions any time, provided only that I am not actually busy with someone else.

Computing language: MATLAB

Auditing: Permitted with preliminary approval

Schedule According to the department's panel
Preface

The purpose of this course is to provide an introduction to the methods used in numerical modeling of the atmosphere.

Numerical weather modeling is one of several approaches to study of the atmosphere. The others are observational studies of the real atmosphere through field measurements and remote sensing, laboratory studies, and theoretical studies. Each of these four approaches has both strengths and weaknesses. In particular, both numerical modeling and theory involve approximations. In theoretical work, the approximations often involve extreme idealizations, e.g. a dry atmosphere on a beta plane, but on the other hand solutions can sometimes be obtained in closed form with a pencil and paper. In numerical modeling, less idealization is needed, but in most cases no closed form solution is possible. Both types of work are subject to errors in the old-fashioned human sense.

Perhaps the most serious weakness of numerical modeling, as a research approach, is that it is possible to run a numerical model built by someone else without having the foggiest idea how the model works or what its limitations are. Unfortunately, this kind of thing happens all the time, and the problem is becoming more serious in this era of “community” models with large user groups. One of the purposes of this course is to make it less likely that you, the students, will use a model without having any understanding of it.

This introductory survey of numerical methods in the atmospheric sciences is designed to be a practical, “how to” course, which also conveys sufficient understanding so that after completing the course students are able to understand the properties of schemes that they may encounter out there in the world.

I make every effort to ensure that there are no errors in the text. However, no one is perfect, and mistakes do occur. If you find an error somewhere in the lectures, like a spelling mistake or faulty piece scientific notes, I would be very grateful for your feedback.
Lecture (1)

Some Introductory Principles

1.1 What is a model?

The atmospheric science researchers carry out measurements of the atmosphere. The researchers may do the following:

- Instrument development,
- algorithm development,
- data collection,
- data reduction,
- and data analysis.

The data by themselves are just numbers. In order to make physical sense of the data, some sort of model is needed. The model may be:

Most models in atmospheric science are formulated by starting from basic physical principles, such as conservation of mass, conservation of momentum, and conservation of thermodynamic energy. Many of these equations are prognostic, which means that they involve time derivatives. A simple example is the continuity equation, which expresses conservation of mass:

$$\frac{\partial \rho}{\partial t} = - \nabla \cdot (\rho \vec{V}) \quad (1.1)$$

Here $t$ is time, $\rho$ is density, $\vec{V}$ is velocity vector. This equation is prognostic in which $\rho$ is a prognostic variable.
A model that contains prognostic equations is solved by time integration, and the prognostic variables of such a model must be assigned initial conditions. Any variable that is not prognostic is called *diagnostic*.

### 1.2 Elementary models

These are the models that are essentially direct applications of the physical principles (such as conservation of mass, conservation of momentum, and conservation of thermodynamic energy) to phenomena that occur in the atmosphere. These models are “elementary” in the sense that they form the conceptual foundation for other modeling work. Elementary models tend to deal with microscale phenomena (e.g., the evolution of individual cloud droplets or optical properties of ice crystals). Elementary models are usually *analytical*. This means that the results that they produce consist of equations. As a simple example, consider the ideal gas law

$$ p = \rho RT \quad (1.2) $$

Eq. (2) can be derived using the kinetic theory of gases, which is an analytical model; the ideal gas law can be called a “result” of the model. This simple formula can be used to generate numbers, of course; for example, given the density and temperature of the air, and the gas constant, we can use Equation (2) to compute the pressure. This particular formula is sufficiently simple that we can understand what it means just by looking at it.

### 1.3 Numerical models

The results of a numerical model consist of numbers, which represent particular “cases” or “realizations.”. For example, we can “run” a numerical model to create a weather forecast. The forecast consists of a large set of numbers. To perform a new forecast, for a different initial condition, we have to run the model again, generating a new set of numbers. In order to see everything that the (model) atmosphere can do, we would have to run the model for infinitely many cases. In this way, numerical models are quite different from analytical models, which can describe all possibilities in a single formula. We cannot understand the results of a numerical simulation just by looking at the computer code.

Simply, the numerical model can be defined: is a set of discredited mathematical equations, solved on a computer, which represent the behavior of a physical system.
1.4 Physical and mathematical errors

All models contain errors. It is useful to distinguish between physical errors and mathematical errors. For example, we often consider the Navier-Stokes equations to be an exact description of the fluid dynamics of air. For various reasons, we are unable to obtain exact solutions to the Navier-Stokes equations. To simplify the problem, we introduce physical approximations. For example, we may introduce Reynolds averaging along with closures that can be used to determine turbulent and convective fluxes. Physical approximations are useful beyond merely simplifying the equations. A second motivation for making physical approximations is that the approximate equations may describe the phenomena of interest more directly, omitting or “filtering” phenomena of less interest, and so yielding a set of equations that is more focused on and more appropriate for the problem at hand. For example, we may choose approximations that filter sound waves (e.g., the anelastic approximation) or even gravity waves (e.g., the quasigeostrophic approximation). These physical approximations introduce physical errors, which may or may not be considered acceptable for the intended application of a model.

At this point, we have chosen the equation system of the model. Once we have settled on a suitable set of physical equations, we must devise mathematical methods to solve them. The mathematical methods are almost always approximate, which is another way of saying that they introduce errors. This course focuses on the mathematical methods for models in the atmospheric sciences, and especially the mathematical errors that these methods entail.

1.5 Discretization

Numerical models are “discrete.” This simply means that a numerical model involves a finite number of numbers. The process of approximating a continuous model by a discrete model is called “discretization.”

There are multiple approaches to discretization. This course emphasizes grid-point methods, which are sometimes called finite-difference methods. The fields of the model are defined at the discrete points of a grid. The grid can and usually does span time as well as space. Derivatives are then approximated in terms of differences involving neighboring grid-point values. A finite-difference equation (or set of equations) that approximates a differential equation (or set of equations) is called a finite-difference scheme, or a grid-point scheme. Grid-point schemes can be derived by various approaches, and the derivation methods themselves are sometimes given names. Examples include finite-volume methods, finite-element methods, and semi-Lagrangian methods.
The major alternative to the finite-difference method is the spectral method, which involves expanding the fields of the model in terms of weighted sums of continuous, and therefore differentiable, basis functions, which depend on horizontal position. Simple examples would include Fourier expansions, and spherical harmonic expansions. In a spectral model, the basis functions are global, which means that they each basis function is defined over the entire horizontal domain.

In a continuous model, infinitely many basis functions are needed to represent the spatial distribution of a model field. The basis functions appear inside a sum, each weighted by a coefficient that measures how strongly it contributes to the field in question. In a discrete model, the infinite set of basis functions is replaced by a finite set, and the infinite sum is replaced by a finite sum. In addition, the coefficients are defined on a discrete time grid, and almost always on a discrete vertical grid as well. Spectral models use grid-point methods to represent the temporal and vertical structures of the solution.

In a spectral model, horizontal derivatives are computed by differentiating the continuous basis functions. Although the derivatives of the individual basis functions are computed exactly, the derivatives of the model fields are only approximate since they are represented by finite (rather than infinite) sums.

Even after we have chosen a grid-point method or a spectral method, there are many additional choices to make:

- Design of a model entails choosing which variables to prognose.
- If we choose a grid-point method, then we have to choose the shapes of the grid cells. Possibilities include rectangles, triangles, and hexagons.
- Having chosen the shapes of the grid cells, we must choose where to locate the predicted quantities on the grid. In many cases, different quantities will be located in different places. This is called “staggering.” We will discuss systematic approaches to identifying the best staggering choices.
- For any given grid shape and staggering, we can make numerical schemes that are more accurate or less accurate. The meaning of accuracy will be discussed later. More accurate schemes have smaller errors, but less accurate schemes are simpler and faster. As usual, there are trade-offs. We have to make choices.

1.6 Review of the vector mathematics

Scalars: are variables such as temperature and air pressure that have magnitude but not direction. Vectors: are variables such as velocity that have magnitude and direction.

- The velocity vector:
  \[ \vec{V} = iu + jv + kw \]  
  (total vector)  

(1.3)
\[
\vec{V}_h = iu + jv \quad \text{(horizontal vector)} \tag{1.4}
\]

And, \( u, v, w \) are the scalar components of velocity:

\[
u = \frac{dx}{dt}, \quad v = \frac{dy}{dt}, \quad w = \frac{dz}{dt} \tag{1.5}
\]

The magnitude of the wind is its speed. The total and horizontal wind speeds are defined as:

\[
|\vec{V}| = \sqrt{u^2 + v^2 + w^2}, \quad |\vec{V}_h| = \sqrt{u^2 + v^2} \tag{1.6}
\]

- The dot product: It is a product of two vectors gives a scalar.

Let \( \vec{A} \) and \( \vec{B} \) are two vectors,

\[
\vec{A} \cdot \vec{B} = A_xB_x + A_yB_y + A_zB_z \tag{1.7} \quad \text{(How?)}
\]

\[
\vec{A} \cdot \vec{B} = |\vec{A}| |\vec{B}| \cos \theta \tag{1.8}
\]

**Example 1.1**: Let \( \vec{A} = 2i - 1/2j - 3k \), and \( \vec{B} = -3i + j - 1/2k \) Find \( \vec{A} \cdot \vec{B} \) and the angle between the two vectors.

Solution: Using equation (1.7)

\[
\vec{A} \cdot \vec{B} = -6 + (-1/2) + 3/2 = -6.5 + 1.5 = -5
\]

\[
|\vec{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2} = \sqrt{4 + 1/4 + 9} = \sqrt{13.75}
\]

\[
|\vec{B}| = \sqrt{B_x^2 + B_y^2 + B_z^2} = \sqrt{9 + 1 + 1/4} = \sqrt{10.25}
\]

Using eqn. 1.8:

\[
\cos \theta = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|} = \frac{-5}{\sqrt{13.75} \sqrt{10.25}} = \frac{-5}{11.52} = -0.434
\]

- The cross product: it is product of vectors gives a vector:

\[
\vec{A} \times \vec{B} = \vec{C} \tag{1.9}
\]

\[
|\vec{C}| = |\vec{A} \times \vec{B}| = |\vec{A}| |\vec{B}| \sin \theta \tag{1.10}
\]

The direction of \( \vec{C} \) is perpendicular on the plane of \( \vec{A} \) and \( \vec{B} \).\n
\[
\vec{A} \times \vec{B} = \vec{C} = (A_yB_z - A_zB_y)i + (A_zB_x - A_xB_z)j + (A_xB_y - A_yB_x)k \tag{1.11}
\]

- Del Operator: is a vector differential operator denoted by the symbol \( \vec{\nabla} \):

\[
\vec{\nabla} = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z} \tag{1.12}
\]

a. Gradient of scalar (pressure)
\[ \vec{\nabla}p = i \frac{\partial p}{\partial x} + j \frac{\partial p}{\partial y} + k \frac{\partial p}{\partial z} \]  \hspace{1cm} (1.13)

b. Divergence of a Vector (velocity)
\[ \vec{\nabla}.\vec{V} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \]  \hspace{1cm} (scalar) \hspace{1cm} (1.14)

c. Curl of a Vector (velocity)
\[ \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u & v & w \end{vmatrix} = \left( \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right)i + \left( \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right)j + \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)k \]  \hspace{1cm} (vector) \hspace{1cm} (1.15)

- Laplacian Operator
If \( Q \) is any quantity then,
\[ \vec{\nabla}\cdot\vec{\nabla}Q \equiv \vec{\nabla}^2Q \]  \hspace{1cm} (1.16)

where \( \vec{\nabla}^2 \) (del squared) is the scalar differential operator:
\[ \vec{\nabla}^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \]  \hspace{1cm} (1.17)

\( \vec{\nabla}^2Q \) is called the Laplacian of \( Q \) and appears in several important partial equations of mathematical physics.

**Exercises**

Q1. What are the possible works of the atmospheric science researchers to get and handle measurements?

Q2. List the main types of models, then explain what are they developed for?

Q3. Define: Numerical model

Q4. What is the deference between the physical error and mathematical error? Give examples.

Q5. Give the expressions of dot product, cross product, and Laplacian operator.

**MATLAB Work**

Write Matlab code to do the calculations in Example 1.

**Homework**

Prove in detail answering (How?) questions in the lecture.
2.1 One dimension finite difference method

Consider the derivative \( \frac{df}{dx} \), where \( f = f(x) \), and \( x \) is the independent variable (which could be either space or time). Finite-difference methods represent the continuous function \( f(x) \) by a set of values defined at a finite number of discrete points in a specified (spatial or temporal) region. Thus, we usually introduce a “grid” with discrete points where the variable \( f \) is defined.

![Figure 2.1 An example of a grid, with uniform spacing \( \Delta x \) (grid size).](image)

We assume that the grid spacing is constant, then \( x_i = i\Delta x \), where \( I \) is the index used to identify the grid points. Using the notation \( f_i = f(x_i) = f(i\Delta x) \), we can define

The forward difference at the point \( i \) by \( f_{i+1} - f_i \), (2.1)

The backward difference at the point \( i \) by \( f_i - f_{i-1} \), (2.2)

The centered difference quotient at the point \( i + \frac{1}{2} \) by \( f_{i+1} - f_i \), (2.3)

Note that \( f \) itself is not defined at the point \( i + \frac{1}{2} \). From (2.1) - (2.3) we can define the following “finite-difference quotients:” the forward-difference quotient at the point \( i \):

\[
\left( \frac{df}{dx} \right)_{i, \text{approx}} = \frac{f_{i+1} - f_i}{\Delta x};
\]

(2.4)

The backward – difference quotient at the point \( i \):
\[
\left( \frac{df}{dx} \right)_{i, \text{approx}} = \frac{f_i - f_{i-1}}{\Delta x}; \quad (2.5)
\]

and the centered-difference quotient at the point \( i + \frac{1}{2} \):
\[
\left( \frac{df}{dx} \right)_{i + \frac{1}{2}, \text{approx}} = \frac{f_{i+1} - f_i}{\Delta x}; \quad (2.6)
\]

In addition, the centered-difference quotient at the point \( i \) can be defined by
\[
\left( \frac{df}{dx} \right)_{i, \text{approx}} = \frac{f_{i+1} - f_{i-1}}{2\Delta x}; \quad (2.7)
\]

Since (2.4) and (2.5) employ the values of \( f \) at two points, and give an approximation to \( \frac{df}{dx} \) at one of the same points, they are sometimes called *two-point approximations*. On the other hand, (2.6) and (2.7) are *three-point approximations*, because the approximation to \( \frac{df}{dx} \) is defined at a location different from the locations of the two values of \( f \) on the right-hand side of the equals sign.

How accurate are these finite-difference approximations?

“Accuracy” can be measured in a variety of ways, as we shall see. One measure of accuracy is truncation error. The term refers to the truncation of an infinite series expansion. As an example, consider the forward difference quotient
\[
\left( \frac{df}{dx} \right)_{i, \text{approx}} = \frac{f_{i+1} - f_i}{\Delta x} \equiv \frac{f[[i+1]\Delta x] - f[i\Delta x]}{\Delta x}; \quad (2.8)
\]

Expand \( f \) in a Taylor series about the point \( x_i \), as follows:
\[
f_{i+1} = f_i + \Delta x \left( \frac{df}{dx} \right)_i + \frac{(\Delta x)^2}{2!} \left( \frac{d^2f}{dx^2} \right)_i + \frac{(\Delta x)^3}{3!} \left( \frac{d^3f}{dx^3} \right)_i + \cdots + \frac{(\Delta x)^{n-1}}{(n-1)!} \left( \frac{d^{n-1}f}{dx^{n-1}} \right)_i + \cdots \quad (2.9)
\]

Eq. (2.9) can be rearranged to
\[
\frac{f_{i+1} - f_i}{\Delta x} = \left( \frac{df}{dx} \right)_i + \varepsilon \quad \quad \quad (2.10)
\]

where
\[
\varepsilon \equiv \frac{(\Delta x)^2}{2!} \left( \frac{d^2f}{dx^2} \right)_i + \frac{(\Delta x)^3}{3!} \left( \frac{d^3f}{dx^3} \right)_i + \cdots + \frac{(\Delta x)^{n-1}}{(n-1)!} \left( \frac{d^{n-1}f}{dx^{n-1}} \right)_i + \cdots \quad (2.11)
\]

is called the *truncation error*. 
If $\Delta x$ is small enough, the leading term on the right-hand side of Eq (2.11) will be the largest part of the error. The lowest power of $\Delta x$ that appears in the truncation error is called the “order of accuracy” or “order of approximation” of the corresponding difference quotient. For example, the leading term of (2.11) is of order $\Delta x$, abbreviated as $O(\Delta x)$, and so we say that (2.10) is a first order approximation or an approximation of first-order accuracy. Obviously (2.5) is also first-order accurate. Just to be as clear as possible, a first-order scheme for the first derivative has the form

$$\left(\frac{df}{dx}\right)_{i,\text{approx}} \approx \left(\frac{df}{dx}\right)_{i} + O(\Delta x),$$

where $\left(\frac{df}{dx}\right)_{i,\text{approx}}$ is an approximation to the first derivative and $\left(\frac{df}{dx}\right)_{i}$ is the true first derivative.

Similarly, a second- order scheme for the first derivative has the form

$$\left(\frac{df}{dx}\right)_{i,\text{approx}} \approx \left(\frac{df}{dx}\right)_{i} + O[(\Delta x)^2],$$

and so on for higher orders of accuracy.

Similar analyses of (2.6) and (2.7) show that they are of second-order accuracy.

For example, We can write

$$f_{i-1} = f_{i} + \left(\frac{df}{dx}\right)_{i} (-\Delta x) + \frac{1}{2!} \left(\frac{d^2f}{dx^2}\right)_{i} \frac{(-\Delta x)^2}{2!} + \frac{1}{3!} \left(\frac{d^3f}{dx^3}\right)_{i} \frac{-(\Delta x)^3}{3!} + \cdots \quad (2.12)$$

Subtracting (2.12) from (2.9) gives

$$f_{i+1} - f_{i-1} = 2 \left(\frac{df}{dx}\right)_{i} (\Delta x) + \frac{2}{3!} \left(\frac{d^3f}{dx^3}\right)_{i} [(\Delta x)^3] + \cdots \quad \text{odd powers only}, \quad (2.13)$$

which can be rearranged to

$$\left(\frac{df}{dx}\right)_{i} = \frac{f_{i+1} - f_{i-1}}{2\Delta x} - \frac{1}{2} \left(\frac{d^3f}{dx^3}\right)_{i} \frac{\Delta x^2}{3!} + O[(\Delta x)^4]. \quad (2.14)$$

Similarly,

$$\left(\frac{df}{dx}\right)_{i+\frac{1}{2}} \approx \frac{f_{i+1} - f_{i}}{\Delta x} - \frac{1}{2} \left(\frac{d^3f}{dx^3}\right)_{i+\frac{1}{2}} \frac{(\Delta x/2)^2}{3!} + O[(\Delta x)^4]. \quad (2.15)$$
From (2.14) and (2.15), we see that

\[
\frac{\text{Error of (2.14)}}{\text{Error of (2.15)}} \cong \frac{(\frac{d^3f}{dx^3})_i \Delta x^2}{3!} = \frac{4(\frac{d^3f}{dx^3})_i}{(\Delta x/2)^2} = \frac{4(\frac{d^3f}{dx^3})_{i+1/2}}{3!} \cong 4 \quad (2.16)
\]

This shows that the error of (2.14) is about four times as large as the error of (2.15), even though both finite-difference quotients have second-order accuracy.

The point is that the “order of accuracy” tells how rapidly the error changes as the grid is refined, but it does not tell how large the error is for a given grid size. It is possible for a scheme of low-order accuracy to give a more accurate result than a scheme of higher-order accuracy, if a finer grid spacing is used with the low-order scheme.

### Exercises

Q1. Define: Finite difference Method, Truncation error, Order of accuracy.
Q2. (IMP.) Prove that the error in forward-difference quotient is four times the error in the centered-difference quotient at the point $+\frac{1}{2}$.

### MATLAB Work

Write Matlab script to calculate the air pressure for the following temperatures (20, 15, 10, 5, 0, -5, -10, -20, -40, -60) °C by using iteration loop, then draw the results. (Note: Use ideal gas law in Lecture 1).

### Homework

1. The centered-difference quotient (Eq 2.7) is more accurate than the forward- and backward-difference quotients (Eq 2.4 and 2.5, respectively). Prove that.
2. Are the equations (2.4-2.16) ordinary differential equations or partial differential equations and why?
Lecture (3)

Finite Difference Methods (Part 2)

3.1 Two dimensions finite difference method

So far, we have discussed one-dimensional finite differences, which implicate ordinary differential equations. However, most atmospheric sciences problems require using two-dimensional formulae which necessarily contain partial differential equations. The Laplacian operator is an example and which is existing in the diffusion equation:

$$\frac{\partial f}{\partial t} = K \nabla^2 f$$  \hspace{1cm} (3.1)

The two-dimensional Taylor series is:

$$f(x_{i+1}, y_{j+1}) \equiv f_{i+1,j+1}$$

$$f_{i,j} + \left[(\delta x)_{i,j} \frac{\partial}{\partial x} + (\delta y)_{i,j} \frac{\partial}{\partial y}\right] f + \frac{1}{2!} \left[(\delta x)_{i,j} \frac{\partial}{\partial x} + (\delta y)_{i,j} \frac{\partial}{\partial y}\right]^2 f$$

$$+ \frac{1}{3!} \left[(\delta x)_{i,j} \frac{\partial}{\partial x} + (\delta y)_{i,j} \frac{\partial}{\partial y}\right]^3 f$$

$$+ \frac{1}{4!} \left[(\delta x)_{i,j} \frac{\partial}{\partial x} + (\delta y)_{i,j} \frac{\partial}{\partial y}\right]^4 f \ldots$$  \hspace{1cm} (3.2)

which can be written out in a different form as:

$$f_{i+1,j+1} = f_{i,j} + \left[(\delta x)_{i,j} f_x + (\delta y)_{i,j} f_y\right]$$

$$+ \frac{1}{2!} \left[(\delta x)^2_{i,j} f_{xx} + 2(\delta x)_{i,j}(\delta y)_{i,j} f_{xy} + (\delta y)^2_{i,j} f_{yy}\right]$$

$$+ \frac{1}{3!} \left[(\delta x)^3_{i,j} f_{xxx} + 3(\delta x)^2_{i,j}(\delta y)_{i,j} f_{xxy} + 3(\delta x)_{i,j}(\delta y)^2_{i,j} f_{xyy} + (\delta y)^3_{i,j} f_{yyy}\right]$$

$$+ \left[(\delta x)^3_{i,j} f_{yxy}\right] + \ldots$$  \hspace{1cm} (3.3)

Here we use the notation:

$$(\delta x)_{i,j} \equiv x_{i+1,j} - x_{i,j} \quad \text{and} \quad (\delta y)_{i,j} \equiv y_{i,j+1} - y_{i,j}$$  \hspace{1cm} (3.4)

and it is understood that all of the derivatives are evaluated at the point $(x_i, y_j)$. 

13
3.2 Second Order Differential Equations

The mathematical formulation for most atmospheric sciences problems implies change rates with respect to two or more independent variables, mostly time and space. This will lead to partial differential equation or set of these equations. Let us take the following general second order equation with two dimensions $x$ and $y$:

$$a \frac{\partial^2 \phi}{\partial x^2} + b \frac{\partial^2 \phi}{\partial x \partial y} + c \frac{\partial^2 \phi}{\partial y^2} + d \frac{\partial \phi}{\partial x} + e \frac{\partial \phi}{\partial y} + f \phi + g = 0$$  \hspace{1cm} (3.5)

Where $a$, $b$, $c$, $d$, $e$, $f$ and $g$ are either constants or more generally functions of $x$ and $y$ and of the dependent variable $\phi$.

This equation is,

1. Elliptic when $b^2 - 4ac < 0$
2. Parabolic when $b^2 - 4ac = 0$
3. Hyperbolic when $b^2 - 4ac > 0$

3.3 Three Important Equations

From equation (3.5):

1. If we put $\phi=T$, $x=x$, $y=t$, $d=k$, $e=1$, and the other coefficients $a=b=c=f=g=0$ we have the advection equation:

$$\frac{\partial T}{\partial t} + k \frac{\partial T}{\partial x} = 0$$  \hspace{1cm} (3.6)

If we put $\phi=u$, $x=x$, $y=t$, $a=c^2$, $c=-1$, and the other coefficients $b=d=e=f=g=0$ we have the wave equation:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$  \hspace{1cm} (3.7)

If we put $\phi=T$, $x=x$, $y=t$, $a=k$, $e=-1$, and the other coefficients $b=c=d=f=g=0$ we have the diffusion equation:

$$\frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2}$$  \hspace{1cm} (3.8)

The above three equations have great importance in science and can be classified into elliptic, parabolic, and hyperbolic (how?).

Determination of the equation type is essential because of its relation to the nature of the initial and boundary conditions in one hand, and how each PDE can be solved on the other hand.
3.3 Representing the dependent variables on the grid

Representation of variables on the rectangular mesh (or grid) is useful to understand and employ the finite difference approach. The equations 3.6-3.8 are PDE’s and contain two independent variables, $x$ and $t$.

To make a two dimensional grid, do:

1. Divide the plane of $x$ and $t$ into sets of rectangles.
2. Assume $\Delta x = h$, $\Delta t = k$.
3. Hence $x = ih$, $t = jk$ where $i = 0,1,2, \ldots$ $j = 0,1,2, \ldots$

If the dependent variable is $u$ we can write the following Taylor expansions:

$$u_{i+1,j} = u_{i,j} + \left(\frac{\partial u}{\partial x}\right)_{i,j} h + \frac{1}{2} \left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} h^2 + \frac{1}{6} \left(\frac{\partial^3 u}{\partial x^3}\right)_{i,j} h^3 + \cdots \tag{3.9}$$

and

$$u_{i-1,j} = u_{i,j} - \left(\frac{\partial u}{\partial x}\right)_{i,j} h + \frac{1}{2} \left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} h^2 - \frac{1}{6} \left(\frac{\partial^3 u}{\partial x^3}\right)_{i,j} h^3 + \cdots \tag{3.10}$$

It is clear that the two equations applies on $x$-axis. Now we can get the following equations from eq 3.9 and eq 3.10:

$$\left(\frac{\partial u}{\partial x}\right)_{i,j} \approx \frac{u_{i+1,j} - u_{i,j}}{h} \quad \text{forward difference} \tag{3.11}$$

$$\left(\frac{\partial u}{\partial x}\right)_{i,j} \approx \frac{u_{i,j} - u_{i-1,j}}{h} \quad \text{backward difference} \tag{3.12}$$

Note that the $3^{rd}$ and $4^{th}$ terms on the right hand side were truncated with a truncation error of order $O(h^2)$. That means the terms that contain the second order derivative and above were truncated.

Now, by subtracting eq 3.10 from 3.9 and neglecting the terms that contain $3^{rd}$ derivative and above we can get:

$$\left(\frac{\partial u}{\partial x}\right)_{i,j} \approx \frac{u_{i+1,j} - u_{i-1,j}}{2h} \quad \text{centered difference} \tag{3.13}$$

Adding the equations 3.9 to 3.10 yields a formula of second order derivative:

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} \approx \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \tag{3.14}$$
Note that the same equations can be written to the other coordinate (t). The finite difference grid can be graphically represented as in figure (3.1).

**Exercises**

Q1. From the general equation (4.1) derive each of advection, wave, and diffusion equations by making some suitable assumptions. Then classify these equations according to the type.

Q2. Why the determination of the type of PDE is important?

Q3. Speaking about Taylor expansion, what do we mean by truncation error of the order $O(h^2)$?

**MATLAB Work**

Try writing a matlab script to solve equation 3.14. (Use internet)

**Homework**

3. Rewrite the equations 3.9-3.14 with respect to the time (t) coordinate.

4. Classify the equations 3.6-3.8 to their types.
Lecture (4)

Approximating the Solution of a Differential Equation

4.1 Introduction

In this lecture, we will discuss the truncation error and the accuracy of finite-difference scheme, which is defined as a finite-difference equation that approximates, term-by-term, a differential equation. Our goal is to find an approximation to the solution of the differential equation.

As an example, we will analyze the one-dimensional advection equation, given by:

$$\frac{\partial A}{\partial t} + c \frac{\partial A}{\partial x} = 0$$

where $A=A(x,t)$. The physical meaning of Eq.(4.1) is that $A$ remains constant at the position of a particle that moves in the x-direction with speed $c$.

Suppose that:

$$A(x,0) = F(x) \quad \text{for } -\infty < x < \infty \quad (4.2)$$

This is an “initial condition.” Our goal is to determine $u(x,t)$. This is a simple example of an initial value problem.

For making a comparison, we first find out the analytical solution of Eq.(4.1). Define:

$$\xi \equiv x - ct \quad \quad (4.3)$$

Using the chain rule, we write:

$$\left( \frac{\partial A}{\partial x} \right)_\xi = \left( \frac{\partial A}{\partial x} \right)_t + \left( \frac{\partial A}{\partial t} \right)_x \frac{\partial x}{\partial \xi} = \left( \frac{\partial A}{\partial x} \right)_t + \left( \frac{\partial A}{\partial t} \right)_x \frac{1}{c}$$

$$= 0 \quad \quad (4.4)$$

The first line of eq(4.4) can be understood by reference to Fig. 4.1. The last line comes by use of Eq.(4.1).
Similarly, 

\[
\frac{\partial A}{\partial t}\bigg|_{\xi} = \left(\frac{\partial A}{\partial t}\right)_x + \left(\frac{\partial A}{\partial x}\right)_t \frac{\partial x}{\partial t}\bigg|_{\xi}
\]

\[
= \left(\frac{\partial A}{\partial t}\right)_x + \left(\frac{\partial A}{\partial x}\right)_t c\n\]

\[
= 0 \quad (4.5)
\]

Again, the last equality follows from eq 4.1. We can interpret \(\frac{\partial A}{\partial t}\bigg|_{\xi}\) as the time rate of change of \(A\) that we would see if we were riding on the moving particle of air, so the advection equation describe what happens as particles of air move around without changing their values of \(A\).

From Eqns. (4.4) and (4.5), we conclude that:

\[
A = f(\xi) \quad (4.6)
\]

is the general solution to Eq. (4.1). This means that \(A\) depends only on \(\xi\), in the sense that if you tell me the value of \(\xi\), that is all the information I need to tell you the value of \(A\). (Note that “\(A\) depends only on \(\xi\)” does not mean that \(A\) is independent of \(x\) for fixed \(t\), or of \(t\) for fixed \(x\).)

The initial condition is:

\[
\xi \equiv x \quad and \quad A(x) = f(x) \quad at \quad t = 0 \quad (4.7)
\]
i.e., the shape of $f(\xi)$ is determined by the initial condition. Eq. (4.6) means that $A$ is constant along the line $\xi=constant$. In order to satisfy the initial condition, we chose $f \equiv F$ [See Eq. (4.2)]. Referring to Eq.(4.6), we see that $A(\xi) = F(\xi) \equiv F(x - ct)$ is the solution to Eq. (4.1) that satisfies the initial condition Eq. (4.7). An initial value simply “moves along” the lines of constant $\xi$, which are called characteristics. The initial shape of $A(x)$, namely $F(x)$, is just carried along by the wind. From a physical point of view this is obvious. Partial differential equations whose solutions are constant along characteristics are called *hyperbolic* equations. The advection equation is hyperbolic.

### 4.2 The Upstream Scheme

Keeping in mind the exact solution, we now investigate the solution of one possible numerical scheme for Eq. (4.1). We construct a grid, as in Fig. 4.2. One of the infinitely many possible finite difference approximations to Eq. (4.1) is:

$$\frac{A_{i,j+1} - A_{i,j}}{\Delta t} + c \left( \frac{A_{i,j} - A_{i-1,j}}{\Delta x} \right) = 0 \quad (4.8)$$

For $c>0$, Eq. (4.8) is called the “upstream” scheme. It is one-sided or asymmetric in both space and time. It seems naturally suited to modeling advection, in which air comes from one side and goes to the other, as time passes by.

![Fig. 4.2: A grid for the solution of the one-dimensional advection equation](image-url)
Because:

\[
\frac{A_{i,j+1} - A_{i,j}}{\Delta t} \rightarrow \frac{\partial A}{\partial t} \quad \text{as } \Delta t \rightarrow 0 \quad (4.9)
\]

and

\[
\frac{A_{i,j} - A_{i-1,j}}{\Delta x} \rightarrow \frac{\partial A}{\partial x} \quad \text{as } \Delta x \rightarrow 0 \quad (4.10)
\]

we conclude that Eq. (4.8) does approach Eq. (4.1) as \(\Delta t\) and \(\Delta x\) both approaches zero. In view of Eq. (4.9) and Eq. (4.10), it may seem obvious that the solution of Eq. (4.8) approaches the solution of Eq. (4.1) as \(\Delta x \rightarrow 0\) and \(\Delta t \rightarrow 0\). This “obvious” conclusion is not necessarily true.

### 4.3 The Truncation Error of a Finite-Difference Scheme

Let \(A(x,t)\) denote the (exact) solution of the differential equation [Analytical sol], so that \(A(i\Delta x, j\Delta t)\) is the value of this exact solution at the discrete point \((i\Delta x, j\Delta t)\) on the grid shown in Fig. 4.2.

We use the notation \(A_{i,j}\) to denote the “exact” solution of a finite-difference equation, at the same point.

In general, \(A_{i,j} \neq A(i\Delta x, j\Delta t)\). We wish that they were equal!

A measure of the accuracy of the finite-difference scheme can be obtained by substituting the solution of the differential equation into the finite-difference equation. For the upstream scheme given by Eq. (4.8), we get:

\[
\left\{ \frac{A[i\Delta x, (j + 1)\Delta t] - A(i\Delta x, j\Delta t)}{\Delta t} \right\} + c \left\{ \frac{A(i\Delta x, j\Delta t) - A[(i - 1)\Delta x, j\Delta t]}{\Delta x} \right\} = \varepsilon \quad (4.11)
\]

where \(\varepsilon\) is called the “truncation error” of the scheme. The truncation error of the scheme is a measure of how accurately the solution \(A(x,t)\) of the original differential equation (4.1), satisfies the finite-difference equation, (4.8). It is far from a perfect measure of accuracy, however, as will become evident.

Because \(A_{i,j}\) is defined only at discrete points, it is not differentiable, and so we cannot substitute \(A_{i,j}\) into the differential equation. Because of this, we cannot measure how accurately \(A_{i,j}\) satisfies the differential equation.
If we obtain the terms of (4.11) from a Taylor Series expansion of $A(x,t)$ about the point $(i\Delta x, j\Delta t)$, and use the fact that $A(x,t)$ satisfies (4.1), we find that:

$$\varepsilon = \left( \frac{1}{2!} \Delta t \frac{\partial^2 A}{\partial t^2} + \cdots \right) + c \left( - \frac{1}{2!} \Delta x \frac{\partial^2 A}{\partial x^2} + \cdots \right)$$

(4.12)

We say this is a “first-order scheme” because the lowest powers of $\Delta t$ and $\Delta x$ in Eq. (4.12) are 1. The notations $O(\Delta t, \Delta x)$ or $O(\Delta t) + O(\Delta x)$ can be used to express this. The upstream scheme is first order accurate in both space and time. A scheme is said to be “consistent” with the differential equation if the truncation error of the scheme approaches zero as $\Delta t$ and $\Delta x$ approach zero. The upstream scheme is consistent. Consistency is necessary for a good scheme, but it is nowhere near sufficient.

**Exercises**

Q1. Write a formula of advection differential equation and write the corresponding finite difference equation. Then find out the general solution of this equation and express the initial condition for this differential equation.

Q2. Discuss how the finite difference equation approaches its corresponding advection differential equation.

Q3. Explain the truncation error of a finite-difference scheme.

**MATLAB Work**

Try writing a matlab script to import data from excel worksheet and doing the calculations of pressure in the matlab example in Lecture 2. Then export the results to another excel worksheet.

**Homework**

Derive Eq. (4.12) from Eq. (4.11) and Taylor expansion.
Lecture (5)

Finite Difference of Parabolic Equation

5.1 An Explicit Method of Solution

An الطريقة الصريعة للحل هي الطريقة التي تعبر عن قيمة مجهولة عند نقطة شبكة معينة بدلالة قيم معلومة عند نقاط مجاورة. باستخدام المعادلات (4.7) و (4.10) في معادلة الانتشار (التي هي من نوع

\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \] (parabolic)

\[ u_{i,j+1} - u_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{k} \]

And this can be written as:

\[ u_{i,j+1} = u_{i,j} + r(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}) \] (5.1)

where \( r = \frac{k}{h^2} \)

وعليه نستطيع حساب قيمة \( u \) على طول الصف الزمني الأول بدالة قيم حدود وقيم أولية على طول.

Example: Consider a metal rod is heated for a long time in the center and keeping the ends in contact with blocks of melting ice and that the initial temperature distribution in non-dimensional form is:

(a) \( u = 2x, \quad 0 \leq x \leq \frac{1}{2} \)

(b) \( u = 2(1 - x), \quad \frac{1}{2} \leq x \leq 1 \).

In other words we are seeking a numerical solution of \( \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \) which satisfies,

1. \( u = 0 \) when \( x = 0 \) and 1 for all \( t \). (the boundary conditions)
2. \( u = 2x \) for \( 0 \leq x \leq \frac{1}{2} \), and

\( u = 2(1 - x) \) for \( \frac{1}{2} \leq x \leq 1 \) \( t = 0 \). (the initial conditions)

For \( h=1/10 \), the initial values and boundary values are as shown in in table (5.1). The problem is symmetric with respect to \( x=1/2 \) so we need the solution only for \( 0 \leq x \leq \frac{1}{2} \).
Table 5.1

<table>
<thead>
<tr>
<th>j=0</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>j=1</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>j=2</td>
<td>0</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>j=3</td>
<td>0</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>j=4</td>
<td>0</td>
<td>0.5</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
<td>0.8</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Take \( h=1/10, \ k=1/1000 \) so \( r=1/10 \) equation (5.1) then reads as:

\[
\frac{u_{i,j+1}}{10} = \frac{1}{10} (u_{i-1,j} - 8u_{i,j} + u_{i+1,j})
\]

Table 5.2

<table>
<thead>
<tr>
<th>i=0</th>
<th>i=1</th>
<th>i=2</th>
<th>i=3</th>
<th>i=4</th>
<th>i=5</th>
<th>i=6</th>
</tr>
</thead>
<tbody>
<tr>
<td>x=0</td>
<td>t=0.000</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>(j=0)</td>
<td>t=0.001</td>
<td>0.2000</td>
<td>0.4000</td>
<td>0.6000</td>
<td>0.8000</td>
<td>1.0000</td>
</tr>
<tr>
<td>(j=2)</td>
<td>t=0.002</td>
<td>0.2000</td>
<td>0.4000</td>
<td>0.6000</td>
<td>0.7960</td>
<td>0.9280</td>
</tr>
<tr>
<td>(j=4)</td>
<td>t=0.004</td>
<td>0.2000</td>
<td>0.4000</td>
<td>0.5996</td>
<td>0.7896</td>
<td>0.9016</td>
</tr>
<tr>
<td>(j=5)</td>
<td>t=0.005</td>
<td>0.2000</td>
<td>0.3999</td>
<td>0.5971</td>
<td>0.7732</td>
<td>0.8597</td>
</tr>
<tr>
<td>(j=10)</td>
<td>t=0.000</td>
<td>0.1996</td>
<td>0.3968</td>
<td>0.5822</td>
<td>0.7281</td>
<td>0.7867</td>
</tr>
<tr>
<td>(j=20)</td>
<td>t=0.020</td>
<td>0.1938</td>
<td>0.3781</td>
<td>0.5373</td>
<td>0.6486</td>
<td>0.6891</td>
</tr>
</tbody>
</table>

Fig. (5.2) Molecular representation
ان الحل التحليلي للمعادلة التفاضلية الجزئية الذي يحقق هذه الشروط هو:

\[ u = \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \left( \sin \frac{1}{2} n \pi \right) \left( \sin n \pi x \right) \exp \left( -n^2 \pi^2 t \right) \]

ان مقارنة هذا الحل التحليلي مع الفرق المحدد عند \( x=0.3 \) ، كما في ادناء، يظهر ان حل الفرق المحدد ذو دقة مقبولة. ان الخطأ المنوي هو فرق الحلول معبراً عنه بنسبة منوية للحل التحليلي للمعادلة التفاضلية الجزئية.

<table>
<thead>
<tr>
<th>Table 5.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time step</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>t=0.005</td>
</tr>
<tr>
<td>t=0.01</td>
</tr>
<tr>
<td>t=0.02</td>
</tr>
<tr>
<td>t=0.10</td>
</tr>
</tbody>
</table>

### 5.2 Exercises

Q1. For the following PDE: \( \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \), find the numerical solution at the grid points in the first time row by using the explicit method. Consider the following boundary conditions: \( u=4.56 \) at \( x=0 \) and \( u=1.56 \) at \( x=1 \)  \( t \geq 0 \)

\( u=3(1.52 – x) \) for \( 0 \leq x \leq 1 \)  \( t = 0 \)

knowing that \( h=2/10 \) , \( k=3/1000 \).

Then, suppose that \( u=0.668 \) (Numerical solution) and \( u=0.665 \) (Analytical solution) calculate the difference and percentage error.

Q2. Consider the differential equation, \( \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \) with the following boundary conditions: \( u=3x+1 \) for \( 0 \leq x \leq 1.2 \) find the numerical solution at the grid points in the first row ( Take \( \Delta x = h = 3/10 \), \( \Delta t = k = 1/1000 \) )
6.1 Criteria of optimal numerical solution

The exact solution of a numerical problem can be obtained by a finite-difference method if the partial differential equations are well posed and the numerical solution converges to the exact solution as the grid spacing is reduced. The convergence of the solution is governed by the consistency and stability of the numerical scheme.

Consistency requires that the finite-difference analog of the differential equation converges to the differential equation as the grid spacing is reduced.

\[
\frac{\partial N}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta N}{\Delta x}
\]  

(6.1)

This ensures that the truncation error is small.

Consistency ensures that the order of approximation is equal to the order of the differential equation.

\[
\lim_{\Delta x \to 0} \left\| TE \left( \frac{\Delta N}{\Delta x} \right) \right\| = 0
\]  

(6.2)

where \( TE \) is the truncation error of the approximation \( \Delta N/\Delta x \).

Stability requires that the solution does not grow unboundedly as the solution progresses through time.

Stability is ensured by the spectral radius of the matrix representation of the linear system being less than or equal to one.

\[
R = \frac{1}{N} \text{max} \left| \sum_{n=1}^{N} A_{n,n} \right|
\]

(6.3)

where \( A \) is the matrix of the linear system.

A stable numerical scheme ensures that the solution is well-behaved and does not oscillate or diverge.

Stability is ensured by the spectral radius of the matrix representation of the linear system being less than or equal to one.

\[
R = \frac{1}{N} \text{max} \left| \sum_{n=1}^{N} A_{n,n} \right|
\]

(6.3)

where \( A \) is the matrix of the linear system.

A stable numerical scheme ensures that the solution is well-behaved and does not oscillate or diverge.

Stability is ensured by the spectral radius of the matrix representation of the linear system being less than or equal to one.

\[
R = \frac{1}{N} \text{max} \left| \sum_{n=1}^{N} A_{n,n} \right|
\]

(6.3)

where \( A \) is the matrix of the linear system.

A stable numerical scheme ensures that the solution is well-behaved and does not oscillate or diverge.

Stability is ensured by the spectral radius of the matrix representation of the linear system being less than or equal to one.

\[
R = \frac{1}{N} \text{max} \left| \sum_{n=1}^{N} A_{n,n} \right|
\]

(6.3)

where \( A \) is the matrix of the linear system.

A stable numerical scheme ensures that the solution is well-behaved and does not oscillate or diverge.

Stability is ensured by the spectral radius of the matrix representation of the linear system being less than or equal to one.

\[
R = \frac{1}{N} \text{max} \left| \sum_{n=1}^{N} A_{n,n} \right|
\]

(6.3)
Numerical W. Modeling / 4th Year/ Atm. Sci. Dept

Dr. Thaer Obaid Roomi

A high order approximation can be made if the error is small enough. However, the high order approximation requires more variables to be calculated than the low order approximation. For the high order approximation, more variables are required to describe the same order of magnitude. However, because the high order approximation contains more variables, it requires more calculations than the low order approximation. The results of the high order approximation are more accurate than the low order approximation. However, if the high order approximation is not convergent, it is not useful. The fourth order approximation is convergent when the time-step size is small enough. However, if the time-step size is not small enough, the fourth order approximation is not convergent. The fifth order approximation is convergent if the time-step size is small enough. However, if the time-step size is not small enough, the fifth order approximation is not convergent. Therefore, the fifth order approximation is convergent if the time-step size is small enough. However, if the time-step size is not small enough, the fifth order approximation is not convergent. The time-step size is a constant.

where C is a constant.
6.2 Exercises

Q1. Discuss numerical convergency.

Q2. Discuss numerical stability.

Q3. Discuss numerical consistency.

Q4. Define order of approximation.

Q5. Define truncation error.

Q6. Discuss the relationship between the truncation error and order of approximation.

Q7. Why do we truncate Taylor Expansion though we know that this will lessen the accuracy?
Lecture (7)

The Effect of Changing Time Step

(Explicit scheme)

7.1 The Effect of Changing Time Step

In this lecture, we will discuss the effect of changing time step on the solution of the partial differential equations, using the explicit scheme.

Example (5.2) and the time step length \( h = \frac{1}{10} \), \( k = \frac{5}{1000} \):

\[ u_{i,j+1} = u_{i,j} + r(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}) \quad (7.1) \]

This becomes:

\[ u_{i,j+1} = \frac{1}{2}(u_{i-1,j} + u_{i+1,j}) \quad (7.2) \]

The obtained solution by applying the differentiated equations for these with boundary and initial conditions:

The first row is recorded in Table (7.1):

<table>
<thead>
<tr>
<th></th>
<th>( i=0 )</th>
<th>( i=1 )</th>
<th>( i=2 )</th>
<th>( i=3 )</th>
<th>( i=4 )</th>
<th>( i=5 )</th>
<th>( i=6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x=0 )</td>
<td>( u_{0,0} )</td>
<td>( u_{0.1} )</td>
<td>( u_{0.2} )</td>
<td>( u_{0.3} )</td>
<td>( u_{0.4} )</td>
<td>( u_{0.5} )</td>
<td>( u_{0.6} )</td>
</tr>
<tr>
<td>( t=0.000 )</td>
<td>0.8000</td>
<td>1.0000</td>
<td>0.8000</td>
<td>0.6000</td>
<td>0.4000</td>
<td>0.2000</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.005</td>
<td>0.7000</td>
<td>0.8000</td>
<td>0.7000</td>
<td>0.5500</td>
<td>0.3750</td>
<td>0.2000</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.010</td>
<td>0.6250</td>
<td>0.7000</td>
<td>0.6250</td>
<td>0.5500</td>
<td>0.3750</td>
<td>0.2000</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.015</td>
<td>0.2778</td>
<td>0.3071</td>
<td>0.2778</td>
<td>0.2484</td>
<td>0.1717</td>
<td>0.0949</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.020</td>
<td>0.2778</td>
<td>0.3071</td>
<td>0.2484</td>
<td>0.2778</td>
<td>0.3071</td>
<td>0.2778</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Table (7.2)

<table>
<thead>
<tr>
<th>t</th>
<th>Finite–difference Solution (x=0.3)</th>
<th>Analytical solution (x=0.3)</th>
<th>Difference</th>
<th>Percentage error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.005</td>
<td>0.6000</td>
<td>0.5966</td>
<td>0.0034</td>
<td>0.57</td>
</tr>
<tr>
<td>0.01</td>
<td>0.6000</td>
<td>0.5799</td>
<td>0.0201</td>
<td>3.5</td>
</tr>
<tr>
<td>0.02</td>
<td>0.5500</td>
<td>0.5334</td>
<td>0.0166</td>
<td>3.1</td>
</tr>
<tr>
<td>0.1</td>
<td>0.2484</td>
<td>0.2444</td>
<td>0.0040</td>
<td>1.6</td>
</tr>
</tbody>
</table>

In Table (7.2), observe that the finite-difference solution is not a good approximation for the analytical solution at x=0.3.

Table (7.3)

<table>
<thead>
<tr>
<th>i</th>
<th>x = 0</th>
<th>i = 1</th>
<th>i = 2</th>
<th>i = 3</th>
<th>i = 4</th>
<th>i = 5</th>
<th>i = 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>t=0.000</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>0.6</td>
<td>0.6</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.4</td>
<td>1.0</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
<td>1.2</td>
<td>-0.2</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>0.0</td>
<td>0.0</td>
<td>1.4</td>
<td>-1.2</td>
<td>2.6</td>
<td>-1.2</td>
<td></td>
</tr>
</tbody>
</table>

Table (7.3) shows the values of u at different times and locations for the explicit method.

### Finite Difference Solution

The finite difference solution uses the formula:

$$u_{i,j+1} = u_{i-1,j} - u_{i,j} + u_{i+1,j}$$

This formula is used to calculate the values of $u$ at the next time step.

### Analytical Solution

The analytical solution for the equation is:

$$\frac{d^2u}{dx^2} = \frac{1}{r} \frac{du}{dt}$$

For $r=1$, the analytical solution is:

$$u(x,t) = \sin(x) \sin(t)$$

### Percentage Error

The percentage error is calculated as:

$$\text{Percentage Error} = \frac{|\text{Analytical solution} - \text{Finite difference solution}|}{\text{Analytical solution}} \times 100$$

The table above shows the percentage error for different times.

### Discussion

In Table (7.2), it is observed that the finite difference solution is not a good approximation for the analytical solution at x=0.3. However, the percentage error is relatively small, indicating that the finite difference solution is a reasonable approximation for practical purposes.

### Explicit Method

The explicit method is a simple and straightforward numerical method for solving differential equations. It is based on the finite difference approximation of the derivatives. The implicit method, on the other hand, requires solving a system of linear equations, which can be computationally more demanding.

### Conclusion

The finite difference method is a useful tool for approximating solutions to differential equations, especially when an analytical solution is not available or is too complex. The explicit method is particularly useful for problems where the solution changes rapidly over time or space, as it allows for simpler and faster computations.
Figure (7.1) Analytical and finite – difference solutions

7.2 Exercise

Q1. Explain the effect of changing the time step on the accuracy of the finite difference solution.
8.1 Grank – Nicolson Implicit Method

رغم أن الطريقة الصريحة الصعبة للحساب، فإنها سهلة حسابياً. ان خطوة الزمن يجب أن تكون $k \leq \frac{1}{2} h^2$ وأن $h$ يجب أن يكون صغيرة جداً لأن العملية تكون قبلاً فقط لـ $0 \leq k/h^2 \leq \frac{1}{2}$ أي أن تكون صغيرة لغرض احترام دقة المعولمة. اقترح واستخدم كرانك ونيكولسون عام 1947 طريقة لتقلص الحجم الكلي للحسابات وهي قبلاً (valid) أي أنها منظمة ومستقرة (convergent and stable) لجميع القيم المحددة لـ $r$. لقد تم استبدال العلاقة السابقة للفرفقات المحصلة كالآتي:

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

$$\frac{u_{i,j+1} + u_{i,j}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \text{ replaced by}$$

$$\frac{u_{i,j+1} + u_{i,j}}{k} = \frac{1}{2} \left\{ \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2} + \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \right\} \text{ giving}$$

$$-ru_{i-1,j+1} + (2 + 2r)u_{i,j+1} - ru_{i,j+1} = ru_{i-1,j} + (2 - 2r)u_{i,j} + ru_{i+1,j} \quad (8.1)$$

عموماً الجانب الأيسر من المعادلة (8.1) يحتوي ثلاثة م الجهاز والجانب الأيمن يحتوي ثلاثة معايير، لقيم المحورية (الشكل 8.1) إذا وجدت $N$ من نقاط الشبكة الداخلية على طول كل صفح زمني وعدد لكل صفح زمني $j=0$ and $i=1,2,\ldots,N$ تم إعطاء المعادلة (8.1) تعطي $N$ من المعادلات الأدنى للقيم المحورية المحولة على طول الصف الزمني الأول بدلاً من القيم الأولية والقيم الحدودية المعولمة. وبالمثل، صفح زمني ثاني بدلاً من القيم المحولة على طول $u$ يعبر عن $N$ من القيم المحولة لـ $j=1$ وتظل هذه الطريقة حيث حساب القيم المحورية المحولة تتطلب حل لمجموعة من المعادلات الآلية توصف بأنها طريقة ضمن.

![](Fig. (8.1))
Example 8.1: Use the Grank-Nicolson method to calculate a numerical solution of the previous worked example, namely, \( \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \) \((0 \leq x \leq 1)\),

where (i) \( u = 0 , x = 0 \) and \( t \geq 0 \),

(ii) \( u = 2x , 0 \leq x \leq \frac{1}{2} \), \( t = 0 \),

(iii) \( u = 2(1 - x) , \frac{1}{2} \leq x \leq 1 \), \( t = 0 \),

Take \( h = 1/10 \). Although the method is valid for all finite values of \( r = k/h^2 \), a large value will yield an inaccurate approximation for \( \frac{\partial u}{\partial t} \). A suitable value is \( r = 1 \) and has the advantage of making the coefficient of \( u_{i,j} \) zero in (8.1).

The eq. (8.1) then reads as:

\[-u_{i-1,j+1} + 4u_{i,j+1} - u_{i+1,j+1} = u_{i-1,j} + u_{i+1,j} \] (8.2)

The computational molecule corresponding to equation (8.2) is shown in Fig. 8.2. Denote \( u_{i,j+1} \) by \( u_i \) \((i = 1, 2, ..., 9)\). For this problem, because of symmetry, \( u_6 = u_4 \), \( u_7 = u_3 \), etc. Fig (8.3).

The values of \( u \) for the first time step then satisfy:

\[-0 + 4u_1 - u_2 = 0 + 0.4,\]
\[-u_1 + 4u_2 - u_3 = 0.2 + 0.6,\]
\[-u_2 + 4u_3 - u_4 = 0.4 + 0.8,\]
\[-u_3 + 4u_4 - u_5 = 0.6 + 1.0,\]
\[-2u_4 + 4u_5 = 0.8 + 0.8.\]

These equations are easily solved by systematic eliminations to give:

\( u_1 = 0.1989, \ u_2 = 0.3956, \ u_3 = 0.5834, \ u_4 = 0.7381, \ u_5 = 0.7691\)

Hence the equations for the pivotal values of \( u \) along the next time row are:
-0 + 4u_1-u_2 = 0 +0.3956,
-u_1 +4u_2-u_3 =0.1989+0.5834,
-u_2+4u_3-u_4 = 0.3956+0.7381,
-u_3+4u_4-u_5 = 0.5834+0.7691,
-2u_4+4u_5 = 2×0.7381

The solution of these equations is given in Table (8.1). There is a comparison with the solution of the partial differential equations at t=0.1. The numerical solution is clearly a good one.

<table>
<thead>
<tr>
<th>x = 0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>t=0.00</td>
<td>0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>t=0.01</td>
<td>0</td>
<td>0.1989</td>
<td>0.3956</td>
<td>0.5834</td>
<td>0.7381</td>
</tr>
<tr>
<td>t=0.02</td>
<td>0</td>
<td>0.1936</td>
<td>0.3789</td>
<td>0.5400</td>
<td>0.6461</td>
</tr>
<tr>
<td>Analytical solution</td>
<td>t=0.10</td>
<td>0</td>
<td>0.0948</td>
<td>0.1803</td>
<td>0.2482</td>
</tr>
</tbody>
</table>

Table (8.2) below displays both solutions at x=0.5 for various values of t. It was found that the accuracy of this implicit method over the time range taken is about the same as for the explicit method which uses ten times as many time steps.

| t=0.01 | 0.7691 | 0.7743 | -0.0052 | -0.7 |
| t=0.02 | 0.6921 | 0.6809 | +0.0112 | +1.6 |
| t=0.10 | 0.3069 | 0.3021 | 0.0048 | 1.6 |

**Example 8.1:** Use Crank-Nicolson method to calculate the numerical solution of the following PDE (at the first row only):
\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}
\]

where u=4.5 at x=0 and u=1.5 at x=1, t ≥ 0

\[u=3(1.5-x)\] at t=0, take h=0.2, r=1.

Sol. 1. We write the above equation in finite difference method:

\[-r u_{i-1,j+1} + (2 + 2r) u_{i,j+1} - r u_{i+1,j+1} = r u_{i-1,j} + (2 - 2r) u_{i,j} + r u_{i+1,j}\]

where \(r=k/h^2\). At \(r=1\), we get:
2. We draw the grid and the molecule

3. Establishing initial values

\[ u = 3(1.5 - 0.2) = 3.9 \]
\[ u = 3(1.5 - 0.4) = 3.3 \]
\[ u = 3(1.5 - 0.6) = 2.7 \]
\[ u = 3(1.5 - 0.8) = 2.1 \]

4. Set the instantaneous equs.

\[ -4.5 + 4u_1 - u_2 = 4.5 + 3.3 \quad \ldots (1) \]
\[ -u_1 + 4u_2 - u_3 = 3.9 + 2.7 \quad \ldots (2) \]
\[ -u_2 + 4u_3 - u_4 = 3.3 + 2.1 \quad \ldots (3) \]
\[ -u_3 + 4u_4 - 1.5 = 2.7 + 1.5 \quad \ldots (4) \]

5. Finding u-values

From eq. 4

\[ 4u_4 = u_3 + 5.7 \quad \Rightarrow u_4 = 0.25u_3 + 1.425 \quad \ldots (5) \]
we substitute the value of \( u_4 \) in eq. 3

\[ -u_2 + 4u_3 - 0.25u_3 - 1.425 = 5.4 \quad \Rightarrow -u_2 + 3.75u_3 = 6.825 \quad \Rightarrow u_3 = 0.2666u_2 + 1.82 \quad \ldots (6) \]

Sub. in eq2

\[ -u_1 + 4u_2 - 0.2666u_2 - 1.82 = 6.6 \quad \Rightarrow -u_1 + 3.7334u_2 = 8.42 \quad \Rightarrow u_2 = 0.2678u_1 + 2.2553 \quad \ldots (7) \]
Sub. in eq1

\[ -4.5 + 4u_1 - 0.2678u_1 + 2.2553 = 7.8 \quad \Rightarrow 4.5 + 3.7322u_1 = 10.0553 \quad \Rightarrow 3.7322u_1 = 14.5553 \quad \Rightarrow u_1 = 3.8999 \]
substituted in eq. 7 \( u_2 = 0.2678 \times 3.8999 + 2.2553 = 3.2996 \), Sub. in eq6

\[ u_3 = 0.2666 \times 3.2996 + 1.82 = 2.6996 \], Sub. in eq.5

\[ u_4 = 0.25 \times 2.6996 + 1.425 = 2.0999 \]
Lecture (9)

Time – Stepping Schemes and CFL Criterion

9.1 The differences between the explicit and the implicit solution of finite- difference method

طريقة الحل الصريح Explicit هي التعبير عن قيمة مجاهلة عند نقطة شبكة معينة بدلالة قيمة معينة عند نقاط شبكة نقاط متتابعة للحل يمكن الحصول بها على قيمة المتغير المعتمد (الحل) عند نقاط مجاورة، وهي طريقة بسيطة حسابياً للحل يمكن الحصول بها على قيمة المتغير المعتمد (الحل) بشكل مباشر. هذه الطريقة سلبية وهي أن خطوة الزمن يجب أن تكون بالضرورة صغيرة جدًا لأن العملية يجب أن تكون صغيرة لفرض احترام دقة معقولة وهذا يؤدي إلى أن $h$ وأن $k \leq \frac{1}{2} h^2$ وذات القوة فقط لـ $k$ تكون العملية الحسابية كبيرة جداً.

طريقة الحل الضمني Implicit هي التعبير عن الحل على شكل معادلات آنية بعدد نقاط الشبكة عند السطر الأول بدالة القيم الأولية والقيم الحدودية المعلومة. هذه الطريقة تلزم حل المعادلات الآنية للحصول على الحل العددي. أن هذه الطريقة تقتصر الحجم الكلي للحسابات وهي مقبولة (أي أنها متقاربة ومستقرة) لجميع قيم $r = k/h^2$. هذه الطريقة لا يتوجب فيها أن تكون $k$ صغيرة جداً قياساً بـ $h$ ويمكن أن يتزايد.

9.2 Advection-Diffusion Equations

معادلات التأقث – الانتشار

وهي نوع شائع من المعادلات التفاضلية الجزئية التي تستخدم في علوم الجو وفي أنظمة النتربة الجوي العددي. ويمكن كتابة معادلة الاستمرارية النوعية بصيغة معادلة تأقث – الانتشار على النحو التالي:

\[
\frac{\partial N}{\partial t} + \frac{\partial (uN)}{\partial x} - \frac{\partial}{\partial y} \left( K_{h,xx} \frac{\partial N}{\partial x} \right) = 0 \quad (9.1) \quad \text{west - east eq.} \\
\frac{\partial N}{\partial t} + \frac{\partial (vN)}{\partial y} - \frac{\partial}{\partial y} \left( K_{h,yy} \frac{\partial N}{\partial y} \right) = 0 \quad (9.2) \quad \text{south - north eq.} \\
\frac{\partial N}{\partial t} + \frac{\partial (wN)}{\partial z} - \frac{\partial}{\partial z} \left( K_{h,zz} \frac{\partial N}{\partial z} \right) = 0 \quad (9.3) \quad \text{vertical eq.}
\]

where $N$ is the gas concentration,
Numerical W. Modeling / 4th Year / Atm. Sci. Dept

Dr. Thaer Obaid Roomi

$K_{h,xx}, K_{h,yy}$ and $K_{h,zz}$ are eddy diffusion coefficients in x-, y-, and z-directions, respectively. The subscript h indicates that the diffusion coefficient for energy (eddy thermal diffusivity) is used.

9.3 Courant-Friedrichs-Lewy Stability Criterion (CFL criterion)

معيار كورانت فريدريك ليوي للاستقرارية

بينما تكون الأنظمة الصريحة مستقرة شرطياً (مستقرة عند أي خطوة زمنية تحت قيمة معينة) فإن بعض الأنظمة الضمنية ربما تكون مستقرة لا شرطياً (مستقرة بغض النظر عن خطوة الزمن).

ان حدود خطوة الزمن لحل صريحة لمعادلة التفاعل (بالتوجه west-east) يمكن ان تستند باستخدام معيار CFL، من خلال رزالة حد الانتشار من معادلة 9.1. عندما تحل المعادلة الناتجة بشكل صريح، فإن الاستقرارية تضمن عموماً عندما يتحقق معيار CFL التالي:

$$k < \frac{\Delta x_{\text{min}}}{|u_{\text{max}}|}$$ (9.4)

where $|u_{\text{max}}|$ is the maximum west-east wind speed, and $\Delta x_{\text{min}}$ is the minimum west-east grid – cell length in the domain.

**Homework:** If the maximum wind speed is 20 m/s and minimum grid-cell length is 5 km. Find the maximum time step that the CFL criterion predicts for maintaining stability?

والآن إذا أهمل حد التفاعل في معادلة 9.1 فأنه تبسط الى معادلة الانتشار ويكون معيار CFL لها كالآتي:

$$k < \frac{\Delta x_{\text{max}}^2}{|K_{\text{max}}|}$$ (9.5)

where $K_{\text{max}}$ is the largest eddy diffusion coefficient in the domain.

لمعامل انتشار دوامي رأسي نموذجي قيمته (50 m$^2$ s$^{-1}$) فأن معيار الاستقرارية يقترح أن خطوة الزمن لانتشار الدوامي خلال ارتفاع 100 m ينبغي ان تكون أقل من 200s. لمعامل انتشار دوامي افقي نموذجي (2500 m$^2$ s$^{-1}$) فأن خطوة الزمن لانتشار دوامي عبر خليه طولها 5km يجب ان تكون أقل من 10000s.
Lecture (10)
Discretization Error and Convergence (Part 1)

10.1 Discretization Error and Convergence

\[ \Delta t \text{ and } \Delta x \]

It can be shown that a smaller value of \( \Delta x \) and \( \Delta t \) are not always desirable, as this will only make the problem more computationally complex. A consistent \( (x,t) \) is a requirement for this. This is a characteristic of round off error. This is a result of the finite precision of the computer used. This error increases with the number of iterations, but not necessarily.

If we assume that the errors are not independent, we can observe that the errors in the solution are not independent. The discretization error is defined as:

\[ E = A(i,j) - A(i\Delta x, j\Delta t) \]

This error is called discretization error. This error is significant, especially when the space is not uniform. This error may be increased with the number of iterations.

As we approach zero, the solution converges to a solution that is independent of the grid size. The upstream scheme is used in this case.

The domain of dependence of the scheme is illustrated in the figure (10.1). In the case of the upstream scheme, the solution is independent of the grid size. The domain of dependence of the scheme is the region where the solution is not affected by the grid size.

\[ \frac{A(i+1,j) - A(i,j)}{\Delta t} + c \left( \frac{A(i,j) - A(i-1,j)}{\Delta x} \right) = 0 \]
الفئة (10.1) المنطقة المظللة تمثل "مجال اعتماد" حل نظام ضد التيار في النقطة \((i \Delta x, j \Delta t)\).

تستطيع زيادة دقة النظام بتصنيف \(\Delta x, \Delta t\) ، ولكن مجال العدم لا يتغير ، بغض النظر عن مقدار كثافة الشبكة ما دامت النسبة \(\frac{c \Delta t}{\Delta x}\) تبقى ثابتة. وهذا دليل على أن هذه كمية مهمة.

افرض بأن الخط الذي يمر بالنقطة \((i \Delta x, j \Delta t)\) أي \(x_0, 0\) حيث \(x - ct = x_0\) ، حيث هو مقدار ثابت ، لا يقع ضمن مجال الاعتماد. وهذه الحال ظاهرة في الشكل (10.1). عموماً ، لا أمل من الحصول على خطأ تردد True أصغر ، مما أصبحت \(c \Delta t/\Delta x\) غير متغيرة ، وذلك لأن الحل الصحيح solution لم يعد صحيح (حل عددي صحيح) يعتمد فقط على القيمة الأولية لـ \(A\) في النقطة المنفردة \((x_0, 0)\) الذي لا يستطيع التأثير على

\(A_{i,j}\) ، ولكن الحل المحصور \(A(i \Delta x, j \Delta t)\) وبالتالي الحل المضبوط 

سيبقى كما هو. في مثل هذه الحالة الخطأ في الحل لن يتناقص بتنعيم الشبكة. وهذا يوضح بان لو كانت قيمة \(c\) بحيث خارج مجال الاعتماد ، فإن من غير الممكن لحل معادلة الفروقات المحددة (العدي) أن يقترب من حل المعادلة التفاضلية (التحليلي) ، لا يهم كيف ستصبح الشبكة داعمة. معادلة الفروقات

38
المقددة تقترب من المعادلة التفاضلية، ولكن حل المعادلة التفاضلية لا يقترب من حل المعادلة التفاضلية.

خطأ القطع يذهب إلى الصفر ولكن خطأ التفرید لا.

المناقشة اعلاه تظهر بأن:

\[
0 \leq \frac{c\Delta t}{\Delta x} \leq 1 \quad (10.1)
\]

هو شرط ضروري لتقرب نظام ضد التيار (upstream scheme) لاحظ بأنه إذا كانت \( c \) سالبة (downstream scheme معطية ما يسمى نظام مع التيار)، الخط الفاقد يقع خارج مجال الاعتماد الظاهر في الشكل. وبالطبع إذا كانت \( c < 0 \) نستطيع الحصول على:

\[
\frac{A_{i,j+1} - A_{i,j}}{\Delta t} + c \left( \frac{A_{i,j+1} - A_{i,j}}{\Delta x} \right) = 0 \quad (10.2)
\]

بدل معادلة نظام ضد التيار (4.8).

كخلاصة: خطأ القطع يقيس دقة التقريب للمعادلة التفاضلية. وهو قياس للدقة التي تقرب بها حدود المعادلة التفاضلية. بينما خطأ التفرید يقيس الدقة التي تقرب بها حل المعادلة التفاضلية. تقلص خطأ القطع الى مستويات مقبولة يكون سهلًا في العادة ولكن تقلص خطأ التفرید يكون أكثر صعوبة.
Lecture (11)

Numerical Stability

11.1 Introduction

Numerical Stability

An estimation of the numerical error is essential for the stability of the numerical model. The numerical model is influenced by the computational errors. If the errors are constant, the system is said to be stable. If the errors increase with time, it is unstable.

The estimation of the errors depends on the time. Numerical stability ensures that the method produces a bounded solution. If the errors are small, the numerical solution is close to the exact solution.

True or exact solution: True solution

Numerical solution: Numerical solution

The estimation of the errors depends on the time. Numerical stability ensures that the method produces a bounded solution. If the errors are small, the numerical solution is close to the exact solution.
هناك ثلاثة طرق لمعرفة الاستقرار:

1. الطريقة المباشرة
2. الطريقة الطاقة
3. طريقة فون نيومان

11.2 Direct Method

في كل نظم أن الحل الصحيح هو حل مقيّد وعليه يكون كافياً لاختيار تقييدية الحل العددي. من الممكن كتابة معادلة التأذن:

\[ \frac{u_{i+1,j} - u_{ij}}{\Delta t} + \frac{c}{\Delta x} (u_{ij} - u_{i-1,j}) = 0 \] ... (11.2)

بالشكل التالي:

\[ u_{i+1,j} = (1 - \mu)u_{ij} + \mu u_{i-1,j} \] ... (11.3)

حيث

\[ \mu = \frac{c\Delta t}{\Delta x} \] ... (11.4)

إذا كان 0 ≤ \( \mu \) ≤ 1 والذي يعتبر كشرط ضروري للننقار (\( c\Delta t \leq \Delta x \)) يكون لدينا:

\[ |u_{i+1,j}| \leq (1 - \mu)|u_{ij}| + \mu |u_{i-1,j}| \] ... (11.5)

سنطبق هذه العلاقة عند النقطة 1 عند المستوى 1+1 وحيث تكون

\[ |u_{i,j+1}| \]

اعظم ما يكون أي عند:

\[ Max_{(i)} |u_{j+1}| \]

\[ Max_{(i)} |u_{i+1,j}| \leq Max_{(i)} (1 - \mu)|u_{ij}| + Max_{(i)} \mu |u_{i-1,j}| \]

إذا نحن نعلم أن الحل الصحيح هو حل مقيّد وعليه يكون كافياً لاختيار تقييدية الحل العددي. من الممكن كتابة معادلة التأذن:

\[ \frac{u_{i+1,j} - u_{ij}}{\Delta t} + \frac{c}{\Delta x} (u_{ij} - u_{i-1,j}) = 0 \] ...

والتي يعتبر كشرط ضروري للننقار (\( c\Delta t \leq \Delta x \)) يكون لدينا:

\[ |u_{i+1,j}| \leq (1 - \mu)|u_{ij}| + \mu |u_{i-1,j}| \] ...

إذا كان 0 ≤ \( \mu \) ≤ 1 والذي يعتبر كشرط ضروري للننقار (\( c\Delta t \leq \Delta x \)) يكون لدينا:

\[ Max_{(i)} |u_{j+1}| \]

\[ Max_{(i)} |u_{i+1,j}| \leq Max_{(i)} (1 - \mu)|u_{ij}| + Max_{(i)} \mu |u_{i-1,j}| \]

إذا نحن نعلم أن الحل الصحيح هو حل مقيّد وعليه يكون كافياً لاختيار تقييدية الحل العددي. من الممكن كتابة معادلة التأذن:

\[ \frac{u_{i+1,j} - u_{ij}}{\Delta t} + \frac{c}{\Delta x} (u_{ij} - u_{i-1,j}) = 0 \] ...

ورغم أن الطريقة المباشرة تتطلب اختيار تيتيكية حل عددي (11.2) لهذا النظام، أصبح شرط الاستقرار نفسه شرط التقارب. بتعبير آخر، إذا كان النظام متقارب فإنه مستقر، والعكس بالعكس. رغم أن الطريقة المباشرة هي طريقة بسيطة إلا أن هناك حاجة لعدد محدود فقط من الأنظمة.
11.3 طريقة الطاقة

إذا علمنا أن الحل الصحيح هو مقبَّل، سوف نختبر فيما إذا كان كل $\sum_i (u_{i,j+1}^2)$ مقيَّداً أيضاً وبالتالي فإن كل $u_{i,j}$ ينبغي أن يكون مقيَّداً عليه يتم إثبات استقرار النظام. أن هذه الطريقة تسمى بطريقة الطاقة بسبب أن

في التطبيقات الفيزيائية $u^2$ غالباً تتناسب مع صيغة معينة للطاقة.

وتتربع المعادلة (11.3) والجمع على $i$ نحصل على:

$$\sum_i (u_{i,j+1})^2 = \sum_i [(1 - \mu)^2 (u_{i,j})^2 + 2\mu (1 - \mu) u_{i,j} u_{i-1,j}$$

$$+ \mu^2 (u_{i-1,j})^2] \quad \ldots \quad (11.7)$$

والتي نفترض أن الجمع يكون على دورات كاملة واحدة $x$، وأفترض بأن $u$ هي دورية للتبسيط، فنفترض بأن $u$ هي دورية في $x$، وأفترض بأن الجمع يكون على دورات كاملة واحدة $x$، وعندئذ $

$$\sum_i (u_{i-1,j})^2 = \sum_i (u_{i,j})^2 \quad , \quad \ldots \quad (11.8)$$

Now we use Schwartz's inequality which states that:

$$\sum a b \leq \sqrt{\sum a^2} \sqrt{\sum b^2} \quad \ldots \quad (11.9)$$

باستخدام معادلة (11.8) و (11.10) وإذا كانت $0 - 1 \geq -1$ و (11.10) تعطي الامساواة التالية:

$$\sum_i u_{i,j} u_{i-1,j} \leq \sqrt{\sum_i (u_{i,j})^2} \sqrt{\sum_i (u_{i-1,j})^2} = \sum_i (u_{i,j})^2 . \quad \ldots \quad (11.10)$$

وباستخدام معادلة (11.8) و (11.10) وإذا كانت $0 - 1 \geq -1$ و (11.10) تعطي الامساواة التالية:

$$\sum_i (u_{i,j+1})^2 \leq [(1 - \mu)^2 + 2\mu (1 - \mu) + \mu^2] \sum_i (u_{i,j})^2 , \quad \ldots \quad (11.11)$$

or

$$\sum_i (u_{i,j+1})^2 \leq \sum_i (u_{i,j})^2 . \quad \ldots \quad (11.12)$$

وعليه $0 - 1 \geq -1$، سواءً مع شرط الحدود الدوري، يكون شرطاً كافياً لاستقرارية (11.3).
11.4 Von Neumann's Method

وهي الأكثر استخداماً. على أي حال هي لا نستخدم Fourier series في المعادلات الغير خطية. ولكن نستطيع استخدامها في المعادلات التي تحولت إلى خطية من معادلات غير خطية. حل المعادلات الخطية يمكن أن يعبر عنه على شكل سلسلة فوريير. يمكن صياغة سلسلة فوريير بدالة الجيب والجيب تمام ولكن من الأسهل جبرياً استبدالها بما يكافئها من الصيغة الأسيوية العقدية. سوف نستبدل ترميزنا الشائع \( u_{i,j} \) إلى

\[ u(\phi, qk) = u_{p,q} \]

وبدالة هذا الترميز:

\[ A_n e^{in\pi x/l} = A_n e^{in\pi \phi/Nh} = A_n e^{i\beta_n \phi}, \quad ...(11.13) \]

where \( \beta_n = \frac{n\pi}{Nh} \) and \( Nh = l. \)

لتقصي عن انتشار الخطأ مع زيادة t يحتاج لأجواء حل لمعادلة الفروقات المحددة والذي يختصر إلى

للخطأ مع زيادة t سوف لن يزداد بزيادة t بشرط أن يكون:

\[ |\lambda| \leq 1. \]

Ex (1). Investigate the stability of the fully-implicit finite -difference equation,

\[ \frac{u_{p,q+1} - u_{p,q}}{k} = \left( \frac{u_{p-1,q+1} - 2u_{p,q+1} + u_{p+1,q+1}}{h^2} \right) \quad ...(11.15) \]

approximating the parabolic equation \( \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}. \)

الحل: بما أن دالة الخطأ \( u_{p,q} \) تحقق نفس معادلات الفرق كما هو الحال مع

\[ e^{i\beta \phi} \chi^{q+1} - e^{i\beta \phi} \chi^q = r \{ e^{i\beta (p-1)^h} \chi^{q+1} - 2e^{i\beta \phi} \chi^{q+1} + e^{i\beta (p+1)^h} \chi^{q+1} \}, \]

where \( r = k/h^2. \) Division by \( e^{i\beta \phi} \chi^q \) leads to:

\[ \lambda - 1 = r\lambda(e^{-i\beta h} - 2 + e^{i\beta h}) \]
\begin{align*}
\lambda &= \frac{1}{1 + 4r \sin^2 \left(\frac{\beta h}{2}\right)}
\end{align*}

Homework: write the above example in details.

Ex (2). The hyperbolic equation \( \partial^2 u/\partial t^2 = \partial^2 u/\partial x^2 \) is approximated by the explicit scheme:

\[
\frac{u_{p+1,q} - 2u_{p,q} + u_{p-1,q}}{k^2} = \frac{u_{p+1,q} - 2u_{p,q} + u_{p-1,q}}{h^2} \quad \text{... (*)}
\]

investigate its stability.

\[
\lambda^2 - 2A\lambda + 1 = 0
\]

where \( A = 1 - 2r^2 \sin^2 \left(\frac{\beta h}{2}\right) \), \( r = k/h \) \( \text{... (**)} \)

Hence the values of \( \lambda \) are:

\[
\lambda_1 = A + (A^2 - 1)^{\frac{1}{2}} \quad \text{and} \quad \lambda_2 = A - (A^2 - 1)^{\frac{1}{2}}
\]

For stability \( |\lambda| \leq 1 \)

As \( r, k, \beta \) are real, \( A \leq 1 \) by eq. (**)

When \( A < -1 \), \( |\lambda_2| > 1 \), giving instability.

When \( -1 \leq A \leq 1 \), \( A^2 \leq 1 \), \( \lambda_1 = A + i(1 - A^2)^{\frac{1}{2}} \), \( \lambda_2 = A - i(1 - A^2)^{\frac{1}{2}} \).

hence \( |\lambda_1| = |\lambda_2| = \{|A^2 + (1 - A^2)|^{\frac{1}{2}} = 1\}

proving that equation (*) is stable for \( -1 \leq A \leq 1 \). By eq.(**), we then have:

\[-1 \leq 1 - 2r^2 \sin^2 \left(\frac{\beta h}{2}\right) \leq 1 \]

The only useful inequality is:

\[-1 \leq 1 - 2r^2 \sin^2 \left(\frac{\beta h}{2}\right) \]

giving \( r \leq 1 \)
Lecture (12)

Series Expansion Methods

طرق مفكوك السلسلة

إن الحلول التقريبية للمعادلات التفاضلية الجزئية PDE مثل معادلات التافق-الانتشار، يمكن الحصول عليها بواسطة طرق الفروقات المحددة أو طرق الحجوم المحددة. الغرض من استخدام التقريب هو تقليص فضاء الحل لكل معادلة تفاضلية مستمرة من عدد غير محدود إلى عدد محدد من العقد التفاضلية أو الزمنية لgrese زيادة سرعة الحسابات للمعادلات التفاضلية.

في طريقة الفروقات المحددة التي درسناها سابقاً، التقريب يتضمن استبدال كل معامل تفاضلي مستمر (Δd) بنظير فرق منفصل (Δd) discrete difference analog بنظير مكتوب بدلالة عدد محدد من القيم عند كل عقدة زمنية أو فضائية.

بما أن السرعة العددية غرب - شرق هي دالة مستمرة، فالغير يعبر عنها من دالة function يعبر عنها من دالة discretized grid، كما يظهر في الشكل (12.1).

Fig.(12.1) Disretization of a continuous west-east scalar velocity u.

PDE في طريقة مفكوك السلسلة series expansion method، فالتغير المعتدل في PDE (u,v,w,or N) يستبدل بسلسلة محددة تقرب قيمته.

اذا كانت ال PDE الناتجة من معادلة تفاق غرب-شرق عند عقدة i هي: 

45
\[
\frac{\partial N_i}{\partial t} + \frac{\partial (uN)_i}{\partial x} = 0
\]  
(1)

\[
N_i \approx N_i(x) = \sum_j N_j e_j(x)
\]  
(2)

Note: In the finite element method, instead of solving the problem for the entire body in one operation, we formulate the equations for each finite element and combine them to obtain the solution of the whole body.

Basis Function: In mathematics, a **basis function** is an element of a particular basis for a function space. Every continuous function in the function space can be represented as a linear combination of basis functions, just as every vector in a **vector space** can be represented as a linear combination of basis vectors.

In numerical analysis and approximation theory, basis functions are also called **blending functions**, because of their use in interpolation: In this application, a mixture of the basis functions provides an interpolating function (with the “blend” depending on the evaluation of the basis functions at the data points).
Lecture (13)

Semi-Lagrangian Advection Scheme (Part 1)

13.1 Introduction

There are two methods to describe flow:

- **Eulerian Flow:** remained in a place and saw the flow.

\[ \frac{\partial V}{\partial t} = x, y, \text{ and } z \text{ fixed} \]

In a local sense:

- **Lagrangian Flow:** flow with the flow and watch where it goes.

\[ \frac{\partial T}{\partial t} \text{ Change at a fixed location.} \]

Now, we have come close, then the flow in the river. The temperature changes in a fixed point is:

\[ \frac{dV}{dt} \text{ parcel of fluid fixed} \]

Now, we have come close, then the flow in the river. The temperature changes in a fixed point is:

\[ \frac{dV}{dt} \text{ Change for a material parcel.} \]

What is the relationship between the velocity and time?

\[ \text{V} = V(x(t), y(t), z(t), t). \]
The chain rule gives:

\[
\frac{dV}{dt} = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial V}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial V}{\partial z} \cdot \frac{dz}{dt}
\]

\[
= \frac{\partial V}{\partial t} + u \cdot \frac{\partial V}{\partial x} + v \cdot \frac{\partial V}{\partial y} + w \cdot \frac{\partial V}{\partial z}
\]

\[
= \frac{\partial V}{\partial t} + \vec{V} \cdot \vec{V}.
\]

This conditionally stable condition holds for all changes.

- For high spatial resolution (small \( \Delta x \)) this severely limits the maximum time step \( \Delta t \) that is allowed.
- In numerical weather prediction (NWP), timeliness of the forecast is of the essence.
- In this lecture, we study an alternative approach to time integration, which is unconditionally stable and so, free from restrictions of the CFL condition.

13.2 The Basic Idea of Semi-Lagrangian Scheme

The semi-Lagrangian scheme for advection is based on the idea of approximating the Lagrangian time derivative.

It is so formulated that the numerical domain of dependence always includes the physical domain of dependence. This necessary condition for stability is satisfied automatically by the scheme.

In a fully Lagrangian scheme, the trajectories of actual physical parcels of fluid would be followed throughout the motion.

The problem with this approach is that the distribution of representative parcels rapidly becomes highly non-uniform.

In the semi-Lagrangian scheme the individual parcels are followed only for a single time-step. After each step, we revert to a uniform grid.

The semi-Lagrangian algorithm has enabled us to integrate the primitive equations using a time step of 15 minutes. This can be compared to a typical time step of 2.5 minutes for conventional schemes.
The consequential saving of computation time means that the operational numerical guidance is available to the forecasters much earlier than would otherwise be the case.

Semi-Lagrangian advection schemes are now in widespread use in all the main Numerical Weather Prediction centres.

13.3 The Eulerian and the Lagrangian Approach

We consider the linear advection equation which describes the conservation of a quantity $Y(x, t)$ following the motion of a fluid flow in one space dimension with constant advecting velocity $c$.

This may be written in either of two alternative forms:

$$\frac{\partial Y}{\partial t} + c \frac{\partial Y}{\partial x} = 0 \iff \text{Eulerian Form}$$

$$\frac{dY}{dt} = 0 \iff \text{Lagrangian Form} \quad (13.1)$$

The general solution is $Y = Y(x - ct)$.

To develop numerical solution methods, we may start from either the Eulerian or the Lagrangian form of the equation. For the semi-Lagrangian scheme, we choose the latter.

Since the advection equation is linear, we can construct a general solution from Fourier components

$$Y = a \exp[ik(x - ct)]; \quad k = \frac{2\pi}{L}$$

This expression may be separated into the product of a function of space and a function of time:

$$Y = a \times \exp(-i\omega t) \times \exp(ikx); \quad \omega = kc$$

Therefore, in analyzing the properties of numerical schemes, we seek a solution of the form

$$Y_{p,q} = a \times \exp(-i\omega q \Delta t) \times \exp(ikp \Delta x) = aA_q \exp(ikp \Delta x)$$

where $A = \exp(-i\omega \Delta t)$.

The character of the solution depends on the modulus of $A$:

- If $|A| < 1$, the solution decays with time.
- If $|A| = 1$, the solution is neutral with time.
- If $|A| > 1$, the solution grows with time.

In the third case (growing solution), the scheme is unstable.
Semi-Lagrangian Advection Scheme (Part 2)

14.1 Numerical Domain of Dependence

In a numerical domain of dependence, $Y_{p,q}$ at time $k \Delta t$ and space location $p \Delta x$, we are given a value $Y_{p,q}$ at that location and time. All values in the domain of dependence are computed based on this value, and the domain of dependence is defined as the set of all locations and times where $Y_{p,q}$ is computed. If a value outside of this domain is computed, it has no effect on the solution and is discarded. The solution to the problem is bounded, and the physical trajectory is bounded, ensuring that the solution is physically meaningful.

Figure 14.1 Numerical domain of dependence
14.2 Parcel coming from outside domain of dependence

Suppose that the parcel is coming from outside the domain of dependence (3), (as shown in Figure 14.2). The value at point \( (p, q) \) is represented by \( Y_{p,q} \). However, if the parcel is outside the domain for the first time, the system cannot adjust, and the system is not able to obtain a correct solution. The focus of the numerical scheme is to simulate the energy flux [13.1]:

\[
\frac{dy}{dt} = 0 = \left[ q \Delta t, (q + 1) \Delta t \right],
\]

where \( Y \) takes a constant value in the calculation period. By implementing the equation for the current time period, we obtain:

\[
\left( \frac{Value \ of \ Y \ at \ point \ p \Delta x \ at \ \text{time} \ (q + 1) \Delta t}{Value \ of \ Y \ at \ \text{departure} \ point \ at \ time \ q \Delta t} \right)
\]

14.3 Interpolation Using Surrounding Points

Interpolation of the surrounding points is used to estimate the value at the current time. The value at point \( (p, q) \) is represented by \( Y_{p,q} \), where \( Y_{p,q+1} = Y_{p,q} \), if it is not an interpolation point. If it is an interpolation point, the value at point \( (p, q) \) is estimated using the surrounding points.

![Figure 14.2 Parcel trajectory](image-url)
The distance travelled in time $\Delta t$ is $s = c\Delta t$.
We define the integer and fractional parts of $s$ as follows:

$\gamma = \lfloor s \rfloor = \text{Integral part of } s$

$\alpha = s - \gamma = \text{Fractional part of } s$

Note that, by definition, $0 \leq \alpha < 1$.

So, the departure point falls between the grid points $p - \gamma - 1$ and $p - \gamma$. In the figure (14.4), $\gamma = 1$ and $\alpha \approx 2/3$. A linear interpolation gives:

$$Y_{p,q} = \alpha Y_{p-\gamma-1,q} + (1 - \alpha)Y_{p-\gamma,q}$$

14.4 Numerical Stability of the Scheme

الاستقرارية العددية للنظام

المعادلة المنقطعة يمكن تكتب بالشكل:

$$Y_{p,q+1} = \alpha Y_{p-\gamma-1,q} + (1 - \alpha)Y_{p-\gamma,q}$$

دعنا نبحث عن حل من الشكل:

$$Y_{p,q} = AA_q \exp(ikp\Delta x).$$

$$aA_{q+1} \exp(ikp\Delta x) = aaA_q \exp[ik(p - \gamma - 1)\Delta x] + (1 - \alpha)A_q \exp[ik(p - \gamma)\Delta x]$$

أو، باستبعاد الحدود الشائعة،

$$A = \alpha \exp[ik(-\gamma - 1)\Delta x] + (1 - \alpha)\exp[ik(-\gamma)\Delta x]$$

$$= \exp(-ik\gamma\Delta x) \cdot [(1 - \alpha) + \alpha \exp(-ik\Delta x)]$$

الآن ننخد بنظر الاعتبار مربع المعامل $A$.
\[ |A|^2 = | \exp(ik\Delta x)|^2. |(1 - \alpha) + \alpha \exp(-ik\Delta x)|^2 \\
= | (1 - \alpha) + \alpha \cos k\Delta x - i\alpha \sin k\Delta x|^2 \\
= [ (1 - \alpha) + \alpha \cos k\Delta x]^2 + [\alpha \sin k\Delta x]^2 \\
= (1 - \alpha)^2 + 2(1 - \alpha)\alpha \cos k\Delta x + \alpha^2 \cos^2 k\Delta x + \alpha^2 \sin^2 k\Delta x \\
= 1 - 2\alpha(1 - \alpha)[1 - \cos k\Delta x] \\
\]

We note that \( 0 \leq (1 - \cos \theta) \leq 2 \). Taking the largest value of \( 1 - \cos k\Delta x \) gives:
\[ |A|^2 = 1 - 4\alpha(1 - \alpha) = (1 - 2\alpha)^2 < 1 \]
Taking the smallest value of \( 1 - \cos k\Delta x \) gives:
\[ |A|^2 = 1 \]
In either case, \( |A|^2 = 1 \), so there is numerical stability.
Historical Background of Numerical Weather Prediction (NWP)

الخلفية التاريخية للتنبؤ الجوي العددي

15.1 Introduction

The story goes that in 1904 Wilhelm Bjerknes was the first to point out that the future state of the atmosphere could be predicted by integrating the partial differential equations that govern the behaviour of the atmosphere, using as initial fields the observed state of the atmosphere at a particular time. However, the equations are too complicated for analytic solutions to be found and one must resort to numerical methods. We refer now to such integrations as numerical weather prediction, commonly abbreviated NWP.

Bjerknes proposed a “graphical calculus,” based on weather maps, for solving the equations. Although his methods continued to be used and developed until the 1950s, both the lack of faster calculating methods and the dearth of accurate observational data limited their success as forecasting techniques.

The first attempt at NWP was carried out by Lewis Frey Richardson during the First World War. At this time the calculations had to be carried out by hand and were very tedious and time-consuming. The result was a spectacular failure, but the details were published in what has become one of the most famous books in meteorology (Richardson, L. F. (1922). Weather prediction by numerical process.).

Richardson estimated that it would require a team of 64,000 persons to carry out a 24 hour forecast in 24 hours. This together with the unrealistic nature of his calculation, which predicted a surface pressure change of 145 mb in 6 hours, cast doubt on the practicality of the method! Several later developments changed this pessimistic view. Courant, Friedrichs and Lewy (1928) found that space and time increments chosen to discretize the differential equations have to meet a certain stability requirement. Later, mainly through the work of Rossby and his collaborators in the late 1930s, it was found that the large-scale motions in the atmosphere could be represented approximately by a rather simple equation expressing the conservation of absolute vorticity following the motion of air columns. Finally, at the end of World War II the first electronic computer ENIAC (Electronic Numerical Integrator and Computer) was constructed. This computer was used by Charney, Fjørtoft and von Neumann in the late 1940s for the first successful numerical forecast, based on integration of the absolute vorticity conservation equation.

The Princeton mathematician John von Neumann was among the earliest computer pioneers. Engaged in computer simulations of nuclear weapons explosions, he immediately saw parallels to weather prediction. (Both are non-linear problems of fluid dynamics.) In 1946, soon after the ENIAC became
Operational, von Neumann began to advocate the application of computers to weather prediction. As a committed opponent of Communism, Von Neumann hoped that weather modeling might lead to weather control, which might be used as a weapon of war. Soviet harvests, for example, might be ruined by a US-induced drought.

In the last four decades, progress in the accuracy and sophistication of NWP models has been swift, partly through the development of improved numerical algorithms for solving the equations and also because of the astonishing technological developments in computer engineering.

15.2 The Intelligence of Charney

Jule G. Charney (1917-1981) was one of the giants in the history of numerical weather prediction, revealed that the huge error of Richardson was due mostly to the fact that the initial conditions were not balanced, and therefore included fast moving gravity waves which masked the initial rate of change of the meteorological signal in the forecast (Figure 15.1). Moreover, if the integration had been continued, it would have suffered "computational blow-up" due to the Courant- Friedrichs-Lewy (CFL) condition.

Fig 15.1 Schematic of forecast with slowly varying weather-related variations and superimposed high-frequency gravity waves

Charney analysed the primitive equations using the technique of scale analysis, and was able to simplify them in such a way that the gravity wave solutions were completely eliminated. The resulting equations are known as the quasi-geostrophic system. In the special case of horizontal flow with constant static stability, the vertical variation can be separated out and the quasi-geostrophic potential vorticity equation reduces to a form equivalent to the nondivergent barotropic vorticity equation

\[
\frac{d(\zeta + f)}{dt} = 0
\]  

(15.1)

where \( \zeta \) is the vorticity of the motion and \( f \) is the planetary vorticity. The barotropic equation had, of course, been used by Rossby in his analytical study of atmospheric waves, but nobody seriously believed that it was capable of producing a quantitatively accurate prediction of atmospheric flow.

15.3. From barotropic to multi-level models and primitive equations

Several baroclinic models were developed in the few years after the ENIAC forecast. They were all based on the quasi-geostrophic system of equations. The limitations of the filtered equations were recognized at an early stage. In his 1951
paper ‘The mechanism of meteorological noise’, Karl-Heinz Hinkelmann [14] tackled the issue of suitable initial conditions for primitive equations integrations. Hinkelmann had been convinced from the outset that the best approach was to use these equations. He knew that they would simulate the atmospheric dynamics and energetics more realistically than the filtered equations. Moreover, he felt certain, from his studies of noise, that high frequency oscillations could be controlled by appropriate initialization.

Routine numerical forecasting was introduced in the Deutscher Wetterdienst in 1966; this was the first ever use of the primitive equations in an operational setting. A six-level primitive equation model was introduced into operations at the National Meteorological Center in Washington in June, 1966. In 1972 a 10-level primitive equation model was introduced.

Previous models had been essentially dynamical, without any adequate representation of the wide range of ‘physical’ processes that determine the weather. The new model incorporated a sophisticated parameterization of physical processes and with it the first useful forecasts of precipitation were produced.

15.4. General circulation models and climate modeling

A declaration issued at the World Economic Forum in Davos, Switzerland in 2000 read: Climate change is the greatest global challenge facing humankind in the 21st century. There is no doubt that the study of climate change and its impacts is of enormous importance for our future. Global climate models are the best means we have of anticipating likely changes.

Phillips (1956) carried out the first long-range simulation of the general circulation of the atmosphere. He used a two-level quasi-geostrophic model on a beta-plane channel with rudimentary physics. The computation, done on the IAS computer (MANIAC I), used a spatial grid of $16 \times 17$ points, and the simulation was for a period of about one month.

15.5. Numerical weather prediction today

It is no exaggeration to describe the advances made over the past half century as revolutionary. Thanks to this work, meteorology is now firmly established as a quantitative science, and its value and validity are demonstrated daily by the acid test of any science, its ability to predict the future. Operational forecasting today uses guidance from a wide range of models. In most centres a combination of global and local models is used.

References
Lecture (16)

Initial Conditions and Data Assimilation

16.1 Introduction

Mathematically, numerical weather prediction (NWP) can be viewed as solving an initial-boundary value problem in which the governing equations of fluid system are integrated forward in time in a finite domain.

- Therefore, in addition to the boundary conditions, we must also provide suitable initial conditions for the model.
- For idealized numerical simulations, the initial conditions may be prescribed by known functions or values.

  If the Coriolis force is included in the model, then the initial basic state should be in geostrophic balance. Otherwise, the initial state will be adjusted to reach a new balanced state by the model.

- For real data mesoscale modeling or numerical weather prediction, the observational data must be modified to be accurate and dynamically consistent with the governing equations of the model.

  Strictly speaking, the process in producing initial conditions may be classified as the following four components:

  (i) **quality control**, 
  (ii) **objective analysis**, 
  (iii) **initialization**, and 
  (iv) **initial guess from a short-range forecast by an NWP model**.

  Reason for quality control: The errors associated with the data may be misrepresented (*nonlinear aliasing*) and amplified by the model.

  To reduce the errors in the sounding data, the following steps of quality control have been taken in numerical weather prediction:

  (a) **plausibility check**,
(b) *contradiction check*,

(c) *gross check*, and

(d) *buddy check*.

In *plausibility check*, data values cannot possibly occur in the real atmosphere or extremely exceed climatological mean are rejected.

For example, positive temperatures in Celsius at 300 hPa are rejected.

☐ In *contradiction check*, data values of two or more parameters at the same location contradicting to each other are removed.

For example, the occurrence of rain in the absence of clouds is removed.

☐ In *gross check*, observations with large deviations from the first guess field forecast by an operational model are removed.

☐ In *buddy check*, observations not agreeing with neighboring observations are removed.

- Observational data are often not regularly spaced, which are not ready for use as initial fields for a mesoscale or NWP model because they do not match the model grid mesh.
- In some areas, such as over ocean, observational data are sparse.

Therefore, in order to use the observational data as initial fields for NWP model, one needs to interpolate or extrapolate the data to fit into the grid mesh of the model and to apply some balance relations, such as geostrophy and mass continuity, to make the data dynamically consistent. This procedure is called *objective analysis*.

○ In an objective analysis, it is desirable to:

1. filter out scales of motion that cannot be resolved by the model,

2. use a first guess field or background field provided by an earlier forecast from the same model, which help avoid the extrapolation of observation data in data sparse areas and introduce dynamically consistency, and

3. make use of our knowledge of the probable errors associated with each observation, which may be weighted based on past records of accuracy.
When the maximum information from data sources, including the observations, climatological records, space correlation among the meteorological variables, etc., are extracted statistically, the approach is called optimal interpolation.

This often requires knowledge of the statistical structure of the fields of the variables. The variables may be analyzed separately or simultaneously, which is referred to as univariate analysis or multivariate analysis, respectively.

- The objective analysis procedure generally does not provide fields of mass and motion that are consistent with model dynamics to initiate a forecast.

Thus, the use of such objectively analyzed data to initialize an NWP model may generate large, spurious inertial-gravity wave modes. Theoretically, these inertial-gravity wave modes will be dispersed, dissipated or propagate out of the domain due to redistribution of mass and wind fields. However, these modes or noise, as often referred to by NWP modelers, cannot be dissipated locally because of the relatively low resolution in NWP models.

Therefore, an additional procedure, called initialization, is required to force the data after objective analysis to be dynamically consistent with the model dynamics, and to allow the model to integrate forward in time with a minimum of noise and maximum accuracy of the forecasts.

Historically, a number of initialization techniques have been developed and used in mesoscale and NWP models, such as:

(a) damping method,

(b) static initialization,

(c) variational method,

(d) normal mode initialization, and

(e) dynamic initialization.

16.2 Exercise

Q1. Plot a suitable graph summarizes the above lecture.
Initialization Techniques

**damping method**

A simple and straightforward way to reducing the *gravity wave mode* is to dampen or filter the *inertial-gravity wave ‘noise’* by adding a divergence damping term to the horizontal momentum equation.

In this way, the local rate of change of the divergence will be diffused according to

\[
\frac{\partial D}{\partial t} = -\frac{1}{\partial} \nabla^2 p + v \nabla^2 D + \cdots
\]

Eq. (16.1)

This approach in initializing the data is called **damping method**.

**Static initialization**

Another way to adjust the data at a single time level, usually to conform to some dynamical constraints in order to reduce or eliminate the generation of inertial-gravity wave ‘noise’ is the static initialization.

For example, in an isobaric model, one may

a. estimate the geopotential field (\(\phi\)) from the pressure-height data and the geostrophic wind relations,

b. calculate the streamfunction (\(\psi\)) from analyzed \(\phi\) fields on the isobaric surfaces, and then,

c. compute the rotational wind component from the following relationship,

\[
V_\psi = k \times \nabla \psi ,
\]

Eq. (17.1) may be written as an elliptic function of \(\phi\), which may become hyperbolic in some areas.

**Variational method**

It is another approach to initializing the data, in which one or more of the conservation relations are applied to minimize the variance of the difference between the observations and the objectively analyzed fields. In performing the variational method, the concept of principles of variational calculus is applied.
For example, the difference may be minimized in a least-square sense subjected to one or more dynamical constraints, such as the balance equation, hydrostatic relation, and steady state momentum equation. The static initialization described above is based on the distinction between gravity wave modes with relatively high divergence and other meteorological modes of the quasi-geostrophic type with small divergence and relatively high vorticity.

- **Normal mode initialization**

  However, in reality the separation is far less clear cut in some occasions. Thus, it has been proposed to keep some normal modes if they can be represented by the model grid resolution. Retaining these gravity wave modes are important since some severe weather have been found to be induced by gravity waves.

  Unlike applying the balance equation constraint, the **normal mode initialization** produces a divergent component as well.

  Thus, this type of normal mode initialization makes an optimal use of the observed data by adjusting both mass and motion fields while achieving dynamical consistency through appropriate constraints.

- **Dynamic initialization**

  Since the normal mode initialization is performed separately right after the objective analysis, the initialized fields may no longer fit the observations as closely as possible. Therefore, the **dynamic initialization** is propose. The basic idea of dynamic initialization is to let the NWP or mesoscale model to do the job by itself because any primitive equation models are supposed to inherently possess the mechanism for the geostrophic adjustment process.

  Indeed, the mass and velocity fields do mutually adjust to one another and tend toward a quasi-geostrophic state when they are executed in an NWP model.

  In this way, observations are inserted intermittently or continuously over a period of time. In this type of initialization, the model is integrated forward and backward about the initial time and let the model adjust itself before starting the forecast.
Lecture (18)

Nondivergent Barotropic Model

18.1 Barotropy

The barotropic model is of historical interest because it was used in the first successful attempt at NWP. In this model, the forecast equation involves only one dependent variable and is applied at a single pressure level in the vertical. In the barotropic atmosphere the density is a function of pressure alone; therefore, the density, temperature, and pressure surfaces are parallel; hence only one level needs to be forecasted. The thermal wind equation

$$\frac{\partial \vec{V}_g}{\partial \ln P} = - \frac{R}{f} \hat{k} \times \nabla_p T$$

therefore becomes $\frac{\partial \vec{V}_g}{\partial \ln P} = 0$, which states that the geostrophic wind is independent of height in a barotropic atmosphere. Here $\vec{V}_g$ is the geostrophic wind, $P$ is air pressure, $R$ is gas constant, $f$ is the Coriolis parameter, $k$ is the unit vector in $z$-axis, $T$ is air temperature. Thus, barotropy provides a very strong constraint on the motions in a rotating fluid; the large-scale motion can depend only on horizontal position and time, not on height.

An atmosphere in which density depends on both temperature and pressure, $\rho = \rho(P,T)$, is referred to as a baroclinic atmosphere. In a baroclinic atmosphere the geostrophic wind is generally has vertical shear, and this shear is related to the horizontal temperature gradient by the thermal wind equation (18.1). Obviously, the baroclinic atmosphere is of primary importance in dynamic meteorology.

18.2 Dynamic of Barotropic Model

This model based on the conservation of absolute vorticity. The vorticity equation can be written in the form:

$$\frac{D}{Dt} (\xi + f) = -(\xi + f) \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) - \left( \frac{\partial w}{\partial x} \frac{\partial v}{\partial z} - \frac{\partial w}{\partial y} \frac{\partial u}{\partial z} \right) + \frac{1}{\rho^2} \left( \frac{\partial \rho}{\partial x} \frac{\partial p}{\partial y} - \frac{\partial \rho}{\partial y} \frac{\partial p}{\partial x} \right)$$

(18.2)
where \( \xi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \) is the relative vorticity. Equation (18.2) states that the rate of changing of the sum of the absolute vorticity following the motion is given by the sum of the three terms on the right, called the divergence term, the tilting or twisting term, and the solenoidal term, respectively.

The order of magnitude analysis proves that the second and the third terms on the right in this equation are sufficiently small so that they can be neglected for synoptic scale motion. Hence the vorticity equation may be rewritten as:

\[
\frac{D}{Dt} (\xi + f) = - (\xi + f) \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \quad (18.3)
\]

18.3 Vorticity in Barotropic Fluids

The continuity equation for a homogeneous incompressible fluid, simplifies to:

\[
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial d} = - \frac{\partial w}{\partial o} \quad (18.4)
\]

so that the vorticity equation (18.3) may be written as:

\[
\frac{D}{Dt} (\xi + f) = (\xi + f) \left( \frac{\partial w}{\partial x} \right) \quad (18.5)
\]

If the flow is purely horizontal \((w=0)\), as is the case for barotropic flow in a fluid of constant depth, the divergence term vanishes in (18.5) and one obtains the barotropic vorticity equation:

\[
\frac{D}{Dt} (\xi + f) = 0 \quad (18.6)
\]

which states that absolute vorticity is conserved following the horizontal motion.

More generally, absolute vorticity is conserved for any fluid layer in which the divergence of the horizontal wind vanishes, without the requirement that the flow be geostrophic. For horizontal motion that is nondivergent \(\left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \right)\), the flow field can represented by a streamfunction \(\psi(x, y)\) defined so that the velocity components are given as \(u = - \frac{\partial \psi}{\partial y}, \quad v = + \frac{\partial \psi}{\partial x}\). The vorticity is then given by:

\[
\xi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \equiv \nabla^2 \psi \quad (18.7)
\]
Thus, the velocity field and the vorticity can both be represented in terms of the variation of the single scalar field $\psi(x,y)$, and (18.6) can be written as a prognostic equation for vorticity in the form:

$$\frac{\partial}{\partial t} \nabla^2 \psi = -\vec{V}_\psi \cdot \nabla (\nabla^2 \psi + f)$$  \hspace{1cm} (18.8)

where $\vec{V}_\psi \equiv \vec{k} \times \nabla \psi$ is a nondivergent horizontal wind. Equation (18.8) states that the local tendency of relative vorticity is given by the advection of absolute vorticity. This equation can be solved numerically to predict the evolution of streamfunction, and hence of the vorticity and wind fields. Since the flow in the mid-troposphere is often nearly nondivergent on the synoptic scale, (18.8) provides a surprisingly good model for short-term forecasts of the synoptic scale 500-hPa flow field. Equation (18.8) can be expressed as:

$$\frac{\partial}{\partial t} \nabla^2 \psi = \frac{\partial \psi}{\partial y} \frac{\partial^2 \psi}{\partial x \partial y} - \frac{\partial \psi}{\partial x} \frac{\partial^2 \psi}{\partial y^2} - \beta \frac{\partial \psi}{\partial x}$$  \hspace{1cm} (18.9)

or simply,

$$\frac{\partial}{\partial t} \nabla^2 \psi = -J(\psi, \nabla^2 \psi) - \beta \frac{\partial \psi}{\partial x}$$  \hspace{1cm} (18.10)

where $\beta = \frac{\partial f}{\partial y}$ and $J$ is the Jacobian operator which is a common operator. The finite-difference form of this Jacobian can be written as:

$$J = \frac{1}{4h^2} \left[ (\psi_{i+1,j} - \psi_{i-1,j}) \cdot (\xi_{i,j+1} - \xi_{i,j-1}) - (\psi_{i,j+1} - \psi_{i,j-1}) \cdot (\xi_{i+1,j} - \xi_{i-1,j}) \right]$$  \hspace{1cm} (18.11)

Use of the nonlinear balance equation is of great usefulness to close the system:

$$\nabla^2 g z = \nabla \cdot f \nabla \psi + 2J \left( \frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y} \right)$$  \hspace{1cm} (18.12)

This equation defines the height field from the streamfunction, and together with (18.10) constitutes a closed system from the evolution of the barotropic nondivergent flow. Equation (18.12) represents actually a nonlinear reverse balance law since it determines the geopotential height from the wind field.

Reference:

20.1 Preamble

A model's vertical structure is as important as the horizontal structure and model type.
- To represent the vertical structure of the atmosphere properly requires selection of a suitable vertical coordinate and sufficient vertical resolution.
- Unlike the horizontal structure of models where discrete or continuous (grid point or spectral) configurations can be used, virtually all operational models use discrete vertical structures.

20.2 Sigma Vertical Coordinate

- The equations of motion have their simplest form in pressure coordinates.
- Unfortunately, pressure coordinate systems are not particularly suited to solving the forecast equations because, like height surfaces, they can intersect mountains and consequently 'disappear' over parts of the forecast domain.
- To deal with this problem Phillips (1957) developed a terrain-following coordinate called the sigma (σ) coordinate.
- The sigma coordinate or variants are used in the NGM, GFS, ECMWF, NOGAPS, and UKMET models and appear in some mesoscale models, such as MM5, COAMPS, and RAMS.
- The sigma coordinate is defined by

\[
\sigma = \frac{p}{p_s} \quad (20.1)
\]

where \(p\) is the pressure on a forecast level within the model and \(p_s\) is the pressure at the earth's surface, not mean sea level pressure.
- The lowest coordinate surface (usually labeled \(\sigma = 1\)) follows a smoothed version of the actual terrain. Note that the terrain slopes used in sigma models are always smoothed to some degree.
- The other sigma surfaces gradually transition from being nearly parallel to the smoothed terrain at the bottom of the model (\(\sigma = 1\)) to being nearly horizontal to the constant pressure surface at the top of the model (\(\sigma = 0\)).
• The sigma vertical coordinate can also be formulated with respect to height \((z)\), rather than pressure.

\[
\eta_s = \frac{P_r(z_s) - P_t}{P_r(z = 0) - P_t} \quad (20.2)
\]

where \(P_t\) is pressure at the model top, \(P_r(z = 0)\) is the standard atmosphere MSL pressure (1013 hPa), \(P_r(z_s)\) is the standard atmosphere pressure at the model terrain level \(z_s\).

20.3 Eta Vertical Coordinate

• The eta coordinate \((\eta)\) was created in the early 1980s in an effort to reduce the errors incurred in calculating the pressure gradient force using sigma coordinate models.

• The eta coordinate is, in fact, another form of the sigma coordinate, but uses mean sea level pressure instead of surface pressure as a bottom reference level. As such, eta is defined as
20.4 Calculating Eta Surfaces

First, the heights at each model level must be defined. In this example, we have defined a model with 10 eta layers (only 3 shown) distributed evenly with respect to pressure from sea level to the top of the atmosphere. Standard atmosphere pressures are then determined at each of these heights.
At point 1, the actual terrain elevation is 848m. This is closest to the 1000-m height defined for the first eta level. The standard atmosphere pressure at that height is 900 hPa. What, then, is the eta level closest to this point?

Using the eta equation,
\[ \eta = \frac{900 - 0}{1000 - 0} = 0.9 \]

If we go to point 2, the actual terrain height is 1126 m and is also closest to the 1000-m height. Therefore, the eta level closest to this point is again 0.9.

However, if we go up to point 3 (1832 m), the nearest eta surface in the model is at 2000 m. Here the standard atmospheric pressure is 800 hPa. The nearest eta level is therefore
\[ = 1 \times \left( \frac{800 - 0}{1000 - 0} \right) = 0.8 \]

This has been a simplified example. In reality, it is necessary to choose the intervals between eta levels in a way that both depicts the planetary boundary layer (PBL) with sufficient detail and yet represents the average changes in elevation over the entire forecast domain.

Note that the eta levels are predefined and the model topography is set to the nearest eta surface even if it does not quite match the average or smoothed terrain height in the grid box.

Eta usually is labeled from 0 to 1 from the top of the model domain to mean sea level.

Some of the model's grid cubes are located underground in areas where the surface elevation is notably above sea level. This requires special numerical formulations to model flow near the earth's surface.

The Eta coordinate systems allows the bottom atmospheric layer of the model to be represented within each grid box as a flat "step," rather than sloping like sigma in steep terrain. This configuration eliminates nearly all errors in the PGF
calculation and allows models using the eta coordinate to have extreme differences in elevation from one grid point to its neighbor.

Eta coordinate models can therefore develop strong vertical motions in areas of steep terrain and thus more accurately represent many of the blocking effects that mountains can have on stable air masses.

Even when the step-like Eta is used as the vertical coordinate, model terrain is still much coarser than real terrain, but the topographic gradients are less smoothed than in sigma models.

Note that the Eta levels are predefined and the model topography is set to the nearest eta surface even if it does not quite match the average or smoothed terrain height in the grid box.

20.5 Sigma vs Eta Surfaces

Figure (20.4) Sigma vs Eta surfaces
20.5 Limitations of the Eta Vertical Coordinate

1. The step nature of the eta coordinate makes it difficult to retain detailed vertical structure in the boundary layer over the entire model domain, particularly over elevated terrain.
2. Eta models do not accurately depict gradually sloping terrain.
3. Eta models have difficulty predicting extreme downslope wind events.
4. Eta models must broaden valleys a few grid boxes across or fill them in.
5. Eta coordinates can create spurious waves at step edges.

20.6 Other coordinates

1. Isentropic vertical coordinate.
   Since flow in the free atmosphere is predominantly isentropic, potential temperature (θ) can be very useful as a vertical coordinate system.
   However, non-adiabatic processes dominate in the boundary layer and isentropic surfaces intersect the earth's surface. For these reasons, potential temperature alone is not currently used as a vertical coordinate in any operational numerical model system. However, isentropic coordinates do form an essential part of many hybrid vertical coordinate systems.

2. Hybrid vertical coordinates
   Different hybrid combinations are currently in use, e.g.: hybrid isentropic-sigma vertical coordinates, hybrid sigma-pressure coordinate, hybrid isentropic-sigma coordinate.
Lecture (21)

Grid Staggering

21.1 Introduction

There are two kinds of grid arrangements: unstaggered grids and staggered grids. On a staggered grid the scalar variables (pressure, density, total enthalpy etc.) are stored in the cell centers of the control volumes, whereas the velocity or momentum variables are located at the cell faces. This is different from a collocated grid arrangement (unstaggered grids), where all variables are stored in the same positions.

Using a staggered grid is a simple way to avoid odd-even decoupling between the pressure and velocity. Odd-even decoupling is a discretization error that can occur on collocated grids and which leads to checkerboard patterns in the solutions.

The disadvantage of using staggered grids is that different variable are stored at different places and this makes it more difficult to handle different control volumes for different variables.

The use of unstaggered grids greatly reduces the required storage memory and shortens the computational time in three-dimensional calculations. However, they are prone to produce a false pressure field-checkerboard pressure. For this reason, in the 1980s and before, unstaggered grids were rarely used in the primitive variable method for incompressible flow. However, since 1983 the unstaggered grid (or collocated grid) has been used again more and more widely, after the scientists Rhie and Chow (1983) proposed a momentum interpolation method to eliminate the checkerboard pressure.

Figure (3.2) shows a one-dimensional schematic of an approach of staggering. For the unstaggered grid shown in figure (3.2a), calculation of an advection term such as $u \partial \theta / \partial x$ with a centered, three-point method would require differencing across a $2\Delta x$ interval. For the staggered grid in figure 3.2b, the derivative can be calculated by differencing across a $1\Delta x$ interval. This halves the effective grid increment for such terms, increasing the spatial resolution and decreasing the effects of truncation error on the solution.
Figure 3.2. Schematic of one-dimensional unstaggered (a) and staggered (b) grids. For the staggered grid, the mass-field variable ($\theta$) is offset by one-half grid increment from the momentum variable ($u$).

Let's consider the shallow water equations (SWE) in two dimensions:

\[
\frac{\partial h}{\partial t} = -\left[h\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) - u \frac{\partial h}{\partial x} - v \frac{\partial h}{\partial y}\right] \tag{3.48}
\]

\[
\frac{\partial u}{\partial t} = -\left[g \frac{\partial h}{\partial x} + fv\right] - u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} \tag{3.49}
\]

\[
\frac{\partial v}{\partial t} = -\left[g \frac{\partial h}{\partial y} + fu\right] - u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} \tag{3.50}
\]

The terms in square brackets in (3.48-3.50) are the dominant terms for the geostrophic and the inertia-gravity wave dynamics. These terms are computed in different ways depending on the type of grid used. The advective terms are less affected by the choice of alternative (staggered) grids.

In two dimensions there are several possibilities for staggered grids (Arakawa and Lamb, 1977), which are shown in Figure (3.3). Grid A (unstaggered) has several advantages and disadvantages. The advantages are its simplicity, and, because all variables are available at all the grid points, it is easy to construct a higher order accuracy scheme. Grid A tends to be favored by proponents of the philosophy “accuracy is more important than conservation”. Its main disadvantage
is that all differences occur on distances $2\Delta$, and that neighboring points are not coupled for the pressure and convergence terms. This can give rise in time to a horizontal uncoupling (checkerboard pattern), which needs to be controlled by using a high order diffusion.

Grid C (Arakawa C-Grid) has the advantage that the convergence and pressure terms in square brackets in (3.48-3.50) are computed over a distance of only $1\Delta$, which is equivalent to doubling the resolution of grid A. For this reason geostrophic adjustment (the dispersion of gravity waves generated when the fields are not in geostrophic balance) is computed much more accurately.

The B and E grids carry both wind components at the same points, and are referred to as semi-staggered. The C and D grids provide two fully staggered layouts in which there are different locations for the two wind components and for height. Note that the E grid can be viewed as the B grid rotated through 45° (Simmons, 1994). The staggered grid has hence the advantage that: (i) the computational time is half as long or the number of grid points is halved, (ii) the
truncation error is half as big with $\Delta x \to \Delta x/2$, (iii) waves with $k\Delta x > \pi/2$, which are shorter than 4 grid cells, are eliminated. They are the ones that have a big error in the phase speed (Döös, 2011). The problems on the staggered grid arise due to the averaging of the velocity components in the Coriolis force terms (Janjic et al. 2010). The momentum components (u and v) are located together on what will be called the momentum or velocity grid. Surface pressure, temperature, and an arbitrary number of tracers are located together on what will be called the mass or temperature grid. The momentum and mass grids are rectangular in shape, with equal spacing in longitude along the x-axis and latitude along the y-axis. The grids are diagonally shifted from each other, such that, the center of a momentum grid box is located at the corner where four mass grid boxes intersect. Auxiliary grids can be defined for computing additional quantities. The zonal mass flux grid (or U grid) has grid boxes centered on the east and west faces of a mass grid box. The meridional mass flux grid (or V grid) has grid boxes centered on the north and south faces of a mass grid box (Wyman, 2003). It is well known for finite-difference models that the geostrophic adjustment and Rossby wave propagation are better reproduced by the unstaggered grid combined with the vorticity-divergence formulation than by the standard $u-v$ formulation on the staggered C grid. The unstaggered grid applied to the semi-Lagrangian model also allows use of a single set of trajectories for all variables (while for a staggered grid it is necessary either to use the multiple set of trajectories or to make additional interpolations). On the other hand, the unstaggered grid makes it possible to apply easily compact high-order finite differences. Otherwise, one would need to apply high-order interpolation between half- and integer nodes of the grid (for example, to calculate the Coriolis term on the C grid) (Tolstyk, 2002).
22.1 The Governing System of Equations in Spherical Coordinates

These equations can serve as a basis for most numerical weather prediction models. The following set of equations is a closed system of seven equations and six unknowns (u, v, w, T, ρ, and q_v) in spherical coordinates:

\[
\begin{align*}
\frac{\partial u}{\partial t} & = -u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} - w \frac{\partial u}{\partial z} + \frac{uv \tan \phi}{a} - \frac{1}{\rho} \frac{\partial P}{\partial x} - 2\Omega (w \cos \phi - v \sin \phi) + F_{r_x} \quad (1) \\
\frac{\partial v}{\partial t} & = -u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} - w \frac{\partial v}{\partial z} + \frac{u^2 \tan \phi}{a} - \frac{uw}{a} - \frac{1}{\rho} \frac{\partial P}{\partial y} - 2\Omega u \sin \phi + F_{r_y} \quad (2) \\
\frac{\partial w}{\partial t} & = -u \frac{\partial w}{\partial x} - v \frac{\partial w}{\partial y} - w \frac{\partial w}{\partial z} + \frac{u^2 + v^2}{a} - \frac{1}{\rho} \frac{\partial P}{\partial z} + 2\Omega u \cos \phi - g + F_{r_z} \quad (3) \\
\frac{\partial T}{\partial t} & = -u \frac{\partial T}{\partial x} - v \frac{\partial T}{\partial y} + (\gamma - \gamma_d)w + \frac{1}{c_p} \frac{\partial H}{\partial z} \quad (4) \\
\frac{\partial \rho}{\partial t} & = -u \frac{\partial \rho}{\partial x} - v \frac{\partial \rho}{\partial y} - w \frac{\partial \rho}{\partial z} - \rho \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) \quad (5) \\
\frac{\partial q_v}{\partial t} & = -u \frac{\partial q_v}{\partial x} - v \frac{\partial q_v}{\partial y} - w \frac{\partial q_v}{\partial z} + Q_v \quad (6) \\
P & = \rho RT \quad (7)
\end{align*}
\]

The momentum equations for spherical Earth (eqs. 1-3) represent Newton's second law of motion. The thermodynamic energy equation (eq. 4) accounts for various effects, both adiabatic and diabatic, on temperature. The continuity equation for total mass (eq.5) states that mass is neither gained nor destroyed, and equation (6) is analogous, but applies only on water vapor. The ideal gas law (eq.7) relates temperature, pressure, and density. The variables u, v, and w are the Cartesian velocity components, P is pressure, ρ is density, T is temperature, q_v is specific humidity, Ω is the rotational frequency of Earth, φ is latitude, a is the radius of Earth, γ is the lapse rate of temperature, γ_d is the dry adiabatic lapse rate, c_p is the specific heat of air at constant pressure, g is the acceleration of gravity, H represents a gain or loss of heat, Q_v is the gain or loss of water vapor.
through phase changes, and $F_r$ is a generic friction term in each coordinate direction. These equations are a set of prognostic, coupled, nonlinear, non-homogeneous partial differential equations. The equations are called the primitive-equations and the models that based on these equations are called the primitive equations models. This terminology is used to distinguish these models from ones that are based on differentiated versions of the equations, such as the vorticity equation.

### 22.2 Shallow-Water Equations (SWE)

Many NWP models nowadays are based on the primitive equations, which in general provide two different types of motion. One is the low-frequency meteorologically significant Rossby waves. The other contains the relatively high frequency gravity inertia waves, which are found to be a minor component of the observed atmospheric flow and are considered as noise in the forecast problem. Many methods are used to filter out the unwanted meteorological noise, for instance by using such drastic approximations as those in the shallow water equations model.

Shallow means that the vertical depth is much smaller than the typical horizontal depth which justifies hydrostatic approximation. The shallow water equations are the simplest form of the equations of motion that can be used to describe the horizontal structure of the atmosphere.

The major physical difference with this model and reality is the absence of density stratification that is present in the real fluids such as Earth's atmosphere or oceans. The hydrostatic approximation also allows $\rho$ to vary with $z$, but we will consider $\rho$ constant for this model.

There are various forms of this set of equations but in all cases the fluid is assumed to be homogeneous, incompressible, hydrostatic, and inviscid. The homogeneity condition means that the density does not vary in space, and incompressibility means that density does not change in time following a parcel. The equations from which one can begin the derivation are:
\[-\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - fv + \frac{1}{\rho} \frac{\partial p}{\partial x} = 0 \quad (8)\]

\[-\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + fu + \frac{1}{\rho} \frac{\partial p}{\partial y} = 0 \quad (9)\]

\[-\frac{\partial p}{\partial z} = -\rho g \quad (10)\]

\[-\frac{\partial \rho}{\partial t} + \rho \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) = 0 \quad (11)\]

Now, incompressibility and homogeneity imply:

\[-\frac{\partial \rho}{\partial t} = 0 \quad (12)\]

\[\rho = \rho_0, \text{ for } \rho_0 \text{ a constant, and} \]

\[-\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad (13)\]

The hydrostatic equation can thus be written,

\[-\frac{\partial p}{\partial z} = -\rho_0 g \quad (14)\]

Differentiating equation (14) with respect to \(x\), and using the fact that the right side is a constant, yield:

\[-\frac{\partial}{\partial x} \left( \frac{\partial p}{\partial z} \right) = \frac{\partial}{\partial z} \left( \frac{\partial p}{\partial x} \right) = 0, \quad (15)\]

which means that there is no horizontal variation of the vertical pressure gradient or vertical variation of the horizontal pressure gradient (the definition of barotropy). Since the pressure-gradient force generates the wind, the resulting Coriolis force, all forces are invariant with height. Integrating equation (14) over the depth of the fluid,

\[-\int_{z(P_S)}^{z(P_T)} \frac{\partial p}{\partial z} dz = -\rho_0 g \int_{z(P_S)}^{z(P_T)} dz, \quad (16)\]

where \(P_T\) and \(P_S\) represent the pressure at the top and bottom boundaries of the fluid respectively, yields

\[-P_S - P_T = \rho_0 gh, \quad (17)\]
for $h$ equal to the depth of the fluid. If $P_T = 0$ or $P_T \ll P_S$,

$$\frac{P_S}{\rho_0} = gh \quad (18)$$

$$\frac{1}{\rho_0} \frac{\partial P_S}{\partial x} = g \frac{\partial h}{\partial x} \quad (19)$$

This states that the horizontal pressure gradient at the bottom of the fluid is proportional to the gradient in the depth of the fluid provides a new form of the pressure-gradient term in equation 8 and 9. The incompressible continuity equation (eq.13) can also be rewritten by integrating it with respect to $z$:

$$\int_0^z \frac{\partial w}{\partial z} dz = w_z - w_s = - \int_0^z \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dz. \quad (20)$$

If $u$ and $v$ are not initially a function of $z$, they will remain so because the pressure gradient is not a function of. Furthermore, because $u$ and $v$ are not a function of $z$, neither are their derivatives, so that,

$$w_h - w_s = - \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) h \quad (21)$$

For $z=h$. For a horizontal lower boundary, the kinematic boundary condition $w_s = 0$ prevail. Recognizing that,

$$w_h = \frac{dh}{dt} \quad (22)$$

leads to a new continuity equation. There are now three equations in three variables, $u$, $v$, and $h$.

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} - f v + g \frac{\partial h}{\partial x} = 0, \quad (23)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + f u + g \frac{\partial h}{\partial y} = 0, \quad (24)$$

$$\frac{\partial h}{\partial t} + u \frac{\partial h}{\partial x} + v \frac{\partial h}{\partial y} + h \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0. \quad (25)$$

The above three equations represent the governing equations of shallow water equations model.