Lecture 5: Numerical Stability

Part 1

Numerical and Simulation Methods

- Objective is to develop and understand numerical methods to solve (Navier-Stokes) eqns.
- Considered forms of the equations.

- PDE's with $\frac{\partial}{\partial t}$, $\frac{\partial}{\partial x,y,z}$, $\frac{\partial^2}{\partial x^2,y^2,z^2}$

- Intoduced model equation, $\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$.
- Introduced the concept of 'time-marching'

- 1. Guess u(x, t=0)
- 2. Apply boundary conditions (solid surface, farfield)
- 3. Evaluate $\frac{\partial u(x,t)}{\partial x}$
- 4. $\frac{\partial u(x,t)}{\partial t} = -c\frac{\partial u(x,t)}{\partial x}$ $\Rightarrow u(x, t + \Delta t)$
- 5. If $\frac{\partial u}{\partial x}$ small, solution converged
- 6. Else go to 2

- Store solution at 'discrete' points in space and time $\Rightarrow u(i\Delta x, n\Delta t) = u_i^n$.
- Considered methods to approximate gradients
- Methods that give $u_i^{n+1} = f(.., u_{i-1}^n, u_i^n, u_{i+1}^n, ...)$ are called EXPLICIT schemes.
- For example to time-march 10 points using explicit first order scheme

$$u_i^{n+1} = u_i^n - \frac{c\Delta t}{\Delta x}(u_i^n - u_{i-1}^n)$$

(unew[] and uold[] are arrays)

```
DO NT=1,NTIMESTEPS
    Apply boundary conditions at I=1 and I=10
    DO I=1,10          ( What happens at I=1 ??)
        unew[i] = uold[i] - c*deltat*(uold[i]
                                   - uold[i-1])/deltax
    ENDDO
    DO I=1,10
        uold[i] = unew[i]
    ENDDO
ENDDO
```

TODAY

- Consider stability of numerical schemes.

Discrete Numerical Schemes

By considering a 'discrete' solution, stored (and updated) only at discrete points in space and time, we have derived two possible methods to update the solution.
We have obtained the finite-difference analogue of the original partial differential equation.

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

can be approximated by

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{c}{2\Delta x}(u_{i+1}^n - u_{i-1}^n) + O(\Delta t, (\Delta x)^2) = 0.$$

This uses a 'central' difference in space. Alternatively we can use a one-sided difference, i.e.

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{c}{\Delta x}(u_{i+1}^n - u_i^n) + O(\Delta t, \Delta x) = 0$$

or

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{c}{\Delta x}(u_i^n - u_{i-1}^n) + O(\Delta t, \Delta x) = 0$$

Examination of Numerical Schemes

The scheme derived above (FTCS) is one of many possible ways we can discretise the equation numerically. We have derived the FTCS scheme by a rational mathematical approach, but how do we know that it will actually solve our equation?
It is obviously inefficient to examine every possible finite-difference stencil by coding it and testing on the computer. We need some measures:

- 1. The **ORDER OF ACCURACY** of the method is formally defined as the lowest power of Δx and Δt in the truncation error.
  Hence the FTCS scheme is 1st order accurate in time, and 2nd order accurate in space.
- 2. **CONSISTENCY** of the finite-difference analogue (FDA): a scheme is a formally consistent discretisation of the original PDE if the truncation error $\to 0$ as $\Delta x, \Delta t \to 0$, in any way.
- 3. The **STABILITY** of a numerical scheme describes whether errors are amplified or reduced by the discrete scheme (FDA); a stable scheme damps out errors whereas an unstable scheme amplifies them.

Fourier Stability Analysis

We can check the stability of our scheme by applying Fourier stability analysis, also known as Von Neumann stability analysis.

- **Step 1:** Write our numerical solution as Fourier series
  - Because the Fourier series is linear, the stability of a single mode is the same as the stability for the entire series; hence **we only need to consider the behaviour of a single Fourier mode** for the rest of the analysis
- **Step 2:** Define the amplification factor λ: *the relationship between consecutive iterations of our numerical scheme*
  - If $|\lambda| > 1$, then our scheme will amplify errors (**UNSTABLE**)
  - If $|\lambda| \leq 1$, then our scheme will damp out errors (**STABLE**)
- **Step 3:** Derive an expression for the amplification factor by substituting the Fourier mode into our FDA update relation
- **Step 4:** Calculate the magnitude of the amplification factor

Step 1: Write our numerical solution as a Fourier series

We know that a particular set of Fourier modes are exact solutions of our equation. Hence, a similar Fourier mode should be an exact solution of our FDA.
Suppose we say, for a particular value of k and ω,

$$u(x,t) = ae^{I(kx+\omega t)}$$

where $I = \sqrt{-1}$, and a is a constant, then

$$\frac{\partial u}{\partial t} = aI\omega e^{I(kx+\omega t)}$$

$$\frac{\partial u}{\partial x} = aIke^{I(kx+\omega t)}$$

Hence,

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = a(\omega + ck)Ie^{I(kx+\omega t)} = 0$$

only if

$$\omega = -kc$$

Of course we can superimpose an infinite number of these modes. We take

$$k = m\pi$$

and then

$$u(x,t) = \sum_{m=-\infty}^{+\infty} a_m e^{I(m\pi x - m\pi c t)}$$

For our FDA we can then say

$$u_i^n = \sum_{m=-\infty}^{+\infty} a_m e^{I(m\pi i\Delta x - m\pi c n\Delta t)}$$

Step 2: Define the amplification factor λ

Only one mode need be considered for a stability analysis. All other modes follow similarly. The relationship between u_i^{n+1} and u_i^n is of interest. Using the Fourier mode above (remember $e^{A+B+C} = e^A e^B e^C$ etc.)

$$u_i^n = a_m e^{I(m\pi i\Delta x - m\pi cn\Delta t)} = a_m e^{Im\pi i\Delta x} e^{-Im\pi cn\Delta t}$$

$$u_i^{n+1} = a_m e^{I(m\pi i\Delta x - m\pi c(n+1)\Delta t)} = a_m e^{Im\pi i\Delta x} e^{-Im\pi cn\Delta t} e^{-Im\pi c\Delta t}$$

$$u_i^{n+1} = u_i^n e^{-Im\pi c\Delta t}$$

We say that

$$u_i^{n+1} = \lambda^m u_i^n$$

where λ^m is called the amplification factor. Clearly this needs to ≤ 1 or the Fourier mode will grow with time, and the FDA will be unstable. (λ^m can be a function of n as well as m if the time step changes).

Step 3: Derive an expression for the amplification factor

We substitute the Fourier mode into our FDA relation and rearrange for λ.
Inserting the Fourier mode into the **FTCS update relation** (drop m superscript from λ from hereon)

$$u_i^{n+1} = u_i^n - \frac{c\Delta t}{2\Delta x}(u_{i+1}^n - u_{i-1}^n)$$

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{2\Delta x}(a_m e^{Im\pi(i+1)\Delta x} e^{-Im\pi cn\Delta t} - a_m e^{Im\pi(i-1)\Delta x} e^{-Im\pi cn\Delta t})$$

Splitting the exponential terms we obtain :

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{2\Delta x}(a_m e^{Im\pi i\Delta x} e^{-Im\pi cn\Delta t} e^{Im\pi\Delta x} - a_m e^{Im\pi i\Delta x} e^{-Im\pi cn\Delta t} e^{-Im\pi\Delta x})$$

or

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{2\Delta x} u_i^n (e^{Im\pi\Delta x} - e^{-Im\pi\Delta x})$$

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{2\Delta x} u_i^n (e^{Im\pi\Delta x} - e^{-Im\pi\Delta x})$$

Remembering

$$e^{I\theta} = \cos(\theta) + I\sin(\theta) \qquad \text{and} \qquad e^{-I\theta} = \cos(\theta) - I\sin(\theta)$$

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{2\Delta x} u_i^n 2I\sin(m\pi\Delta x)$$

Dividing through by u_i^n leaves

$$\lambda = 1 - \frac{c\Delta t}{\Delta x} I\sin(m\pi\Delta x)$$

Step 4: Calculate the magnitude of the amplification factor

$$\lambda = 1 - \frac{c\Delta t}{\Delta x} I \sin(m\pi\Delta x)$$

It is the MAGNITUDE of λ that is important, not its phase, so this is computed as

$$|\lambda| = \sqrt{1 + \frac{c^2(\Delta t)^2}{(\Delta x)^2} \sin^2(m\pi\Delta x)}$$

Hence,

$$|\lambda| > 1$$

for every mode except $m = 0$, and so the FDA is unstable as disturbances are not damped, but grow exponentially.
The amplification is the effective dissipation of the scheme. If $|\lambda| < 1$ the dissipation coefficient can be thought of as positive, if $|\lambda| > 1$ it is negative. Negative damping is clearly unstable.

Lecture 5: Numerical Stability

Part 2

University of Bristol

Fourier Stability Analysis - continued

The amplification factor derived for the FTCS scheme is:

$$\lambda = 1 - \frac{c\Delta t}{\Delta x} I \sin(m\pi\Delta x)$$

with magnitude:

$$|\lambda| = \sqrt{1 + \frac{c^2(\Delta t)^2}{(\Delta x)^2}\sin^2(m\pi\Delta x)}$$

Hence,

$$|\lambda| > 1$$

for every mode except $m = 0$, and so the FDA is unstable as disturbances are not damped, but grow exponentially.

Hence, although the FTCS scheme is consistent, **it is always unstable**.

Consistency and stability are the VITAL requirements. Order of accuracy is important, but not vital, it determines how small Δx and Δt have to be before the FDA solution $\rightarrow$ the real PDE solution.

Can we find stable, consistent methods, even if they are of lower order accuracy ?

Consider again
$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$$
As before we will take a first order temporal approximation
$$\left.\frac{\partial u}{\partial t}\right|_i^n = \frac{u_i^{n+1} - u_i^n}{\Delta t} + O(\Delta t)$$
but now consider the BACKWARD spatial difference
$$\left.\frac{\partial u}{\partial x}\right|_i^n = \frac{u_i^n - u_{i-1}^n}{\Delta x} + O(\Delta x)$$
and so
$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{c}{\Delta x}(u_i^n - u_{i-1}^n) = 0 + O(\Delta x, \Delta t)$$
or, as $\Delta x, \Delta t \to 0$ the scheme is
$$u_i^{n+1} = u_i^n - \frac{c\Delta t}{\Delta x}(u_i^n - u_{i-1}^n)$$
Considering the truncation error it is clear that the scheme is first order accurate in both space and time, i.e. less accurate than the FTCS scheme. It is consistent, so is it stable ?

Apply the same stability analysis as before (NOTE: from hereon we will always set $a_m = 1$ for simplicity, as it drops out anyway)

$$u_i^{n+1} = u_i^n - \frac{c\Delta t}{\Delta x}(u_i^n - u_{i-1}^n)$$

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{\Delta x}(e^{Im\pi i\Delta x}e^{-Im\pi cn\Delta t} - e^{Im\pi(i-1)\Delta x}e^{-Im\pi cn\Delta t})$$

Splitting the exponential terms we obtain

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{\Delta x}(e^{Im\pi i\Delta x}e^{-Im\pi cn\Delta t}$$
$$-e^{Im\pi i\Delta x}e^{-Im\pi cn\Delta t}e^{-Im\pi\Delta x})$$

or

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{\Delta x}u_i^n(1 - e^{-Im\pi\Delta x})$$

$$\lambda u_i^n - u_i^n = -\frac{c\Delta t}{\Delta x}u_i^n(1 - \cos(m\pi\Delta x) + I\sin(m\pi\Delta x))$$

We shall call

$$\nu = \frac{c\Delta t}{\Delta x}$$

then dividing through by u_i^n leaves

$$\lambda = 1 - \nu(1 - \cos(m\pi\Delta x)) - I\nu\sin(m\pi\Delta x)$$

or

$$|\lambda| = \sqrt{((1-\nu) + \nu\cos(m\pi\Delta x))^2 + (\nu\sin(m\pi\Delta x))^2}$$

Hence,

$$|\lambda|^2 = 1 - 4\nu(1-\nu)\sin^2(\frac{1}{2}m\pi\Delta x)$$

$$|\lambda|^2 = 1 - 4\nu(1-\nu)\sin^2(\frac{1}{2}m\pi\Delta x)$$

Clearly

$$|\lambda| \leq 1$$

for all modes (all m), as long as

$$0 \leq \frac{c\Delta t}{\Delta x} \leq 1$$

Hence, for $c < 0$ the scheme is always unstable, and for $c > 0$ we have the condition

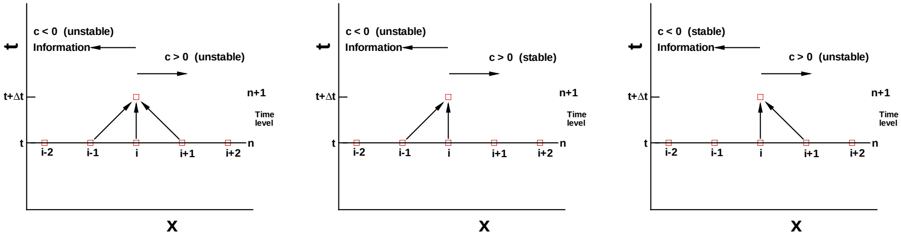
$$\boxed{\frac{c\Delta t}{\Delta x} \leq 1}$$

which is a physical limit on the time-step. This condition is of fundamental importance, and is known as the CFL condition, from a famous paper from Courant, Friedrichs, and Lewy 1928.

Three issues to consider:

1. Why is FTCS always unstable ?
2. Why is one-sided backward difference stable for $c > 0$, unstable for $c < 0$, and one-sided forward difference stable for $c < 0$, unstable for $c > 0$?
3. For stable one-sided difference, why is there a stability limit ?

Consider the stencils for the three schemes considered so far, and the signal propagation in each. Hence, it is clearly physically incorrect to propagate information against the flow direction, and the stability analysis has shown this to be incorrect. The correct propagation of signals forms the basis of UPWIND methods. In these methods information may only propagate from "upwind" or "upstream", i.e. all information must travel with the local wave direction.



The stability analysis shows that the scheme must be

$$u_i^{n+1} = u_i^n - \frac{c\Delta t}{\Delta x}(u_i^n - u_{i-1}^n) \quad \text{if} \quad c \geq 0$$

$$u_i^{n+1} = u_i^n - \frac{c\Delta t}{\Delta x}(u_{i+1}^n - u_i^n) \quad \text{if} \quad c \leq 0$$

i.e. the upwind scheme must switch the direction of spatial differencing depending on the sign of the wavespeed.

Many schemes commonly used do not satisfy this upwind condition, but to overcome the stability problem they have extra dissipation added, called "explicit dissipation". The scheme would then look like

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

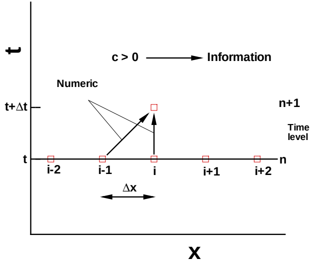
The RHS term artificially stabilises the scheme, so is also known as "artificial dissipation". This will be considered in more detail later.

Physical Interpretation of CFL Condition

Consider progressing the solution of

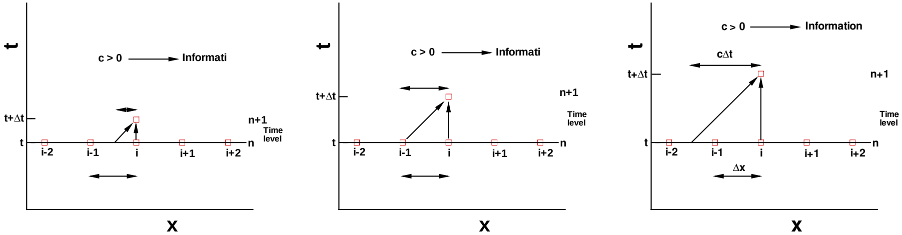
$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$$

from time level n to $n+1$ for $c > 0$. The stable upwind scheme (backward difference) uses the stencil



i.e. only u_i^n and u_{i-1}^n can influence u_i^{n+1} over one time step. However, over time Δt the solution will propagate a distance $c\Delta t$. So the physical domain of dependence is $c\Delta t$. The numerical domain of dependence is limited to Δx as only the information within one computational cell influences each point.

Consider three time step sizes, and the resulting influence



Hence, for the scheme to be stable, i.e. the correct information to influence each point, the numerical domain of dependence must be at least as large as the physical one,

$$\Delta x \geq c\Delta t \quad \text{or more commonly} \quad \frac{c\Delta t}{\Delta x} \leq 1$$

This places a limit on the numerical time step. Clearly, the smaller we make the mesh spacing (Δx) the more accurate our scheme becomes. But this then means the time-step becomes smaller and we need more time-steps to reach a converged solution. Hence, increasing the number of mesh points increases the accuracy of the scheme, but costs twice - more time-steps on more points.

Summary

• Discretisation is the concept of representing a real continuous funtion, $u(x, t)$, with a discrete solution, i.e. only stored/known at discrete points in time and space, $u(i\Delta x, n\Delta t)$.
• Taylor Series expansions used to derive numerical approximations to spatial and temporal gradients in our discrete solution, leading to a discrete time-marching scheme.
• Fourier analysis used to to investigate stability of the numerical scheme. This gives amplification factor
- FTCS scheme unconditionally unstable
- First order 'upwind' scheme has amplification factor

$$|\lambda|^2 = 1 - 4\nu(1 - \nu)\sin^2(\frac{1}{2}m\pi\Delta x)$$

• In general we have

$$\lambda = 1 - F(\Delta x^p, \Delta t^q)$$

where p and q are determined by the gradient approximation used.

• For explicit schemes, we must also satisfy the CFL condition. This results because the numerical domain of dependence MUST contain the physical one.

$$\frac{c\Delta t}{\Delta x} \leq 1$$

• The discrete scheme must model correctly the signal propagation, otherwise will be unstable - the 'upwind' principle. Non-upwind schemes can be stabilised by adding an 'artificial dissipation' term.

NEXT LECTURE: Consider more fundamental details of numerical methods.