Homework 2 correction

Q4) Implement the numerical approximation of:

\[
\frac{\partial \rho}{\partial t} + 3 \frac{\partial \rho}{\partial x} = 0
\]

\[
\rho(x, t = 0) = e^{-x^2}
\]

Geometry:

\[
x_1 = -4, x_N = 4, x_i = \left( \frac{N - i}{N - 1} \right) x_1 + \left( \frac{i - 1}{N - 1} \right) x_N
\]

Scheme:

\[
\bar{\rho}^{n+1}_i = (1 - \lambda) \bar{\rho}^n_i + (\lambda) \bar{\rho}^n_{i-1}
\]

\[
\lambda = 3 \frac{dt}{dx}
\]

Initial Condition:

\[
\bar{\rho}^0_i = \rho \left( \left( \frac{x_i + x_{i+1}}{2} \right), t = 0 \right)
\]

Boundary Condition:

\[
\bar{\rho}^0_0 = 0
\]
Homework 2 cont

Q4 cont)

For N=10,40,160,320,640,1280 run to $t=10$, with:

- $dx = \frac{8}{(N-1)}$
- $dt = dx/6$

Use

$$\rho_0^{n+1} = \left(1 - \frac{u}{dx} \frac{dt}{dx}\right)\rho_0^n + \left(\frac{u}{dx} \frac{dt}{dx}\right)\rho_{N-1}^n$$

On the same graph, plot $t$ on the horizontal axis and error on the vertical axis. The graph should consist of a sequence of 6 curves – one for each choice of $dx$.

Comment on the curves.

NOTE: For the purposes of this test we define error as:

$$error^n = \max_{1 \leq i \leq N} \left|\rho_i^n - \rho\left(\frac{x_i + x_{i+1}}{2}, ndt\right)\right|$$
Lecture 5

- We will define stability for a numerical scheme and investigate stability for the upwind scheme.

- We will compare this scheme with a finite difference scheme.

- We will consider alternative ways to approximate the flux functions.
Recall

Basic Upwind Finite Volume Method

\[
\frac{d}{dt}(\bar{q}_i dx) = -\bar{u}q(x_{i+1}, t) + \bar{u}q(x_i, t)
\]

Approximate fluxes with upwind flux

\[
\frac{d}{dt}(\bar{q}_i dx) \approx -\bar{u}q_i + \bar{u}q_{i-1}
\]

Approximate time derivative and look for solution \( \bar{\rho}_i^n \approx \bar{q}_i^n \)

\[
(\bar{\rho}_i^{n+1} - \bar{\rho}_i^n) dx = -\bar{u} \bar{\rho}_i^n + \bar{u} \bar{\rho}_{i-1}^n
dt
\]

\[
\lambda = \bar{u} \frac{dt}{dx}
\]

\[
\bar{\rho}_i^{n+1} = (1 - \lambda) \bar{\rho}_i^n + (\lambda) \bar{\rho}_{i-1}^n
\]

Note we must supply a value for the left most average at each time step: \( \bar{\rho}_0^n \)
Convergence

• We have constructed a physically reasonable numerical scheme to approximate the advection equation.

• However, we need to do some extra analysis to determine how good at approximating the true PDE the discrete scheme is.

• Let us suppose that the i’th subinterval cell average of the actual solution to the PDE at time T=n*dt is denoted by

\[
\overline{q}_i^n = \frac{1}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} q(x, n dt) \, dx = \frac{1}{dx} \int_{x_i}^{x_{i+1}} q(x, n dt) \, dx
\]

where \( q \) satisfies:

\[
\frac{d}{dt} \int_{x_i}^{x_{i+1}} q(x, t) \, dx = \frac{d}{dt}(dx \overline{q}_i) = -\overline{u} q(x_{i+1}, t) + \overline{u} q(x_i, t)
\]
Error Equation

- The goal is to estimate the difference of the exact solution and the numerically obtained solution at some time \( T=n\cdot dt \).

- So we are interested in the error: 
  \[ E_i^n = \left| \bar{q}_i^n - \bar{\rho}_i^n \right| \text{, } n = \frac{T}{dt} \]

- For the given finite volume scheme \( dt \) and \( dx \) will be related in a fixed manner (i.e. \( dt = Cdx \) for some \( C \), independent of \( dx \)).

- Suppose we let \( dt \rightarrow 0 \) and \( E^n = O\left( dt^s \right) \) then the scheme is said to be of order \( s \).
Norms and Definitions

• We define the **discrete** $p$-norms:

$$
\|E\|_p = \left( dx \sum_{i=-\infty}^{i=+\infty} |E_i|^p \right)^{1/p}
$$

• We say that the scheme is convergent at time $T$ in the norm $\|\cdot\|$ if:

$$
\lim_{\begin{cases} dt \to 0 \\ ndt = T \end{cases}} \|E^n\| = 0
$$

• It is said to be accurate of order $s$ if:

$$
\|E^n\| = O\left(dt^s\right) \text{ as } dt \to 0
$$
Local Truncation Error

- Suppose at the beginning of a time step we actually have the exact solution -- one question we can ask is how large is the error committed in the evaluation of the approximate solution at the end of the time step.

- i.e. choose
  \[ \bar{\rho}_i^n = \bar{q}_i^n \]

- Then
  \[ \bar{\rho}_i^{n+1} = \left(1 - \frac{dt}{dx} \bar{u}\right) \bar{q}_i^n + \left(\frac{dt}{dx} \bar{u}\right) \bar{q}_{i-1} \]

- We expand \( \bar{q}_{i-1} \), \( \bar{q}_i \) about \( x_i, t_n \) with Taylor series:
  \[ \bar{q}_{i-1}^n = \bar{q}_i^n - dx \left( \frac{\partial q}{\partial x} \right) + \frac{dx^2}{2} \left( \frac{\partial^2 q}{\partial x^2} \right) + O(dx^3) \]
  \[ \bar{q}_i^{n+1} = \bar{q}_i^n + dt \left( \frac{\partial q}{\partial t} \right) + \frac{dt^2}{2} \left( \frac{\partial^2 q}{\partial t^2} \right) + O(dt^3) \]
Estimating Truncation Error

• Inserting the formulas for the expanded q’s:

\[ R_i^n := \frac{1}{dt} \left( \left( 1 - \frac{dt}{dx} \bar{u} \right) \bar{q}_i^n + \left( \frac{dt}{dx} \bar{u} \right) \bar{q}_{i-1}^n - \bar{q}_i^{n+1} \right) \]

\[
= \frac{1}{dt} \left( \frac{dt}{dx} \bar{u} \right) \left\{ \bar{q}_i^n - dx \left( \frac{\partial \bar{q}}{\partial x} \right) + \frac{dx^2}{2} \left( \frac{\partial^2 \bar{q}}{\partial x^2} \right) + O(dx^3) \right\}
\]

\[
- \left\{ \bar{q}_i^n + dt \left( \frac{\partial \bar{q}}{\partial t} \right) + \frac{dt^2}{2} \left( \frac{\partial^2 \bar{q}}{\partial t^2} \right) + O(dt^3) \right\}
\]
Estimating Truncation Error

• Removing canceling terms:

\[
R_i^n := \frac{1}{dt} \left\{ \left( 1 - \frac{dt}{dx} \frac{\partial}{\partial x} \right) \bar{q}_i^n \right. \\
+ \left( \frac{dt}{dx} \frac{\partial}{\partial x} \right) \left\{ \bar{q}_i^n - dx \left( \frac{\partial \bar{q}}{\partial x} \right) + dx^2 \left( \frac{\partial^2 q}{\partial x^2} \right) + O(dx^3) \right\} \\
- \left\{ \frac{\partial}{\partial t} \left( \frac{\partial \bar{q}}{\partial t} \right) + \frac{dt^2}{2} \left( \frac{\partial^2 q}{\partial t^2} \right) + O(dt^3) \right\}
\]
Estimating Truncation Error

• Simplifying:

\[
R_i^n := \frac{1}{dt} \left\{ \left. \frac{dx}{dt} \right| \partial_q \right\} - dt \left( \frac{\partial q}{\partial t} \right) + \frac{dt^2}{2} \left( \frac{\partial^2 q}{\partial t^2} \right) + O(dt^3) \right\}
\]

\[
R_i^n := \left\{ \left( \frac{\partial q}{\partial t} \right) + \bar{u} \left( \frac{\partial q}{\partial x} \right) \right\} - \left( \frac{dt}{2} \left( \frac{\partial^2 q}{\partial t^2} \right) + O(dt^2) - \left( \bar{u} \right) \frac{dx}{2} \left( \frac{\partial^2 q}{\partial x^2} \right) + O(dx^2) \right\}
\]
Final Form

- Using the definition of \( q \)

\[
R_i^n := \left\{ \left\{ \frac{\partial q}{\partial t} + \bar{u} \frac{\partial q}{\partial x} \right\} - \left\{ \frac{dt}{2} \left( \frac{\partial^2 q}{\partial t^2} \right) + O(dt^2) - (\bar{u}) \frac{dx}{2} \left( \frac{\partial^2 q}{\partial x^2} \right) + O(dx^2) \right\} \right\}
\]

Using that:

\[
\frac{\partial^2 q}{\partial t^2} = \bar{u}^2 \frac{\partial^2 q}{\partial x^2}
\]

\[
R_i^n := \frac{\bar{u}dx}{2} \left( 1 - \frac{\bar{u}dt}{dx} \right) \left( \frac{\partial^2 q}{\partial x^2} \right) + O(dx^2)
\]
Interpretation of Consistency

\[ R^n_i := \frac{\bar{u}dx}{2} \left( 1 - \frac{\bar{u}dt}{dx} \right) \left( \frac{\partial^2 q}{\partial x^2} \right) + O(dx^2) \]

So the truncation error is \( O(dx) \) under the assumption that \( dt/dx \) is a constant..

This essentially implies that the numerical solution diverges from the actual solution by an error of \( O(dx) \) every time step.

If we assume that the solution \( q \) is smooth enough then the truncation error converges to zero with decreasing \( dx \). This property is known as consistency.
**Error Equation**

- We define the error variable: \[ \rho_i^n = q_i^n + E_i^n \]

- We next define the numerical iterator \( N: \) \[ \rho^{n+1} = N \rho^n \]

- Then:

\[
E^{n+1} = N \left( q^n + E^n \right) - q^{n+1}
= N \left( q^n + E^n \right) - N \left( q^n \right) + N \left( q^n \right) - q^{n+1}
= \left\{ N \left( q^n + E^n \right) - N \left( q^n \right) \right\} + dtR^n
\]

- So the new error consists of the action of the numerical scheme on the previous error and the error committed in the approximation of the derivatives.
Abstract Scheme

• Without considering the specific construction of the scheme suppose that the numerical $N$ operator satisfies:

$$\|N(P) - N(Q)\| \leq \|P - Q\|$$

• i.e. $N$ is a contraction operator in some norm then…
Estimating Error in Terms of Initial Error and Cumulative Truncation Error

\[ E^{n+1} = \left\{ N \left( \bar{q}^n + E^n \right) - N \left( \bar{q}^n \right) \right\} + dtR^n \]

\[ \downarrow \]

\[ \| E^{n+1} \| \leq \| N \left( \bar{q}^n + E^n \right) - N \left( \bar{q}^n \right) \| + dt \| R^n \| \quad \text{triangle inequality} \]

\[ \downarrow \]

\[ \| E^{n+1} \| \leq \| \bar{q}^n + E^n - \bar{q}^n \| + dt \| R^n \| \quad \text{contraction property of } N \]

\[ \downarrow \]

\[ \| E^{n+1} \| \leq \| E^n \| + dt \| R^n \| \]

\[ \leq \left\{ \| E^{n-1} \| + dt \| R^{n-1} \| \right\} + dt \| R^n \| \]

\[ \leq \ldots \]

\[ \downarrow \]

\[ \| E^{n+1} \| \leq \| E^0 \| + dt \sum_{m=1}^{m=n} \| R^m \| \quad \text{by induction} \]
**Error at a Time T (independent of dx,dt)**

\[
\|E^{n+1}\| \leq \|E^0\| + dt \sum_{m=1}^{m=n} \|R^m\| \\
\downarrow \\
\|E^{n+1}\| \leq \|E^0\| + T \max_{m=1,\ldots,n} \left(\|R^m\|\right)
\]

- If the method is consistent (and the actual solution is smooth enough) then:

\[
\|E^{n+1}\| \leq \|E^0\| + T \, O(dx)
\]

- As \(dx \to 0\) the initial error \(\to 0\) and consequently the numerical error at time \(T\) tends to zero with decreasing \(dx\) (and \(dt\)).
Specific Case: Stability and Consistency for the Upwind Finite Volume Scheme

• We already proved that the upwind FV scheme is consistent.

• We still need to prove stability of:

\[ \overline{\rho}_i^{n+1} = (1 - \lambda) \overline{\rho}_i^n + (\lambda) \overline{\rho}_{i-1}^n \]
Stability in the Discrete 1-norm

\[ \bar{\rho}_{i}^{n+1} = (1 - \lambda) \bar{\rho}_{i}^{n} + (\lambda) \bar{\rho}_{i-1}^{n} \]

\[ \downarrow \]

\[ \left\| \bar{\rho}^{n+1} \right\|_{1} = dx \sum_{i=1}^{N-1} \left| \bar{\rho}_{i}^{n+1} \right| \]

\[ = dx \sum_{i=1}^{N-1} \left| (1 - \lambda) \bar{\rho}_{i}^{n} + (\lambda) \bar{\rho}_{i-1}^{n} \right| \quad \text{triangle inequality} \]

\[ \leq dx \sum_{i=1}^{N-1} (1 - \lambda) \left| \bar{\rho}_{i}^{n} \right| + dx \sum_{i=1}^{N-1} \lambda \left| \bar{\rho}_{i-1}^{n} \right| \quad \text{assuming } 0 \leq \lambda \leq 1 \]

\[ \leq dx \lambda \left| \bar{\rho}_{0}^{n} \right| + dx \sum_{i=1}^{N-1} \left| \bar{\rho}_{i}^{n} \right| \]

\[ \leq dx \lambda \left| \bar{\rho}_{0}^{n} \right| + \left\| \bar{\rho}^{n} \right\|_{1} \]

So here’s the interesting story. In the case of a zero boundary condition then we automatically observe that the operator is a contraction operator.
Boundary Condition

• Suppose \( \bar{\rho}, \bar{\sigma} \) are two numerical solutions with \( \bar{\rho}_0^n = \bar{\sigma}_0^n \)

• Then:

\[
\left\| \rho^{n+1} - \sigma^{n+1} \right\|_1 \leq dx \lambda \left( \rho^n_0 - \sigma^n_0 \right) + \left\| \rho^n - \sigma^n \right\|_1 \\
\leq \left\| \rho^n - \sigma^n \right\|_1
\]

• i.e. if we are spot on with the left boundary condition the \( N \) iterator is indeed a contraction.
Relaxation on Stability Condition

• Previous contraction condition on the numerical iterator $N$
  \[
  \|N(P) - N(Q)\| \leq \|P - Q\|
  \]

• A less stringent condition is:
  \[
  \|N(P) - N(Q)\| \leq (1 + \alpha dt)\|P - Q\|
  \]

• Where alpha is a constant independent of $dt$ as $dt \rightarrow 0$
Relaxation on Stability Condition

- In this case the stability analysis yields:

\[
E^{n+1} = \left\{ N\left( \bar{q}^n + E^n \right) - N\left( \bar{q}^n \right) \right\} + dtR^n
\]

\[
\downarrow
\]

\[
\|E^{n+1}\| \leq \|\bar{q}^n + E^n - \bar{q}^n\| + dt\|R^n\|
\]

\[
\downarrow
\]

\[
\|E^{n+1}\| \leq (1 + \alpha dt)\|E^n\| + dt\|R^n\|
\]

\[
\leq (1 + \alpha dt)\left\{ (1 + \alpha dt)\|E^{n-1}\| + dt\|R^{n-1}\| \right\} + dt\|R^n\|
\]

\[
\leq ....
\]

\[
\downarrow
\]

\[
\|E^{n+1}\| \leq (1 + \alpha dt)^{n+1}\|E^0\| + dt\sum_{m=1}^{m=n} (1 + \alpha dt)^{n-m}\|R^m\|
\]

\[
\leq e^{\alpha T} \left( \|E^0\| + T \max_{m=1,...,n} \left( \|R^m\| \right) \right) Ndt=T
\]
Interpretation

\[ \| E^{n+1} \| \leq e^{\alpha T} \left( \| E^0 \| + T \max_{m=1,..,n} (\| R^m \|) \right), \text{ Ndt=T} \]

- Relaxing the stability yields a possible exponential growth – but this growth is independent of T so if we reduce dt (and dx) then the error will decay to zero for fixed T.
Next Lecture (6)

- Alternative flux formulations
- Alternative time stepping schemes