

Finite Differences: Consistency, Stability and Convergence

Varun Shankar

March 1, 2016

1 Introduction

Now that we have tackled our first space-time PDE, we will take a quick detour from presenting new FD methods, and discuss the convergence of the FD method on the heat equation (and for other PDEs).

We will center our discussion on the *Lax Equivalence Theorem*. This theorem states that a consistent FD method on a well-posed linear IVP is convergent if and only if it is stable.

2 Convergence

If U is the approximate solution to the PDE and u is the true solution, convergence is the requirement that

$$\lim_{\Delta t, h \rightarrow 0} \|U - u\| = 0. \quad (1)$$

In other words, we require that the numerical solution recover the approximate solution as our discrete differential operators become more accurate.

Practically speaking, we generally are not given u . How would one go about measuring or ascertaining convergence of U in this scenario? There are two possibilities: for simple linear problems, one could show stability and consistency, and this would imply convergence. In the more general case, one typically runs a convergence study using a *proxy* for u : simply generate an FD solution to the PDE with a very small $\Delta t, h$ pair, then use that as the “true” solution in a convergence study. All solutions with larger $\Delta t, h$ pairs are then compared to this “true” solution, and convergence is ascertained by gradually reducing $\Delta t, h$, either together or separately.

3 Consistency

Consider a PDE of the form $\mathcal{L}u = 0$, where $\mathcal{L}$ is some linear (space-time) differential operator. Consistency is the requirement that the discrete differential operator $\mathcal{L}_{\Delta t, h}$ approach the continuous operator $\mathcal{L}$ as $\Delta t, h \rightarrow 0$.

If we are deriving a new FD scheme from a polynomial interpolation approach, we can show consistency by showing that the error in the derivative of the interpolant vanishes as $\Delta t, h \rightarrow 0$. If we are, however, presented with an FD scheme without any error terms, we can show consistency by Taylor expanding the spatial approximations and temporal approximations separately, and then showing that the total truncation error goes to zero.

A natural consequence of this is that an FD scheme must be at least first-order to be consistent ($O(\Delta t) + O(h)$).

4 Stability

Stability is typically harder to ascertain for PDEs. First, recall that stability is the requirement that the error not grow from iteration to iteration of the time-stepping scheme. We would either like the error to remain constant, or, ideally, we would like the error to reduce. In other words, we would like:

$$\|e^{n+1}\| \leq \|e^n\|, \quad (2)$$

$$\implies \|U^{n+1}\| \leq \|U^n\|. \quad (3)$$

Obviously, this is only possible if the PDE problem is well-posed (and possibly also well-behaved). The discrete version of the above conditions amounts to:

$$\|U^{n+1}\| \leq \|U^n\|. \quad (4)$$

Clearly, this is only possible if the spectral radius of the time-stepping matrix is less than 1, *i.e.*, $\rho(A) \leq 1$. We can show this when we know the eigenvalues of the matrix A . However, if the eigenvalues of A are hard to find, this approach does not generalize.

Often, it is possible to use an alternative form of analysis in place of this matrix norm analysis. This method is called *Von Neumann analysis* or *Fourier analysis*. We will focus on this method of stability analysis for the remainder of this chapter, with the caveat that this analysis is equivalent to the matrix norm analysis on linear, constant coefficient IVPs with periodic or Dirichlet boundary conditions.

4.1 Stability of Forward Euler on the 1D Heat Equation

Fourier analysis proceeds as follows: express the numerical solution function $U(x, t)$ using a Fourier series. Then, substitute all (or even just one) of the Fourier modes into the FD scheme, and attempt to obtain an expression for the *growth factor* G , defined as $U^{n+1} = GU^n$. Let us write out this Fourier expansion:

$$U(x, t) = \sum_k \hat{U}(t) e^{ikx}. \quad (5)$$

Just taking one Fourier mode and evaluating on the grid, we get

$$U_j^n = \hat{U}(t_n) e^{ikjh} = \hat{U}^n e^{ikjh}. \quad (6)$$

For stability, we need $|G| \leq 1$. In practice, conditional stability manifests as a condition on Δt and h , derived from the expression $|G| \leq 1$. This will make more sense as we proceed.

Let us now write out the Forward Euler Centered Difference scheme (often called FTCS, forward in time, centered in space) for a single grid point.

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = D \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2}. \quad (7)$$

Plugging in the Fourier mode, we get

$$\frac{G \hat{U}^n e^{ikjh} - \hat{U}^n e^{ikjh}}{\Delta t} = D \frac{\hat{U}^n e^{ik(j+1)h} - 2\hat{U}^n e^{ikjh} + \hat{U}^n e^{ik(j-1)h}}{h^2}. \quad (8)$$

We can divide through by $\hat{U}^n e^{ikjh}$, leaving us with

$$\frac{G - 1}{\Delta t} = D \frac{e^{ikh} - 2 + e^{-ikh}}{h^2}. \quad (9)$$

We can use Euler's formula to simplify this. We know that $e^{ikh} + e^{-ikh} = 2 \cos(kh)$. Rearranging for G , we get

$$G = 1 - \alpha(2 \cos(kh) - 2), \quad (10)$$

where $\alpha = \frac{D \Delta t}{h^2}$. Using the half angle formula, we can further simplify this to

$$G = 1 - 4\alpha \sin^2\left(\frac{kh}{2}\right). \quad (11)$$

We require $|G| \leq 1$. This establishes a bound on α . More specifically, we have

$$-1 \leq 1 - 4\alpha \sin^2\left(\frac{kh}{2}\right) \leq 1. \quad (12)$$

Simplifying, we have $\alpha \leq \frac{1}{2}$. This means that given a D and h value, we have an upper limit on Δt for stability. This is called the CFL condition (Courant, Friedrich and Lewy). This not surprising: taking a small time-step with Forward Euler lets us stay within its stability region given the spread of the eigenvalues of L .

4.2 Stability of Backward Euler on the 1D Heat Equation

To see how things may change for an implicit method, let us perform a Fourier analysis on the BTCS (Backward in Time, Centered in Space) scheme for the Heat equation. The scheme is as follows:

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = D \frac{U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}}{h^2}. \quad (13)$$

Plugging in the Fourier mode, we have:

$$\frac{G\hat{U}^n e^{ikjh} - \hat{U}^n e^{ikjh}}{\Delta t} = DG \frac{\hat{U}^n e^{ik(j+1)h} - 2\hat{U}^n e^{ikjh} + \hat{U}^n e^{ik(j-1)h}}{h^2}. \quad (14)$$

Dividing through by $\hat{U}^n e^{ikjh}$, we are left with

$$\frac{G - 1}{\Delta t} = \frac{DG}{h^2} (e^{ikh} - 2 + e^{-ikh}). \quad (15)$$

Using Euler's formula and the half angle formula, we can simplify this to

$$\frac{G - 1}{\Delta t} = -4 \frac{DG}{h^2} \sin^2 \left(\frac{kh}{2} \right). \quad (16)$$

Multiplying through by Δt and re-arranging for G , we get

$$G + 4G\alpha \sin^2 \left(\frac{kh}{2} \right) = 1. \quad (17)$$

Re-arranging, we get

$$G = \frac{1}{1 + 4\alpha \sin^2 \left(\frac{kh}{2} \right)}. \quad (18)$$

Clearly, $|G| \leq 1$. This implies that the BTCS scheme is unconditionally stable on the heat equation.

4.3 Stability of Forward Euler on the 2D Heat Equation

What if we are dealing with the 2D heat equation? In this scenario, we have *two* spatial variables and one temporal variable. Consequently, our Fourier mode must also contain two exponentials, one for each spatial variable. Let us work this out for the FTCS scheme for a single Fourier mode. Assuming again that $h_x = h_y = h$, the scheme is:

$$\frac{U_{j,\ell}^{n+1} - U_{j,\ell}^n}{\Delta t} = D \frac{4U_{j,\ell}^n - U_{j-1,\ell}^n - U_{j+1,\ell}^n - U_{j,\ell-1}^n - U_{j,\ell+1}^n}{h^2}. \quad (19)$$

Looking at a single Fourier mode (we can now leave out the coefficient, since it will get divided out anyway), we have $U_{j,\ell}^n \sim e^{ikjh}e^{ip\ell h}$, where k and p are the wave numbers in the x and y directions, and j and ℓ are the grid indices in the x and y directions. Plugging in and dividing by $e^{ikjh}e^{ip\ell h}$, we have

$$\frac{G-1}{\Delta t} = \frac{D}{h^2}(e^{ikh} - 2 + e^{-ikh} + e^{iph} - 2 + e^{-iph}). \quad (20)$$

We can use Euler's formula and the half-angle formulae on both the k and p terms. Thus, we have

$$G = 1 - 4\alpha \left(\sin^2 \left(\frac{kh}{2} \right) + \sin^2 \left(\frac{ph}{2} \right) \right) \quad (21)$$

We require $|G| \leq 1$. We can use this to show that $\alpha \leq \frac{1}{4}$. Notice that this is a *stricter* time-step restriction than in the 1D case! It will get worse in 3D ($\alpha \leq \frac{1}{8}$).

4.4 Stability of Backward Euler on the 2D Heat Equation

We know the pattern now. We can start with the Fourier mode plugged in, the coefficients divided out, and the $e^{ikjh}e^{ip\ell h}$ divided out as well. We then get

$$\frac{G-1}{\Delta t} = \frac{DG}{h^2}(e^{ikh} - 2 + e^{-ikh} + e^{iph} - 2 + e^{-iph}). \quad (22)$$

Using Euler's formula and the half-angle formulae, this becomes

$$\frac{G-1}{\Delta t} = -4\frac{DG}{h^2} \left(\sin^2 \left(\frac{kh}{2} \right) + \sin^2 \left(\frac{ph}{2} \right) \right). \quad (23)$$

Multiplying through by Δt and collecting the G terms to the left side, we get

$$G + 4\alpha G \left(\sin^2 \left(\frac{kh}{2} \right) + \sin^2 \left(\frac{ph}{2} \right) \right) = 1. \quad (24)$$

This implies that

$$G = \frac{1}{1 + 4\alpha \left(\sin^2 \left(\frac{kh}{2} \right) + \sin^2 \left(\frac{ph}{2} \right) \right)}. \quad (25)$$

Clearly, once again $|G| \leq 1$. Backward Euler is unconditionally stable on the 2D heat equation as well.