Mathematical Tripos Part II: Michaelmas Term 2015

Numerical Analysis – Lecture 5

Revision 2.4 We consider the solution of the *diffusion equation*

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}\,, \qquad 0 \le x \le 1, \quad t \ge 0,$$

with *initial conditions* $u(x,0) = u_0(x)$ for $t = 0$ and Dirichlet *boundary conditions* $u(0,t) = \phi_0(t)$ at $x = 0$ and $u(1,t) = \phi_1(t)$ at $x = 1$. Taylor's expansion motivates the numerical scheme for approximation $u_m^n \approx u(x_m, t_n)$ on the rectangular mesh $(x_m, t_n) = (mh, nk)$:

$$u_m^{n+1} = u_m^n + \mu\left(u_{m-1}^n - 2u_m^n + u_{m+1}^n\right), \qquad m = 1...M\,. \tag{2.4}$$

Here $h = \frac{1}{M+1}$ and $\mu = \frac{k}{h^2} = \frac{\Delta t}{(\Delta x)^2}$ is the so-called *Courant number.* With μ being fixed, we have $k = \mu h^2$, so that the local truncation error of the scheme is $\mathcal{O}(h^4)$. Substituting whenever necessary initial conditions u_m^0 and boundary conditions u_0^n and u_{M+1}^n, we possess enough information to advance in (2.4) from $\boldsymbol{u}^n := [u_1^n, \ldots, u_M^n]$ to $\boldsymbol{u}^{n+1} := [u_1^{n+1}, \ldots, u_M^{n+1}]$.

Similarly to ODEs or Poisson equation, we say that the method is *convergent* if, for a fixed μ, and for every $T > 0$, we have

$$\lim_{h\to 0} |u_m^n - u(x_m, t_n)| = 0 \quad \text{uniformly for} \quad (x_m, t_n) \in [0,1]\times[0,T]\,.$$

In the present case, however, a method has an extra parameter μ, and it is entirely possible for a method to converge for some choice of μ and diverge otherwise. In particular, from Theorem 2.2 we know that if $\mu \le \frac{1}{2}$, then method (2.4) converges.

Definition 2.5 (Stability in the context of time-stepping methods for PDEs of evolution) A numerical method for a PDE of evolution is *stable* if (for zero boundary conditions) it produces a uniformly bounded approximation of the solution in any bounded interval of the form $0 \le t \le T$ when $h \to 0$ and the generalized Courant number $\mu = k/h^r$, with r being the maximum degree of the differential operator, is constant. This definition is relevant not just for the diffusion equation but for every PDE of evolution which is *well-posed*, i.e. such that its exact solution depends (in a compact time interval) in a uniformly bounded manner on the initial conditions. Thus, "stability" is nothing but the statement that well-posedness is retained under discretization, uniformly for $h \to 0$. Most PDEs of practical interest are well posed.

Theorem 2.6 (The Lax equivalence theorem) *Suppose that the underlying PDE is well-posed and that it is solved by a numerical method with an error of $\mathcal{O}(h^{p+r})$, $p \ge 1$, where r is the maximum degree of the differential operator. Then stability $\Leftrightarrow$ convergence.*

Problem 2.7 (Stability of (2.2)) Although we can deduce from the theorem that $\mu \le \frac{1}{2}$ implies stability, we will prove directly that stability $\Leftrightarrow \mu \le \frac{1}{2}$. Let $\boldsymbol{u}^n = [u_1^n, \ldots, u_M^n]^T$. We can express the recurrence (2.4) as

$$\boldsymbol{u}_h^{n+1} = A_h \boldsymbol{u}_h^n, \qquad A_h = \begin{bmatrix} 1-2\mu & \mu & & \\ \mu & \ddots & \ddots & \\ & \ddots & \ddots & \mu \\ & & \mu & 1-2\mu \end{bmatrix}_{M\times M}.$$

Here A_h is TST, hence $\lambda_\ell(A_h) = (1-2\mu) + 2\mu\cos\pi\ell h$, so that its spectrum lies within the interval $[\lambda_M, \lambda_1] = [1 - 4\mu\cos^2\frac{\pi h}{2}, 1 - 4\mu\sin^2\frac{\pi h}{2}]$. Since A_h is symmetric, we have

$$\|A_h\|_2 = \rho(A_h) = \begin{cases} |1 - 4\mu\sin^2\frac{\pi h}{2}| \le 1\,, & \mu \le \frac{1}{2}, \\ |1 - 4\mu\cos^2\frac{\pi h}{2}| > 1, & \mu > \frac{1}{2} \quad (h \le h_\mu)\,. \end{cases}$$

We distinguish between two cases.

1) $\mu \le \frac{1}{2}$: $\|\boldsymbol{u}^n\| \le \|A\| \cdot \|\boldsymbol{u}^{n-1}\| \le \cdots \le \|A\|^n \|\boldsymbol{u}^0\| \le \|\boldsymbol{u}^0\|$ as $n \to \infty$, for every $\boldsymbol{u}^0$.

2) $\mu > \frac{1}{2}$: Choose $\boldsymbol{u}^0$ as the eigenvector corresponding to the largest (in modulus) eigenvalue, λ, say. Then $\boldsymbol{u}^n = \lambda^n \boldsymbol{u}^0$, becoming unbounded as $n \to \infty$.

Technique 2.8 (Semidiscretization) Let $u_m(t) = u(mh,t)$, $m = 1...M$, $t \ge 0$. Approximating $\partial^2/\partial x^2$ as before, we deduce from the PDE that the *semidiscretization*

$$\frac{du_m}{dt} = \frac{1}{h^2}(u_{m-1} - 2u_m + u_{m+1}), \qquad m = 1...M \tag{2.5}$$

carries an error of $\mathcal{O}(h^2)$. This is an ODE system, and we can solve it by any ODE solver. Thus, Euler's method yields (2.2), while backward Euler results in

$$u_m^{n+1} - \mu(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}) = u_m^n.$$

This approach is commonly known as *the method of lines.* Much (although not all!) of the theory of finite-difference methods for PDEs of evolution can be presented as a two-stage task: first semidiscretize, getting rid of space variables, then use an ODE solver. Typically, each stage is conceptually easier than the process of discretizing in unison in both time and in space (so-called *full discretization*).

Method 2.9 (The Crank–Nicolson scheme) Discretizing the ODE (2.5) with the trapezoidal rule, we obtain

$$u_m^{n+1} - \tfrac{1}{2}\mu(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}) = u_m^n + \tfrac{1}{2}\mu(u_{m-1}^n - 2u_m^n + u_{m+1}^n), \qquad m = 1...M. \tag{2.6}$$

Thus, each step requires the solution of an $M \times M$ TST system. The error of the scheme is $\mathcal{O}(k^3 + kh^2)$, so basically the same as with Euler's method. However, as we will see, Crank–Nicolson enjoys superior stability features, as compared with the method (2.2).

Note further that (2.6) is an *implicit* method: advancing each time step requires to solve a linear algebraic system. However, the matrix of the system is TST and its solution by sparse Cholesky factorization can be done in $\mathcal{O}(M)$ operations.

Method 2.10 (Eigenvalue analysis of stability) Suppose that a numerical method (with zero boundary conditions) can be written in the form

$$\boldsymbol{u}_h^{n+1} = A_h \boldsymbol{u}_h^n,$$

where $\boldsymbol{u}_h^n \in \mathbb{R}^M$ are vectors, $A_h \in \mathbb{R}^{M\times M}$ is a matrix, and $h = \frac{1}{M+1}$. Then $\boldsymbol{u}_h^n = (A_h)^n \boldsymbol{u}_h^0$, and

$$\|\boldsymbol{u}_h^n\| = \|(A_h)^n \boldsymbol{u}_h^0\| \le \|(A_h)^n\| \cdot \|\boldsymbol{u}_h^0\| \le \|A_h\|^n \cdot \|\boldsymbol{u}_h^0\|\,,$$

for any vector norm $\|\cdot\|$ and the induced matrix norm $\|A\| = \sup \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$. If we define stability as preserving the boundedness of $\boldsymbol{u}_h^n$ with respect to the norm $\|\cdot\|$, then, from the inequality above,

$$\|A_h\| \le 1 \text{ as } h \to 0 \quad \Rightarrow \quad \text{the method is stable.}$$

Usually, the norm of $\boldsymbol{u}_h$ is set to be an averaged Euclidean length, namely, $\|\boldsymbol{u}_h\|_h := [h\sum_{i=1}^M |u_i|^2]^{1/2}$, and that does not change the Euclidean matrix norm. The reason for the factor $h^{1/2}$ is to ensure that, because of the convergence of Riemann sums, we obtain

$$\|\boldsymbol{u}_h\|_h := \left[h\textstyle\sum_{i=1}^M |u_i|^2\right]^{1/2} \to \left[\int_0^1 |u(x)|^2 \mathrm{d}x\right]^{1/2} =: \|u\|_{L_2} \qquad (h \to 0),$$

provided that u is a square-integrable function such that $\boldsymbol{u}_h = u(x)|_{\Omega_h}$.