7. Finite difference methods for the heat equation

We begin by considering the approximation of the initial boundary value problem for the heat equation in one space dimension, i.e., Find $u(x,t)$ satisfying

$$Lu \equiv \frac{\partial u}{\partial t} - \sigma \frac{\partial^2 u}{\partial x^2} = f(x,t), \quad a < x < b, \quad t > 0,$$
$$u(a,t) = 0, \quad u(b,t) = 0, \ t > 0 \qquad u(x,0) = \phi(x), \quad a < x < b.$$

To approximate problems of this type by finite difference methods, we place a mesh on the rectangle $[a,b] \times [0,T]$ of width h in the x direction and width k in the t direction. We then replace the differential equation by a difference equation and look for an approximation to $u(x,t)$ at the mesh points. From the study of elliptic problems, we know that a simple approximation to $\partial^2 u/\partial x^2(x,t)$ is given by

$$\frac{\partial^2 u}{\partial x^2}(x,t) = [u(x+h,t) - 2u(x,t) + u(x-h,t)]/h^2 + O(h^2).$$

If we approximate $\partial u/\partial t(x,t)$ by the forward difference approximation

$$\frac{\partial u}{\partial t}(x,t) = [u(x,t+k) - u(x,t)]/k + O(k)$$

and define U_j^n to be an approximation to the true solution $u(a+jh, nk)$, then we are led to the difference equation

$$[U_j^{n+1} - U_j^n]/k = \sigma[U_{j+1}^n - 2U_j^n + U_{j-1}^n]/h^2 + f_j^n,$$

where $f_j^n = f(a+jh, nk)$. This is an example of an explicit scheme, i.e., a scheme that involves only one point at the advanced time level. Since $\phi(x)$ is given, u is known at the initial time level. Hence, we have a marching scheme in time, whose solution is easily computed.

By contrast, an implicit scheme is one that involves more than one point at the advanced time level. A simple example is obtained by considering the equation at time $t+k$ and then using a backward difference approximation to $\partial u/\partial t(x, t+k)$. This leads to the difference equation:

$$[U_j^{n+1} - U_j^n]/k = \sigma[U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}]/h^2 + f_j^{n+1}.$$

This equation can no longer be solved explicitly, since there are now 3 unknown values at time $t+k$. Instead, we must use an equation at each mesh point $(a+jh, (n+1)k)$ at the advanced time level and solve a linear system of equations to simultaneously determine an approximation to u at each spatial mesh point at this time level. For example, if $h = (b-a)/J$ with $J = 4$, then the unknowns at $t = k$ would be U_1^1, U_2^1, U_3^1. The values $U_0^1 = U_4^1 = 0$ are known boundary values and the values $U_j^0 = \phi(a+jh) = \phi_j$ are the given initial values. So in this case, we get the following system of 3 equations for the 3 unknowns.

$$(U_1^1 - U_1^0)/k = \sigma[U_2^1 - 2U_1^1 + U_0^1]/h^2 + f_1^1,$$
$$(U_2^1 - U_2^0)/k = \sigma[U_3^1 - 2U_2^1 + U_1^1]/h^2 + f_2^1,$$
$$(U_3^1 - U_3^0)/k = \sigma[U_4^1 - 2U_3^1 + U_2^1]/h^2 + f_3^1.$$

In matrix form, we get after multiplication by k and setting $\lambda = \sigma k/h^2$,

$$\begin{pmatrix} 1+2\lambda & -\lambda & 0 \\ -\lambda & 1+2\lambda & -\lambda \\ 0 & -\lambda & 1+2\lambda \end{pmatrix} \begin{pmatrix} U_1^1 \\ U_2^1 \\ U_3^1 \end{pmatrix} = \begin{pmatrix} U_1^0 + kf_1^1 \\ U_2^0 + kf_2^1 \\ U_3^0 + kf_3^1 \end{pmatrix}.$$

This is a tridiagonal system and hence easy to solve.

If we average the two formulas, we get the Crank-Nicholson scheme, i.e.,

$$\frac{U_j^{n+1} - U_j^n}{k} = \frac{\sigma}{2h^2}\left[U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1} + U_{j+1}^n - 2U_j^n + U_{j-1}^n\right] + \frac{1}{2}\left[f_j^{n+1} + f_j^n\right].$$

More generally, we could take a weighted average to get

$$\frac{U_j^{n+1} - U_j^n}{k} = \frac{\sigma}{h^2}\left\{(1-\theta)[U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}] + \theta[U_{j+1}^n - 2U_j^n + U_{j-1}^n]\right\}$$
$$+ (1-\theta)f_j^{n+1} + \theta f_j^n, \qquad 0 \le \theta \le 1.$$

All these are examples of two level schemes, i.e., there are only two time levels represented in the formula. Note that by taking $\theta = 0, 1/2$, or 1, we reproduce the three previous formulas.

An example of a 3-level scheme is obtained by replacing $\partial u/\partial t(x,t)$ by the centered difference approximation $[u(x,t+k) - u(x,t-k)]/(2k)$. This leads to the difference method

$$[U_j^{n+1} - U_j^{n-1}]/2k = \sigma[U_{j+1}^n - 2U_j^n + U_{j-1}^n]/h^2 + f_j^n.$$

As we shall see later, this scheme is not a good one. Another example of a 3-level scheme is one by Dufort and Frankel (1953).

$$[U_j^{n+1} - U_j^{n-1}]/2k = \sigma[U_{j+1}^n - U_j^{n+1} - U_j^{n-1} + U_{j-1}^n]/h^2 + f_j^n.$$

8. Analysis of some basic schemes for the heat equation

To analyze these schemes, recall some of the ideas from the analysis of finite difference methods for Poisson's equation, i.e.,

$$\Delta_h U = f \quad \text{in } \Omega_h, \qquad U = g \quad \text{on } \partial\Omega_h.$$

To analyze this problem, we first established the stability result that for all mesh functions v,

$$\max_{\Omega_h \cup \partial\Omega_h} |v| \le \max_{\partial\Omega_h} |v| + \frac{1}{2}\max_{\Omega_h} |\Delta_h v|.$$

We then applied this result to the function $u - U$, where u is the exact solution of

$$\Delta u = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega.$$

Then

$$\Delta_h u - \Delta_h U = \Delta_h u - f = \Delta_h u - \Delta u.$$

The last term is the consistency error, i.e., the local truncation error of the method. Inserting a bound for this quantity, we obtain an error estimate.

We now consider a similar approach to analyze the class of θ methods discussed above for the heat equation, first deriving a stability result for this class of difference schemes. Define

$$L_{h,k}v_j^n = \frac{v_j^{n+1} - v_j^n}{k} - \frac{\sigma}{h^2}\left\{(1-\theta)[v_{j+1}^{n+1} - 2v_j^{n+1} + v_{j-1}^{n+1}] + \theta[v_{j+1}^n - 2v_j^n + v_{j-1}^n]\right\}.$$

Let $b - a = Jh$ and

$$\Omega_{h,k}^{0,m} = \{(a + jh, nk), 1 \le j \le J-1, 0 \le n \le m-1\}.$$

Theorem 9. *Suppose* $0 \le \theta \le 1$, $0 \le \sigma k/h^2 \le 1/(2\theta)$, *and* $Nk = T$. *Then for any mesh function* v *satisfying* $v_0^n = v_J^n = 0$, *we have*

$$\max_{0\le j\le J} |v_j^n| \le \max_{0\le j\le J} |v_j^0| + T \max_{\Omega_{h,k}^{0,N}} |L_{h,k}v_j^n|, \quad 0 \le n \le N.$$

Proof. Letting $\lambda = \sigma k/h^2$, and multiplying by k, we get

$$[1 + 2\lambda(1-\theta)]v_j^{n+1} = \lambda(1-\theta)[v_{j+1}^{n+1} + v_{j-1}^{n+1}] + [1 - 2\lambda\theta]v_j^n + \lambda\theta[v_{j+1}^n + v_{j-1}^n] + kL_{h,k}v_j^n.$$

Let $V^n = \max_{0\le j\le J} |v_j^n|$. Now for $0 \le \theta \le 1$, if $0 \le \lambda \le 1/(2\theta)$, then

$$1 + 2\lambda(1-\theta), \quad \lambda(1-\theta), \quad 1 - 2\lambda\theta, \quad \lambda\theta$$

are all non-negative. Hence

$$\begin{aligned}[1 + 2\lambda(1-\theta)]|v_j^{n+1}| &\le \lambda(1-\theta)[|v_{j+1}^{n+1}| + |v_{j-1}^{n+1}|] \\ &\quad + [1 - 2\lambda\theta]|v_j^n| + \lambda\theta[|v_{j+1}^n| + |v_{j-1}^n|] + k|L_{h,k}v_j^n| \\ &\le 2\lambda(1-\theta)V^{n+1} + V^n + k \max_{1\le j\le J-1} |L_{h,k}v_j^n|.\end{aligned}$$

Since $v_0 = v_J = 0$, taking the maximum over all $1 \le j \le J-1$ gives

$$[1 + 2\lambda(1-\theta)]V^{n+1} \le 2\lambda(1-\theta)V^{n+1} + V^n + k \max_{1\le j\le J-1} |L_{h,k}v_j^n|.$$

Hence

$$V^{n+1} \le V^n + k \max_{1\le j\le J-1} |L_{h,k}v_j^n|.$$

Iterating this equation, we obtain

$$V^m \le V^0 + k \sum_{n=0}^{m-1} \max_{1\le j\le J-1} |L_{h,k}v_j^n| \le V^0 + mk \max_{\Omega_{h,k}^{0,m}} |L_{h,k}v_j^n|.$$

Finally, for $0 \le m \le N$, where $Nk = T$, we get

$$V^m \le V^0 + T \max_{\Omega_{h,k}^{0,N}} |L_{h,k}v_j^n|,$$

which is just a restatement of the theorem. □

Note that to obtain this stability result, we have assumed that $0 \le \sigma k/h^2 \le 1/(2\theta)$. For the purely implicit scheme, $\theta = 0$, this is no restriction, so we say the method is unconditionally stable. For the purely explicit scheme, $\theta = 1$, and we get the stability condition $0 \le \sigma k/h^2 \le 1/2$. In fact, while this result is precise for the choices $\theta = 0$ and $\theta = 1$, one can prove a better result for $0 < \theta < 1$, i.e., we get unconditional stability for

$1/2 \le \theta \le 1$, and for $0 \le \theta < 1/2$, we have stability if $\sigma k/h^2 \le 1/[2(1-2\theta)]$. In particular, the Crank-Nicholson scheme is unconditionally stable.

To obtain an error estimate, we apply the stability result to $v = u - U$, where u is the exact solution of the original initial boundary value problem for the heat equation. Then $u - U = 0$ at boundary mesh points and at mesh points for which $t = 0$. Hence, if we let

$$\Omega_{h,k}^N = \{(a + jh, nk), 0 \le j \le J, 0 \le n \le N-1\},$$

then we easily conclude from the theorem that

$$\max_{\Omega_{h,k}^N} |u - U| \le T \max_{\Omega_{h,k}^{0,N}} |L_{h,k}(u - U)|.$$

As in the case of the elliptic problem, we have

$$\begin{aligned} L_{h,k}(u_j^n - U_j^n) &= L_{h,k}u_j^n - (1-\theta)f_j^{n+1} - \theta f_j^n \\ &= L_{h,k}u_j^n - (1-\theta)Lu(a + jh, [n+1]k) - \theta Lu(a + jh, nk), \end{aligned}$$

so the error in the method is bounded by the maximum of the local truncation error of the method. For the purely explicit method or purely implicit method, this local truncation error is of order $O(k) + O(h^2)$. For the Crank-Nicholson method, one can show that the local truncation error is $O(k^2) + O(h^2)$.