## 11. Finite difference methods for the heat equation

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

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

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 two-point boundary value 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 $\psi(x)$ 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 $h = (b-a)/N$ with $N = 4$, 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 $U_j^0 = \psi(a+jh) = \psi_j$ 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).

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

## 12. 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 two-point boundary value problems, e.g.,

$$Lu \equiv -u'' = f \quad a < x < b, \qquad u(a) = g_a, \quad u(b) = g_b.$$

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

$$\max_{0 \le j \le N} |v_j| \le \max(|v_0|, |v_N|) + \frac{(b-a)^2}{2} \max_{1 \le j \le N-1} |L_h v_j|.$$

We then applied this result to the error $u - u_j$, and used the fact that

$$L_h u - L_h u_j = L_h u - f(x_j) = L_h u - Lu.$$

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

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

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

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 35.** *Let $V_j^n$ be a function defined on the mesh points $(a+jh, nk)$. For $0 \le \theta \le 1$, $0 \le \sigma k/h^2 \le 1/(2\theta)$, and $Nk = T$,*

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

*Proof.* Letting $\lambda = \sigma k/h^2$, and multiplying by $k$, we get from the definition of $L_{h,k}$,

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

Let $W^n = \max_{0\le j\le J} |V_j^n|$. 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

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

Since $U_0 = U_J = 0$, taking the maximum over all $1 \le j \le J-1$ gives

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

Hence

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

Iterating this equation, we obtain

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

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

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

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 $0 \le \sigma k/h^2 \le 1/2$.

To obtain an error estimate, we apply the stability result to $V_j^n = u(a + jh, nk) - U_j^n$, 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

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

If we let $L_{h,k}^E$ and $L_{h,k}^I$ denote the difference operators corresponding to the explicit and implicit methods defined above, then $L_{h,k} = (1-\theta)L_{h,k}^I + \theta L_{h,k}^E$. Hence, for $(x, t) = (a+jh, nk)$ a mesh point, we have

$$\begin{aligned}
L_{h,k}(u - U)(x, t) &= (1 - \theta)L_{h,k}^I u(x, t) + \theta L_{h,k}^E u(x, t) - (1 - \theta)f^{n+1} - \theta f^n \\
&= (1 - \theta)L_{h,k}^I u(x, t) + \theta L_{h,k}^E u(x, t) - (1 - \theta)Lu(x, t + k) - \theta Lu(x, t) \\
&= (1 - \theta)[L_{h,k}^I u(x, t) - Lu(x, t + k)] + \theta[L_{h,k}^E u(x, t) - Lu(x, t)],
\end{aligned}$$

which can be bounded by the local truncation error of these methods. 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 $O(k^2) + O(h^2)$.