(2.37),

$$-1 \leqslant 1-2r^2 \sin^2(\beta h/2) \leqslant 1.$$

The only useful inequality is

$$-1 \leqslant 1-2r^2 \sin^2(\beta h/2),$$

giving

$$r = k/h \leqslant 1.$$

In reference 25, this condition is also shown to be sufficient.

The global rounding error

For simplicity, assume that all boundary values are zero so that the finite-difference equations approximating the initial-value differential equation in the solution domain $0<x<1$, $t>0$, can be written as

$$\mathbf{u}_j = \mathbf{A}\mathbf{u}_{j-1},$$

where $\mathbf{u}_0 = \mathbf{U}_0$ is the vector of known initial values and $\mathbf{A}$ is a square matrix of known elements of order $(N-1)$.

In general, the computer will not store the initial value $u_{i,0}$ exactly, but a numerical approximation $N_{i,0}$, so that

$$N_{i,0} = u_{i,0} - r_{i,0}, \quad \text{i.e.} \quad \mathbf{N}_0 = \mathbf{u}_0 - \mathbf{r}_0,$$

where $\mathbf{r}_0$ is the vector of initial rounding errors. As rounding errors will be introduced at every stage of the calculations the numerical solution values calculated by the computer at the first time-level will be

$$\mathbf{N}_1 = \mathbf{A}\mathbf{N}_0 - \mathbf{r}_1 = \mathbf{A}\mathbf{u}_0 - \mathbf{A}\mathbf{r}_0 - \mathbf{r}_1.$$

Finally, at the jth time-level, the computed solution will be

$$\mathbf{N}_j = \mathbf{A}^j\mathbf{u}_0 - \mathbf{A}^j\mathbf{r}_0 - \mathbf{A}^{j-1}\mathbf{r}_1 - \ldots - \mathbf{r}_j.$$

If there were no rounding errors the exact solution of the difference equations would be

$$\mathbf{u}_j = \mathbf{A}^j\mathbf{u}_0.$$

Hence the difference between the exact solution and the computed solution, i.e., the global rounding error $\mathbf{R}_j$, at the jth time-level is

$$\mathbf{u}_j - \mathbf{N}_j = \mathbf{A}^j\mathbf{r}_0 + \mathbf{A}^{j-1}\mathbf{r}_1 + \ldots + \mathbf{r}_j.$$

This shows that the local rounding error vector at each time-level propagates forward in the same way as the exact solution vector at that time-level. As proved earlier, the effect of each local rounding error will diminish with increasing j if $\max_i|\lambda_i|<1$, where λ_i, $i=1(1)(N-1)$, are the eigenvalues of $\mathbf{A}$, but the global rounding error cannot possibly tend to zero because of the terms $\mathbf{r}_j$, $\mathbf{A}\mathbf{r}_{j-1}, \ldots$.

Lax's equivalence theorem

Given a properly posed linear initial-value problem and a linear finite-difference approximation to it that satisfies the consistency condition, stability is the necessary and sufficient condition for convergence.

The proof of this theorem is beyond the scope of this book and interested readers should consult reference 25.

A simple example demonstrating the relationship between convergence, stability, and consistency

The classical explicit approximation to the heat conduction equation provides a comparatively simple illustration of Lax's equivalence theorem.

In general, a problem is properly posed if:

(i) The solution is unique when it exists.

(ii) The solution depends continuously on the initial data.

(iii) A solution always exists for initial data that is arbitrarily close to initial data for which no solution exists. (In heat flow problems, for example, discontinuous temperature distributions can be approximated by a sum of N continuous functions whose limiting value, as N tends to infinity, equals the discontinuous distribution except at the points of discontinuity.)

Let U satisfy the equation

$$\frac{\partial U}{\partial t}=\frac{\partial^2 U}{\partial x^2}, \quad 0<x<1, \quad t>0, \tag{2.38}$$

have known continuous initial values when $t=0$, $0\leqslant x\leqslant 1$, and known continuous boundary values at $x=0$ and 1, $t>0$.

The classical explicit approximation to (2.38) is

$$\frac{u_{i,j+1}-u_{i,j}}{k}=\frac{u_{i-1,j}-2u_{i,j}+u_{i+1,j}}{h^2}, \quad i=1(1)(N-1), \quad (2.39)$$

where $x=ih$, $t=jk$, and $Nh=1$. The local truncation error of this difference scheme is defined by

$$T_{i,j}=\frac{U_{i,j+1}-U_{i,j}}{k}-\frac{U_{i-1,j}-2U_{i,j}+U_{i+1,j}}{h^2},$$

which may be written as

$$U_{i,j+1}=kT_{i,j}+rU_{i-1,j}+(1-2r)U_{i,j}+rU_{i+1,j}, \quad (2.40)$$

where $r=k/h^2$. For $i=1(1)(N-1)$, eqns (2.40) can be written in matrix form as

$$\begin{bmatrix} U_{1,j+1} \\ U_{2,j+1} \\ \vdots \\ U_{N-1,j+1} \end{bmatrix} = k\begin{bmatrix} T_{1,j} \\ T_{2,j} \\ \vdots \\ T_{N-1,j} \end{bmatrix} + \begin{bmatrix} (1-2r) & r & & \\ r & (1-2r) & r & \\ & \ddots & \ddots & \ddots \\ & & r & (1-2r) \end{bmatrix} \times \begin{bmatrix} U_{1,j} \\ U_{2,j} \\ \vdots \\ U_{N-1,j} \end{bmatrix} + \begin{bmatrix} rU_{0,j} \\ 0 \\ \vdots \\ 0 \\ rU_{N,j} \end{bmatrix}$$

i.e. as

$$\mathbf{U}_{j+1}=k\mathbf{T}_j+\mathbf{A}\mathbf{U}_j+\mathbf{c}_j, \quad (2.41)$$

where $\mathbf{c}_j$ is a vector of known boundary values. Applying this recursively,

$$\begin{aligned}\mathbf{U}_{j+1}&=k\mathbf{T}_j+\mathbf{A}(k\mathbf{T}_{j-1}+\mathbf{A}\mathbf{U}_{j-1}+\mathbf{c}_{j-1})+\mathbf{c}_j\\ &=k(\mathbf{T}_j+\mathbf{A}\mathbf{T}_{j-1})+\mathbf{A}^2\mathbf{U}_{j-1}+(\mathbf{c}_j+\mathbf{A}\mathbf{c}_{j-1})\\ &=\dots\\ &=k(\mathbf{T}_j+\mathbf{A}\mathbf{T}_{j-1}+\dots+\mathbf{A}^j\mathbf{T}_0)+\mathbf{A}^{j+1}\mathbf{U}_0\\ &\quad+(\mathbf{c}_j+\mathbf{A}\mathbf{c}_{j-1}+\dots+\mathbf{A}^j\mathbf{c}_0).\end{aligned} \quad (2.42)$$

As the boundary and initial values for u are the same as for U, it follows from eqn (2.39), which can be written as

$$u_{i,j+1} = ru_{i-1,j} + (1-2r)u_{i,j} + ru_{i+1,j}, \; i = 1(1)(N-1),$$

that

$$\mathbf{u}_{j+1} = \mathbf{A}\mathbf{u}_j + \mathbf{c}_j.$$

This leads as before to

$$\mathbf{u}_{j+1} = \mathbf{A}^{j+1}\mathbf{u}_0 + (\mathbf{c}_j + \mathbf{A}\mathbf{c}_{j-1} + \ldots + \mathbf{A}^j\mathbf{c}_0). \tag{2.43}$$

Subtraction of (2.43) from (2.42) shows that

$$\mathbf{U}_{j+1} - \mathbf{u}_{j+1} = k(\mathbf{T}_j + \mathbf{A}\mathbf{T}_{j-1} + \ldots + \mathbf{A}^j\mathbf{T}_0) + \mathbf{A}^{j+1}(\mathbf{U}_0 - \mathbf{u}_0).$$

But $\mathbf{U}_0 = \mathbf{u}_0$ = vector of initial values. Hence

$$\mathbf{U}_{j+1} - \mathbf{u}_{j+1} = k(\mathbf{T}_j + \mathbf{A}\mathbf{T}_{j-1} + \ldots + \mathbf{A}^j\mathbf{T}_0). \tag{2.44}$$

This equation shows that the difference between the *exact* solution of the partial differential equation and the *exact* solution of the approximating difference equation at, say, the $(i, j+1)$th mesh point depends on the local truncation errors at certain mesh points on every preceding time-level and on the difference scheme used. Whether or not the accumulative effect at the $(i, j+1)$th mesh point of these preceding errors is a catastrophic build-up or a hoped-for decay as j increases depends clearly on the matrix $\mathbf{A}$ and the nature of the $T_{i,j}$, $i = 1(1)N$, $j = 0(1)J$, if the field of integration of the differential equation is the rectangle $0 \leqslant x \leqslant 1$, $0 \leqslant t = jk \leqslant T = Jk$, T finite.

Hence eqn (2.44) gives

$$\|\mathbf{U}_{j+1} - \mathbf{u}_{j+1}\| \leqslant k\{\|\mathbf{T}_j\| + \|\mathbf{A}\|\,\|\mathbf{T}_{j-1}\| + \ldots + \|\mathbf{A}^j\|\,\|\mathbf{T}_0\|\}, \quad 0 \leqslant j \leqslant J-1.$$

By the Lax–Richtmyer definition of stability,

$$\|\mathbf{A}^j\| \leqslant M, \quad j = 1(1)J,$$

where M is a positive number independent of j, h and k. Therefore,

$$\|\mathbf{U}_{j+1} - \mathbf{u}_{j+1}\| \leqslant k\,\|\mathbf{T}_j\| + kM(\|\mathbf{T}_{j-1}\| + \ldots + \|\mathbf{T}_0\|).$$

If the maximum of all of the moduli of the components of $\mathbf{T}_0$, $\mathbf{T}_1, \ldots, \mathbf{T}_j$ is C, then the infinity norm of each $\mathbf{T}_s$, $s = 1(1)j$, will be $\leqslant C$. Using this norm it follows that

$$\|\mathbf{U}_{j+1} - \mathbf{u}_{j+1}\| \leqslant kC + kMjC = kC + MtC.$$

But $jk = t \leqslant T = Jk$ is finite and $k \to 0$ as $J \to \infty$, so the first term of the right-hand side tends to zero irrespective of the magnitudes of M and C. The second term, however, tends to zero if and only if $C = \max_{i,j} |T_{i,j}|$ tends to zero as k and $h = (k/r)^{\frac{1}{2}}$ tend to zero, which, by definition, is the condition for consistency. This proves, in this particular case, that the difference scheme is convergent when it is stable and consistent.

Finite difference approximations to $\partial U/\partial t = \nabla^2 U$ in cylindrical and spherical polar co-ordinates

The non-dimensional form of the equation for heat-conduction in three dimensions is $\partial U/\partial t = \nabla^2 U$, which, in cylindrical polar coordinates (r, θ, z) is

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2} + \frac{\partial^2 U}{\partial z^2}.$$

Assuming, for simplicity, that U is independent of z, this reduces to the two-dimensional equation

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2}. \tag{2.45}$$

For non-zero values of r there is no difficulty in expressing each derivative in terms of standard finite-difference approximations, as shown in Chapter 5, but at $r = 0$ the right side appears to contain singularities. This complication can be dealt with by replacing the polar co-ordinate form of $\nabla^2 U$ by its Cartesian equivalent which transforms eqn (2.45) to the equation

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2}. \tag{2.46}$$

Now construct a circle of radius δr, centre the origin, and denote the four points in which Ox, Oy meet this circle by 1, 2, 3, 4. Denote the corresponding function values by u_1, u_2, u_3, and u_4 and the value at the origin by u_0. Then

$$\nabla^2 U = \frac{(u_1 + u_2 + u_3 + u_4 - 4u_0)}{(\delta r)^2} + O\{(\delta r)^2\}.$$

Rotation of the axes through a small angle clearly leads to a similar equation. Repetition of this rotation and the addition of all such equations then gives

$$\nabla^2 U = \frac{4(u_M - u_0)}{(\delta r)^2} + O\{(\delta r)^2\},$$

where u_M is a mean value of U round the circle. The best mean value available is given, of course, by adding all values and dividing by their number.

When a two-dimensional problem in cylindrical co-ordinates possesses circular symmetry, then $\partial^2 U/\partial\theta^2 = 0$, and eqn (2.45) simplifies to

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r}\frac{\partial U}{\partial r}. \tag{2.47}$$

Assuming $\partial U/\partial r = 0$ at $r = 0$, which it will be if the problem is symmetrical with respect to the origin, it is seen that $(1/r)\partial U/\partial r$ assumes the indeterminate form 0/0 at this point.

By Maclaurin's expansion,

$$U'(r) = U'(0) + rU''(0) + \tfrac{1}{2}r^2 U'''(0) + \ldots,$$

but $U'(0) = 0$, so the limiting value of $(1/r)\partial U/\partial r$ as r tends to zero is the value of $\partial^2 U/\partial r^2$ at $r = 0$. Hence eqn (2.47) at $r = 0$ can be replaced by

$$\frac{\partial U}{\partial t} = 2\frac{\partial^2 U}{\partial r^2}. \tag{2.48}$$

This result can also be deduced from eqn (2.46) because $\partial^2 U/\partial x^2 = \partial^2 U/\partial y^2$ from the circular symmetry, and we can make the x-axis coincide with the direction of r. The finite-difference representation of (2.48) is further simplified by the condition $\partial U/\partial r = 0$ at $r = 0$ because this gives $u_{-1,j} = u_{1,j}$. For example, the explicit approximation

$$\frac{(u_{0,j+1} - u_{0,j})}{\delta t} = \frac{2(u_{1,j} - 2u_{0,j} + u_{-1,j})}{(\delta r)^2}$$

to eqn (2.48) simplifies to

$$\frac{(u_{0,j+1} - u_{0,j})}{\delta t} = \frac{4(u_{1,j} - u_{0,j})}{(\delta r)^2}.$$ (See Example 2.13)

A complication identical to the one above also arises at $r=0$ with the spherical polar form of $\nabla^2 U$, namely

$$\frac{\partial^2 U}{\partial r^2}+\frac{2}{r}\frac{\partial U}{\partial r}+\frac{\cot}{r}\frac{\partial U}{\partial \theta}+\frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 U}{\partial \phi^2}.$$

By the same argument as in the two-dimensional case, this can be replaced at $r=0$ by $\partial^2 U/\partial x^2+\partial^2 U/\partial y^2+\partial^2 U/\partial z^2$ and approximated by $6(u_M-u_0)/(\delta r)^2$, where u_M is the mean value of U over the sphere of radius δr, centre the origin. The factor 6 occurs because Ox, Oy, Oz meet the sphere in six points. If, however, the problem is symmetrical with respect to the origin, i.e. independent of θ and ϕ, $\nabla^2 U$ reduces to $\partial^2 U/\partial r^2+(2/r)\partial U/\partial r$, with $\partial U/\partial r$ zero at $r=0$. By either of the previous arguments it follows that the heat conduction equation at $r=0$ becomes

$$\frac{\partial U}{\partial t}=3\frac{\partial^2 U}{\partial r^2}.$$

It is of interest to note that symmetrical heat flow problems for hollow cylinders and spheres that *exclude* $r=0$ can be solved by simpler equations than those considered because the change of independent variable defined by $R=\log r$ transforms the cylindrical equation

$$\frac{\partial U}{\partial t}=\frac{\partial^2 U}{\partial r^2}+\frac{1}{r}\frac{\partial U}{\partial r}\quad\text{to}\quad e^{2R}\frac{\partial U}{\partial t}=\frac{\partial^2 U}{\partial R^2},$$

and the change of dependent variable given by $U=w/r$ transforms the spherical equation

$$\frac{\partial U}{\partial t}=\frac{\partial^2 U}{\partial r^2}+\frac{2}{r}\frac{\partial U}{\partial r}\quad\text{to}\quad\frac{\partial w}{\partial t}=\frac{\partial^2 w}{\partial r^2}.$$

Example 2.13

The function U is a solution of the equation

$$\frac{\partial U}{\partial t}=\frac{\partial^2 U}{\partial x^2}+\frac{2}{x}\frac{\partial U}{\partial x},\quad 0<x<1,\, t>0,$$

and satisfies the initial conditions

$$U=1-x^2 \text{ when } t=0,\quad 0\leqslant x\leqslant 1,$$

and the boundary conditions

$$\frac{\partial U}{\partial x}=0 \text{ at } x=0, \quad t>0; \quad U=0 \text{ at } x=1, \quad t>0.$$

Using a rectangular grid defined by $\delta x=0.1$ and $\delta t=0.001$, calculate a finite-difference solution to 4D by an explicit method at the points $(0, 0.001)$, $(0.1, 0.001)$ and $(0.9, 0.001)$ in the x–t plane. (See Chapter 2, Exercise 23, for the stability of the difference scheme.) At $x=0$, $(2/x)(\partial U/\partial x)$ is indeterminate. As

$$\lim_{x\to 0}\frac{2}{x}\frac{\partial U}{\partial x}=\lim_{x\to 0} 2\frac{\partial^2 U}{\partial x^2},$$

the equation can be replaced at $x=0$ by

$$\frac{\partial U}{\partial t}=3\frac{\partial^2 U}{\partial x^2}.$$

This may be approximated by the difference equation

$$\frac{u_{0,j+1}-u_{0,j}}{\delta t}=\frac{3(u_{-1,j}-2u_{0,j}+u_{1,j})}{(\delta x)^2}. \qquad (2.49)$$

If $(\partial U/\partial x)_{i,j}$ is approximated by $(u_{i+1,j}-u_{i-1,j})/2(\delta x)$, it follows that $u_{-1,j}=u_{1,j}$ since $(\partial U/\partial x)_{0,j}=0$. Hence eqn (2.49) reduces to

$$u_{0,j+1}=u_{0,j}+3r(2u_{1,j}-2u_{0,j})=(1-6r)u_{0,j}+6ru_{1,j},$$

where $r=\delta t/(\delta x)^2=0.1$ in this example. Therefore

$$u_{0,j+1}=\tfrac{1}{5}(2u_{0,j}+3u_{1,j}). \qquad (2.50)$$

At $x\neq 0$ the differential equation can be approximated by

$$\frac{1}{\delta t}(u_{i,j+1}-u_{i,j})=\frac{1}{(\delta x)^2}(u_{i-1,j}-2u_{i,j}+u_{i+1,j})+\frac{2}{2i(\delta x)^2}(u_{i+1,j}-u_{i-1,j})$$

giving

$$u_{i,j+1}=r\left(1-\frac{1}{i}\right)u_{i-1,j}+(1-2r)u_{i,j}+r\left(1+\frac{1}{i}\right)u_{i+1,j}.$$

Therefore

$$u_{1,j+1}=\tfrac{1}{5}(4u_{1,j}+u_{2,j}) \qquad (2.51)$$

and

$$u_{9,j+1}=\tfrac{4}{5}(\tfrac{1}{9}u_{8,j}+u_{9,j}) \text{ since } u_{10,j}=0. \qquad (2.52)$$

By eqns (2.50), (2.51), and (2.52) the solution values at the points in question are as shown below

$x=$	0	0.1	0.2	...	0.8	0.9
$t=0$	1.0000	0.9900	0.9600	...	0.3600	0.1900
$t=0.001$	0.9940	0.9840				0.1840

Exercises and solutions

F.D.S. *means finite difference solution.*
A.S. *means analytical solution of the partial differential equation.*

1. Calculate a finite-difference solution of the equation

$$\frac{\partial U}{\partial t}=\frac{\partial^2 U}{\partial x^2} \quad (0<x<1, t>0)$$

satisfying the initial condition

$$U=\sin \pi x \text{ when } t=0 \text{ for } 0\leqslant x\leqslant 1,$$

and the boundary condition

$$U=0 \text{ at } x=0 \text{ and } 1 \text{ for } t>0,$$

using an explicit method with $\delta x=0.1$ and $r=0.1$.

Show by the method of separation of the variables, or merely verify, that the analytical solution is $U=e^{-\pi^2 t}\sin \pi x$. Hence check the accuracy of the numerical solution for $t=0.005$.

Solutions (Tables 2.18–2.21)

Solutions for $r=0.1$ and 0.5 and comparisons with the analytical solution at $x=0.5$ are given overpage. They show clearly that when the initial function and all its derivatives are continuous, and the boundary values at $(0,0)$ and $(1,0)$ remain equal to the initial values at these points, then the finite-difference solution can be very accurate indeed. Only one-half of the solution is shown because the problem is symmetric with respect to $x=0.5$.

2. Calculate a numerical solution of the equation $\partial U/\partial t = \partial^2 U/\partial x^2$, $0<x<1$, satisfying the initial condition $U=1$ when $t=0$, $0<x<1$, and the boundary condition $U=0$ at $x=0$ and 1, $t\geqslant 0$.