C. FINITE DIFFERENCE METHODS

The analytical solution of the B-S eqtn. is quite complicated and for other models, or where other effects are incorporated, an analytical solution is not generally possible. We now, therefore, consider how to solve such systems numerically.

1. Finite difference approximation

We first consider how to approximate derivatives. When a function $U(x)$ and its derivatives are single-valued, finite and continuous functions of x, by Taylor's theorem

$$U(x+\Delta x) = U(x) + \Delta x\, U'(x) + \tfrac{1}{2}(\Delta x)^2 U''(x) + \tfrac{1}{6}(\Delta x)^3 U'''(x) + \ldots \quad (49)$$

$$U(x-\Delta x) = U(x) - \Delta x\, U'(x) + \tfrac{1}{2}(\Delta x)^2 U''(x) - \tfrac{1}{6}(\Delta x)^3 U'''(x) + \ldots \quad (50)$$

Add:

$$U(x+\Delta x) + U(x-\Delta x) = 2U(x) + (\Delta x)^2 U''(x) + O((\Delta x)^4)$$

where $O((\Delta x)^4)$ denotes terms containing fourth and higher powers of Δx. Assuming these are negligible compared with lower powers of Δx it follows that

$$U''(x) = \left(\frac{d^2U}{dx^2}\right)_{x=x} \simeq \frac{1}{(\Delta x)^2}\left\{U(x+\Delta x) - 2U(x) + U(x-\Delta x)\right\} \quad (51)$$

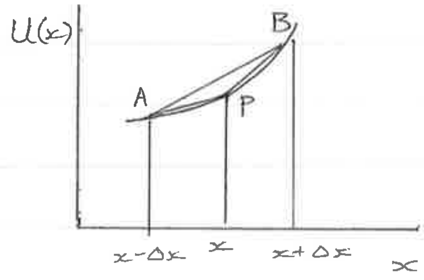
with an error of order $(\Delta x)^2$ on the rhs.

Similarly, subtraction of (50) from (49) and neglect of terms of order $(\Delta x)^3$ gives

$$U'(x) = \left(\frac{dU}{dx}\right)_{x=x} \approx \frac{1}{2\Delta x}\left\{U(x+\Delta x) - U(x-\Delta x)\right\} \quad (52)$$

with an error of order $(\Delta x)^2$.

Eqtn (52) approximates the slope of the tangent at P by the slope of the chord AB and is called a CENTRAL DIFFERENCE APPROXIMATION.



We can also approximate the slope of this tangent at P by either the slope of the chord PB giving the forward difference formula

$$\frac{dU}{dx} \simeq \frac{1}{\Delta x}\left\{ U(x+\Delta x) - U(x) \right\} \qquad (53)$$

or the slope of the chord AP giving the backward difference formula

$$\frac{dU}{dx} \simeq \frac{1}{\Delta x}\left\{ U(x) - U(x-\Delta x) \right\} \qquad (54)$$

It is seen that these can also be obtained from (49) and (50) (respectively) but that here the error on the rhs is order Δx rather than order $(\Delta x)^2$ for the central difference approx^n. Hence (52) is a better approx^n for small Δx.

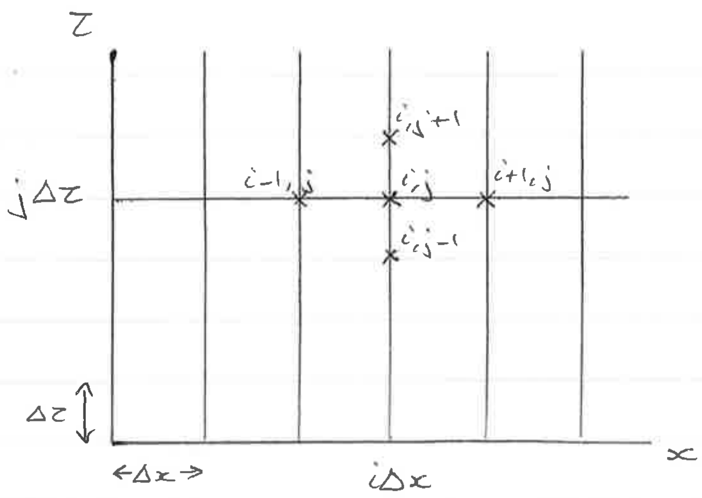
Notation for functions of two variables

Typically we shall be dealing with functions $U(x, z)$ of two variables x, z and finding them at 'grid points' (i, j) in the x, z plane, defined such that

$$x = i\Delta x \quad , \quad i = 0, 1, 2 \ldots$$
$$z = j\Delta z \quad , \quad j = 0, 1, 2 \ldots$$

where Δx and Δz are the spacings of the grid lines (or 'step lengths') in the x and z directions respectively.

The value of $U(x, z)$ at the grid point i, j will be denoted by $U_{i,j}$.



Then, for example, eqtn (51) gives the partial derivative at (i,j):

$$\frac{\partial^2 U}{\partial x^2}(x,\tau) = \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} \approx \frac{1}{(\Delta x)^2}\left\{U_{i+1,j} - 2U_{i,j} + U_{i-1,j}\right\}$$

with an ~~leading~~ error of order $(\Delta x)^2$.

Similarly, a forward difference approx^n for $\frac{\partial U}{\partial \tau}$ at (i,j) is

$$\frac{\partial U}{\partial \tau}(x,\tau) = \left(\frac{\partial U}{\partial \tau}\right)_{i,j} \approx \frac{1}{\Delta\tau}\left\{U_{i,j+1} - U_{i,j}\right\}$$

with an error of order $\Delta\tau$.

We shall denote the <u>numerical approximation</u> to U at the grid point i,j by $u_{i,j}$

2. Methods for parabolic systems

Consider the heat equation

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2} \qquad (0<x<1,\ \tau>0) \qquad (55)$$

with boundary conditions

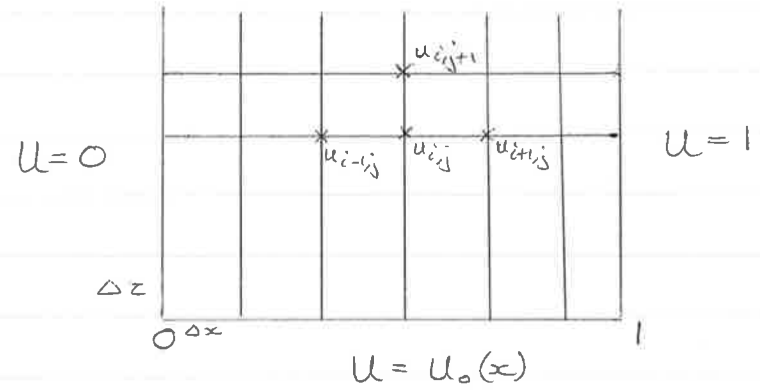
$$U = 0 \quad (x=0), \qquad U = 1 \quad (x=1) \qquad (56)$$

and initial condition

$$U = U_0(x) \quad \text{at} \quad \tau = 0 \qquad (57)$$

where U_0 is a specified function of x.

(a) Explicit method of solution



One finite difference approx^n. to (55) at (i,j) is obtained by using a forward difference approx^n. for $\frac{\partial U}{\partial \tau}$, i.e. set $\left(\frac{\partial U}{\partial \tau}\right)_{i,j} = \frac{u_{i,j+1} - u_{i,j}}{\Delta\tau}$

and a central difference approx^n. for $\frac{\partial^2 U}{\partial x^2}$, ie set

$$\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2}$$

where $x = i\Delta x$, $\tau = j\Delta\tau$. Note that these are 'exact' equals because we insert the numerical approx^n u on the rhs rather than U. We thus obtain in place of (55), at grid pt. i,j:

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

for $i = 0, 1, \ldots N$ (where $N\Delta x = 1$) and $j = 0, 1, 2, \ldots$.
This can be written as

$$u_{i,j+1} = \alpha u_{i-1,j} + (1-2\alpha)u_{i,j} + \alpha u_{i+1,j} \qquad (58)$$

(note explicit formula for $u_{i,j+1}$ - hence name of method)
where $\alpha = \Delta\tau/(\Delta x)^2$, and gives a formula for ~~[illegible]~~ U at the 'new' time step $j+1$ in terms of the solution at the previous time step j. The formula can be used at all internal grid points, i.e. $i = 1, 2 \ldots N-1$. The end values $u_{0,j+1}$ and $u_{N,j+1}$ are provided by the b.cs (56):

$$u_{0,j+1} = 0, \qquad u_{N,j+1} = 1 \qquad (59)$$

Thus a solution can be found by 'marching forwards' in time, starting from the known initial profile at $\tau = 0$:

$$u_{i,0} = U_0(i\Delta x)$$

and then applying (58) with $j = 0, j = 1, \ldots$ to determine successive rows of the solution.

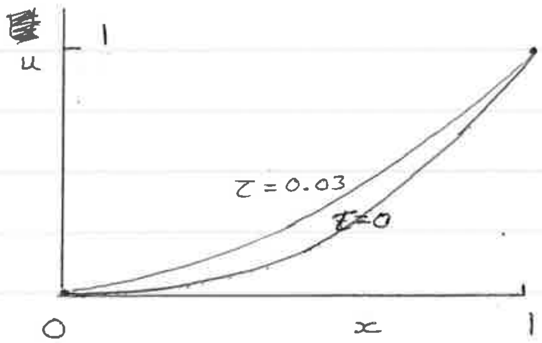
e.g. $U_0(x) = x^2$

Take $\Delta x = 0.2$, $\Delta\tau = 0.01$ so that $N = 5$ and $\alpha = \Delta\tau/(\Delta x)^2 = 1/4$. Thus (58) becomes

$$u_{i,j+1} = \tfrac{1}{4}(u_{i-1,j} + u_{i+1,j}) + \tfrac{1}{2}u_{i,j} \qquad (i = 1, 2, 3, 4)$$

		$i=0$, $x=0$	$i=1$, $x=0.2$	$i=2$, 0.4	$i=3$, 0.6	$i=4$, 0.8	$i=5$, 1
$j=0$	$\tau=0$	0	0.04	0.16	0.36	0.64	1
$j=1$	$\tau=.01$	0	0.06	0.18	0.38	0.66	1
$j=2$	$\tau=.02$	0	0.075	0.20	0.40	0.675	1
$j=3$	$\tau=.03$	0	0.0875	0.21875	0.41875	0.6875	1
⋮							

NUMERICAL SOLUTION VALUES $u_{i,j}$



Exact solution: first note that as $\tau \to \infty$ we expect $U \to x$ because if $\partial U/\partial\tau = 0$ we have $\frac{\partial^2 U}{\partial x^2} = 0$ with $U = 0$ $(x = 0)$, $U = 1$ $(x = 1)$ so $U = ax + b$ with $b = 0$, $a = 1$. Thus we write

$$U = x + \bar{U}(x, \tau)$$

Then (55), (56), (57) become

$$\frac{\partial \bar{U}}{\partial \tau} = \frac{\partial^2 \bar{U}}{\partial x^2}$$

$$\bar{U} = 0 \text{ at } x = 0 \text{ and } x = 1$$

$$\bar{U} = x^2 - x \text{ at } \tau = 0$$

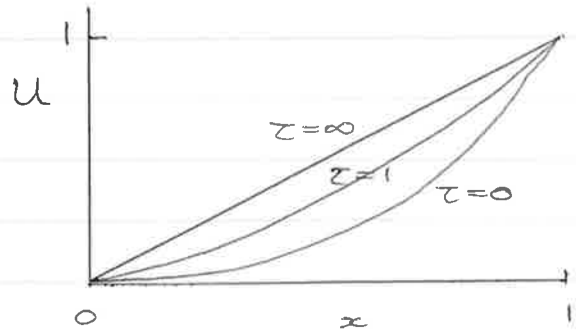
We solved this problem earlier (by sep^n. variables) but with $x - x^2$ in place of $x^2 - x$. Thus here

$$\bar{U} = -\frac{8}{\pi^3}\sum_{m=0}^{\infty} \frac{e^{-(2m+1)^2\pi^2\tau}}{(2m+1)^3} \sin(2m+1)\pi x$$

so

$$U = x - \frac{8}{\pi^3}\sum_{m=0}^{\infty} \frac{e^{-(2m+1)^2\pi^2\tau}}{(2m+1)^3} \sin(2m+1)\pi x$$

We see that $U \to x$ as $\tau \to \infty$



Comparison with numerical soln : from above formula

$$U(0.4, 0.03) = 0.21867$$

so good agreement obtained by numerical method

$u_{2,3} = 0.21875$.

<u>Effect of time step in numerical scheme</u>

Now suppose we take a larger time step in the explicit method e.g $\Delta\tau = 0.04$ so that $\alpha = \Delta\tau/(\Delta x)^2 = 1$. Then

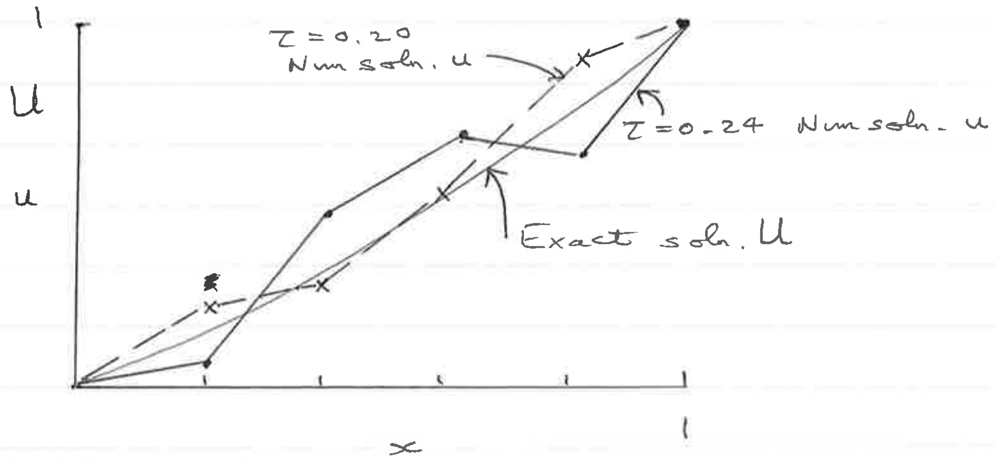
$$u_{i,j+1} = u_{i+1,j} + u_{i-1,j} - u_{i,j}$$

and we obtain:

		$i=0$ $x=0$	1 0.2	2 0.4	3 0.6	4 0.8	5 1
$j=0$	$t=0$	0	0.04	0.16	0.36	0.64	1
1	.04	0	0.12	0.24	0.44	0.72	1
2	.08	0	0.12	0.32	0.52	0.72	1
3	.12	0	0.20	0.32	0.52	0.80	1
4	.16	0	0.12	0.40	0.60	0.72	1
5	.20	0	0.28	0.32	0.52	0.88	1
6	.24	0	0.04	0.48	0.68	0.64	1

Num Soln. Values

In this case the numerical solution does not reproduce the exact solution, with large oscillations occurring about the actual solution as τ increases.



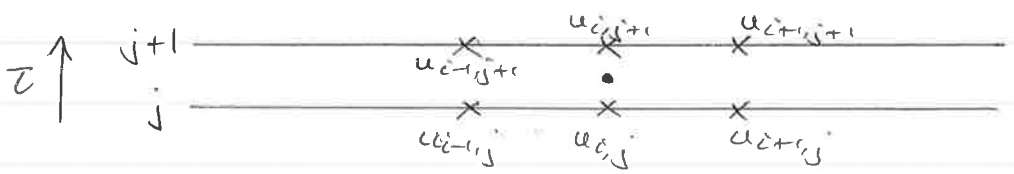
Thus it appears that the time step cannot be taken too large (later we shall show that it is necessary that $\alpha = \Delta\tau/(\Delta x)^2 \le 1/2$).

In general, of course, we expect better accuracy by taking both Δx and $\Delta\tau$ smaller.

(b) Crank-Nicolson implicit method

In 1947 Crank and Nicolson proposed a method based on a central difference approx^n to $\partial U/\partial\tau$ with the idea that this should improve accuracy and allow larger time steps to be taken.
We now discretize the equation at $(i, j+\frac{1}{2})$:

$$\left(\frac{\partial U}{\partial \tau}\right)_{i,j+\frac{1}{2}} = \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j+\frac{1}{2}}$$



using central differences on each side to get

$$\frac{u_{i,j+1} - u_{i,j}}{\Delta\tau} = \frac{1}{(\Delta x)^2}\left\{u_{i+1,j+\frac{1}{2}} - 2u_{i,j+\frac{1}{2}} + u_{i-1,j+\frac{1}{2}}\right\}$$

On the rhs we recover u at grid points by using the average

$$u_{i,j+\frac{1}{2}} = \frac{1}{2}(u_{i,j} + u_{i,j+1}) \quad \text{etc.}$$

to get

$$\frac{u_{i,j+1} - u_{i,j}}{\Delta \tau} = \frac{1}{2(\Delta x)^2}\left\{(u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}) + (u_{i+1,j} - 2u_{i,j} + u_{i-1,j})\right\}$$

Now setting $\alpha = \Delta\tau/(\Delta x)^2$, and taking unknowns ($j+1$ terms) on to lhs, and knowns (j terms) on to rhs

$$-\alpha u_{i-1,j+1} + 2(1+\alpha)u_{i,j+1} - \alpha u_{i+1,j+1} = \alpha u_{i-1,j} + 2(1-\alpha)u_{i,j} + \alpha u_{i+1,j}$$

$$(i = 1, 2, \ldots N-1) \qquad (60)$$

At a given time level j, (60) is a set of $N-1$ ~~difference~~ eqtns for $u_{i,j+1}$ $(i = 1, \ldots N-1)$ which can be written in matrix form as

$$\begin{array}{l} i=1: \\ i=2: \\ i=3: \\ \vdots \\ i=N-2: \\ i=N-1: \end{array}
\underbrace{\begin{bmatrix} 2(1+\alpha) & -\alpha & 0 & \cdots & & 0 \\ -\alpha & 2(1+\alpha) & -\alpha & 0 & \cdots & 0 \\ 0 & -\alpha & 2(1+\alpha) & -\alpha & 0 & 0 \\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & -\alpha & 2(1+\alpha) & -\alpha \\ 0 & \cdots & & 0 & -\alpha & 2(1+\alpha) \end{bmatrix}}_{\text{Matrix of coeffs}}
\underbrace{\begin{bmatrix} u_{1,j+1} \\ u_{2,j+1} \\ u_{3,j+1} \\ \vdots \\ u_{N-2,j+1} \\ u_{N-1,j+1} \end{bmatrix}}_{\text{Unknowns}}
=
\underbrace{\begin{bmatrix} \alpha u_{0,j} + 2(1-\alpha)u_{1,j} + \alpha u_{2,j} + \alpha u_{0,j+1} \\ \alpha u_{1,j} + 2(1-\alpha)u_{2,j} + \alpha u_{3,j} \\ \alpha u_{2,j} + 2(1-\alpha)u_{3,j} + \alpha u_{4,j} \\ \vdots \\ \alpha u_{N-3,j} + 2(1-\alpha)u_{N-2,j} + \alpha u_{N-1,j} \\ \alpha u_{N-2,j} + 2(1-\alpha)u_{N-1,j} + \alpha u_{N,j} + \alpha u_{N,j+1} \end{bmatrix}}_{\text{Known}}$$

where all terms on rhs are known from j time step and b.cs ($u_{0,j+1}$ and $u_{N,j+1}$).

Each time step now requires the soln of the matrix eqtn

$$A\underline{u} = \underline{d}$$

for the unknown column vector $\underline{u} = \begin{bmatrix} u_{1,j+1} \\ u_{2,j+1} \\ \vdots \\ u_{N-1,j+1} \end{bmatrix}$.

The method is now <u>implicit</u> because it no longer gives a single explicit formula for each unknown $u_{i,j+1}$ $(i = 1, \ldots N-1)$.

Gauss elimination method for solving $A\underline{u} = \underline{d}$

A special ~~treatment~~ is possible making use of the fact that A is tri-diagonal. Consider

$$\begin{bmatrix} b_1 & -c_1 & & & 0 \\ -a_2 & b_2 & -c_2 & & \\ & -a_3 & b_3 & -c_3 & \\ & & \ddots & \ddots & \ddots \\ 0 & & & -a_{N-1} & b_{N-1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{N-1} \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{N-1} \end{bmatrix}$$

Use first row to eliminate a_2 in second row.
then new second row " " a_3 " third " etc.
i.e. strip off lower diagonal of A.
The c's are unaffected by this process, giving

$$\begin{bmatrix} \beta_1 & -c_1 & & & & \\ 0 & \beta_2 & -c_2 & & & \\ 0 & 0 & \beta_3 & -c_3 & & \\ & & & \ddots & \ddots & \\ 0 & \cdots & 0 & \beta_{N-2} & -c_{N-2} \\ 0 & & \cdots & 0 & \beta_{N-1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{N-1} \end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_{N-1} \end{bmatrix}$$

where

$$\beta_1 = b_1$$

$$\beta_i = b_i - \frac{a_i c_{i-1}}{\beta_{i-1}} \qquad (i = 2, .. N-1)$$

$$S_1 = d_1$$

$$S_i = d_i + \frac{a_i S_{i-1}}{\beta_{i-1}} \qquad (i = 2, .. N-1)$$

Now we solve by backward substitution, ~~starting~~ from the final row i.e.

$$u_{N-1} = S_{N-1} / \beta_{N-1}$$

and then

$$u_i = (S_i + c_i u_{i+1}) / \beta_i \qquad (i = N-2, N-3, \ldots 1)$$

completes the solution for $\underline{u}$.

(c) Weighted average method

A generalisation of the previous approach is to approximate the eqtn by

$$\frac{u_{i,j+1}-u_{i,j}}{\Delta\tau} = \frac{1}{(\Delta x)^2}\Big\{\theta\left(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}\right)$$
$$+(1-\theta)\left(u_{i+1,j}-2u_{i,j}+u_{i-1,j}\right)\Big\}$$

where the parameter θ is normally such that $0\leq\theta\leq 1$.

$\theta = 0$ corresponds to the explicit scheme.

$\theta = \frac{1}{2}$ " " " Crank-Nicolson scheme.

$\theta = 1$ " " an (implicit) backward difference scheme.

It turns out that the approx^n is useful for any α when $\frac{1}{2}\leq\theta\leq 1$ but that for $0\leq\theta<\frac{1}{2}$ it is necessary that $\alpha\leq\frac{1}{2(1-2\theta)}$ (in particular $\alpha\leq\frac{1}{2}$ for the explicit scheme). This will be discussed later when ideas of convergence and stability have been introduced.

Note: methods (a)-(c) can also be adapted to other 2nd order parabolic equations.

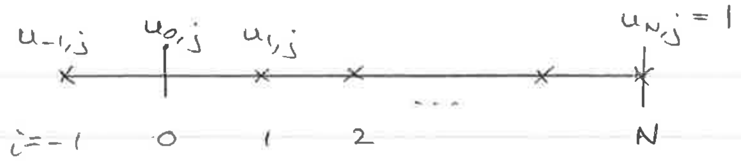
Derivative boundary conditions

Suppose we have

$$\frac{\partial U}{\partial x} = 0 \quad \text{at} \quad x = 0$$

in place of $U=0$ at $x=0$ (if U is the temp. of a metal rod, this is equivalent to no heat escaping at the end $x=0$).

This can be handled in the explicit numerical scheme as follows. Introduce a fictitious grid point $i=-1$:



Explicit scheme: $\qquad u_{i,j+1} = \alpha u_{i-1,j} + (1-2\alpha) u_{i,j} + \alpha u_{i+1,j}$ (61)

Discretization now used up to end of rod i.e.

$$i = 0, 1, \ldots \quad N-1$$

giving N eqtns for $u_{0,j+1}, \ldots u_{N-1,j+1}$
which involve $u_{-1,j}, \ldots u_{N,j}$.

But $u_{N,j} = 1$ from b.c at $x = 1$ and a central difference approx^n of $\partial U/\partial x = 0$ at $x = 0$ gives

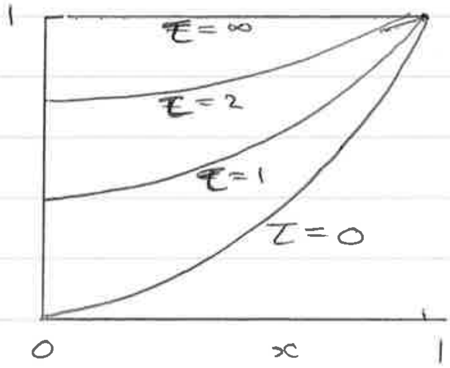
$$\frac{u_{1,j} - u_{-1,j}}{2\Delta x} = 0$$

i.e. $\qquad u_{-1,j} = u_{1,j}$

$\therefore$ All terms on rhs of (61) are known, as required.

$\alpha = 1/4$.
$\Delta \tau = .01$
$\Delta x = .2$

		$i=-1$	$i=0$	1	2	3	4	5
		$x = -.2$	$x = 0$	.2	.4	.6	.8	1
$j = 0$	$\tau = 0$	.04	0	.04	.16	.36	.64	1
1	.01	.06	.02	.06	.18	.38	.66	1
2	.02	.08	.04	.08	.20	.40	.675	1
3	.03	.10	.06	.10	.22	.41875	.6875	1



Exact solution

In this case $U \to 1$ as $\tau \to \infty$.

The C-N scheme can also be adapted, using $u_{-1,j+1} = u_{1,j+1}$ and solving $A\underline{u} = \underline{d}$ to find $\underline{u} = \begin{bmatrix} u_{0,j+1} \\ u_{1,j+1} \\ \vdots \\ u_{N-1,j+1} \end{bmatrix}$ in this case.

3. Accuracy and stability

(a) Local truncation error

Let $F_{i,j}(u) = 0$ represent the difference eqtn approximating the pde at the i,j grid point, with exact solution of this difference eqtn u.

If u is replaced by U at the grid points of the difference eqtn, where U is the exact soln. of the pde, the value of $F_{i,j}(U)$ is called the LOCAL TRUNCATION ERROR $T_{i,j}$ at the i,j grid point.

$T_{i,j}$ measures the amount by which the pde is not satisfied exactly by the numerical solution.

<u>Example:</u> Explicit scheme for the heat equation

Here $$F_{i,j}(u) = \frac{u_{i,j+1} - u_{i,j}}{\Delta \tau} - \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{(\Delta x)^2} = 0$$

$$\therefore \; T_{i,j} = F_{i,j}(U) = \frac{U_{i,j+1} - U_{i,j}}{\Delta \tau} - \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{(\Delta x)^2} \qquad (62)$$

By Taylor expansion,

$$\begin{aligned} U_{i+1,j} &= U((i+1)\Delta x, j\Delta \tau) \\ &= U_{i,j} + \Delta x \left(\frac{\partial U}{\partial x}\right)_{i,j} + \frac{1}{2}(\Delta x)^2\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} + \frac{1}{6}(\Delta x)^3\left(\frac{\partial^3 U}{\partial x^3}\right)_{i,j} + \dots \end{aligned}$$

$$\begin{aligned} U_{i-1,j} &= U((i-1)\Delta x, j\Delta \tau) \\ &= U_{i,j} - \Delta x \left(\frac{\partial U}{\partial x}\right)_{i,j} + \frac{1}{2}(\Delta x)^2\left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} - \frac{1}{6}(\Delta x)^3\left(\frac{\partial^3 U}{\partial x^3}\right)_{i,j} + \dots \end{aligned}$$

$$\begin{aligned} U_{i,j+1} &= U(i\Delta x, (j+1)\Delta \tau) \\ &= U_{i,j} + \Delta \tau \left(\frac{\partial U}{\partial t}\right)_{i,j} + \frac{1}{2}(\Delta \tau)^2\left(\frac{\partial^2 U}{\partial \tau^2}\right)_{i,j} + \frac{1}{6}(\Delta \tau)^3\left(\frac{\partial^3 U}{\partial \tau^3}\right)_{i,j} + \dots \end{aligned}$$

Substituting these in (62) $\Rightarrow$

$$T_{i,j} = \left(\frac{\partial U}{\partial \tau} - \frac{\partial^2 U}{\partial x^2}\right)_{i,j} + \tfrac{1}{2}\Delta\tau\left(\frac{\partial^2 U}{\partial \tau^2}\right)_{i,j} - \tfrac{1}{12}(\Delta x)^2\left(\frac{\partial^4 U}{\partial x^4}\right)_{i,j}$$

$$+ \tfrac{1}{6}(\Delta\tau)^2\left(\frac{\partial^3 U}{\partial \tau^3}\right)_{i,j} - \frac{1}{360}(\Delta x)^4\left(\frac{\partial^6 U}{\partial x^6}\right)_{i,j} + \cdots$$

But U satisfies the pde so $\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2}$ at (i,j)

Thus the PRINCIPAL PART (ie the dominant part) of the truncation error is

$$\left(\tfrac{1}{2}\Delta\tau\,\frac{\partial^2 U}{\partial \tau^2} - \tfrac{1}{12}(\Delta x)^2\frac{\partial^4 U}{\partial x^4}\right)_{i,j}$$

Note:

1. When $\Delta\tau = \alpha(\Delta x)^2$ with α finite, $T_{i,j}$ is $O(\Delta\tau)$ or $O((\Delta x)^2)$, as one would expect from the error in our original approx^ns to $\partial U/\partial\tau$ (order $\Delta\tau$) and $\partial^2 U/\partial x^2$ (order $(\Delta x)^2$).

2. Since

$$T_{i,j} = \tfrac{1}{12}(\Delta x)^2\left\{6\frac{\Delta\tau}{(\Delta x)^2}\frac{\partial^2 U}{\partial\tau^2} - \frac{\partial^4 U}{\partial x^4}\right\}_{i,j} + O((\Delta\tau)^2)$$

$$+ O((\Delta x)^4)$$

and since $\frac{\partial^2 U}{\partial \tau^2} = \frac{\partial^4 U}{\partial x^4}$ from the original pde,

if we put $\alpha = \Delta\tau/(\Delta x)^2 = 1/6$ the truncation error is reduced to

$$T_{i,j} = O((\Delta\tau)^2) + O((\Delta x)^4)$$

Since $\alpha = 1/6$ is quite small, this requires a fairly large number of time steps, but gives improved accuracy over the standard explicit scheme with $\alpha \neq 1/6$.

Consistency

The difference eqtn approximating a given pde is said to be CONSISTENT if the local truncation error tends to zero as $\Delta x \to 0$, $\Delta \tau \to 0$.

Our previous example shows that the explicit scheme (58) ~~method~~ is a consistent scheme for the heat eqtn.

<u>Example</u>: Consider the following difference scheme for the heat eqtn:

$$F_{i,j}(u) = \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta\tau} - \frac{u_{i+1,j} - 2\{\theta u_{i,j+1} + (1-\theta)u_{i,j-1}\} + u_{i-1,j}}{(\Delta x)^2} = 0$$

(i.e. $u_{i,j}$ replaced by $u_{i,j-1} + \theta(u_{i,j+1} - u_{i,j-1})$ and a central difference for time derivative)

The l.t.e. is

$$T_{i,j} = F_{i,j}(U)$$

$$= \frac{U_{i,j+1} - U_{i,j-1}}{2\Delta\tau} - \frac{U_{i+1,j} - 2\{\theta U_{i,j+1} + (1-\theta)U_{i,j-1}\} + U_{i-1,j}}{(\Delta x)^2}$$

where, again,

$$U_{i,j\pm1} = U_{i,j} \pm \Delta\tau \left(\frac{\partial U}{\partial \tau}\right)_{i,j} + \cdots$$

$$U_{i\pm1,j} = U_{i,j} \pm \Delta x \left(\frac{\partial U}{\partial x}\right)_{i,j} + \tfrac{1}{2}(\Delta x)^2 \left(\frac{\partial^2 U}{\partial x^2}\right)_{i,j} + \cdots$$

Substitute $\Longrightarrow$

$$T_{i,j} = \left(\frac{\partial U}{\partial \tau} - \frac{\partial^2 U}{\partial x^2}\right)_{i,j} + \left\{ \frac{1}{6}(\Delta\tau)^2 \frac{\partial^3 U}{\partial \tau^3} - \frac{(\Delta x)^2}{12}\frac{\partial^4 U}{\partial x^4} + (2\theta - 1)\frac{2\Delta\tau}{(\Delta x)^2}\frac{\partial U}{\partial \tau} + \frac{(\Delta\tau)^2}{(\Delta x)^2}\frac{\partial^2 U}{\partial \tau^2} \right\}_{i,j} + O\left(\frac{(\Delta\tau)^3}{(\Delta x)^2}, (\Delta x)^4, (\Delta\tau)^4\right)$$

Thus the truncation error in representing the heat eqtn $\partial U/\partial \tau = \partial^2 U/\partial x^2$ is

$$T_{i,j} = \left\{ \tfrac{1}{6}(\Delta\tau)^2 \frac{\partial^3 U}{\partial \tau^3} - \frac{(\Delta x)^2}{12}\frac{\partial^4 U}{\partial x^4} + (2\theta-1)\frac{2\Delta\tau}{(\Delta x)^2}\frac{\partial U}{\partial \tau} + \frac{(\Delta\tau)^2}{(\Delta x)^2}\frac{\partial^2 U}{\partial \tau^2} \right\}_{i,j}$$
$$+ O\left(\frac{(\Delta\tau)^3}{(\Delta x)^2}, (\Delta x)^4, (\Delta\tau)^4 \right)$$

As $\Delta\tau, \Delta x \to 0$ the first two terms tend to zero as required.
However the second two only tend to zero if $\Delta\tau \ll (\Delta x)^2$ and $\Delta\tau \ll \Delta x$.
i.e it is required that $\Delta\tau/(\Delta x)^2 \to 0$ as $\Delta x \to 0$ for this scheme to be consistent for the heat eqtn. (e.g $\Delta\tau = (\Delta x)^3$ would do).

Note that if $\alpha = \Delta\tau/(\Delta x)^2$ is finite the scheme would approximate the wrong eqtn i.e.

$$\frac{\partial U}{\partial \tau} - \frac{\partial^2 U}{\partial x^2} + 2(2\theta-1)\alpha\frac{\partial U}{\partial \tau} = 0$$

unless $\theta = \frac{1}{2}$.

In fact the scheme with $\theta = \frac{1}{2}$ is called the DUFORT-FRANKEL 3-time level scheme which can be used for all $\alpha > 0$ (although accuracy requires α not too large). It has the advantage that it is explicit and has second order accuracy in time, due to the use of the central difference approxn for $\partial U/\partial \tau$, i.e. the lte is (as shown above)

$$T_{i,j} = O((\Delta\tau)^2) + O((\Delta x)^2),$$

instead of $O((\Delta\tau)) + O((\Delta x)^2)$ for the standard explicit scheme.

Convergence

This addresses the possible error in the solution as $\Delta x \to 0$, as opposed to the truncation error present in the representation of the equation.

Let the exact soln of the pde be U.

Let the exact soln of the finite difference eqtn be u.

Let $e = U - u$ denote the error in the soln.

Consider the explicit scheme for the heat eqtn:

$$\frac{u_{i,j+1} - u_{i,j}}{\Delta\tau} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{(\Delta x)^2} \qquad (63)$$

At the grid pts,

$$u_{i,j} = U_{i,j} - e_{i,j}, \quad u_{i,j+1} = U_{i,j+1} - e_{i,j+1} \quad \text{etc.}$$

and substitution into (63) gives

$$e_{i,j+1} = \alpha e_{i-1,j} + (1-2\alpha)e_{i,j} + \alpha e_{i+1,j} + U_{i,j+1} - U_{i,j} + \alpha(2U_{i,j} - U_{i-1,j} - U_{i+1,j}) \qquad (64)$$

By Taylor's ~~expansion~~ theorem

$$U_{i+1,j} = U(i\Delta x + \Delta x, j\Delta\tau)$$
$$= U_{i,j} + \Delta x\left(\frac{\partial U}{\partial x}\right)_{i,j} + \frac{(\Delta x)^2}{2!}\frac{\partial^2 U}{\partial x^2}(i\Delta x + \theta_1\Delta x, j\Delta\tau) \qquad (0 < \theta_1 < 1)$$

$$U_{i-1,j} = U(i\Delta x - \Delta x, j\Delta\tau)$$
$$= U_{i,j} - \Delta x\left(\frac{\partial U}{\partial x}\right)_{i,j} + \frac{(\Delta x)^2}{2!}\frac{\partial^2 U}{\partial x^2}(i\Delta x - \theta_2\Delta x, j\Delta\tau) \qquad (0 < \theta_2 < 1)$$

$$U_{i,j+1} = U(i\Delta x, j\Delta\tau + \Delta\tau)$$
$$= U_{i,j} + \Delta\tau\left(\frac{\partial U}{\partial \tau}\right)(i\Delta x, j\Delta\tau + \theta_3\Delta\tau) \qquad (0 < \theta_3 < 1)$$

Substitution into (64) gives

$$e_{i,j+1} = \alpha e_{i-1,j} + (1-2\alpha)e_{i,j} + \alpha e_{i+1,j} + \Delta\tau\left\{\frac{\partial U}{\partial \tau}(i\Delta x, j\Delta\tau + \theta_3\Delta\tau) - \frac{\partial^2 U}{\partial x^2}(i\Delta x + \theta_4\Delta x, j\Delta\tau)\right\}$$

where $-1 < \theta_4 < 1$ (65)

Let E_j denote the max. value of $|e_{i,j}|$ along the jth time row.

Let M denote the max. modulus of $\{\ \}$ for all i,j.

When $\alpha \leq \frac{1}{2}$, all the coefficients of $e_{i,j}$ in (65) are +ve or zero.

$$\therefore |e_{i,j+1}| \leq \alpha|e_{i-1,j}| + (1-2\alpha)|e_{i,j}| + \alpha|e_{i+1,j}| + \Delta\tau M$$
$$\leq \alpha E_j + (1-2\alpha)E_j + \alpha E_j + \Delta\tau M$$
$$= E_j + \Delta\tau M$$

As this is true for all i it is true for $\max|e_{i,j+1}|$. Hence

$$E_{j+1} \leq E_j + \Delta\tau M \leq (E_{j-1} + \Delta\tau M) + \Delta\tau M$$
$$= E_{j-1} + 2\Delta\tau M \quad \text{etc.}$$
$$\therefore E_j \leq E_0 + j\Delta\tau M$$

But $E_0 + j\Delta\tau M = \tau M$ because the initial values ($j=0$) for u and U are the same, i.e. $E_0 = 0$ (because the initial state is specified by the initial condition).

Hence $\qquad E_j \leq \tau M \qquad (66)$

Now since $M = \left\{\frac{\partial U}{\partial \tau}(i\Delta x, j\Delta\tau + \theta_3\Delta\tau) - \frac{\partial^2 U}{\partial x^2}(i\Delta x + \theta_4\Delta x, j\Delta\tau)\right\}$ and $\Delta\tau = \alpha(\Delta x)^2$ tends to zero as $\Delta x \to 0$ it follows that

$$M \to \left(\frac{\partial U}{\partial \tau} - \frac{\partial^2 U}{\partial x^2}\right)_{i,j} \quad \text{as } \Delta x \to 0$$

$\therefore M \to 0$ as $\Delta x \to 0$ since U is the solution of the pde.

$\therefore$ From (66) $E_j \to 0$ as $\Delta x \to 0$ for all j, provided τ is finite and $\alpha \leq \frac{1}{2}$.

But $|U_{i,j} - u_{i,j}| \leq E_j$

Hence u converges to U as $\Delta x \to 0$ provided $\alpha \leq \frac{1}{2}$ and τ is finite.

Vector and matrix norms

(a) Vector norms

The norm of a vector $\underline{x}$ is a real +ve number giving a measure of the 'size' of the vector and is denoted by $\|\underline{x}\|$. It must satisfy the following axioms:

(i) $\|\underline{x}\| > 0$ if $\underline{x} \neq \underline{0}$ and $\|\underline{x}\| = 0$ if $\underline{x} = \underline{0}$.

(ii) $\|c\underline{x}\| = |c|\,\|\underline{x}\|$ for a real or complex scalar c.

(iii) $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$.

Examples for a vector $\underline{x}$ with n components.

"1-norm" $\quad \|\underline{x}\|_1 \equiv |x_1| + |x_2| + \dots |x_n| = \sum_{i=1}^{n} |x_i|$

"∞-norm" $\quad \|\underline{x}\|_\infty \equiv \max_i |x_i|$

(b) Matrix norms

A $n \times n$ matrix

The norm of a matrix A is a real +ve number giving a measure of the 'size' of the matrix and must satisfy the following axioms:

(i) $\|A\| > 0$ if $A \neq 0$ and $\|A\| = 0$ if $A = 0$.

(ii) $\|cA\| = |c|\,\|A\|$ for a real or complex scalar c.

(iii) $\|A + B\| \leq \|A\| + \|B\|$

(iv) $\|AB\| \leq \|A\|\,\|B\|$

Compatible norms

Where matrices and vectors occur together it is essential that they satisfy a condition equivalent to (iv), i.e.

(v) $\|A\underline{x}\| \leq \|A\|\,\|\underline{x}\|$, $\underline{x} \neq \underline{0}$.

Thus there are appropriate matrix norms, corresponding to the 2 vector norms defined above, which are arranged to satisfy (v):

"1-norm" : $\|A\|_1$ is the max column sum of the moduli of elts of A : $\max_j \{\sum_{i=1}^n |a_{ij}|\}$

"∞-norm" : $\|A\|_\infty$ is the max row sum of the moduli of elts of A: $\max_i \{\sum_{j=1}^n |a_{ij}|\}$

Example : $A = \begin{pmatrix} -1 & 1 \\ 3 & -2 \end{pmatrix}$

Then $\|A\|_1 = 1+3 = 4$

$\|A\|_\infty = 3+2 = 5$

Proof of compatibility for the ∞-norm

Let $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ have ∞ norm $\|\underline{x}\|_\infty = |x_K|$ say.

Let $A = \begin{bmatrix} a_{11} \cdots & a_{1n} \\ \vdots & \\ a_{n1} & a_{nn} \end{bmatrix}$ have ∞ norm $\|A\|_\infty = |a_{I1}| + |a_{I2}| + \ldots + |a_{In}|$ say

Note

$$A\underline{x} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{i1}x_1 + \cdots + a_{in}x_n \\ a_{n1}x_1 + \cdots + a_{nn}x_n \end{bmatrix}$$

Consider

$$\begin{aligned}
\|A\|_\infty \|\underline{x}\|_\infty &= (|a_{I1}| + |a_{I2}| + \cdots |a_{In}|)|x_K| \\
&\geq (|a_{i1}| + \cdots |a_{in}|)|x_K| \quad \text{for all } i = 1, \ldots n
\end{aligned}$$

since $i = I$ is max mod row sum of A

$$\begin{aligned}
&\geq |a_{i1}|x_K| + \cdots + a_{in}|x_K|| \\
&\geq |a_{i1}|x_1| + \cdots + a_{in}|x_n|| \quad \text{since } |x_K| \geq |x_i|
\end{aligned}$$

for all $i = 1, \ldots n$ since $|x_K|$ is ∞-norm of $\underline{x}$.

$$\begin{aligned}
&\geq |a_{i1}x_1 + \cdots + a_{in}x_n| \quad \text{for all } i = 1, \ldots n \\
&\geq \max_i |a_{i1}x_1 + \cdots + a_{in}x_n| = \|A\underline{x}\|_\infty
\end{aligned}$$

as required

Stability of a numerical scheme

This addresses whether the scheme will remain accurate as Z increases i.e whether the numerical solution of the f. d. eqtns will remain close to the exact solution of the pde as τ increases.

The ~~expl~~ solution of the f. d. eqtns advances the solution from the initial line to time levels $\Delta\tau, 2\Delta\tau, \ldots$

If we introduce a small error at $\tau = 0$, will this remain small as τ increases, or will it amplify?

Assume that the f. d. eqtns relating time level j to $j+1$ can be expressed as

$$b_{i-1}u_{i-1,j+1} + b_i u_{i,j+1} + b_{i+1}u_{i+1,j+1} = c_{i-1}u_{i-1,j} + c_i u_{i,j} + c_{i+1}u_{i+1,j}$$

where the coeffs are indep. of j (equivalent to coeffs indep. of t in the pde).

If the b.cs at $i = 0$ and $i = N$ are known, this can be written in matrix form as:

$$\begin{bmatrix} b_1 & b_2 & & & \\ b_1 & b_2 & b_3 & & \\ & & \ddots & & \\ & & b_{N-3} & b_{N-2} & b_{N-1} \\ & & & b_{N-2} & b_{N-1} \end{bmatrix} \begin{bmatrix} u_{1,j+1} \\ u_{2,j+1} \\ \vdots \\ \\ u_{N-1,j+1} \end{bmatrix} =$$

$$\begin{bmatrix} c_1 & c_2 & & & \\ c_1 & c_2 & c_3 & & \\ & & \ddots & & \\ & & c_{N-3} & c_{N-2} & c_{N-1} \\ & & & c_{N-2} & c_{N-1} \end{bmatrix} \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ \\ u_{N-1,j} \end{bmatrix} + \begin{bmatrix} c_0 u_{0,j} - b_0 u_{0,j+1} \\ 0 \\ \vdots \\ 0 \\ c_N u_{N,j} - b_N u_{N,j+1} \end{bmatrix}$$

i.e.

$$B\,\underline{u}_{j+1} = C\,\underline{u}_j + \underline{d}_j$$

where B, C are $(N-1)\times(N-1)$ matrices, $\underline{u}_j$, $\underline{u}_{j+1}$, $\underline{d}_j$ are column vectors. Hence

$$\underline{u}_{j+1} = B^{-1}C\,\underline{u}_j + B^{-1}\underline{d}_j$$

$$\therefore \quad \underline{u}_{j+1} = A\underline{u}_j + \underline{f}_j \qquad (67)$$

where $A = B^{-1}C$, $\underline{f}_j = B^{-1}\underline{d}_j$.

(67) provides the method of advancing from the initial state $\underline{u}_0$, with A indep. of j, and $\underline{f}_0, \ldots \underline{f}_j \ldots$ vectors of known boundary values.

To investigate stability, introduce a small pert^n. to the initial data so that $\underline{u}_0$ becomes

$$\underline{u}_0 + \underline{e}_0 \quad \text{say}.$$

Then (67) $\Rightarrow$

$$\underline{u}_{j+1} + \underline{e}_{j+1} = A(\underline{u}_j + \underline{e}_j) + \underline{f}_j \qquad (68)$$

Subtract (67) from (68):

$$\underline{e}_{j+1} = A\underline{e}_j$$

and, applying this in succession, $\underline{e}_j = A^j \underline{e}_0$.

Thus $\qquad \|\underline{e}_{j+1}\| = \|A\underline{e}_j\| \leq \|A\|\|\underline{e}_j\|$

$$\therefore \quad \frac{\|\underline{e}_{j+1}\|}{\|\underline{e}_j\|} \leq \|A\| \qquad j = 0, 1, \ldots$$

Thus for the 'error norm' to decrease with j it is sufficient that

$$\underline{\|A\| < 1}$$

Then
$$\begin{aligned}\|\underline{e}_j\| = \|A^j \underline{e}_0\| &\leq \|A^j\|\|\underline{e}_0\| \\ &= \|A A^{j-1}\|\|\underline{e}_0\| \\ &\leq \|A\|\|A^{j-1}\|\|\underline{e}_0\| \\ &\ldots \\ &\leq \|A\|^j\|\underline{e}_0\|\end{aligned}$$

so that with $\|A\| < 1$ the error decreases to zero as $j \to \infty$.

If $\|A\| = 1$ then the error remains bounded as $\mathcal{I}$ increases, which is generally acceptable (errors are introduced in practice at every time step due to rounding in the arithmetic). In general then stability is guaranteed if

$$\|A\| \leq 1 \qquad (69)$$

Example   Explicit method for heat eqtn.

$$u_{i,j+1} = \alpha u_{i-1,j} + (1-2\alpha)u_{i,j} + \alpha u_{i+1,j}$$

where $\alpha = \Delta\tau/(\Delta x)^2 >$

Clearly here

$$A = \begin{bmatrix} 1-2\alpha & \alpha & & & 0 \\ \alpha & 1-2\alpha & \alpha & & \\ & \alpha & 1-2\alpha & \alpha & \\ & & & \ddots & \\ 0 & & & \alpha & 1-2\alpha \end{bmatrix}$$

assuming $u_{0,j}$, $u_{N,j}$ known from b.c.s.

If $\alpha \leq \frac{1}{2}$, $\quad 1-2\alpha > 0$ so

$$\|A\|_\infty = \alpha + (1-2\alpha) + \alpha = 1 \quad : \text{acceptable}$$

If $\alpha > \frac{1}{2}$, $\quad 1-2\alpha < 0$ so

$$\|A\|_\infty = \alpha + 2\alpha - 1 + \alpha = 4\alpha - 1 > 1 :$$ will generally lead to growth of perturbations.

Hence the explicit method can only be used in practice for $\alpha \leq \frac{1}{2}$; otherwise it is unstable.

This result is related to the earlier result showing convergence of the scheme for $\alpha \leq \frac{1}{2}$.

Generally stability and convergence are related by LAX'S EQUIVALENCE THEOREM: for a linear i.v. problem and a linear f.d. approx^n to it that satisfies the consistency condition, stability is a necessary and sufficient condition for convergence. (see Smith for further details)

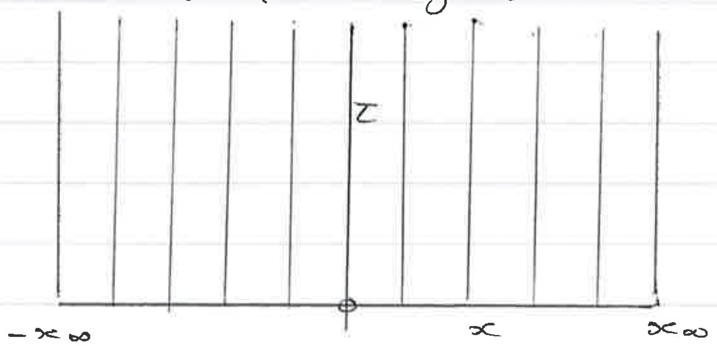
4. Black-Scholes equation

Method 1 Use the transformed version, for example for the European call option we solve for $U(x,\tau)$ where

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2}$$

$$U(x,0) = U_0(x) = e^{\frac{1}{2}(k+1)x} - e^{\frac{1}{2}(k-1)x}, \quad x > 0$$
$$= 0, \quad x < 0$$
$$U \sim e^{\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\tau}, \quad x \to \infty$$
$$U \to 0, \quad x \to -\infty.$$

In practice we use an outer grid boundary x_∞ and solve in the region $-x_\infty < x < x_\infty$, $\tau > 0$



using either the explicit, C-N or similar scheme. x_∞ must be chosen sufficiently large to ensure an accurate solution.

Method 2 Use the actual domain $S \geq 0$. First non-dimensionalise:

$$t = T - \frac{\tau}{\frac{1}{2}\sigma^2}, \quad r = \frac{1}{2}\sigma^2 k$$

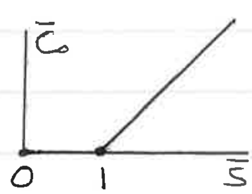
$$S = E\bar{S}, \quad C = E\bar{C}(\bar{S},\tau)$$

so that for a European call

$$\frac{\partial \bar{C}}{\partial \tau} = \bar{S}^2 \frac{\partial^2 \bar{C}}{\partial \bar{S}^2} + k\bar{S}\frac{\partial \bar{C}}{\partial \bar{S}} - k\bar{C} \qquad (70)$$

with initial condition

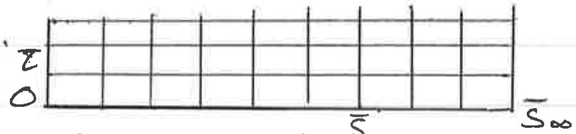
$$\bar{C} = \bar{C}_0(\bar{S}) = \bar{S} - 1, \quad \bar{S} > 1$$
$$= 0, \quad 0 \leq \bar{S} < 1$$



and boundary conditions

$$\bar{C} = \bar{S}, \quad \bar{S} \to \infty$$
$$\bar{C} = 0, \quad \bar{S} = 0$$

We solve on the domain $0 \leq \bar{S} \leq \bar{S}_\infty$, $\tau > 0$ with $\bar{S}_\infty$ chosen sufficiently large to give an accurate solution.



This approach simplifies the b.c.s. and the results can easily be used to give the physically important quantities for financial applications. Either explicit or C–N schemes can be used as follows.

(a) Explicit scheme

Steps $\Delta\tau$, $\Delta\bar{S}$

Let $\bar{S} = i\Delta\bar{S}$, $\tau = j\Delta\tau$

Then (70) becomes

$$\frac{\bar{c}_{i,j+1} - \bar{c}_{i,j}}{\Delta\tau} = (i\Delta\bar{S})^2\left(\frac{\bar{c}_{i+1,j} - 2\bar{c}_{i,j} + \bar{c}_{i-1,j}}{(\Delta\bar{S})^2}\right) + k(i\Delta\bar{S})\left(\frac{\bar{c}_{i+1,j} - \bar{c}_{i-1,j}}{2\Delta\bar{S}}\right) - k\bar{c}_{i,j}$$

where $\bar{c}_{i,j}$ is the numerical approxn. to $\bar{C}$ at grid pt. i,j.

Hence at internal grid pts

$$\bar{c}_{i,j+1} = \bar{c}_{i,j} + \Delta\tau\left\{ i^2(\bar{c}_{i+1,j} - 2\bar{c}_{i,j} + \bar{c}_{i-1,j}) + \frac{ki}{2}(\bar{c}_{i+1,j} - \bar{c}_{i-1,j}) - k\bar{c}_{i,j} \right\} \quad (71)$$

$(i = 1, \ldots N-1)$

and at the two ends

$$\bar{c}_{0,j+1} = 0 \quad , \quad \bar{c}_{N,j+1} = Nh$$

$\bar{c}_{N,j+1} = \bar{c}_{N-1,j+1}$

The initial condition is

$$\bar{c}_{i,0} = \bar{C}_0(i\Delta\bar{S})$$

<u>Note</u> – for stability of the scheme we expect $\frac{\Delta\tau}{(\Delta X)^2} \leq \frac{1}{2}$ in $U(X,\tau)$ variables, and since

$\bar{S} = e^X \Rightarrow X = \ln\bar{S} \Rightarrow \Delta X = \bar{S}^{-1}\Delta\bar{S}$ we will need $\frac{\Delta\tau}{\bar{S}^{-2}(\Delta\bar{S})^2} \leq \frac{1}{2}$

$$\Rightarrow \quad \frac{\Delta\tau}{(\Delta\bar{S})^2} \leq \frac{1}{2}\bar{S}^{-2}$$

Thus we expect to require $\frac{\Delta\tau}{(\Delta\bar{S})^2} \leq \frac{1}{2}\bar{S}_\infty^{-2}$ in practice (since this will be the <u>minimum</u> value of $\bar{S}^{-2}$ over the domain).

(b) Crank–Nicolson scheme

Here the formula (71) will be replaced by using

$$\frac{\bar{c}_{i,j+1} - \bar{c}_{i,j}}{\Delta\tau} = (i\Delta\bar{S})^2\left\{\frac{\bar{c}_{i+1,j+1} - 2\bar{c}_{i,j+1} + \bar{c}_{i-1,j+1}}{2(\Delta\bar{S})^2} + \frac{\bar{c}_{i+1,j} - 2\bar{c}_{i,j} + \bar{c}_{i-1,j}}{2(\Delta\bar{S})^2}\right\}$$

$$+ k\,i\Delta\bar{S}\left\{\frac{\bar{c}_{i+1,j+1} - \bar{c}_{i-1,j+1}}{4\Delta\bar{S}} + \frac{\bar{c}_{i+1,j} - \bar{c}_{i-1,j}}{4\Delta\bar{S}}\right\}$$

$$- k\left(\frac{\bar{c}_{i,j+1} + \bar{c}_{i,j}}{2}\right) \qquad (i = 1, \ldots N-1)$$

with same i.c. and b.cs as in (a).

Thus

$$\bar{c}_{i,j+1} - \frac{i^2}{2}\left\{\bar{c}_{i+1,j+1} - 2\bar{c}_{i,j+1} + \bar{c}_{i-1,j+1}\right\}\Delta\tau - \frac{ki}{4}\left\{\bar{c}_{i+1,j+1} - \bar{c}_{i-1,j+1}\right\}\Delta\tau$$

$$+ \frac{k}{2}\bar{c}_{i,j+1}\Delta\tau = \frac{i^2}{2}\left\{\bar{c}_{i+1,j} - 2\bar{c}_{i,j} + \bar{c}_{i-1,j}\right\}\Delta\tau$$

$$+ \frac{ki}{4}\left\{\bar{c}_{i+1,j} - \bar{c}_{i-1,j}\right\}\Delta\tau - \frac{k}{2}\bar{c}_{i,j}\Delta\tau$$

$$+ \bar{c}_{i,j}$$

We set up a matrix eqtn for the unknowns

$$\underline{\bar{c}} = \begin{bmatrix} \bar{c}_{1,j+1} \\ \vdots \\ \bar{c}_{N-1,j+1} \end{bmatrix}$$

at each time step.

Calculation of Δ:

Once the numerical solution is found ~~using one of these methods~~ we can easily calculate

$$\Delta = \frac{\partial C}{\partial S}$$

using, at grid pt i,j,

$$\Delta_{i,j} = \left(\frac{\partial C}{\partial S}\right)_{i,j} = \left(\frac{\partial \bar{C}}{\partial \bar{S}}\right)_{i,j} \approx \frac{\bar{c}_{i+1,j} - \bar{c}_{i-1,j}}{2\Delta \bar{S}}$$