MSc. course

Numerical Methods for Partial Differential Equations

Syllabus

* Introduction and preliminaries, operators,

* Finite difference formula,

- finite difference formula for linear partial differential equations (parabolic, elliptic, hyperbolic),high order formula for finite difference methods,
- finite difference formula for nonlinear partial differential equations, analysis of finite difference formula (consistency, convergence , stability).

* Finite elements method (FEM),

- properties,
- Lagrange polynomial
- relation between interpolation polynomial and shape function ,
- varitional methods, Rayleigh-Ritz method(RRM) ,relation between RRM and FEM.
- weighted residual method, collocation method ,least square method and Galerkin method.

* Differential quadrature method (DQM),

- analysis of DQM constructor ,
- applications.

* Other methods(Adomian decomposition method, variaitional iteration method, homotopy method,…)with applications.

References:

1- simth G.D."Numerical Solution of Partial Differential Equations,Finite Difference Methods" London.1978

2- Noye B.J. " Numerical Solution of Partial Differential Equations " North-Australia,Holand,1981

3- 3-RaoS.S. " The finite Element in Engineering" U.S.A 1982

4- Michel A.R." Computational Methods in Partial Differential Equations " London 1976

5- Grossmann C. and Roos H-G." Numerical Treatment of Partial Differential Equations" Springer-Verlag Berlin Heidelberg 2007

6- Strikwerda J.C." finite difference schemes and Partial Differential Equations" U.S.A 2004

Abdul-Sattar Jaber Ali Al-Saif (Ph.D)
Professor in Applied Mathematics
(differential equations & computational mathematics

Numerical solution of partial differential equations

Numerical analysis is a branch of applied mathematics; the subject can be standard with a good skill in basic concepts of mathematics. This subject has many applications and wide uses in the area of applied sciences such as, physics, engineering, Biological, …ect. So, when any body wants to study this subject, should be to get answers, which do not agree with experiment or observation data. This is because there always has to be careful choice of the mathematical model that is to be used to describe a particular phenomenon. The problems of the real subject of P.D.Es are possible great complexity involving many physical effects (or other sciences) and a considerable set of non-linear equations. These problems can not be solved either by advanced techniques or by putting then on the computer. The techniques do not exist and the machines are neither powerful enough nor sophisticated enough (to reject spurious solutions). the problem only be omitting, after much careful thought, perhaps and special case can be dealt with analytically , and this will show what sort of calculation the machines must be programmed for more general case. After determine the mathematical model for the problems, should be try to solve it. For this situation, we need good mathematical procedure to simplified or linearized problems, which are non-linear or involving complex geometries, or both. Here the numerical techniques such as finite difference, finite elements, differential quadrature ,….ect;are play important role to computational of problems are described by a set of linear and/or non-linear equations.

Important examples of the three type equations are the

$$\begin{aligned} u_{xx}+u_{yy}&=0 \quad \textit{Laplace equation}\\ u_t&=u_{xx} \quad \textit{Heat equation}\\ u_{tt}&=u_{xx} \quad \textit{Wave equation}\end{aligned}$$

Before derivation of finite difference formulas, which are using to approximation partial differential equations, we wanted to introduce classification of second order linear partial differential equations

$$Au_{xx}+Bu_{xy}+Cu_{yy}+Du_x+Eu_y+Fu+G=0 \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Where $A, B, C, D, .E, F, \text{ and } G$ are functions $x \text{ and } y$

Now, If

$$B^2-4AC\begin{cases} <0 \Rightarrow & \textit{Elliptic} \quad P..D..E. & & u_{xx}+u_{yy}=0 \quad \textit{Laplace equation}\\ =0 \Rightarrow & \textit{Parabolic} \quad P..D..E. & \textit{examples} & u_t=u_{xx} \quad \textit{Heat equation}\\ >0 \Rightarrow & \textit{Hyperbolic} \quad P..D..E. & & u_{tt}=u_{xx} \quad \textit{Wave equation}\end{cases}$$

And, also we need to give information about u on the boundary (C) of ***R*** (Fig(1)

- given u on C [Dirichlet problem]
- given $\frac{\partial u}{\partial n}$ on C, where n the norm [Neuman problem]
- $\alpha u+\beta\frac{\partial u}{\partial n}$, where α,β are given, [Mixed problem]

Example 1:

One end of a bar $2\,ft$ long .whose sides are insulated, is kept at the temperature 0^0C,while the other end is kept at 10^0C. If the initial temperature distribution is linear along the bar, write down the boundary value problem that governing the temperature in the bar.

The bar has the length $2\,ft$ (i.e. $\Omega=[0,1]$), then by conservation law of energy ,we have

$$s\rho Au_t(x_1,t)\Delta x=KA\left[u_x(x+\Delta x)-u_x(x,t)\right]$$

Where the constant s is the specific heat of the material, ρ is the mass per unit volume, and x_1 is between $x\ and\ x+\Delta x$, K is the thermal conductivity (positive constant), and A is the area of a cross section . Dividing through in this equation by $s\rho A\Delta x$ and then letting Δx approach to zero, we obtain the equation

$$u_t(x,t)\Delta x=\alpha\,u_{xx}(x,t)\quad 0<x<2,\ t>0$$

Where $\alpha=K/s\rho$ is the thermal diffusivity of material.

One end kept at the temperature 0^0C and the other end is kept at 10^0C

$$\Rightarrow\ u(0,t)=0\quad and\quad u(2,t)=10\quad ,t\geq 0$$

The initial temperature distribution is linear along the bar

$$\Rightarrow u(x,0)=5x\quad ,0\leq x\leq 2$$

Therefore, the mathematical model for this problem is

$$u_t=\alpha\,u_{xx}$$
$$u(0,t)=0\quad and\quad u(2,t)=10\quad ,t\geq 0$$
$$u(x,0)=5x\quad ,0\leq x\leq 2$$

Example 2:

A string is stretched between the fixed points $(0,0)\,and\,(1,0)$ and released at rest from the position $u=A\sin(\pi x)$,where A is a constant. Write down the mathematical model that governing the transverse displacement of a string.

The mathematical model for this problem is

$u_{tt} = \alpha^2 u_{xx}$ (**For** derivation you can see Churchill R. 'Fourier sires and Boundary Value Problems' page 5)

$u(0,t) = 0 \, and \, u(1,t) = 0, t \geq 0$ (A string infixed at points $(0,0)$ and $(1,0)$ $\Rightarrow$ there is **no** displacement)

$u(x,0) = A\sin(\pi x) \quad , 0 \leq x \leq 1$ (Initial displacement, at $t = 0$)

Depending on the above information**,** the following is a rough summary of well-posed problems for second-order partial differential equations:

elliptic equation plus boundary conditions
parabolic equation plus boundary conditions with respect to space
plus initial condition with respect to time
hyperbolic equation plus boundary conditions with respect to space

Finite difference methods

One of the greatest needs in applied mathematics is a general and reasonably short method of solving partial differential equations by numerical methods. Several methods have been proposed for meeting this need, but none can be called entirely satisfactory. They are all long and laborious. Certain types of boundary value problems can be solved by replacing the differential equation by the corresponding difference equation and then solving the latter by a process of iteration. This method of solving partial differential equations was devise and first used by Richardson (1910). It was later improved by Liebmann(1918) and further improved more recently by Shortley &Weller (1938).the process is slow, but gives good results on boundary value problems which satisfy Laplace , Poisson, and several other partial differential equations. A strong point in its favor is that an automatic sequence-controlled calculating machine can do the computation.

A somewhat similar method is the relaxation method devised by Southwell. This method is shorter and more flexible than the iteration method, but is not adapted to automatic machine computation. In both of these methods the approximate solution of a partial differential equations with given boundary values, is found by finding the solution of the corresponding partial differential equation.

Operators: it is a mathematical operation on an operated function.

- Shifts (translation) operator $Ef(x) = f(x+h)$
- Difference operator $\Delta f(x) = f(x+h) - f(x)$
- Inverse difference operator $\nabla f(x) = f(x) - f(x-h)$
- Intermediate operator $\delta f(x) = f(x+h/2) - f(x-h/2)$

Properties of operators:

- Linearity of operator $E(f+g)=E(f)+E(g)$
- Product of operator $E\cdot E\cdot E\ f=E^3 f$
- Sum and difference operator $(E\mp D)f(x)=Ef(x)\mp Df(x)$
- Equality of operator $E_1=E_2\Leftrightarrow E_1 f(x)=E_2 f(x)$
- Identity(unit) operator $I\ f(x)=f(x)$
- Null(zero) operator $0\ f(x)=0$

Exercise1: Prove that (a) $\Delta=E-1$
(b) $E^n D=DE^n$
(c) $E=e^{hD}$

Inverse operator: it is a mathematical operator that inverse the original operation.

For example; Shifts operator is $Ef(x)=f(x+h)$, the inverse of it is $E^{-1}f(x)=f(x-h)\quad ((EE^{-1}=1))$

Difference operator $\Delta=E-1$, the inverse of it is $\nabla=1-E^{-1}$

Show that
(1) $\delta_x=E^{1/2}-E^{-1/2}$
(2) $\delta_x^2=E\ +E^{-1}-2$
(3) $E^{-1/2}-\delta/2-\mu=0\quad where\quad \mu=(E^{1/2}+E^{-1/2})/2$
(4) $(E^{-1/2}-\delta_x/2)^2-\delta_x^2/4=1$

Finite difference formulas:

Now, the area of integration $\boldsymbol{R}$ is covering by rectangular meshes $P_{ij}=P(i\Delta x,n\Delta t)$, are called mesh points. For a function u of a single variable, the familiar expression $\dfrac{u(x+\Delta x)-u(x)}{\Delta x}$, is called difference quotient, whose limiting value is the derivative of $u(x)$ with respect to x i.e.

$$u'(x)=\lim_{\Delta x\to 0}\frac{u(x+\Delta x)-u(x)}{\Delta x}\Rightarrow u'(x)\approx\frac{u(x+\Delta x)-u(x)}{\Delta x}\quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots..(2)$$

That is mean; a difference quotient approximates the derivative, the approximation becoming closer as Δx become small.

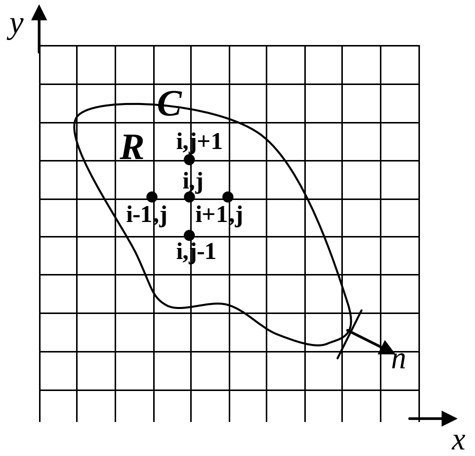


Figure (1)

Suppose we like to solve the parabolic P.D.E.

$$\frac{\partial u}{\partial t} = L(x,t,D,D^2)u$$

$$for\ example \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \qquad L = D^2$$

Using Tayler's series expand $u(x.t+k)$ about (x,t)

$$u(x,t+k) = u(x_i,t_n) + \frac{k}{1!}\frac{\partial u(x_i,t_n)}{\partial t} + \frac{k^2}{2!}\frac{\partial^2 u(x_i,t_n)}{\partial t^2} + \cdots \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3)$$

$$= (1 + \frac{k}{1!}\frac{\partial}{\partial t} + \frac{k^2}{2!}\frac{\partial^2}{\partial t^2} + \cdots\cdots\cdots\cdots\cdots\cdots)u(x_i,t_n)$$

Using the finite difference code; $u_i^n = u(x_i,t_n)$

$$u_i^{n+1} = Exp(k\frac{\partial}{\partial t})u_i^n \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

This is finite difference representation of the parabolic P.D.E.

As special case; let $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \qquad L = D^2$

$$u_i^{n+1} = Exp(kD^2)u_i^n \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

$$hD = \ell n(E) = \ell n(1+\Delta) = 2\ell n\left[\frac{\delta}{2} + (1+(\frac{\delta}{2})^2)^{1/2}\right] \qquad \text{Prove that!}$$

$$D = \frac{1}{h}\sinh^{-1}(\frac{\delta}{2}) \qquad \text{Prove that!}$$

$$D = \frac{1}{h}\left[\delta_x - \frac{1}{2^2 3!}\delta_x^3 + \frac{1}{2^4 5!}\delta_x^5 + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\right] \qquad \ldots\ldots\ldots\ldots\ldots(6)$$

$$D^2 = \frac{1}{h^2}\left[\delta_x - \frac{1}{2^2 3!}\delta_x^3 + \frac{1}{2^4 5!}\delta_x^5 + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\right]^2$$

$$= \frac{1}{h^2}\left[\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6 + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\right] \qquad \ldots\ldots\ldots\ldots\ldots(7)$$

$$u_i^{n+1} = Exp(\frac{k}{h^2}\left[\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6 + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\right])u_i^n$$

$$= Exp(r\delta_x^2)u_i^n \cdot Exp(-\frac{r}{12}\delta_x^4)u_i^n \cdot Exp(\frac{r}{90}\delta_x^6)u_i^n \cdots\cdots\cdots\cdots\cdots\cdots \qquad \ldots\ldots\ldots\ldots\ldots(8)$$

$$= ((1 + r\delta_x^2 + \frac{r}{2}(r-\frac{1}{6})\delta_x^4 + \frac{r}{6}(r^2 - \frac{r}{2} + \frac{1}{15})\delta_x^6 + \cdots\cdots\cdots\cdots)u_i^n$$

$$where \quad r = \frac{k}{h^2},\ in\ which\ k = \Delta t\ and\ h = \Delta x$$

This is the general finite difference representation of the parabolic P.D.E. $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$.

Generally, there are two standard methods:-

** **An explicit methods**

** **An implicit methods**

Definition: An explicit method is one in which one unknown values in the $(n+1)^{th}$ level are specifying in terms of known values in the $(n)^{th}$ level.

Definition: An implicit method is one in which two or more unknown values in the $(n+1)^{th}$ level are specifying in terms of known values in the $(n)^{th}$ level.

Now, an explicit method can be reducing from the above general finite difference representation of the parabolic P.D.E., and may be writing as,

$$u_i^{n+1} = (1+r\delta_x^2)u_i^n = u_i^n + r\delta_x^2 u_i^n = u_i^n + r(u_{i+1}^n - 2u_i^n + u_{i-1}^n)$$

$$\underset{\underset{\text{Unknown}}{\uparrow}}{u_i^{n+1}} = (1-2r)\underbrace{u_i^n + r(u_{i+1}^n + u_{i-1}^n)}_{\text{known}} \qquad \ldots\ldots\ldots\ldots\ldots(10)$$

This expression is calling an explicit method involving four-point formula.

Other an explicit form can be obtaining as,

$$u_i^{n+1} = \frac{1}{2}(2-5r+6r^2)u_i^n + \frac{3}{2}(2-2r)(u_{i+1}^n + u_{i-1}^n) - \frac{1}{12}(1-6r)(u_{i+2}^n + u_{i-2}^n) \ldots\ldots(11)$$

This expression is calling an explicit method involving Six-point formula.

An implicit formula of the parabolic P.D.E. can be deriving as follows;

The finite difference formula $u_i^{n+1} = Exp(kD^2)u_i^n$

$$D^2 = \frac{1}{h^2}\left[\delta_x^2 - \frac{1}{12}\delta_x^4 + \frac{1}{90}\delta_x^6 + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\right] \quad , \quad D^2 \approx \frac{\delta_x^2}{h^2}$$

$$(1+2r)u_i^{n+1} - r(u_{i+1}^{n+1} + u_{i-1}^{n+1}) = u_i^n \qquad \ldots\ldots\ldots\ldots\ldots(12)$$

This equation is call implicit formula of the parabolic P.D.E. $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$. It is suggest by O'Brien et.el , which approximate $\frac{\partial^2 u}{\partial x^2}$ in the $(n+1)^{th}$ time level instead of the $(n)^{th}$ level.

Crank and Nicolson(1947) proposed a method is valid for all values of r. They replaced $\frac{\partial^2 u}{\partial x^2}$ by means of its finite difference representations on the $(n+1)^{th}$ time level and $(n)^{th}$ time level.

$$u_i^{n+1} = Exp(kD^2)u_i^n = Exp(\frac{kD^2}{2})Exp(\frac{kD^2}{2})u_i^n \qquad \ldots\ldots\ldots\ldots\ldots(13)$$

$$Exp(-\frac{kD^2}{2})u_i^{n+1} = Exp(\frac{kD^2}{2})u_i^n$$

$$Exp(-\frac{r}{2}\delta_x^2)u_i^{n+1} = Exp(\frac{r}{2}\delta_x^2)u_i^n$$

$$\left[1-\frac{r}{2}\delta_x^2+\frac{r^2}{(4)\cdot(2!)}\delta_x^4+\cdots\cdots\cdots\right]u_i^{n+1} = \left[1+\frac{r}{2}\delta_x^2+\frac{r^2}{(4)\cdot(2!)}\delta_x^4+\cdots\cdots\cdots\right]u_i^n$$

$$\left(1-\frac{r}{2}\delta_x^2\right)u_i^{n+1} \cong \left(1-\frac{r}{2}\delta_x^2\right)u_i^n$$

$$u_i^{n+1}-\frac{r}{2}(u_{i+1}^{n+1}-2u_i^{n+1}+u_{i-1}^{n+1}) = u_i^n+\frac{r}{2}(u_{i+1}^n-2u_i^n+u_{i-1}^n) \quad \ldots\ldots(14)$$

This equation is call Crank-Nicolson implicit formula of the parabolic P.D.E. $\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$.

<u>**Exercise2**</u>: write down the following parabolic P.D.E. with variable coefficients

(1) $\frac{\partial u}{\partial t}=a(x)\frac{\partial^2 u}{\partial x^2}$ (Hint: put $L=a(x)D^2$)

(2) $\frac{\partial u}{\partial t}=\frac{\partial}{\partial x}(a(x)\frac{\partial u}{\partial x})$

<u>Note:</u> we can approximate the derivatives (first-order &second-order) by finite difference depend on the following definition

$$\frac{\partial u}{\partial x}=\lim_{\Delta x\to 0}\frac{u(x+\Delta x,y)-u(x,y)}{\Delta x} \Rightarrow \frac{\partial u}{\partial x}\approx\frac{u(x+\Delta x,y)-u(x,y)}{\Delta x}$$

Thus, this derivative at point (i, j) can expression as follows,

$$\left.\frac{\partial u}{\partial x}\right|_{i,j}\approx\frac{u_{i+1,j}-u_{i,j}}{\Delta x} \quad \textit{Forward difference} \quad \ldots\ldots\ldots\ldots\ldots(15)$$

$$\left.\frac{\partial u}{\partial x}\right|_{i,j}\approx\frac{u_{i,j}-u_{i-1,j}}{\Delta x} \quad \textit{Backward difference} \quad \ldots\ldots\ldots\ldots\ldots(16)$$

$$\left.\frac{\partial u}{\partial x}\right|_{i,j}\approx\frac{u_{i+1,j}-u_{i-1,j}}{2\Delta x} \quad \textit{Central difference} \quad \ldots\ldots\ldots\ldots\ldots(17)$$

From these relations the second derivative can written as follows

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{i,j}=\frac{\partial}{\partial x}(\left.\frac{\partial u}{\partial x}\right|_{i,j})=(\left.\frac{\partial u}{\partial x}\right|_{i+1,j}-\left.\frac{\partial u}{\partial x}\right|_{i-1,j})/\Delta x$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_{i,j}\approx\frac{u_{i+1,j}-2u_{i,j}+u_{i-1,j}}{(\Delta x)^2} \quad \textit{Central difference} \quad \ldots\ldots\ldots\ldots\ldots(18)$$

We can derivation these formulas by using Taylor's theorem. Apply Taylor's expand to the function u at $x+\Delta x$ and $x-\Delta x$, we obtain

$$u(x+\Delta x,k)=u(x_i,t_n)+\frac{\Delta x}{1!}\frac{\partial u(x_i,t_n)}{\partial x}+\frac{(\Delta x)^2}{2!}\frac{\partial^2 u(x_i,t_n)}{\partial x^2}+\cdots \qquad \text{.........} (19)$$

$$u(x-\Delta x,k)=u(x_i,t_n)-\frac{\Delta x}{1!}\frac{\partial u(x_i,t_n)}{\partial x}+\frac{(\Delta x)^2}{2!}\frac{\partial^2 u(x_i,t_n)}{\partial x^2}+\cdots \qquad \text{......} (20)$$

Use these relations to prove above formulas.

Discuss the truncation errors

From equation(19) , we get

$$\begin{aligned}\frac{\partial u}{\partial x}&=\frac{u_{i+1}^n-u_i^n}{\Delta x}-\frac{\Delta x}{2!}\frac{\partial^2 u}{\partial x^2}-\frac{(\Delta x)^2}{3!}\frac{\partial^3 u}{\partial x^3}+\cdots\cdots\cdots \\ &=\frac{u_{i+1}^n-u_i^n}{\Delta x}+O(\Delta x)\end{aligned} \qquad \text{......}(21)$$

This is the Forward finite difference scheme .it is of the first-order of Δx. Similarly, From equation(20) , we have

$$\frac{\partial u}{\partial x}=\frac{u_i^n-u_{i-1}^n}{\Delta x}+O(\Delta x) \qquad \text{...............}(22)$$

This is the Backward finite difference scheme .it is of the first-order of Δx

Subtracting (19) and (20), we obtain

$$\frac{\partial u}{\partial x}=\frac{u_{i+1}^n-u_{i-1}^n}{2\Delta x}+O((\Delta x)^2) \qquad \text{...........}(23)$$

This is the Central finite difference scheme .it is of the second-order of $(\Delta x)^2$

Adding (19) and (20), we get

$$\frac{\partial^2 u}{\partial x^2}=\frac{u_{i+1}^n-2u_i^2+u_{i-1}^n}{(\Delta x)^2}+O((\Delta x)^2) \qquad \text{.........}(24)$$

This is the Central finite difference scheme .it is of the second-order of $(\Delta x)^2$

Derivative boundary conditions

For solving parabolic P.D.E $\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$ subject to the initial and boundary conditions

$u(x,0)=f(x) \quad for \ 0\le x\le 1$ (where $f(x)$ is known function)

$$\frac{\partial u(0,t)}{\partial x} = u \quad for \ \ t > 0$$

$$\frac{\partial u(1,t)}{\partial x} = -u \quad for \ \ t > 0$$

by using explicit method, we obtain

$$u_i^{n+1} = (1-2r)u_i^n + r(u_{i+1}^n + u_{i-1}^n)$$

For $i = 0 \Rightarrow u_0^{n+1} = (1-2r)u_0^n + r(u_1^n + u_{-1}^n)$

Here appear problem in the computation of the recurrence relation exactly in the term u_{-1}^n. One of the treatments for this problem, one can used central difference scheme (equation (23)) for the boundary condition $\frac{\partial u(0,t)}{\partial x} = u$ gives

$$u_{-1}^n = u_1^n - 2\Delta x \, u_0^n$$

Similarly, For $i = M \Rightarrow u_M^{n+1} = (1-2r)u_M^n + r(u_{M+1}^n + u_{M-1}^n)$, such that problem is appear in the term u_{m+1}^n,so by using the same previous way ,we have

$$u_{M+1}^n = u_{M-1}^n - 2\Delta x \, u_M^n$$

Exercise3:

A bar, with ends at $x = 0$ and $x = a$,with insulated ends, has an initial

temperature distribution $u(x,0) = f(x)$

1) Write down the boundary value problem that corresponding to the physical problem.
2) Approximate the P.D.E. resulting in part(1) by:

a- explicit method **b**-implicit method **c**-Crank-Nicolson method.

Exercise4:

Derive the following expressions:

1) $\left.\frac{\partial^2 u}{\partial x \partial y}\right|_{i,j} \approx \frac{u_{i+1,j+1} + u_{i-1,j-1} - u_{i-1,j+1} + u_{i+1,j-1}}{4(\Delta x \Delta y)}$,which is second order mixed

central difference with respect to $x \, and \, y$.

2) $\left.\frac{\partial u}{\partial y}\right|_{i,j} \approx \frac{1}{6\Delta y}(-11u_{i,j} + 18u_{i,j+1} - 9u_{i,j+2} + 2u_{i,j+3})$, which is third order difference with respect to y.

Two-dimension P.D.E.

(1) the 2D heat equation in the (x, y, t) plane may be written as

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \qquad \text{...............(25)}$$

Here we can define the differential operator as $L = D_x^2 + D_y^2$, thus, equation

(25) can be written as

$$\frac{\partial u}{\partial t} = Lu \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(26)$$

Using the finite difference code; $u_{i,j}^n = u(x_i, y_j, t_n)$, approximation of equation (26) is

$$\begin{aligned} u_{i,j}^{n+1} &= Exp(KL)u_{i,j}^n \\ &= \left(Exp(KD_x^2)\cdot Exp(KD_y^2)\right)u_{i,j}^n \\ &= (Exp(r(\delta_x^2 - \frac{\delta_x^4}{12}\cdots\cdots))\cdot Exp(r(\delta_y^2 - \frac{\delta_y^4}{12}\cdots\cdots)))u_{i,j}^n \quad \ldots\ldots..(27) \\ &= (Exp(r\delta_x^2)\cdot Exp(r\delta_y^2))u_{i,j}^n \\ &= (1 + r\delta_x^2)(1 + r\delta_y^2)u_{i,j}^n \end{aligned}$$

This is explicit scheme for 2D parabolic P.D.E.

(2) Using an explicit method to approximation the2D parabolic P.D.E.

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}(a(x,y)\frac{\partial u}{\partial x}) + \frac{\partial}{\partial y}(b(x,y)\frac{\partial u}{\partial y}) \qquad \ldots\ldots\ldots\ldots(28)$$

.

(3)Use an implicit scheme to approximation the 2D parabolic P.D.E.

Explicit method:

Advantage Relatively simple to set up and program

Disadvantage in terms of above example, for a given $\Delta x, \Delta t$ must be less than some limit imposed by stability constraints. In some cases, Δt must be very small to maintain stability; this can result in long computer running times to make calculations over a given interval of t.

Implicit method:

Advantage stability can be maintained overt much large values of Δt, hence using considerable fewer time steps to make calculations over a given interval of t. This result in less computer time.

Disadvantage more complicated to set up and program

Disadvantage Since massive matrix manipulations are usually required at each time step, the computer time per time step is much larger than in the explicit approach.

Disadvantage Since large Δt can be taken, the truncation error is large, and the use of implicit methods to follow the exact transient (time variation of the independent variable) may not be as accurate as an explicit approach. However, for a time- dependent solution in which the steady state is the desired result, this relative time wise inaccuracy is not important.

Other finite difference approximation methods:

A weighted average approximation method:

we can introduce a weighted factor θ to some finite difference scheme, such as Crank-Nicolson scheme to produce a more general finite difference approximation to $\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}$.

$$\frac{u^{n+1}_{i,j}-u^{n}_{i,j}}{\Delta t}=\frac{1}{(\Delta x)^2}\left[\theta(u^{n+1}_{i+1,j}-2u^{n+1}_{i,j}+u^{n+1}_{i-1,j})+(1-\theta)(u^{n}_{i+1,j}-2u^{n}_{i,j}+u^{n}_{i-1,j})\right] \quad ..(29)$$

If $\theta=0$ then explicit scheme obtain (equ.(10))

$\theta=1$ then fully implicit scheme obtain(equ.(12))

$\theta=\frac{1}{2}$ then Crank-Nicolson scheme obtain(equ.(14))

In practice $0\le\theta\le1$.

Alternating direction implicit method (ADIM):

we consider the parabolic P.D.E.

$$\frac{\partial u}{\partial t}=\alpha(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}) \quad \ldots\ldots\ldots\ldots..(30)$$

Define over a rectangle region $0<x<a \; and \; 0<y<b$,where α, a, b are constants, $u(x,y,t)$ is known on the boundary of the rectangle.define the coordinates (x,y,t) of mesh points as; $x=i\Delta x\,, y=j\Delta y\,, t=n\Delta t$,where $i\,, j\,, n \; are \; +ve$. The implicit Crank-Nicolson finite difference of above parabolic P.D.E. is

$$\frac{u^{n+q}_{i,j}-u^{n}_{i,j}}{\Delta t}=\frac{\alpha}{2}\left[(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2})\Big|^{n}_{i,j}+(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2})\Big|^{n+1}_{i,j}\right]\ldots\ldots\ldots(31)$$

This scheme produces $(M-1)(N-1)$ simultaneous equations for each time step.

However, the best method to solve these simultaneous equations is replaced $\frac{\partial^2 u}{\partial x^2}$ by implicit difference approximation and $\frac{\partial^2 u}{\partial y^2}$ by explicit difference approximation in $(n+1/2)$ time level, as

$$\frac{u_{i,j}^{n+1/2}-u_{i,j}^{n}}{\Delta t/2}=\frac{\alpha}{(\Delta x)^2}(\delta_x^2)u_{i,j}^{n+1/2}+\frac{\alpha}{(\Delta y)^2}(\delta_y^2)u_{i,j}^{n} \quad \ldots\ldots\ldots\ldots.(32)$$

These produce $(N-1)$ equations. For next time step, replace $\frac{\partial^2 u}{\partial x^2}$ by an explicit difference approximation and $\frac{\partial^2 u}{\partial y^2}$ by implicit difference approximation. These produce $(M-1)$ equations in $(M-1)$ unknowns. To obtain the solution at time step $(n+1)$

$$\frac{u_{i,j}^{n+1}-u_{i,j}^{n+1/2}}{\Delta t/2}=\frac{\alpha}{(\Delta x)^2}(\delta_x^2)u_{i,j}^{n+1/2}+\frac{\alpha}{(\Delta y)^2}(\delta_y^2)u_{i,j}^{n+1} \quad \ldots\ldots\ldots\ldots(33)$$

Elliptic P.D.E.

Let the problem of determining the steady –state heat distribution in a thin square metal plate with dimensions 0.5m by 0.5m .two adjacent boundaries are held at $0^o C$,and the heat on the other boundaries increases linearly from $0^o C$ at one corner to $100^o C$ where the sides meet. If we place the sides with the zero boundary conditions along the $x-$ and $y-$axes, the problem is expressed as

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \quad \ldots\ldots\ldots\ldots.(34)$$

for $(x,y)\in R=\{(x,y)|\,0<x<0.5, 0<y<0.5\}$, with the boundary conditions

$u(0,y)=0, u(x,0)=0, u(x,0.5)=200x, u(0.5,y)=200y$.

If $M=N=4$, the problem has 5×5 grid and the finite difference equation

$$4u_{i,j}-\left(u_{i+1,j}+u_{i-1,j}+u_{i,j+1}+u_{i,j-1}\right)=0 \quad \ldots\ldots\ldots\ldots\ldots(35)$$

For $i=1,2,3 \;\; and \;\; j=1,2,3$.

Expressing this in terms of relabeled interior grid points $p_1, p_2,\cdots\cdots p_9$ corresponding to $u_1, u_2,\cdots\cdots u_9$ implies that the equations at the points p_i are

$$\begin{aligned}
&for\;\; j=3,\\
&\quad i=1:\;\; 4u_{1,3}-u_{2,3}-u_{1,2}=u_{0,3}+u_{1,4}\\
&\quad i=2:\;\; 4u_{2,3}-u_{3,3}-u_{1,3}-u_{2,2}=u_{2,4}\\
&\quad i=3:\;\; 4u_{3,3}-u_{1,2}-u_{0,3}=u_{4,3}+u_{3,4}
\end{aligned} \quad \ldots\ldots\ldots\ldots(36a)$$

$$\begin{aligned}
&for\;\; j=2,\\
&\quad i=1:\;\; 4u_{1,2}-u_{2,2}-u_{1,3}-u_{1,1}=u_{0,2}\\
&\quad i=2:\;\; 4u_{2,2}-u_{3,2}-u_{1,2}-u_{2,3}-u_{2,1}=0\\
&\quad i=3:\;\; 4u_{32}-u_{2,3}-u_{3,3}-u_{3,1}=u_{4,2}
\end{aligned} \quad \ldots\ldots\ldots\ldots(36b)$$

$$\begin{aligned}
&for\ j=1,\\
&\quad i=1:\quad 4u_{1,1}-u_{2,1}-u_{1,2}=u_{0,1}+u_{1,0}\\
&\quad i=2:\quad 4u_{2,1}-u_{3,1}-u_{1,1}-u_{2,2}=u_{2,0}\\
&\quad i=3:\quad 4u_{3,1}-u_{2,1}-u_{3,2}=u_{3,0}+u_{4,1}
\end{aligned}\qquad \ldots\ldots\ldots\ldots(36d)$$

Where the right sides of the equations are obtained from the boundary conditions. In fact, the boundary conditions imply that

$$u_{1,0}=u_{2,0}=u_{3,0}=u_{0,1}=u_{0,2}=u_{0,3}=0$$

$$u_{1,4}=u_{4,1}=25 \quad u_{2,4}=u_{4,2}=50 \quad u_{3,4}=u_{4,3}=75$$

The linear system associated with this problem has the matrix form

$$AU=B \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(37)$$

where

$$A=\begin{bmatrix}
4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0\\
-1 & 4 & -1 & 0 & -1 & & & & \\
0 & -1 & 4 & -1 & 0 & -1 & & & \\
-1 & 0 & 0 & 4 & -1 & 0 & -1 & & \\
0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & \\
0 & 0 & -1 & 0 & -1 & 4 & 0 & 0 & -1\\
0 & 0 & 0 & -1 & 0 & 0 & 4 & -1 & 0\\
0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & -1\\
0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4
\end{bmatrix},\quad U=\begin{bmatrix} u_{1,3}\\ u_{2,3}\\ u_{3,3}\\ u_{1,2}\\ u_{2,2}\\ u_{3,2}\\ u_{1,1}\\ u_{2,1}\\ u_{3,1}\end{bmatrix},\quad B=\begin{bmatrix}25\\50\\150\\0\\0\\50\\0\\0\\25\end{bmatrix}$$

Exercise 5: Write the linear system of algebraic equations associated with the problem of **exercise 3** in the matrix form.

Note: the linear system of algebraic equations that is resulting from approximation P.D.Es by finite difference method (finite difference equations), needed good procedure to solve it. There are two procedures to achieve this aim, the first is called direct methods(such as, Gauss elimination , LU factorization,……) and the second is called iterative method (such as;Jacobi iterative, Gauss-seidel iterative, successive over relaxation(SOR) iterative ,……..). For example, system (37) is solving by using Gauss-siedel method and the results are;

i,j	(1,3)	(2,3)	(3,3)	(1,2)	(2,2)	(3,2)	(1,1)	(2,1)	(3,1)
$u_{i,j}$	18.75	37.50	56.25	12.50	25.00	37.50	6.25	12.50	18.75
$u(x,y)=400xy$	18.75	37.50	56.25	12.50	25.00	37.50	6.25	12.50	18.75

This mean the truncation error is zero at each step.

Accuracy of the finite difference equation of the numerical scheme:

Accuracy of finite difference schemes can be determined by many ways such as , theoretically through the order of error of the finite difference equation and experimentally through the measurements of errors(L^2 norm, ….).

There is Richardson's approach to limit the order of accuracy of the finite difference equation. In this method, we need to know the truncation error of the difference equation.

Let u represents the solution of differential equation

And U represents the solution of finite difference equation

The discretization error is

$$u - U = Ak + Bh^2 + Ck^2 + Dh^4 + \cdots\cdots \qquad \ldots\ldots\ldots\ldots(38)$$

If we use k_1 , h_1 as mesh size to produce U_1, and k_2 , h_2 as mesh size to produce U_2 ,then

$$u - U_1 = Ak_1 + Bh_1^2 + Ck_1^2 + Dh_1^4 + \cdots\cdots \qquad \ldots\ldots\ldots(39)$$

$$u - U_2 = Ak_2 + Bh_2^2 + Ck_2^2 + Dh_2^4 + \cdots\cdots \qquad \ldots\ldots(40)$$

Subtracting these two equations, we obtain

$$u = \frac{1}{h_2^2 - h_1^2}(h_2^2 U_1 - h_1^2 U_2) + A\frac{k_1 h_2^2 - k_2 h_1^2}{k_1}$$

If we neglect the term involving A , then

$$u = \frac{1}{h_2^2 - h_1^2}(h_2^2 U_1 - h_1^2 U_2) , \qquad \text{if } h_1 = h_2 \ and \ k_1 = k_2 \text{, then}$$

$$u = \frac{1}{3}(4U_1 - U_2) \text{ the error is } O(h^4)$$

Definition: $u - U$ is the discitization error which be reduced by decreasing $h \ and \ k$.

Definition: Let $F_{i,j}(U) = 0$ represent the finite difference equation at $(i, j)^{th}$ mesh point, then $F_{i,j}(U)$ is called **local truncation** error.

Definition: If N is a numerical solution of finite difference equation that is produce from each calculation is carried up to a finite number of decimal places. Thus, $U - N$ is the **global rounding** error.

Total error= discitization error+ global rounding error$= u_{i,j} - N_{i,j}$.

High accurate formula (high order) for elliptic P.D.E. can be derived by using more terms from the operators series. These formulas are useful when the boundary in highly irregular.

For example: An elliptic P.D.E.(equation(34)) can be approximation by central difference operators as;

$$\frac{1}{(\Delta x)^2}(\delta_x^2)u_{i,j}+\frac{1}{(\Delta y)^2}(\delta_y^2)u_{i,j}=0$$

For high accuracy, using the definition given by equation (7)

$$\frac{1}{h^2}\left[\delta_x^2-\frac{1}{12}\delta_x^4+\cdots\cdots\cdots\right]u_{i.j}+\frac{1}{h^2}\left[\delta_y^2-\frac{1}{12}\delta_y^4+\cdots\cdots\cdots\right]u_{i.j}=0$$

Now, for the first two terms, we have

$$\left[\delta_x^2-\frac{1}{12}\delta_x^4\right]u_{i,j}+\left[\delta_y^2-\frac{1}{12}\delta_y^4\right]u_{i,j}=0 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(41)$$

The first term

$$\left[\delta_x^2-\frac{1}{12}\delta_x^4\right]u_{i,j}=\left(u_{i+1,j}-2u_{i,j}+u_{i-1,j}\right)-\frac{1}{12}\delta_x^2\left(u_{i+1,j}-2u_{i,j}+u_{i-1,j}\right)$$

.

.

.

$$=\frac{4}{3}u_{i+1,j}-\frac{5}{2}u_{i,j}+\frac{4}{3}u_{i-1,j}-\frac{1}{12}u_{i+2,j}-\frac{1}{12}u_{i-2,j}$$

Similarly for the second term

$$\left[\delta_y^2-\frac{1}{12}\delta_y^4\right]u_{i,j}=\frac{4}{3}u_{i,j+1}-\frac{5}{2}u_{i,j}+\frac{4}{3}u_{i,j-1}-\frac{1}{12}u_{i,j+2}-\frac{1}{12}u_{i,j-2}$$

Now, equation (41) becomes

$$u_{i,j}=\frac{1}{5}\left\{\frac{4}{3}\left(u_{i+1,j}+u_{i-1,j}+u_{i,j+1}+u_{i,j-1}\right)-\frac{1}{12}\left(u_{i+2,j}+u_{i-2,j}+u_{i,j+2}+u_{i,j-2}\right)\right\} \quad \ldots.(42)$$

This finite difference equation is called nine-point formula.

<u>Exercise 6</u>:

1) Determine the order of truncation error of equation(42) .
2) Show that the truncation error of the Laplace equation is $TE=\frac{h^2}{12}(\frac{\partial^4 u}{\partial x^4}+\frac{\partial^4 u}{\partial y^4})+\cdots\cdots\cdots$

Stability, Convergence and consistency:

After presented, how to approximate the derivatives that including in P.D.E. to generate the finite difference schemes for its numerical solution? Should be discussing the follow;

- Verity that these schemes are good approximation to the P.D.E. (consistent).
- Verify that the schemes are stable or no.
- Show that the numerical solution converges to the solution of P.D.E.

Let us to define

$$F_{h,k}(U)=b \qquad \ldots\ldots\ldots\ldots\ldots\ldots..(43)$$

Is a finite scheme and,

$$F\,u=b \qquad \ldots\ldots\ldots\ldots\ldots\ldots..(45)$$

is a partial differential equation. Now we need to light up some definition related to the property of finite difference schemes, as follows;

Definition: we say that a finite difference scheme (43) is consistent with P.D.E.(45) of order (k,h) , if for any smooth function

$$Fu-F_{k,h}U=O(k^r,h^s) \qquad \ldots\ldots\ldots\ldots\ldots(46)$$

To verify consistency expand u in Taylor series and make sure equation (46) holds.

Example: If $\dfrac{\partial u}{\partial t}=\dfrac{\partial^2 u}{\partial x^2}+cu$ approximate by explicit finite difference method, then

(1) show that the finite difference equation given as

$$u_i^{n+1}=(1+crh^2+r^2\delta_x^2)u_i^n$$

$$,where \quad r=\frac{k}{h^2},\ in\ which\ k=\Delta t\ and\ h=\Delta x$$

(2) Show that difference equation and P.D.E. are consistent with the truncation error

$$TE=\frac{k}{2}\frac{\partial^2 u}{\partial t^2}-\frac{h^2}{12}\frac{\partial^4 u}{\partial x^4}+\cdots\cdots\cdots\cdots\cdots\cdots..=O(k,h^2)$$

<u>**Solution**</u>: using forward finite difference approximation (15) for the first order time derivative, and central finite difference approximation(18)for the second order spatial derivative, finite difference equation so obtained is

$$\frac{u_{i+1}^{n+1}-u_i^n}{\Delta t}=\frac{u_{i+1}^n-2u_i^n+u_{i-1}^n}{(\Delta x)^2}+cu_i^n \qquad \ldots\ldots\ldots\ldots\ldots\ldots(43)$$

Rearrangement this equation, we have

$$u_i^{n+1} = u_i^n + \frac{\Delta t}{(\Delta x)^2}(u_{i+1}^n - 2u_i^n + u_{i-1}^n) + \Delta t c u_i^n \quad \ldots\ldots\ldots\ldots\ldots\ldots(44)$$

If we put $r = \frac{\Delta t}{(\Delta x)^2}$ and using the definition of the central difference operator, then the finit difference equation becomes

$$u_i^{n+1} = (1 + cr(\Delta x)^2 + r^2\delta_x^2)u_i^n$$

Expand each term in equation(44) ,we obtain

$$u_i^n + \frac{\Delta t}{1!}\frac{\partial u_i^n}{\partial t} + \frac{(\Delta t)^2}{2!}\frac{\partial^2 u_i^n}{\partial t^2} + \frac{(\Delta t)^3}{3!}\frac{\partial^3 u_i^n}{\partial t^3} + \cdots\cdots =$$

$$u_i^n + r(u_i^n + \frac{\Delta x}{1!}\frac{\partial u_i^n}{\partial x} + \frac{(\Delta x)^2}{2!}\frac{\partial^2 u_i^n}{\partial x^2} + \frac{(\Delta x)^3}{3!}\frac{\partial^3 u_i^n}{\partial x^3} + \cdots\cdots$$

$$-2u_i^n + u_i^n - \frac{\Delta x}{1!}\frac{\partial u_i^n}{\partial x} + \frac{(\Delta x)^2}{2!}\frac{\partial^2 u_i^n}{\partial x^2} - \frac{(\Delta x)^3}{3!}\frac{\partial^3 u_i^n}{\partial x^3} + \cdots\cdots) + cr(\Delta x)^2 u_i^n$$

Rearrangement this equation to obtain

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - cu + \frac{(\Delta t)}{2}\frac{\partial^2 u_i^n}{\partial t^2} - \frac{(\Delta x)^2}{12}\frac{\partial^4 u_i^n}{\partial x^4} = 0$$

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - cu + TE = 0 \quad ,where \quad TE = \frac{(\Delta t)}{2!}\frac{\partial^2 u_i^n}{\partial t^2} - \frac{(\Delta x)^2}{12}\frac{\partial^4 u_i^n}{\partial x^4} + \cdots\cdots\cdots\cdots$$

This equation of order $O(k, h^2)$.

Exercise7: Approximation P.D.E. in above example by implicit finite difference method ,then find its order of error

Definition: For a function $v = (\ldots\ldots\ldots, v_{-2}, v_{-1}, v_0, v_1, v_2, \ldots\ldots\ldots\ldots)$ on the grid with step size Δx:

$$norm\, v = \|v\| = \left[\Delta x \sum_{m=-\infty}^{\infty} |v_m|^2\right]^{\frac{1}{2}}$$

And for a function f on the real time

$$norm\ f = \|f\| = \left[\int_{-\infty}^{+\infty} |f(x)|^2 dx\right]^{\frac{1}{2}}$$

Definition: a finite one-step difference scheme (43) for a first order P.D.E. is stable if there exist number $k_0 > 0$ *and* $h_0 > 0$ such that for any for any $T > 0$ there exist a constant C_T such that

$$\|v^n\| \le C_T \|v^0\| \quad , \text{For} \quad 0 < nk \le T, 0 < h \le h_0, 0 < k \le k_0$$

Definition: The initial value problem for the first order P.D.E. is well-posed, if for any time $T \ge 0$, there exist C_T such that any solution $u(x,t)$ satisfies

$$\|u(x,t)\| \le C_T \|u(x,0)\| \text{ for } 0 \le t \le T$$

Definition: A one-step finite difference scheme approximating a P.D.E. is convergent if for any solution to the P.D.E., $u(x,t)$ is approach to numerical solution $u(nh,mk)$ as $h,k \to 0$

Note: A consistent finite difference scheme for a P.D.E. for which the initial value problem well posed is convergent if it is stable.

Definition: Fourier transformation and inversion formula for u defined in region $\Re$ given as;

$$\hat{u}(p) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ipx} u(x)\, dx\,;$$

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ipx} \hat{u}(p)\, dp$$

For a grid function $v = (\ldots\ldots\ldots, v_{-2}, v_{-1}, v_0, v_1, v_2, \ldots\ldots\ldots\ldots)$ with grid spacing Δx

$$\hat{u}(\xi) = \frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{\infty} e^{-imh\xi} u_m\ \Delta x\,;$$

$$u_m = \frac{1}{\sqrt{2\pi}} \int_{-\pi/h}^{\pi/h} e^{-ipx} \hat{u}(\xi)\, d\xi$$

From the Parseval condition

$$\|u(x)\| = \|\hat{u}(p)\|\,, \qquad \text{where } \|\hat{u}(\xi)\|^2 = \int_{-\pi/h}^{\pi/h} |\hat{u}(\xi)|^2\, d\xi$$
$$\|\hat{u}(\xi)\| = \|u_m\|$$

Convergent:

The system of algebraic equations that is resulting from recurrence relation of finite difference schemes, written as

$$\vec{u}^{\,n+1} = M\vec{u}^{\,n} + \vec{c} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\, (45)$$

Let $\hat{u}^n$ be the solution of the finite difference system (45) with a perturbed initial conditions;

$$\hat{\vec{u}}^{\,n+1} = M\hat{\vec{u}}^{\,n} + \vec{c} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\, (46)$$

Let $\qquad \varepsilon = \vec{u}^{\,n} - \hat{\vec{u}}^{\,n}$

From (45) and(46) ,we have

$$\vec{u}^{\,(n+1)} - \hat{\vec{u}}^{\,(n+1)} = M(\vec{u}^{\,(n)} - \hat{\vec{u}}^{\,(n)})$$

$$\Rightarrow \vec{\varepsilon}^{\,(n+1)} = M\vec{\varepsilon}^{\,(n)} \Rightarrow \vec{\varepsilon}^{\,(n)} = M^{(n)}\vec{\varepsilon}^{\,(0)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(47)$$

If $M^n \to 0$, as $n \to \infty$,then $\vec{\varepsilon}^{\,(n)} \to 0$ (i.e. the system is convergent).

Definition: if $M^n \to 0$, as $n \to \infty$,then M is convergent.

Definition: Spectrum radius $\rho(M) = \max\limits_i |\lambda_i|$, where λ_i are the eigenvalues of the matrix M.

Theorem: If M is the matrix coefficients and $\rho(M)$ is spectrum radius, then $\|M\| \geq \rho(M)$.

Proof : Suppose $\rho(M) = \max\limits_i |\lambda_i| = \lambda_1$, then

$$M\vec{u} = \lambda_1 \vec{u} \Rightarrow \|M\vec{u}\| = \|\lambda_1 \vec{u}\|\ , \qquad \|\vec{u}\| \neq 0$$

$$\|M\bar{u}\| = |\lambda_1|\|\bar{u}\| \qquad \text{……………………..(48)}$$

Also, $\|M\bar{u}\| \le \|M\|\|\bar{u}\|$ ……………………...(49)

From equations (48)&(49), we get:

$$\|M\|\|\bar{u}\| \ge |\lambda_1|\|\bar{u}\|$$

Since $\|\bar{u}\| \ne 0 \;\; \Rightarrow \|\bar{u}\| > 0$, this implies that

$$\|M\| \ge |\lambda_1| = \rho(M)$$ ■

Theorem: If $\|M\| < 1$, then M is convergent.

Proof:

$$\left\|M^{(n)}\right\| = \left\|MM^{(n-1)}\right\|$$

$$\left\|MM^{(n-1)}\right\| \le \|M\|^{(n)}$$, Prove that!

From these relations, we get

$$\left\|M^{(n)}\right\| \le \|M\|^{(n)}$$

If $\|M\| < 1$, then $\|M\|^{(n)} \to 0, as\, n \to \infty$, this implies that

$M \to 0$, as $n \to \infty$, from the previous definition ,we have

M is convergent ■

Corollary: If $\|M\| < 1$ for any norm then the iterative process for

$\hat{\bar{u}}^{n+1} = M\hat{\bar{u}}^{n}$ will converge for every $u^{(0)}$.

Note: it is possible that for some norm that $\|M\| > 1$, but M is still convergent.

Theorem: If $\rho(M) \ge 1$, then M is not convergent.

Proof:

Suppose $\rho(M) = \max_i |\lambda_i| = \lambda_1$, then

$$M\bar{u} = \lambda_1\bar{u} \text{ , where } \lambda_1 \ge 0 \,, \bar{u} \ne 0$$

From(47***), we have

$$\vec{\varepsilon}^{(n)} = M^{(n)}\vec{\varepsilon}^{(0)}$$

Let $\bar{u} = \vec{\varepsilon}^{(0)}$, then

$$\vec{\varepsilon}^{(1)} = M\vec{\varepsilon}^{(0)} = M\bar{u} = \lambda_1\bar{u}$$

$$\vec{\varepsilon}^{(n)} = M^{(n)}\bar{u} = \lambda_1^{(n)}\bar{u}$$

Since $\lambda_1 \ge 0$, $\bar{u} \ne 0$, then

$\left\|\vec{\varepsilon}^{(n)}\right\| = |\lambda_1|^{(n)}\|\bar{u}\|$,dose not approach to zero, as $n \to \infty$, thus

M is not convergent ■

Theorem: Necessary and sufficient condition for $M \neq 0$ be convergent iff $\rho(M) < 1$.

Stability of finite difference: the numerical calculation using finite difference formula are done in digital calculators which have round off error, each point calculated there will be error in the finite result which differs from the exact finite difference formula. For stable solution, there should be no large accumulation of round off error.

Von-Neumann (Fourier-series) method:

In each, mesh point (i,n) there will be round off error e_i^n, the sum of which

error the grid points could represent as,

$$E(x,t) = \sum_{i,n}^{M} e_i^n$$

$$= \sum_{i,n}^{M} A_i e^{\sqrt{-1}\beta x} e^{\alpha t} \quad , \quad \beta, \alpha \; are\; arbitrary\; phase\; cons\tan ts.$$

where, $e^{\alpha t}$: growth factor of error with time(amplification factor)

$$E(x,t) = \sum_{i,n}^{M} A_i e^{\sqrt{-1}\beta(i\Delta x)} e^{\alpha(n\Delta t)}$$

If suppose $e^{\alpha t} = \zeta$ is an arbitrary real or complex number, where $\alpha(\Delta t)$ =constant, then

$$E(x,t) = \sum_{i,n}^{M} A_i e^{\sqrt{-1}\beta(i\Delta x)} \zeta^{(n)}$$

If $\zeta > 1$, then $\zeta^{(n)}$ will increase with time(i.e. instability), therefore , the require equation to gives the stability condition is

$$|\zeta| \leq 1 \qquad \text{……………..……….(50)}$$

Simplification: The propagation of error with time is taken for one mesh point rather than the whole mesh

$$\varepsilon_i^n = e^{\sqrt{-1}\beta(i\Delta x)} \zeta^{(n)} \quad \text{(error for single mesh point) …..…..(51)}$$

Theorem: The error term in each mesh point ε_i^n satisfies the same finite difference formula used to calculate value of u at that mesh point.

Proof: For the explicit finite difference formula (equation (10))

$$\bar{u}_i^{n+1} = (1-2r)\bar{u}_i^n + r(\bar{u}_{i+1}^n + \bar{u}_{i-1}^n)$$
$$\bar{u}_1^{n+1} = (1-2r)\bar{u}_1^n + r(\bar{u}_2^n + \bar{u}_0^n)$$
$$.$$
$$.$$
$$\bar{u}_M^{n+1} = (1-2r)\bar{u}_M^n + r(\bar{u}_{M+1}^n + \bar{u}_{M-1}^n)$$

$$\Rightarrow \bar{u}^{(n+1)} = A^{(n)}\bar{u}^{(0)}$$

where A is ($M \times M$) bounded matrix, $\bar{u}^{(0)}$ is initial value at $t=0$. Suppose we introduce an error ε_i^n in initial calculation, we get

$$\bar{\varepsilon}^{(n+1)} = A^{(n)}\bar{\varepsilon}^{(0)} \qquad \text{(Prove that)}$$

Example: Find the stability condition for the explicit finite difference formula.

The explicit finite difference formula is

$$u_i^{n+1} = (1-2r)u_i^n + r(u_{i+1}^n + u_{i-1}^n)$$

From the above theorem, we get

$$\varepsilon_i^{n+1} = (1-2r)\varepsilon_i^n + r(\varepsilon_{i+1}^n + \varepsilon_{i-1}^n)$$

Apply von Neumann analysis (equation (51)) for each term in above equation, we obtain

$$e^{\sqrt{-1}\beta(i\Delta x)}\zeta^{(n+1)} = (1-2r)e^{\sqrt{-1}\beta(i\Delta x)}\zeta^{(n)} + r(e^{\sqrt{-1}\beta(i+1)\Delta x} + e^{\sqrt{-1}\beta(i-1)\Delta x})\zeta^{(n)}$$

Divided by $e^{\sqrt{-1}\beta(i\Delta x)}\zeta^{(n)}$, we have

$$\zeta = (1-2r) + r(e^{\sqrt{-1}\beta\Delta x} + e^{-\sqrt{-1}\beta\Delta x})$$
$$= (1-2r) + 2r\cos\beta\Delta x$$
$$\zeta = 1 - 4r\sin^2\frac{\beta\Delta x}{2}$$

From the stability condition (50), we have $|\zeta| = \left|1-4r\sin^2\frac{\beta\Delta x}{2}\right| \le 1$,and this implies to $0 \le r \le \frac{1}{2}$(give the details to illustrate that)

Exercise8: Find the stability condition for an explicit finite difference formula that is used to approximation $u_t = u_{xx} + u_{yy}$.

Exercise9: Show that an explicit finite difference formula for approximation is stable for $0 \le r \le \frac{1}{6}$.

Exercise10: Consider the finite difference equation

$$u_i^{n+1} - 2u_i^n + u_i^{n-1} = \frac{r^2}{2}\left\{(u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}) + (u_{i+1}^{n-1} - 2u_i^{n-1} + u_{i-1}^{n-1})\right\}$$

(a) Find P.D.E. that is consistent with FDE

(b) Find the stability condition.

Exercise11: (1) Approximate $u_t + vu_x - \alpha u_{xx} = 0$ by ;

(a) Explicit method **(b)** Implicit method **(c)** Crank-Nicolson method

(2) Find the truncation error and stability regions for all above finite difference methods.

(3) Approximating the first derivative in the P.D.E (part 1), by using the weight θ at two time levels, then, find the truncation error and stability condition.

Matrix stability analysis:

Assuming periodic initial data and neglecting the boundary conditions, we have used the von-Neumann method to determine the stability of the difference schemes. We now apply the matrix method, which automatically takes into account the boundary conditions of the problem, to difference schemes for the stability analysis. The two level difference scheme may be written as,

$$A_0\vec{u}^{(n+1)} = A_1\vec{u}^{(n)} + b^n, \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(52)$$

where b^n contains boundary conditions and $|A_0| \neq 0$. For $A_0 = \mathrm{I}$, the difference scheme(52) will be an explicit scheme otherwise an implicit scheme. We now assume that an error is introduced by round-off or some other source in to the solution $\vec{u}^n$ and call it $\vec{u}^{*n}$, then

$$A_0\vec{u}^{*(n+1)} = A_1\vec{u}^{*(n)} + b^n \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(53)$$

Subtracting equation(52) from equation(53), we get

$$A_0\vec{\varepsilon}^{*(n+1)} = A_1\vec{\varepsilon}^{*(n)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(54)$$

,where $\vec{\varepsilon}^{*(n)} = \vec{u}^{*(n)} - \vec{u}^{(n)}$ is the numerical vector error. In the stability analysis by the matrix method, we determine the condition under which the value of the numerical error vector $\|\vec{\varepsilon}^{*(n)}\| = \|\vec{u}^{*(n)} - \vec{u}^{(n)}\|$, where $\|\cdot\|$ denotes a suitable norm, remains bounded as n increases indefinitely, with k remaining fixed.

The equation (54) can be written in the form

$$\vec{\varepsilon}^{*(n+1)} = P\vec{\varepsilon}^{*(n)}$$

where $P = A_0^{-1}A_1$

It is simple to verify that $\vec{\varepsilon}^{*(n+1)} = P^{(n+1)}\vec{\varepsilon}^{*(0)}$

Thus the stability condition in the matrix method depends on the determination of a suitable estimate for $\|P\|$. When P is symmetric or similar to a symmetric matrix then $\|P\|_2$ is given by the spectral radius of P. Now, if the eigenvalues λ_i of P are distinct and the eigenvectors are $V^{(i)}$, we can expand the vector

$$\vec{\varepsilon}^{*(0)}=\sum_{i=1}^{M-1}C_iV^{(i)}$$

Then, we have

$$\vec{\varepsilon}^{*(n+1)}=\sum_{i=1}^{M-1}C_i\lambda_i^{(n+1)}V^{(i)}$$

Moreover, for the stability of difference scheme (52) we required each $|\lambda_i|\le 1$ for all i.

Hence, we get the result that error will not increase exponentially with n provided the eigenvalue with largest modulus has a modulus less than or equal one or

$$\|P\|_2=\max_i|\lambda_i|\le 1$$

It is easy to see that the eigenvalues are the zeros of the characteristic equation

$$|A_1-\lambda A_0|=0$$

For the explicit method, we have

$$A_1=\mathrm{I}+rC,\qquad A_0=\mathrm{I}$$

The eigenvalues and eigenvectors of C are giving by

$$\lambda_i=-4\sin^2\frac{i\pi}{2M},\quad 1\le i\le M-1 \qquad\qquad \text{Prove that!}$$

$$V^{(i)}=\left[\sin\frac{i\pi}{M}\ \ \sin\frac{2i\pi}{M}\ \ \sin\frac{3i\pi}{M}\cdots\cdots\cdots\cdots\cdots\sin\frac{(M-1)i\pi}{M}\right]$$

It follows that the eigenvalues of $\mathrm{I}+rC$ are

$$\lambda_i=1-4\sin^2\frac{i\pi}{2M},\quad 1\le i\le M-1$$

Therefore, the condition for the stability of the explicit method is

$$-1\le 1-4r\sin^2\frac{\beta\Delta x}{2}\le 1$$

Hence, $0\le r\le\frac{1}{2}$. The result obtain, which is identical with that obtained by application of the von-Neumann method.

Exercise12: Use this method to determine the stability of the difference equation that resulting in the previous exercise.

Gersschgorins theorem: The largest of the moduli of the eigenvalues of a square matrix A can not exceed the largest sum of the moduli of the elements along any row or any column.

$$|\lambda|\le|sum\ of\ any\ row\ or\ any\ column|$$

<u>Brours theorem:</u> Let P_i be the sum of the moduli of the elements along the i^{th} row excluding the diagonal elements a_{ii}. Then each eigenvalue of A lies inside or on the boundary of at least one of the circles

$$|\lambda - a_{ii}| \le P_i \text{ , where } P_i \equiv radius, \quad a_{ii} \equiv center.$$

For example ,from Crank-Nicolson formula ,we have

$$B\bar{u}^{(n+1)} = (4\mathrm{I} - B)\bar{u}^{(n)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(55)$$
$$\bar{u}^{(n+1)} = (4B^{-1} - \mathrm{I})\bar{u}^{(n)}$$

Where, $$B = \begin{bmatrix} 2+2r & -r & 0 & 0 & \cdot & \cdot & 0 \\ -r & 2+2r & -r & 0 & \cdot & \cdot & 0 \\ 0 & -r & 2+2r & -r & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & -r & 2+2r & -r \\ 0 & 0 & 0 & 0 & 0 & -r & 2+2r \end{bmatrix}$$

If the eigenvalue of matrix B is λ, then for the system to be stable $\left|\frac{4}{\lambda} - 1\right| \le 1$,for the matrix B: $\max_i P_i = |-r| + |-r| = 2r, \quad a_{ii} = 2 + 2r$ the Brours theorem leads to $2 \le \lambda \le 2 + 4r$, give more details about this application.

<u>Exercise13:</u> Then show that the equations (55) are unconditionally stable for $2 \le \lambda$.

Nonlinear parabolic equation:

The coefficients of the unknowns are functions of the solution .we may solve these equations iteratively after being linearized in some way.

Richtmyer's linearization method:

Consider the P.D.E.

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u^m}{\partial x^2}, \quad m \ge 2$$

Implicit weighted average difference scheme:-

$$\frac{u_{i,n+1} - u_{i,n}}{\Delta t} = \frac{1}{(\Delta x)^2}\left(\theta\delta_x^2\left[u_{i,n+1}^m\right] + (1-\theta)\delta_x^2\left[u_{i,n}^m\right]\right) \qquad (*)$$

$$u_{i,n+1}^m = u_{i,n}^m + \frac{\Delta t}{1!}\frac{\partial u_{i,n}^m}{\partial t} + \cdots\cdots\cdots\cdots$$

$$\begin{aligned} u_{i,n+1}^{m} &= u_{i,n}^{m} + \Delta t \frac{\partial u_{i,n}^{m}}{\partial u_{i,n}} \cdot \frac{\partial u_{i,n}}{\partial t} + \cdots\cdots\cdots\cdots \\ &= u_{i,n}^{m} + \Delta t \cdot m \cdot u_{i,n}^{m-1} \cdot \frac{\partial u_{i,n}}{\partial t} + \cdots\cdots\cdots\cdots \\ &= u_{i,n}^{m} + m \cdot u_{i,n}^{m-1} \cdot (u_{i,n+1} - u_{i,n}) + \cdots\cdots\cdots\cdots \end{aligned}$$

Now, for simplicity we can write this equation as

$$u_{i,n+1}^{m} = u_{i,n}^{m} + m \cdot u_{i,n}^{m-1} \cdot (u_{i,n+1} - u_{i,n})$$

So, $u_{i,n+1}^{m}$ here is a function of linear variable $u_{i,n+1}$. Replace the unknown by a linear approximation in $u_{i,n+1}$. Let $w_u = (u_{i,n+1} - u_{i,n})$,then

$$u_{i,n+1}^{m} = u_{i,n}^{m} + m \cdot u_{i,n}^{m-1} \cdot w_i$$

Substituting in (*), we obtain

$$\begin{aligned} \frac{w_i}{\Delta t} &= \frac{1}{(\Delta x)^2}\left(\theta \delta_x^2 \left[u_{i,n}^{m} + m \cdot u_{i,n}^{m-1} \cdot w_i\right] + (1-\theta)\delta_x^2\left[u_{i,n}^{m}\right]\right) \\ &= \frac{1}{(\Delta x)^2}\left(m\theta \delta_x^2 \left[u_{i,n}^{m-1} \cdot w_i\right] + \delta_x^2\left[u_{i,n}^{m}\right]\right) \end{aligned}$$

Using the definition of the operator δ^2, we obtain

$$\frac{w_i}{\Delta t} = \frac{1}{(\Delta x)^2}\left(m\theta\left[u_{i-1,n}^{m-1} \cdot w_{i-1} - 2u_{i,n}^{m-1} \cdot w_i + u_{i+1,n}^{m-1} \cdot w_{i+1}\right] + \left[u_{i-1,n}^{m} - 2u_{i,n}^{m} + u_{i+1,n}^{m}\right]\right)$$

Which give the set of linear equations for the w_i (when $m = 2$)

$$\begin{bmatrix} 1+4r\theta u_{1,n} & -2r\theta u_{2,n} & 0 & 0 & \cdot & \cdot & 0 \\ -2r\theta u_{1,n} & 1+4r\theta u_{2,n} & -2r\theta u_{3,n} & 0 & \cdot & \cdot & 0 \\ 0 & -2r\theta u_{2,n} & 1+4r\theta u_{3,n} & -2r\theta u_{4,n} & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & -2r\theta u_{M-3,n} & 1+4r\theta u_{M-2,n} & -2r\theta u_{M-1,n} \\ 0 & 0 & 0 & 0 & 0 & -2r\theta u_{M-2,n} & 1+4r\theta u_{M-1,n} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ w_{M-2} \\ w_{M-1} \end{bmatrix}$$

$$= \begin{bmatrix} -2ru_{1,n} & ru_{2,n} & 0 & 0 & \cdot & \cdot & 0 \\ ru_{1,n} & -2ru_{2,n} & ru_{3,n} & 0 & \cdot & \cdot & 0 \\ 0 & ru_{2,n} & -2ru_{3,n} & ru_{4,n} & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & ru_{M-3,n} & -2ru_{M-2,n} & ru_{M-1,n} \\ 0 & 0 & 0 & 0 & 0 & ru_{M-2,n} & -2ru_{M-1,n} \end{bmatrix} \begin{bmatrix} u_{1,n} \\ u_{2,n} \\ \\ \\ \\ u_{M-2,n} \\ u_{M-1,n} \end{bmatrix} + \begin{bmatrix} ru_{0,n} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ ru_{M,n} \end{bmatrix}$$

Exercise: use the Rrichtmyer's method to solve (*) with $m = 3$.

Newton's method: By Taylor's expansion

$$f(x_{n+1}) = f(x_n) + \frac{(x_{n+1} - x_n)}{1!} f'(x_n) + \frac{(x_{n+1} - x_n)^2}{1!} f''(x_n) + \cdots\cdot$$

If $f(x_{n+1})$ is the solution of the equation $f(x)=0$, then

$$0 = f(x_n) + \frac{(x_{n+1}-x_n)}{1!} f'(x_n) \;\Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

In case x_i is a vector $J(x_n)(\vec{x}_{n+1}-\vec{x}_n) = -f(\vec{x}_n)$. If $\vec{x}_{n+1}-\vec{x}_n = \omega$, then $J(x_n)\omega = -f(\vec{x}_n)$, where

$$J(x_{n)} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & . & . & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & . & . & \frac{\partial f_2}{\partial x_n} \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & . & . & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

V_i Known approximation to u_i, thus

$$u_i = V_i + \omega_i$$

The nonlinear equation can be expressed as

$$f_i(u_1,u_2,.........,u_N) = f_i(u_j) = 0 \quad , i,j = 1(1)N$$

$N\ equations\ \ in\ \ N\ unknown\,\mathrm{var}\ iables$

$u_i\ \ arethesolutions$

$V_i\ are\ approximat\!e\ solutionto\ u_i$

$$f_i(V_1,V_2,.........,V_N) + \left[\frac{\partial f_i}{\partial u_1}\omega_1 + \frac{\partial f_i}{\partial u_2}\omega_2 + \cdots\cdots + \frac{\partial f_i}{\partial u_N}\omega_N\right] = 0$$

Example: solve (*) by using Newton linearization method.

By Crank-Nicolson method with $m=2$, we can approximate equation (*) as

$$\frac{u_{i,n+1}-u_{i,n}}{\Delta t} = \frac{1}{2(\Delta x)^2}\left(\left[u^2_{i-1,n+1} - 2u^2_{i,n+1} + u^2_{i+1,n+1}\right] + \left[u^2_{i-1,n} - 2u^2_{i,n} + u^2_{i+1,n}\right]\right) \text{ ....(\&)}$$

Let $\frac{(\Delta x)^2}{\Delta t} = P\ and\ denote u_{i,n+1}\ by\, u_i$. After rearrange equation (&) becomes

$$u^2_{i-1} - 2(u_i^2 + pu_i) + u^2_{i+1} + \left\{\left(u^2_{i-1,n} - 2(u^2_{i,n} - pu_{i,n}) + u^2_{i+1,n}\right)\right\} = 0$$

$$f(u_{i-1},u_i,u_{i+1}) = 0$$

Apply Newton method with $u_i = V_i$ we obtain

$$f(V_{i-1},V_i,V_{i+1}) + \left[\frac{\partial f_i}{\partial u_{i-1}}w_{i-1} + \frac{\partial f_i}{\partial u_i}w_i + \frac{\partial f_i}{\partial u_{i+1}}w_{i+1}\right]_{u_i=V_i} = 0$$

Now,

$$2V_{i-1}\cdot\omega_{i-1} - 2(2V_i + p)\cdot\omega_i + V_{i+1}\cdot\omega_{i+1} + \left\{\left(V^2_{i-1} - 2(V_i^2 + pV_i) + V^2_{i+1}\right)\right\} + \left\{\left(u^2_{i-1,n} - 2(u^2_{i,n} - pu_{i,n}) + u^2_{i+1,n}\right)\right\} = 0$$

Exercise: Write down the set of linear equations for ω_i in matrix form

Irregular boundaries:

When the boundary is curved and intersects the rectangular mesh at points. That are not mesh point, then we cannot use the same formula, which we usually use:- We want to find the finite difference approximations to the derivatives at a point such as O close to the boundary curves figure .

Let the mesh be square and u is known on the curve and Taylor series for u at point O can be written as follows

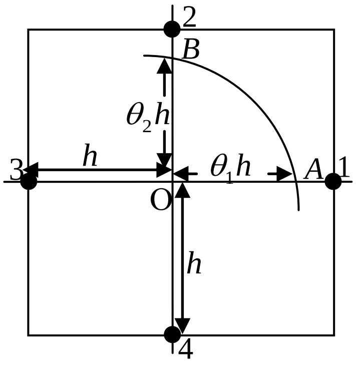


Figure-2- mesh square

$$u_A = u_o + h\theta_1 \frac{\partial u_o}{\partial x} + \frac{1}{2} h^2 (\theta_1)^2 \frac{\partial^2 u_o}{\partial x^2} + O(h^3)$$

$$u_3 = u_o - h\frac{\partial u_o}{\partial x} + \frac{1}{2} h^2 \frac{\partial^2 u_o}{\partial x^2} + O(h^3)$$

Elimination of $\frac{\partial^2 u_o}{\partial x^2}$ gives

$$\frac{\partial u_o}{\partial x} = \frac{1}{h}\left[\frac{1}{\theta_1(1+\theta_1)}u_A - \frac{1-\theta_1}{\theta_1}u_o - \frac{\theta_1}{1+\theta_1}u_3\right] + O(h^2)$$

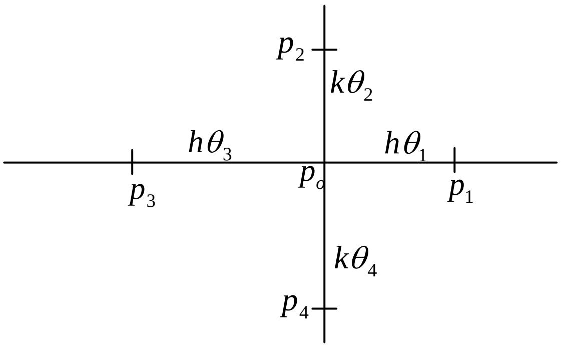
Similarly Elimination of $\frac{\partial u_o}{\partial x}$ gives

$$\frac{\partial^2 u_o}{\partial x^2} = \frac{1}{h^2}\left[\frac{2}{\theta_1(1+\theta_1)}u_A + \frac{2}{1+\theta_1}u_3 - \frac{2}{\theta_1}u_o\right] + O(h)$$

In the same ways approximate $\frac{\partial u_o}{\partial y}$ and $\frac{\partial^2 u_o}{\partial y^2}$.

Exercise: Approximate the elliptic equation $u_{xx} + u_{yy} = -16$.

<u>Exercise</u> 18: if the group of five points whose spacing is non-uniform $h\theta_1$ *and* $h\theta_3$ along $x-$axis, $k\theta_2$ *and* $k\theta_4$ along $y-$axis, arranged as in the figure:



(1) write the finite difference approximation for

$\frac{\partial u}{\partial x}$ at point p_o by(FDS,BDS,and CDS)

(2) show that the approximation formula of $\nabla^2 u(x,y)=0$,can be written as

$$\frac{2}{h^2}\left[\frac{1}{\theta_1+\theta_3}(\frac{u_1}{\theta_1}+\frac{u_3}{\theta_3})-(\frac{1}{\theta_1\theta_3}+\frac{h^2}{k^2\theta_2\theta_4})u_o+\frac{h^2}{k^2(\theta_2+\theta_4)}(\frac{u_2}{\theta_2}+\frac{u_4}{\theta_4})\right]+O(h,k)=0$$

Note: we represent $u_o=u(x_i,y_j)$, $u_1=u(x_i+h\theta_1,y_j)$, $u_3=u(x_i-h\theta_3,y_j)$,
$u_2=u(x_i,y_j+k\theta_2)$, and $u_4=u(x_i,y_j-k\theta_4)$.

Differential quadrature method

Introduction:

In addition to finite difference method, finite elements method and finite volume method there is an efficient discretization technique to obtain accurate numerical solutions. In this technique using a considerably small number of grid points(different point with FDM and FEM),Bellman and his workers (1971, 1972) introduce the method of differential quadrature(DQ) where a partial derivative of a function with respect to a coordinate direction is expressed as a linear weighted sum of all the functional values at mesh points along that direction.The DQ method was initiated from the idea of the integral quadrature(IQ).the key to DQ is to determine the weighting coefficients for the discretization of a derivative of any order .

Bellman et al (1972) use Legendre polynomial to determine the weighting coefficients of the first –order derivative, Civan(1989) improved Bellman approach to determine the weighting coefficients,Quan and Zhang(1989) applied Lagrange interpolated polynomials as test functions,so on.

Concepts and conclusions in DQ:

Differential quadrature method is a numerical method for solving differential equations. It is differs from finite difference method and finite elements method. The derivative along a direction is described into weighting linear combination of functional values at the grid points in differential quadrature method. Because all the information of functional values at the grid points is used in differential quadrature method, it has higher accuracy.

For convenience, we assumed that the function $u(x)$ is sufficiently smooth in the interval $[0,1]$, shown in figure (1).

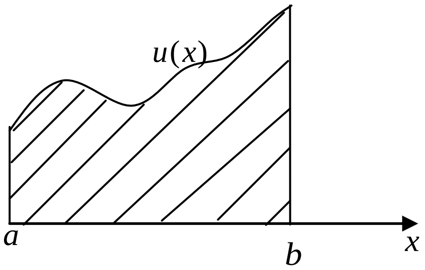


Figure 1- functions u over interval

The integral $\int_a^b u(x)dx$ represents the area under curve $u(x)$.Thus evaluating the integral is equivalent to the approximation of the area. In general, the integral can be approximated by

$$\int_a^b u(x)dx = w_1u_1 + w_2u_2 + \cdots\cdots\cdots + w_nu_n = \sum_{k=1}^{n} w_k u_k \quad \ldots\ldots\ldots\ldots\ldots(64)$$

Where, $w_1, w_2, \cdots\cdots\cdots, w_n$ are the weighting coefficients, $u_1, u_2, \cdots\cdots, u_n$ are the functional values at the discrete points $a = x_1, x_2, \cdots\cdots, x_n = b$ .equation(64) is called integral quadrature, which uses all the functional values in the whole integral domain to approximate an integral over a finite interval. One of these types of integral Trapezoidal rule, Simpson's rule.

By introducing some grids points $a = x_1 \le x_2 \le ..... \le x_N = b$ in the computational domain, Figure (2). The interval $[0,1]$ is divided into sub-intervals.

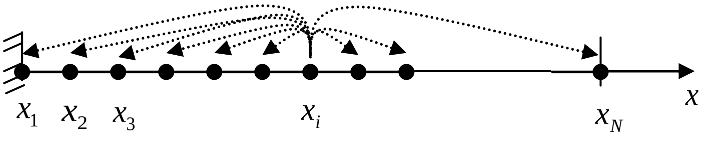


Figure 2- Computational domain stencils.

Assuming that the u_k is a value of function $u(x)$ at $x = x_k$, then the first and second derivatives of $u(x)$ at the grid points x_i is approximated by a linear combination of all functional value as follows;

$$u'(x_i) \cong \sum_{k=1}^{N} C_{ik}^{(1)} u_k \quad , \forall i = 1,2,....., N \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (65)$$

$$u''(x_i) \cong \sum_{k=1}^{N} C_{ik}^{(2)} u_k \quad , \forall i = 1,2,....., N \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (66)$$

where $C_{ik}^{(1)}$ and $C_{ik}^{(2)}$ are the weighting coefficients, and N is the number of grid points in the whole domain. Here the weighting coefficients are different at different location points of x_i. Equations (65) and (66) are called differential quadrature. In the application of the differential quadrature formulae (65) and (66), the choice of grid points and the determination of the weighting coefficients are two key factors. Once the grid points are given the weighting coefficients can be determined by using a set of test functions. There are many kinds of test functions that can be used. For example, striz et al (1995) and Shu and xue (1997) used Harmonic function, Shu (1999) used Fourier series expansion, and Guo and Zhong (2004) used the spline function. The polynomial test functions for determining the weighting coefficients are simply reviewed below.

Determination of the weighting coefficients

The calculation of the differential quadrature coefficients can be accomplished by several methods. In most of these methods, test functions $f_l(x), l = 1,2,....., N$, can be chosen such that:

$$u(x) \cong \sum_{l=1}^{N} \phi_l f_l(x) \quad \text{..} (67)$$

where, ϕ_l are constants to be determined. However, if the differential quadrature coefficients $C_{ik}^{(1)}$ and $C_{ik}^{(2)}$ are chosen such that the equations are represented as;

$$f_l'(x_i) \cong \sum_{k=1}^{N} C_{ik}^{(1)} f_l(x_k) \quad , \forall i, l = 1,2,......, N \quad \text{............................} (68)$$

$$f_l''(x_i) \cong \sum_{k=1}^{N} C_{ik}^{(2)} f_l(x_k) \quad , \forall i, l = 1,2,......, N \quad \text{............................} (69)$$

A relationship between first- and second- order coefficients can be obtained as:

$$f_k''(x_i) = \sum_{m=1}^{N} C_{im}^{(1)} \sum_{k=1}^{N} C_{ik}^{(1)} f_l(x_k) \cong \sum_{k=1}^{N} \sum_{m=1}^{N} C_{im}^{(1)} C_{mk}^{(1)} f_l(x_k) \quad \text{..............} (70)$$

Thus,

$$C_{ik}^{(2)} = \sum_{m=1}^{N} C_{im}^{(1)} C_{mk}^{(1)}, \forall i,k = 1,2,.....N.$$

in matrix notation:

$$\left[C^{(2)}\right] = \left[C^{(1)}\right]^2. \quad \text{..} \quad (71)$$

where

$$\left[C^{(1)}\right] = \begin{bmatrix} c_{11}^{(1)} & c_{12}^{(1)} & ..... & c_{1N}^{(1)} \\ c_{21}^{(1)} & c_{22}^{(1)} & ..... & c_{2N}^{(1)} \\ ..... & ..... & ..... & ..... \\ c_{N1}^{(1)} & c_{N2}^{(1)} & ..... & c_{NN}^{(1)} \end{bmatrix}, \quad \left[C^{(2)}\right] = \begin{bmatrix} c_{11}^{(2)} & c_{12}^{(2)} & ..... & c_{1N}^{(2)} \\ c_{21}^{(2)} & c_{22}^{(2)} & ..... & c_{2N}^{(2)} \\ ..... & ..... & ..... & ..... \\ c_{N1}^{(2)} & c_{N2}^{(2)} & ..... & c_{NN}^{(2)} \end{bmatrix}$$

Equation (71) implies that the values of $C_{ik}^{(2)}$ can be determined by two alternative (but equivalent) procedures, i.e. they can be obtained by directly solving equation (69) or by squaring the first –order matrix $\left[C^{(1)}\right]$.One approach for calculating the entries of$\left[C^{(1)}\right]$ and $\left[C^{(2)}\right]$ (Mingle, 1977; Civan and Sliepcevich, 1984; Naadimuthu et al, 1984; Bellman and Roth, 1986) is to use the test functions:

$$f_l(x) = x^{l-1}, l = 1,2,...., N \quad \text{...} \quad (72)$$

If the polynomials are taken as the test functions, the weighting coefficients ($C_{ik}^{(1)}$ and $C_{ik}^{(2)}$) satisfy the following linear systems

$$\mathbf{V} a_i = Z'(x_k) \quad \text{..} \quad (73)$$

$$\mathbf{V} b_i = Z''(x_k) \quad \text{..} \quad (74)$$

Where

$$a_i = [C_{i1}^{(1)}, C_{i2}^{(1)}, ..., C_{iN}^{(1)}]^{T_r} \ , \ b_i = [C_{i1}^{(2)}, C_{i2}^{(2)}, ..., C_{iN}^{(2)}]^{T_r}, \quad Z = [1, x, ..., x^{N-1}]^{T_r}$$

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & 1 & ..... & 1 \\ x_1 & x_2 & x_3 & ..... & x_N \\ x_1^2 & x_2^2 & x_3^2 & ..... & x_N^2 \\ ..... & ..... & ..... & ..... & ..... \\ x_1^{N-1} & x_2^{N-1} & x_3^{N-1} & ...... & x_N^{N-1} \end{pmatrix}$$

Here $\mathbf{V}$ is called Vandrmonde matrix, which is not singular and

$$\det(\mathbf{V}) = \prod_{k=2}^{n} \prod_{j=1}^{k-1} (x_k - x_j) \neq 0$$

Although the weighting coefficients can be determined by solving the linear system (37), the matrix **V** is highly –ill conditioned as N is large. In order to overcome this difficulty the Legendre interpolation polynomial are used by Bellman et al (1972).the formulations of the weighting coefficients are givens as follows

$$\left.\begin{aligned} C_{ik}^{(1)} &= \frac{L_N^{(1)}(x_i)}{(x-x_k)L_N^{(1)}(x_k)} \\ C_{ii}^{(1)} &= \frac{1-2x_i}{2x_i(x_i-1))} \end{aligned}\right\} \text{...} \quad (75)$$

where $L_N(x)$ and $L_N^{(1)}(x)$ are the Legendre polynomial of degree N and its first order derivative respectively.

Although we can determine the weighting coefficients for the second order derivatives by solving a system (74), the matrices are also highly-ill-conditioned. By using the Lagrange interpolation polynomials as the test function the weighting coefficients of second order derivatives are given by Quan and Chang as follows

$$C_{ik}^{(2)} = \frac{2}{x_k - x_i}\left(\prod_{l=1,l\neq i,k}^{N} \frac{x_i - x_k}{x_k - x_l}\right)\left(\prod_{j=1,j\neq i,k}^{N} \frac{1}{x_i - x_j}\right), for\ \ i \neq k \quad \text{..................} \ (\ 76)$$

$$C_{ii}^{(2)} = 2\sum\nolimits_{l=1,l\neq i}^{N-1}\left(\frac{1}{x_i - x_l}\left(\sum\nolimits_{j=1,j\neq i}^{N}\frac{1}{x_i - x_j}\right)\right) \quad \text{...................................} \ (77)$$

The recurrence formula to compute the weighting coefficients for mth order derivatives are given by Shu's as follows

$$C_{ik}^{(m)} = m\left(C_{ik}^{(1)}C_{ii}^{(m-1)} - \frac{C_{ik}^{(m-1)}}{(x_i - x_k)}\right), for\ \ i,k = 1,2,...N; 2 \le m \le N-1 \ \text{.........} \ (78)$$

$$\sum\nolimits_{k=1}^{N} C_{ik}^{(m)} = 0 \ \text{ or } \ C_{ii}^{(m)} = -\sum\nolimits_{k=1,k\neq i}^{N} C_{ik}^{(m)} \quad \text{................ ....} \ (79)$$

In similarly way with equation (71), a relationship between first- and high- order coefficients can be obtained,

$$\left[C^{(m)}\right] = \left[C^{(1)}\right] \cdot \left[C^{(m-1)}\right] = \left[C^{(m-1)}\right] \cdot \left[C^{(1)}\right], \ m = 2,3,....., N-1 \quad \text{.........} (80)$$

This equation indicates that the weighting coefficients for the high order derivative can be computed by the matrix multiplication of the weighting coefficients of the first order derivative. However, this equation is simple and involves more arithmetic operations as compared to equations (78 and 79). We noted that the calculation of weighting coefficient by equation (80) involves N multiplications and ($N-1$) additions, i.e., a total of ($2N-1$) arithmetic operations. Recurrence relationship (78) only involves two multiplications, one division, and one subtraction, i.e. a total of four arithmetic operations for calculation of each off-diagonal weighting coefficient, which is independent of the number of grid points N. The calculation of each diagonal weighting coefficient from equation (79) involves ($N-1$) subtractions. Thus, the number of arithmetic operations for equation (78) and equation (79) is substantially smaller than what is in equation (80).

Sample of typical grid distributions

Because the described equations obtained by using differential quadrature method are equivalent to one obtained by using quasi-spectrum method, the choice of grid points have a great effect upon accuracy of results. There are two kinds of methods for choosing the mesh points.

The uniform grid points are used in the first kind as follows:

Type (I): By a uniform grid, we mean that the grid has the same sizes.

Thus by

setting $\Delta x = x_2 - x_1 = x_i - x_{i-1} = x_N - x_{N-1}, ...ect.$

The coordinates of the grid points are chosen as

$$x_i = (b-a)\frac{i-1}{N-1} \qquad \text{for } \; i = 1,2,\ldots\ldots, N \; \text{ and } \; x_i \in [a,b].$$

The zeros of orthogonal polynomials such as Chebyshev polynomials are taken as grid points in the second kind as follows:

Type (II) : For this kind, the coordinates of the grid points are chosen as

$$x_i = \frac{b-a}{2}\frac{r_i - r_1}{r_N - r_1}, \quad \text{such that} \qquad r_i = \cos(\frac{i-1}{N-1}\pi)$$

In this field, there are some contribution studies about the effect of grid spacing distribution on the numerical results that were obtained by DQ method. Quan and Chang (1989) compared numerically the performances of the often-used non-uniform meshes and concluded that the grid points originated from the Chebyshev polynomials of the first kind is optimum in all cases examined. Bert and Malik (1996) indicated an important fact that the preferred type of grid points changes with problems of interest and recommended the use of Chebyshev-Gauss-Lobatto grid for structural mechanics computations. Maradi and Taheri (1998) also investigated the effect of various spacing schemes on the accuracy of DQ results for buckling application of composites. They provided insights into the influence of a number of sampling points in conjunctions with various spacing schemes. Chen (1997) and Bert and malik (1996) have provided sensible explanations why a certain type of grid points is superior to the others in the computation of their problems. The details of properties of DQ weighting coefficient matrices for the determination and rank are given by Shu (2000), and we note from this reference that these properties can be derived from the matrices properties in algebraic subject.

Exercise: if weighting coefficients are desired for a range $0 \le x \le 1$, then calculate the weighting coefficients matrices $C_{ik}^{(1)}$ and $C_{ik}^{(2)}$ for $N = 3,4,5$ grid points divided the above range.

Numerical methods to solve DQ resultant equations

It is very important to make simple review about the solution techniques, which are used to update the DQ resultant for the differential equations. In most applications of the DQ method to engineering and physics problems, which are governed by the partial differential equations, considering the second –order partial differential equation as follows:

$$\frac{\partial u}{\partial t} = f\,(t, x, u, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (81)$$

In general, Equation (81) should be specified with proper initial and boundary conditions for the solution to a specific problem. By DQ method at all interior points of whole domain, the original problem, which is defined in equation (81) can be reduced

to a set of N ordinary differential equations(ODEs) as

$$\frac{du(t, x_i)}{dt} \cong f\,(t, x_i, u(t, x_i), \sum_{k=1}^{N} C_{ik}^{(1)} u(t, x_k), \sum_{k=1}^{N} C_{ik}^{(2)} u(t, x_k)) \text{ for } i = 1,2,\ldots\ldots, N \text{ .(82)}$$

When $\frac{\partial u(t, x_i)}{\partial t} = 0$, we can be able to obtain a system of linear algebraic equations. The solution of partial differential equations may not be possible to express in closed-form. Therefore, this solution function can be approximated by polynomial approximation. Rearranging equation (82) to obtain a set of ordinary differential equations as;

$$\frac{d\{u\}}{dt} + L_{dq}\{u\} = \{G\} \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (83)$$

where $\{u\}$is a vector representing a set of unknown functional values at all interior points, $L_{dq}\{u\}$ is a vector resulting from DQ discretization, $\{G\}$ is a vector arising from the given initial and boundary conditions. For time-dependent problems, equation (83) constitutes standard form ordinary differential equations. The time derivative can be approximated by explicit or implicit low order finite difference scheme. From equation (83), we can obtain a system of algebraic equations in the form

$$[H]\cdot\{u\}=\{G\} \quad \text{..} \quad (84)$$

where $\{u\}$is a vector of unknown functional values at all the interior grid points given by

$$\{u\}=\left(u_{2,2},u_{2,3},......,u_{2,M-1},u_{3,2},u_{3,3},......,u_{3,M-1},........,u_{N-1,2},u_{N-1,3},......,u_{N-1,M-1},\right)^{T_r}$$

and $\{G\}$ is a known vector given by

$$\{G\}=\left(G_{2,2},G_{2,3},......,G_{2,M-1},G_{3,2},G_{3,3},......,G_{3,M-1},........,G_{N-1,2},G_{N-1,3},......,G_{N-1,M-1},\right)^{T_r}$$

The dimension of the matrix $[H]$ is $(N-2)(M-2)$ by $(N-2)(M-2)$. Equation (84) can be written alternatively as

$$[C]\ [u]+[u]\ [D]=[G] \quad \text{..} \quad (\ 85\ )$$

this equation is called Lyapunov matrix form and $[C],[D]$are matrices of weighting coefficients for the first and the second-order derivatives have the dimension $(N-2)(N-2),(M-2)(M-2)$ respectively. One can see that the dimensions of $[C]$ and $[D]$are very small compared with the dimensions of$[H]$. To solve this system that is discritized by DQ method, one can adopt direct method or iterative method. To solve the ordinary differential equations that are given in equation (83 or 84), there are different explicit numerical schemes that are used to discritize these equations and compute the results, for example: Euler forward explicit scheme; this is the first order scheme given by

$$u^{n+1}=u^{n}+\Delta t\cdot f^{n} \quad \text{..} \quad (86)$$

The solution techniques that we thought could be possibly used to solve the algebraic equations that are because of employing DQ method in governing equations are divided into two parts. The first part is named direct methods, and the second one is iterative methods.

Direct methods

To solve algebraic equations included in equation (84), there are many standard methods, amongst of them, Gaussian elimination method, LU decomposition approach are used extensively. The details of these methods can be found in textbook of numerical analysis. These methods are very efficient when the dimension of the matrix is not large. However, when the number of grid points increases the dimension of the matrix will increase accordingly. Hence the problem of virtual storage will become critical;

furthermore, the DQ discretization matrix tends to become ill-conditioned when the mesh size is large. This would lead to difficulties in obtaining the solution or even worse, reduce the accuracy of the solutions. The drawbacks of direct methods can be eliminated by using iterative methods. Some of these iterative methods have been used to solve the system of algebraic equations given in form of equation (84).

Iterative methods

If the matrix $[H]$ in equation (84) is composed of two matrices $[A]$ and $[P]$, then we can write it as

$$[H]=[A]+[P] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(87)}$$

by rewriting equation (84) in terms of matrices $[A]$ and $[P]$, we obtain

$$[A]\cdot\{u\}=\{G\}-[P]\cdot\{u\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(88)}$$

The iterative expression for equation (88) can be written as

$$[A]\cdot\left(\{u^{n+1}\}-\{u^{n}\}\right)=\{G\}-[P]\cdot\{u^{n}\}-[A]\cdot\{u^{n}\} \quad \ldots\ldots\ldots\ldots\text{...} \text{(89)}$$

where n represents iterative level and the right side sometimes is called the vector of the residuals. In practical applications, a relaxation factor Θ is introduced on the right hand side of equation (89), and the final iteration expression becomes

$$\{u^{n+1}\} = \{u^n\} + \Theta \cdot [A]^{-1} \cdot \{R^n\} \quad \text{..............................} \; (90)$$

such that

$$R^n = \{G\} - [H] \cdot \{u^n\}$$

Equation (90) is a general iterative expression for equation (84). By using different forms of $[A]$, we can obtain different iterative expressions for equation (84). For the

stability and convergence of iterative method the reader may consult the textbooks Smith (1978) and Rao (2002).

Successive over-relaxation (SOR) iteration method: SOR iteration is used to improve the convergence speed of Jacobi method. It is noted that SOR is a point iteration method. The value of u_i^{n+1} can be evaluated, when the values of $u_k^{n+1}, k = 1,2,....., i-1$ is calculated. These new values at the iteration level $(n+1)$ can then be used to compute the residuals. The residuals of SOR iteration are computed from

$$\{R^n\} = \{G\} - [H_L]\{u^{n+1}\} - ([H_D] + [H_U])\{u^n\}$$

where $[H_L]$ is the lower triangular matrix with diagonal elements being zero, $[H_U]$ is the upper triangular matrix with diagonal elements being zero, and $[H_D]$ is the diagonal matrix with elements being the diagonal elements of $[H]$. It is noted that the elements of

$[H_L],[H_U]$, and $[H_D]$ are equal to those of $[H]$ at the corresponding positions, that is

$$[H]=[H_L]+[H_U]+[H_D]$$

The iterative expression of SOR method is the same as the following equation

$$u^{n+1}=u^n+\Theta\cdot\frac{R_i^n}{a_{ij}}$$

the R_i^n in the SOR method can be expressed as

$$R_i^n=G_i-\sum_{j=1}^{i-1}a_{ij}u_j^{n+1}-\sum_{j=1}^{M}a_{ij}u_j^n$$

SOR iterative methods for the Lyapunov system (2.20) to update the solution can be write as

$$u^{n+1}=u^n+\frac{\Theta}{C_{ii}+D_{jj}}\Big(G_{ij}-\sum_{k=1}^{i-1}C_{ik}u_{kj}^{n+1}-\sum_{k=i}^{N}C_{ik}u_{kj}^{n}-\sum_{k=1}^{j-1}D_{jk}u_{ikj}^{n+1}-\sum_{k=j}^{M}D_{jk}u_{ik}^{n}\Big)\;\ldots\ldots\ldots\ldots\ldots\ldots\ldots(91)$$

with the residuals relation in the form

$$[R^n]=[G]-[C_L][u^{n+1}]-([C_D]+[C_U])[u^n]-[u^{n+1}][D_L]-[u^n]([D_D]+[D_U])$$

where the matrices C's and D's are having the same defined matrices of H's, which are mentioned above.

Gauss-seidel iteration method: it is special case of SOR (successive over-relaxation) iteration when Θ is taken as 1. There are many iterative methods some are related with these methods and others are different like Jacobin method, Jacobin over relaxation iteration method, Richardson iteration method, Conjugate Gradient iteration method…etc.

Error Analysis of DQ Method:

Shu (1991) and Chen(1996) are introduce the error resulting from approximation a function and its derivatives.

Error analysis for the function:

When approximate $u(x)$ by a polynomial of degree $(N-1)$, particularly by the Lagrange interpolation polynomial

$$\mathrm{P}_n u = \sum_{i=1}^{N} u(x_i)\cdot r_i(x) \qquad \ldots\ldots\ldots\ldots\ldots\ldots.(92)$$

where $r_i(x)$ is the Lagrange interpolation polynomial given by

$$\begin{aligned} r_k(x) &= x^{k-1} && ,k=1,2,\cdots\cdots\cdots,N \\ r_k(x) &= \frac{\ell_N(x)}{(x-x_k)\,\ell_N^{(1)}(x)} && ,k=1,2,\cdots\cdots\cdots,N \\ r_k(x) &= \frac{M_N(x)}{(x-x_k)\,M_N^{(1)}(x)} && ,k=1,2,\cdots\cdots\cdots,N \\ r_k(x) &= 1,\ r_k(x)=(x-x_{k-1})\cdot r_{k-1}(x), && ,k=1,2,\cdots\cdots\cdots,N \end{aligned} \qquad \ldots\ldots..(93)$$

Where ℓ is the Legender polynomial of degree N, and $M(x)$ defined as,

$$M(x)=(x-x_1)(x-x_2)\cdots\cdots\cdots(x-x_N) \qquad \ldots\ldots\ldots(94)$$

Thus, $M^{(1)}(x)=\prod_{k=1,k\neq i}^{N}(x_i-x_k)$

The approximate error of $u(x)$ is defined as,

$$\mathrm{E}(u)=u(x)-\mathrm{P}_N u \qquad \ldots\ldots\ldots..(95)$$

If the Nth order derivative of the function $u(x)$ is assumed to be a constant, say k, then $u(x)$ can be expressed as

$$u(x)=m_0+m_1x+m_2x+\cdots\cdots\cdots\cdots+m_{N-1}x^{N-1}+\frac{k\cdot x^N}{N!} \qquad \ldots\ldots\ldots\ldots(96)$$

Since equation(92) is exactly satisfied for a polynomial of degree less than or equal $(N-1)$, we have

$$\mathrm{E}(x^k)=0, \qquad k=0,1,\cdots\cdots\cdots\cdots,N-1 \qquad \ldots\ldots\ldots\ldots(97)$$

Substituting equation(96) into equation(95) and using equation(97), we obtain

$$\mathrm{E}(u)=k\cdot\frac{\mathrm{E}(x^N)}{N!} \qquad \ldots\ldots\ldots\ldots\ldots(98)$$

$$where \quad \mathrm{E}(x^N)=x^N-\sum_{i=1}^{N}x_i^N\cdot r_i(x)$$

On the other hand, substituting the polynomial of degree $(N-1)$, $g(x)=x^N-M(x)$ into equation(95),we obtain

$$\mathrm{E}(g)=x^N-M(x)-\sum\left[x_i^N-M(x_i)\right]\cdot r_i(x)=0 \qquad \ldots\ldots\ldots\ldots\ldots(99)$$

Since, $M(x_i)=0$, equation(99) can be further reduced to

$$x^N - \sum x_i^N \cdot r_i(x) = M(x_i) \qquad \text{................(100)}$$

Finally, substituting equation (100) into equation (98), we get

$$\mathrm{E}(u) = \frac{k \cdot M(x_i)}{N!} \qquad \text{................(101)}$$

In most cases, the Nth-order derivative of the function $u(x)$ is not a constant, but it may be boundad. In this case, we can adopt another method to analyze $\mathrm{E}(u)$. For simplicity, we set $\phi(x) = \mathrm{P}_N u$ and defined a function $U(z)$ as,

$$U(z) = u(z) - \phi(z) - C \cdot M(z) \qquad \text{..................(102)}$$

Clearly, when $z = x_1, x_2, \cdots\cdots, x_N$, $U(z) = 0$.

If we set $U(x) = 0$,we obtain

$$\mathrm{E}(u) = u(x) - \mathrm{P}_N u = u(x) - \phi(z) = C \cdot M(x) \qquad \text{.................(103)}$$

Since $U(z)$ has $N+1$ roots $x_1, x_2, \cdots\cdots, x_N$ in the domain, by repeated application of Roll's theorem, the Nth-order derivative of $U(z)$, $U^{(N)}(z)$ is found to have at least one root lying between x_1 *and* x_N. Denoting this root by ξ,we have

$$U^{(N)}(\xi) = 0 \qquad \text{.................(104)}$$

Not that, $\phi(z)$ is a polynomial of degree $(N-1)$.so from equation(102), we obtain

$$C = \frac{u^{(N)}(\xi)}{N!} \qquad \text{..................(105)}$$

Hence

$$\mathrm{E}(u) = \frac{u^{(N)}(\xi) \cdot M(x)}{N!} \qquad \text{..................(106)}$$

In general ξ is a function of x.

Error analysis for the derivatives:

The error for mth-order derivative approximation can be defined as

$$\mathrm{E}_D^{(m)}(u) = \frac{\partial^m u}{\partial o^m} - \frac{\partial^m (\mathrm{P}_N u)}{\partial o^m} = \frac{\partial^m u}{\partial o^m} - \frac{\partial^m \phi}{\partial o^m} \qquad \text{..................(107)}$$

Where $m = 1,2,\cdots\cdots, N-1$. Using equation (106), equation(107) can be written as

$$\mathrm{E}_D^{(m)}(u) = \frac{\partial^m \left[u^{(N)}(\xi) \cdot M(x)\right]}{N!\ \partial x^m} \qquad \text{..................(108)}$$

Since ξ is an unknown function of x, it is difficult to estimate $\mathrm{E}_D^{(m)}(u)$ using equation(108).as a special case, if we assume that the Nth-order derivative of $u(x)$ is a constant, say k,equation(108) can be simplified to

$$\mathrm{E}_D^m(u(x_i)) = \frac{k \cdot M^{(m)}(x_i)}{N!} \quad \ldots\ldots\ldots\ldots\ldots\ldots(109)$$

For the general case where $u^{(N)}(\xi)$ is not a constant, we can use a similar method as in the analysis of the function approximation to conduct error analysis of the derivative approximation.

Since $g(z) = u(z) - \phi(z)$ has N roots in the domain, according to Roll's theorem, its mth-order derivative $g^{(m)}(z)$ has at least $(N-m)$ roots in the domain, namely $\bar{x}_1, \bar{x}_2, \cdots\cdots, \bar{x}_N$. Thus, the function

$$\begin{aligned} U^{(m)}(z) &= g^{(m)}(z) - \bar{C} \cdot \bar{M}(z) \\ &= u^{(m)}(z) - \phi^{(m)}(z) - \bar{C} \cdot \bar{M}(z) \end{aligned} \quad \ldots\ldots\ldots\ldots(110)$$

where $\bar{M}(z) = (z - \bar{x}_1)(z - \bar{x}_2)\cdots\cdots(z - \bar{x}_{N-m})$,would vanish $\bar{x}_1, \bar{x}_2, \cdots\cdots, \bar{x}_{N-m}$. Now if we set $U^{(m)}(\bar{x}) = 0$,where $\bar{x}$ is different from $\bar{x}_1, \bar{x}_2, \cdots\cdots, \bar{x}_{N-m}$, then $U^{(m)}(z) = 0$ has $(N-m+1)$ roots ,and

$$\mathrm{E}_D^{(m)}[u(\bar{x})] = u^{(m)}(\bar{x}) - \phi^{(m)}(\bar{x}) = \bar{C} \cdot \bar{M}(\bar{x}) \quad \ldots\ldots\ldots\ldots\ldots\ldots(111)$$

Using Roll's theorem repeatedly the $(N-m)th$-order derivative of $U^{(m)}(z)$ is found to have at least one root $\bar{\xi}$. Thus equation(111) can be reduced to

$$\mathrm{E}_D^{(m)}[u(\bar{x})] = \frac{u^{(N)}(\bar{\xi}) \cdot \bar{M}(\bar{x})}{(N-m)!} \quad \ldots\ldots\ldots\ldots\ldots\ldots(112)$$

Equation(112) can be used to estimate the error of the derivative approximation. It is assumed that all the coordinates are in the interval Δx,and the Nth-order derivative of the function $u(x)$ is bounded, then

$\left|u^{(N)}(\bar{\xi})\right| \le C$, where C is a positive constant, and

$$\left|M^{(m)}(x)\right| \le N(N-1)\cdots\cdots(N-m+1)(\Delta x)^{N-m}$$

$$\left|\bar{M}(\bar{x})\right| \le (\Delta x)^{N-m}$$

So, equation(109) and equation(112) can be simplified to

$$\mathrm{E}_D^{(m)}[u(\bar{x})] = \frac{C \cdot (\Delta x)^{N-m}}{(N-m)!}, \quad for\ m = 1, 2, \cdots\cdots, N-1 \quad \ldots\ldots\ldots\ldots\ldots(113)$$

For the general case, the error distribution of the derivative approximation can also be studied using equation (108). The error distributions of the first-, second- , third- and fourth- order approximation have studied by Chen (1996). For the first-order derivative, equation (108) gives

$$\mathrm{E}_D^{(1)}(u) = \frac{u^{(N+1)}(\xi) \cdot \xi_x \cdot M(x) + u^{(N)}(\xi) \cdot M^{(1)}(x)}{N!} \quad \ldots\ldots\ldots\ldots\ldots\ldots(114)$$

Let $k_1 = \max\left\{\left|u^{(N)}(\xi)\right|\right\}$ and note that $M(x_i) = 0$. Applying equation (114) at that grid point x_i gives

$$\mathrm{E}_D^{(1)}\left[u(x_i)\right]=\frac{u^{(N)}(\xi)\cdot M^{(1)}(x_i)}{N!}\leq k_1 e^{(1)}(x_i) \quad , for \;\; i=1,2,\cdots\cdot,N \qquad \ldots\ldots(115)$$

where $e^{(1)}(x_i)=\dfrac{M^{(1)}(x_i)}{N!}$, is the error distribution of the first order derivative approximation. For the second-order derivative

$$\mathrm{E}_D^{(2)}(u)=\frac{2u^{(N+1)}(\xi)\cdot\xi_x\cdot M(x)+u^{(N)}(\xi)\cdot M^{(2)}(x)}{N!}\leq k_2 e^{(2)}(x) \qquad \ldots\ldots\ldots(117)$$

Where $k_2 = \max\left\{\left|u^{(N)}(\xi)\right|, \left|\xi_x u^{(N+1)}(\xi)\right|\right\}$, and $e^{(2)}(x_i)$, is the error distribution of the second order derivative approximation.