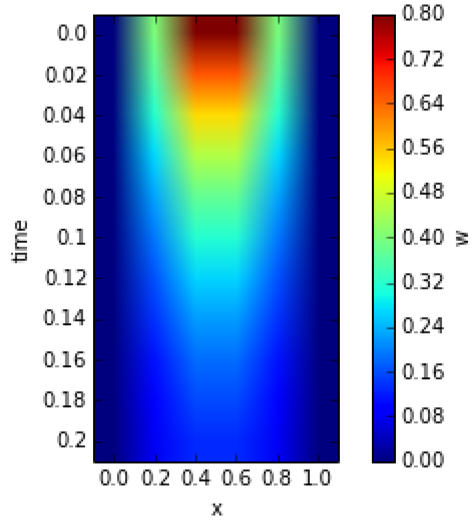
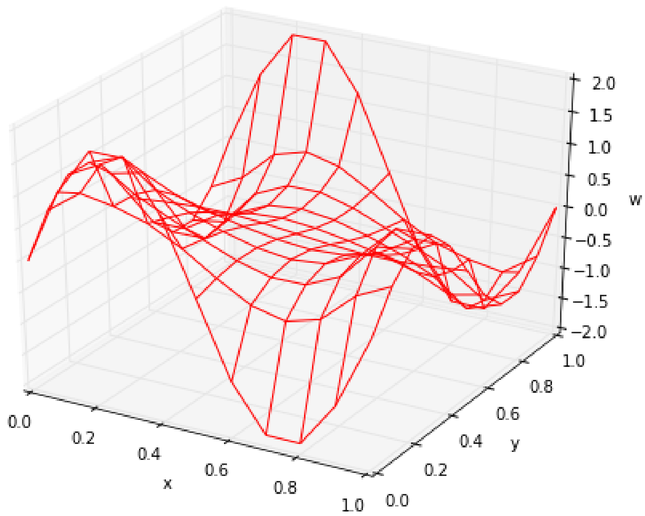
JOHN S BUTLER



Numerical Methods for Differential Equations with Python

Numerical Approximation of the Poisson Equation



CONTENTS

ACRONYMS

IVP Initial Value Problems

BVP Boundary Value Problems

ODE Ordinary Differential Equations

PDE Partial Differential Equations

RK Runge Kutta

Part I

INITIAL VALUE PROBLEMS

1

NUMERICAL SOLUTIONS TO INITIAL VALUE PROBLEMS

Differential equations have numerous applications to describe dynamics from physics to biology to economics.

Initial value problems are subset of Ordinary Differential Equation (ODE's) with the form

$$y' = f(x) \tag{1}$$

f is a function. The general solution to (1) is

$$y = \int f(x)dx + c,$$

containing an arbitrary constant c. In order to determine the solution uniquely it is necessary to impose an initial condition,

$$y(x_0) = y_0. \tag{2}$$

Example 1

Simple Example
The differential equation describes the rate of change of an oscillating input. The general solution of the equation

$$y' = \sin(x) \tag{3}$$

is,

$$y = -\cos(x) + c,$$

with the initial condition,

$$y(0) = 2,$$

then it is easy to find $c = 2$.
Thus the desired solution is,

$$y = 2 - \cos(x).$$

The more general Ordinary Differential Equation is of the form

$$y' = f(x, y), \tag{4}$$

is approached in a similar fashion.
Let us consider

$$y' = a(x)y(x) + b(x),$$

The given functions $a(x)$ and $b(x)$ are assumed continuous for this equation

$$f(x,z) = a(x)z(x) + b(x),$$

and the general solution can be found using the method of integrating factors.

Example 2

General Example
Differential equations of the form

$$y'(x) = \lambda y(x) + b(x), \quad x \geq x_0, \tag{5}$$

where λ is a given constant and $b(x)$ is a continuous integrable function has a unique analytic. Multiplying the equation (5) by the integrating factor $e^{-\lambda x}$, we can reformulate

$$\frac{d(e^{-\lambda x}y(x))}{dx} = e^{-\lambda x}b(x).$$

Integrating both sides from x_0 to x we obtain

$$(e^{-\lambda x}y(x)) = c + \int_{x_0}^{x} e^{-\lambda t}b(t)dt,$$

so the general solution is

$$y(x) = ce^{\lambda x} + \int_{x_0}^{x} e^{\lambda(x-t)}b(t)dt,$$

with c an arbitrary constant

$$c = e^{-\lambda x_0}y(x_0).$$

For a great number of Initial Value Problems there is no known exact (analytic) solution as the equations are non-linear, for example $y' = e^{xy^4}$, or discontinuous or stochastic. There for a numerical method is used to approximate the solution.

1.1 NUMERICAL APPROXIMATION OF DIFFERENTIATION

1.1.1 *Derivation of Forward Euler for one step*

The left hand side of a initial value problem $\frac{df}{dx}$ can be approximated by **Taylors theorem** expand about a point x_0 giving:

$$f(x_1) = f(x_0) + (x_1 - x_0)f^{'}(x_0) + \tau, \tag{6}$$

where τ is the truncation error,

$$\tau = \frac{(x_1 - x_0)^2}{2!}f^{''}(\xi), \qquad \xi \in [x_0, x_1]. \tag{7}$$

Rearranging and letting $h = x_1 - x_0$ the equation becomes

$$f^{'}(x_0) = \frac{f(x_1) - f(x_0)}{h} - \frac{h}{2}f^{''}(\xi).$$

The forward Euler method can also be derived using a variation on the Lagrange interpolation formula called the divided difference. Any function $f(x)$ can be approximated by a polynomial of degree $P_n(x)$ and an error term,

$$\begin{aligned} f(x) &= P_n(x) + error, \\ &= f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1), \\ &\quad +... + f[x_0, ..., x_n]\Pi_{i=0}^{n-1}(x - x_i) + error, \end{aligned}$$

where

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0},$$

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0},$$

$$f[x_0, x_1, .., x_n] = \frac{f[x_1, x_2, ..., x_n] - f[x_0, x_1, ..., x_{n-1}]}{x_n - x_0},$$

Differentiating $P_n(x)$

$$\begin{aligned} P_n^{'}(x) = f[x_0, x_1] + f[x_0, x_1, x_2]\{(x - x_0) + (x - x_1)\}, \\ +... + f[x_0, ..., x_n]\sum_{i=0}^{n-1}\frac{(x - x_0)...(x - x_{n-1})}{(x - x_i)}, \end{aligned}$$

and the error becomes

$$error = (x - x_0)...(x - x_n)\frac{f^{n+1}(\xi)}{(n+1)!}.$$

Applying this to define our first derivative, we have

$$f^{'}(x) = f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0},$$

this leads us other formulas for computing the derivatives

$$f^{'}(x) = \frac{f(x_1) - f(x_0)}{x_1 - x_0} + O(h), \quad \text{Euler,}$$

$$f^{'}(x) = \frac{f(x_1) - f(x_{-1})}{x_1 - x_{-1}} + O(h^2), \quad \text{Central.}$$

Using the same method we can get out computational estimates for the 2nd derivative

$$f^{''}(x_0) = \frac{f_2 - 2f_1 + f_0}{h^2} + O(h^2),$$

$$f^{''}(x_0) = \frac{f_1 - 2f_0 + f_{-1}}{h^2} + O(h^2), \quad \text{central.}$$

Example 3

To numerically solve the first order Ordinary Differential Equation (4)

$$y^{'} = f(x, y),$$

$$a \leq x \leq b,$$

the derivative y' is approximated by

$$\frac{w_{i+1} - w_i}{x_{i+1} - x_i} = \frac{w_{i+1} - w_i}{h},$$

where w_i is the numerical approximation of y at x_i. The Differential Equation is converted to a discrete difference equation with steps of size h,

$$\frac{w_{i+1} - w_i}{h} = f(x_i, w).$$

Rearranging the difference equation gives the equation

$$w_{i+1} = w_i + hf(x_i, w),$$

which can be used to approximate the solution at w_{i+1} given information about y at point x_i.

1.1.1.1 *Simple example ODE* $y' = \sin(x)$

Example 4

Applying the Euler formula to the first order equation with an oscillating input (3)

$$y' = \sin(x),$$

$$0 \leq x \leq 10.$$

The equation can be approximated using the forward Euler as

$$\frac{w_{i+1} - w_i}{h} = \sin(x_i).$$

Rearranging the equation gives the discrete difference equation with the unknowns on the left and the know values of the right

$$w_{i+1} = w_i + h\sin(x_i).$$

The Python code bellow implements this difference equation. The output of the code is shown in Figure 1.1.1.

```
# Numerical solution of a Cosine differential
    equation
import numpy as np
import math
import matplotlib.pyplot as plt

h=0.01
a=0
b=10

N=int(b-a/h)
w=np.zeros(N)
x=np.zeros(N)
Analytic_Solution=np.zeros(N)

# Initial Conditions
w[0]=1.0
x[0]=0
Analytic_Solution[0]=1.0
for i in range (1,N):
    w[i]=w[i-1]+h*math.sin(x[i-1])
    x[i]=x[i-1]+h
    Analytic_Solution[i]=2.0-math.cos(x[i])

fig = plt.figure(figsize=(8,4))

# --- left hand plot
ax = fig.add_subplot(1,3,1)
plt.plot(x,w,color='red')
#ax.legend(loc='best')
plt.title('Numerical Solution')
```

```python

# --- right hand plot
ax = fig.add_subplot(1,3,2)
plt.plot(x,Analytic_Solution,color='blue')
plt.title('Analytic Solution')

#ax.legend(loc='best')
ax = fig.add_subplot(1,3,3)
plt.plot(x,Analytic_Solution-w,color='blue')
plt.title('Error')

# --- title, explanatory text and save
fig.suptitle('Sine Solution', fontsize=20)
plt.tight_layout()
plt.subplots_adjust(top=0.85)
```

Listing 1.1: Python Numerical and Analytical Solution of Eqn 3

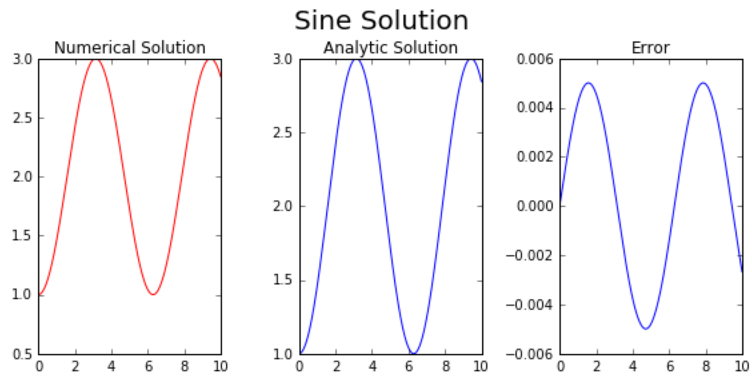


Figure 1.1.1: Python output: Numerical (left), Analytic (middle) and error(right) for $y' = \sin(x)$ Equation 3 with h=0.01

1.1.1.2 *Simple example problem population growth $y' = \varepsilon y$.*

Example 5

Simple population growth can be describe as a first order differential equation of the form:

$$y' = \varepsilon y. \tag{8}$$

This has an exact solution of

$$y = Ce^{\varepsilon x}.$$

Given the initial condition of condition

$$y(0) = 1$$

and a rate of change of

$$\varepsilon = 0.5$$

the analytic solution is

$$y = e^{0.5x}.$$

Example 6

Applying the Euler formula to the first order equation (8)

$$y' = 0.5y$$

is approximated by

$$\frac{w_{i+1} - w_i}{h} = 0.5w_i.$$

Rearranging the equation gives the difference equation

$$w_{i+1} = w_i + h(0.5w_i).$$

The Python code below and the output is plotted in Figure 1.1.2.

```
# Numerical solution of a differential equation
import numpy as np
import math
import matplotlib.pyplot as plt

h=0.01
tau=0.5
a=0
b=10

N=int((b-a)/h)
w=np.zeros(N)
x=np.zeros(N)
Analytic_Solution=np.zeros(N)

Numerical_Solution[0]=1
x[0]=0
w[0]=1

for i in range (1,N):
    w[i]=w[i-1]+dx*(tau)*w[i-1]
    x[i]=x[i-1]+dx
    Analytic_Solution[i]=math.exp(tau*x[i])


fig = plt.figure(figsize=(8,4))
# --- left hand plot
ax = fig.add_subplot(1,3,1)
plt.plot(x,w,color='red')
```

```
#ax.legend(loc='best')
plt.title('Numerical Solution')

# --- right hand plot
ax = fig.add_subplot(1,3,2)
plt.plot(x,Analytic_Solution,color='blue')
plt.title('Analytic Solution')

#ax.legend(loc='best')
ax = fig.add_subplot(1,3,3)
plt.plot(x,Analytic_Solution-Numerical_Solution,
    color='blue')
plt.title('Error')

# --- title, explanatory text and save
fig.suptitle('Exponential Growth Solution',
    fontsize=20)
plt.tight_layout()
plt.subplots_adjust(top=0.85)
```

Listing 1.2: Python Numerical and Analytical Solution of Eqn 8

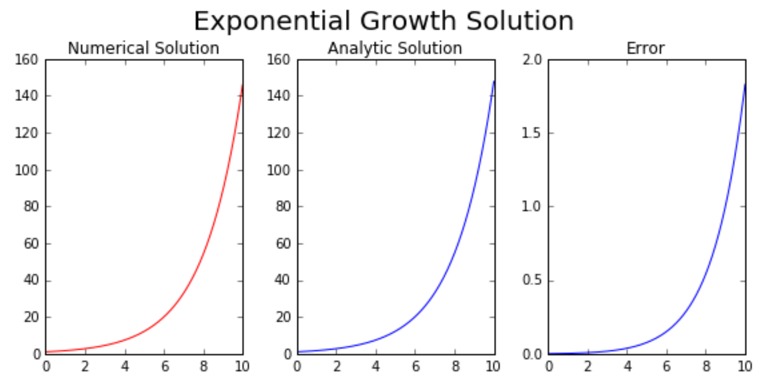


Figure 1.1.2: Python output: Numerical (left), Analytic (middle) and error(right) for $y' = \varepsilon y$ Eqn 8 with h=0.01 and $\varepsilon = 0.5$

1.1.1.3 *Example of exponential growth with a wiggle*

Example 7

An extension of the exponential growth differential equation includes a sinusoidal component

$$y' = \varepsilon(y + y\sin(x)). \tag{9}$$

This complicates the exact solution but the numerical approach is more or less the same. The difference equation is

$$w_{i+1} = w_i + h(0.5w_i + w_i\sin(x_i)).$$

Figure 1.1.3 illustrates the numerical solution of the differential equation.

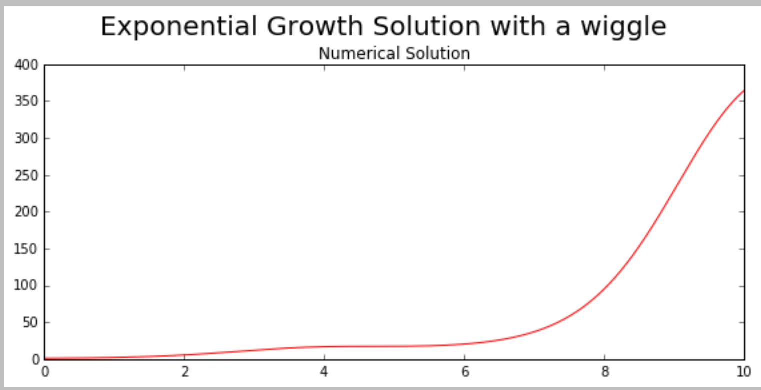


Figure 1.1.3: Python output: Numerical solution for $y' = \varepsilon(y + ysin(x))$ Equation 9 with h=0.01 and $\varepsilon = 0.5$

1.1.2 *Theorems about Ordinary Differential Equations*

Definition A function $f(t, y)$ is said to satisfy a **Lipschitz Condition** in the variable y on the set $D \subset R^2$ if a constant $L > 0$ exist with the property that

$$|f(t, y_1) - f(t, y_2)| < L|y_1 - y_2|,$$

whenever $(t, y_1), (t, y_2) \in D$. The constant L is call the Lipschitz Condition of f.

Definition A set $D \subset R^2$ is said to be convex if whenever $(t_1, y_1), (t_2, y_2)$ belong to D the point $((1 - \lambda)t_1 + \lambda t_2, (1 - \lambda)y_1 + \lambda y_2)$ also belongs in D for each $\lambda \in [0, 1]$.

Theorem 1.1.1. *Suppose $f(t, y)$ is defined on a convex set $D \subset R^2$. If a constant $L > 0$ exists with*

$$\left|\frac{\partial f(t, y)}{\partial y}\right| \leq L,$$

then f satisfies a Lipschitz Condition an D in the variable y with Lipschitz constant L.

Theorem 1.1.2. *Suppose that $D = \{(t, y)|a \leq t \leq b, -\infty < y < \infty\}$, and $f(t, y)$ is continuous on D in the variable y then the initial value problem has a unique solution $y(t)$ for $a \leq t \leq b$.*

Definition The initial-value problem

$$\frac{dy}{dt} = f(t, y), \qquad a \leq t \leq b,$$

with initial condition

$$y(a) = \alpha,$$

is said to be well posed if:

- A unique solution $y(t)$ to the problem exists;
- For any $\varepsilon > 0$ there exists a positive constant $k(\varepsilon)$ with the property that whenever $|\varepsilon_0| < \varepsilon$ and with $|\delta(t)|\varepsilon$ on $[a,b]$ a unique solution $z(t)$ to the problem

$$\frac{dz}{dt} = f(t,z) + \delta(t), \quad a \leq t \leq b, \tag{10}$$

$$z(a) = \alpha + \varepsilon_0,$$

exists with

$$|z(t) - y(t)| < k(\varepsilon)\varepsilon.$$

The problem specified by (10) is called a perturbed problem associated with the original problem.

It assumes the possibility of an error $\delta(\varepsilon)$ being introduced to the statement of the differential equation as well as an error ε_0 being present in the initial condition.

Theorem 1.1.3. *Suppose* $D = \{(t,y)|a \leq t \leq b, -\infty < y < \infty\}$. *If* $f(t,y)$ *is continuous and satisfies a Lipschitz Condition in the variable* y *on the set* D*, then the initial value problem*

$$\frac{dy}{dt} = f(t,y), \quad a \leq t \leq b,$$

with initial condition

$$y(a) = \alpha,$$

is well-posed.

Example 8

$$y'(x) = -y(x) + 1, \quad 0 \leq x \leq b, \quad y(0) = 1$$

has the solution $y(x) = 1$. The perturbed problem

$$z'(x) = -z(x) + 1, \quad 0 \leq x \leq b, \quad z(0) = 1 + \varepsilon,$$

has the solution $z(x) = 1 + \varepsilon e^{-x}$ $x \leq 0$.
Thus

$$y(x) - z(x) = -\varepsilon e^{-x}$$

$$|y(x) - z(x)| \leq |\varepsilon| \quad x \geq 0$$

Therefore the problem is said to be stable. This is illustrated in Figure 1.1.4.

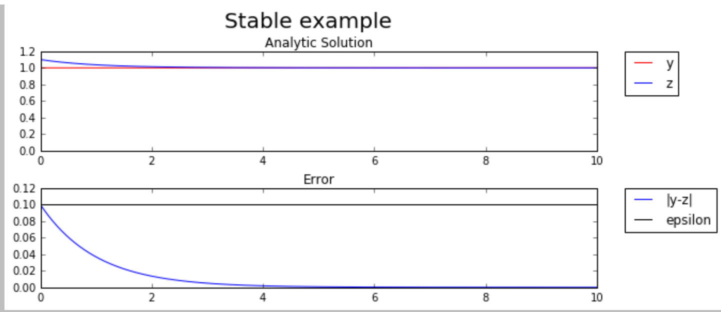


Figure 1.1.4: Python output: Illustrating Stability $y^{'}(x) = -y(x)+1$ with the initial condition $y(0) = 1$ and $z^{'}(x) = -z(x)+1$ with the initial condition $z(0) = 1+\varepsilon$, $\varepsilon = 0.1$

1.2 ONE-STEP METHODS

Dividing $[a,b]$ in to N subsections such that we now have N+1 points of equal spacing $h = \frac{b-a}{N}$. This gives the formula $t_i = a + ih$ for $i = 0, 1, ..., N$. One-Step Methods for Ordinary Differential Equation's only use one previous point to get the approximation for the next point. The initial condition gives $y(a = t_0) = \alpha$, this gives the starting point of our one step method. The general formula for One-step methods is

$$w_{i+1} = w_i + h\Phi(t_i, w_i, h),$$

where w_i is the approximated solution of the Ordinary Differential Equation at the point t_i

$$w_i \approx y_i.$$

1.2.1 *Euler's Method*

The simplest example of a one step method is Euler. The derivative is replaced by the Euler approximation. The Ordinary Differential Equation

$$\frac{dy}{dt} = f(t,y),$$

is discretised

$$\frac{y_i - y_{i-1}}{h} = f(t_{i-1}, y_{i-1}) + T$$

T is the truncation error.

Example 9

Consider the Initial Value Problem

$$y^{'} = -\frac{y^2}{1+t}, \quad a = 0 \leq t \leq b = 0.5,$$

with the initial condition $y(0) = 1$ the Euler approximation is

$$w_{i+1} = w_i - \frac{hw_i^2}{1+t_i},$$

where w_i is the approximation of y at t_i.
Solving, let $t_i = ih$ where $h = 0.05$, from the initial condition we have $w_0 = 1$ at $i = 0$ our method is

$$w_1 = w_0 - \frac{0.05w_0^2}{1+t_0} = 1 - \frac{0.05}{1+0} = 0.95$$

and so forth, each approximation w_i requiring the previous, thus creating a sequence starting at w_0 to w_n. The Table below show the numerical approximation for 10 steps.

i	t_i	w_i
0	0	1
1	0.05	0.95
2	0.1	0.90702381
3	0.15	0.86962871
4	0.2	0.8367481
5	0.25	0.80757529
6	0.3	0.78148818
7	0.35	0.7579988
8	0.4	0.73671872
9	0.45	0.71733463

Lemma 1.2.1. *For all $x \geq 0.1$ and any positive m we have*

$$0 \leq (1+x)^m \leq e^{mx}.$$

Lemma 1.2.2. *If s and t are positive real numbers $\{a_i\}_{i=0}^{N}$ is a sequence satisfying $a_0 \geq \frac{-t}{s}$ and $a_{i+1} \leq (1+s)a_i + t$ then,*

$$a_{i+1} \leq e^{(i+1)s}\left(a_0 + \frac{t}{s}\right) - \frac{t}{s}.$$

Theorem 1.2.3. *Suppose f is continuous and satisfies a Lipschitz Condition with constant L on $D = \{(t,y) | a \leq t \leq b, -\infty < y < \infty\}$ and that a constant M exists with the property that*

$$|y^{''}(t)| \leq M.$$

Let $y(t)$ denote the unique solution of the Initial Value Problem

$$y^{'} = f(t,y), \quad a \leq t \leq b, \quad y(a) = \alpha,$$

and $w_0, w_1, ..., w_N$ be the approx generated by the Euler method for some positive integer N. Then for $i = 0, 1, ..., N$

$$|y(t_i) - w_i| \leq \frac{Mh}{2L}|e^{L(t_i-a)} - 1|.$$

Proof. When $i = 0$ the result is clearly true since $y(t_0) = w_0 = \alpha$. From Taylor we have,

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2}y^{''}(\xi_i),$$

where $x_i \geq \xi_i \geq x_{i+1}$, and from this we get the Euler approximation

$$w_{i+1} = w_i + hf(t_i, w_i).$$

Consequently we have

$$y(t_{i+1}) - w_{i+1} = y(t_i) - w_i + h[f(t_i, y(t_i)) - f(t_i, w_i)] + \frac{h^2}{2}y^{''}(\xi_i),$$

and

$$|y(t_{i+1}) - w_{i+1}| \leq |y(t_i) - w_i| + h|f(t_i, y(t_i)) - f(t_i, w_i)| + \frac{h^2}{2}|y^{''}(\xi_i)|.$$

Since f satisfies a Lipschitz Condition in the second variable with constant L and $|y^{''}| \leq M$ we have

$$|y(t_{i+1}) - w_{i+1}| \leq (1 + hL)|y(t_i) - w_i| + \frac{h^2}{2}M.$$

Using Lemma 1.2.1 and 1.2.2 and letting $a_j = (y_j - w_j)$ for each $j = 0, .., N$ while $s = hL$ and $t = \frac{h^2M}{2}$ we see that

$$|y(t_{i+1} - w_{i+1}| \leq e^{(i+1)hL}(|y(t_0) - w_0| + \frac{h^2M}{2hL}) - \frac{h^2M}{2hL}.$$

Since $w_0 - y_0 = 0$ and $(i+1)h = t_{i+1} - t_0 = t_{i+1} - a$ we have

$$|y(t_i) - w_i| \leq \frac{Mh}{2L}|e^{L(t_i-a)} - 1|,$$

for each $i = 0, 1, ..N - 1$. □

Example 10

$$y^{'} = y - t^2 + 1, \quad 0 \leq t \leq 2, \quad y(0) = 0.5,$$

the Euler approximation is

$$w_{i+1} = w_i + h(w_i - t_i^2 + 1)$$

choosing $h = 0.2$, $t_i = 0.2i$ and $w_0 = 0.5$.
$f(t,y) = y - t^2 + 1$

$$\frac{\partial f}{\partial y} = 1$$

so $L = 1$. The exact solution is $y(t) = (t+1)^2 - \frac{1}{2}e^t$ from this we have

$$y''(t) = 2 - 0.5e^t,$$

$$|y''(t)| \leq 0.5e^2 - 2, \quad t \in [0,2].$$

Using the above inequality we have we have

$$|y_i - w_i| \leq \frac{h}{2}(0.5e^2 - 2)(e^{t_i} - 1).$$

Figure 1.2.1 illustrates the upper bound of the error and the actual error.

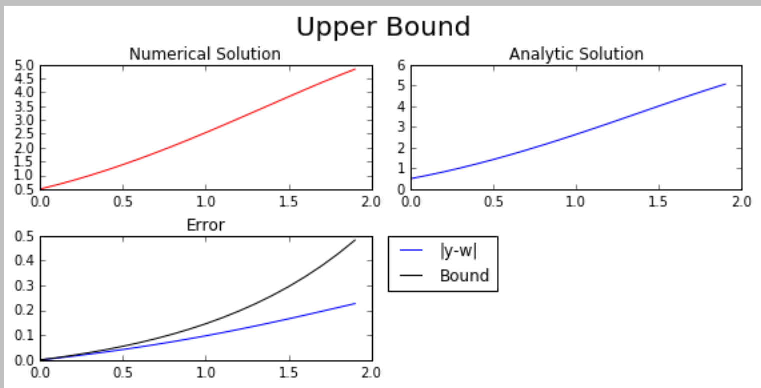


Figure 1.2.1: Python output: Illustrating upper bound $y' = y - t^2 + 1$ with the initial condition $y(0) = 0.5$

Euler method is a typical one step method, in general such methods are given by function $\Phi(t, y; h; f)$. Our initial condition is $w_0 = y_0$, for $i = 0, 1, ..$

$$w_{i+1} = w_i + h\Phi(t_i, w_i : h : f)$$

with $t_{i+1} = t_i + h$.
In the Euler case $\Phi(t, y; h; f) = f(t, y)$ and is of order 1.
Theorem 1.2.3 can be extend to higher order one step methods with the variation

$$|y(t_i) - w_i| \leq \frac{Mh^p}{2L}|e^{L(t_i - a)} - 1|$$

where p is the order of the method.

Definition The difference method $w_0 = \alpha$

$$w_{i+1} = w_i + h\Phi(t_i, w_i),$$

for $i = 0, 1, ..., N-1$ has a local truncation error given by

$$\begin{aligned}\tau_{i+1}(h) &= \frac{y_{i+1} - (y_i + h\Phi(t_i, y_i))}{h}, \\ &= \frac{y_{i+1} - y_i}{h} - \Phi(t_i, y_i),\end{aligned}$$

for each $i = 0, .., N-1$ where as usual $y_i = y(t_i)$ denotes the exact solution at t_i.

For Euler method the local truncation error at the ith step for the problem

$$y' = f(t, y), \quad a \le t \le b, \;\; y(a) = \alpha,$$

is

$$\tau_{i+1}(h) = \frac{y_{i+1} - y_i}{h} - f(t_i, y_i),$$

for $i = 0, .., N-1$.
But we know Euler has

$$\tau_{i+1} = \frac{h}{2} y''(\xi_i), \;\; \xi_i \in (t_i, t_{i+1}),$$

When $y''(t)$ is known to be bounded by a constant M on $[a, b]$ this implies

$$|t_{i+1}(h)| \le \frac{h}{2} M \sim O(h).$$

$O(h)$ indicates a linear order of error. The higher the order the more accurate the method.

1.3 PROBLEM SHEET

1. Show that the following functions satisfy the Lipschitz condition on y on the indicated set D:

   a) $f(t,y) = ty^3$, $D = \{(t,y); -1 \leq t \leq 1, 0 \leq y \leq 10\}$;

   b) $f(t,y) = \frac{t^2y^2}{1+t^2}$, $D = \{(t,y); 0 \leq t, -10 \leq y \leq 10\}$.

2. Apply Euler's Method to approximate the solution of the given initial value problems using the indicated number of time steps. Compare the approximate solution with the given exact solution, and compate the actual error with the theoretical error

   a) $y' = t - y, \quad (0 \leq t \leq 4)$
   with the initial condition $y(0) = 1$,
   $N = 4$, $y(t) = 2e^{-t} + t - 1$,

      The Lipschitz constant is determined on $D = \{(t,y); 0 \leq t \leq 4, y \in \mathbb{R}\}$.

   b) $y' = y - t, \quad (0 \leq t \leq 2)$
   with the initial condition $y(0) = 2$,
   $N = 4$, $y(t) = e^t + t + 1$.

      The Lipschitz constant is determined on $D = \{(t,y); 0 \leq t \leq 2, y \in \mathbb{R}\}$.

2

HIGHER ORDER METHODS

2.1 HIGHER ORDER TAYLOR METHODS

The Taylor expansion

$$y(t_{i+1}) = y(t_i) + hy^{'}(t_i) + \frac{h^2}{2}y^{'}(t_i) + .. + \frac{h^{n+1}}{(n+1)!}y^{n+1}(\xi_i)$$

can be used to design more accurate higher order methods. By differentiating the original Ordinary Differential Equation $y^{'} = f(t, y)$ higher ordered can be derived method it requires the function to be continuous and differentiable.
In the general case of Taylor of order n:

$$w_0 = \alpha$$

$$w_{i+1} = w_i + hT^n(t_i, w_i), \text{ for } i = 0, ..., N-1,$$

where

$$T^n(t_i, w_i) = f(t_i, w_i) + \frac{h}{2}f'(t_i, w_i) + ...\frac{h^{n-1}}{n!}f^{n-1}(t_i, w_i). \quad (11)$$

Example 11

Applying the general Taylor method to create methods of order two and four to the initial value problem

$$y^{'} = y - t^2 + 1, \quad 0 \leq t \leq 2, \quad y(0) = 0.5,$$

from this we have

$$f^{'}(t, y(t)) = \frac{d}{dt}(y - t^2 + 1) = y' - 2t = y - t^2 + 1 - 2t,$$

$$f^{''}(t, y(t)) = y - t^2 - 2t - 1,$$
$$f^{'''}(t, y(t)) = y - t^2 - 2t - 1.$$

From these derivatives we have

$$
\begin{aligned}
T^2(t_i, w_i) &= f(t_i, w_i) + \frac{h}{2} f'(t_i, w_i) \\
&= w_i - t_i^2 + 1 + \frac{h}{2}(w_i - t_i^2 - 2t_i + 1) \\
&= \left(1 + \frac{h}{2}\right)(w_i - t_i^2 - 2t_i + 1) - ht_i
\end{aligned}
$$

and

$$
\begin{aligned}
T^4(t_i, w_i) &= f(t_i, w_i) + \frac{h}{2} f'(t_i, w_i) \\
&\quad + \frac{h^2}{6} f''(t_i, w_i) + \frac{h^3}{24} f''(t_i, w_i) \\
&= \left(1 + \frac{h}{2} + \frac{h^2}{6} + \frac{h^3}{24}\right)(w_i - t_i^2) \\
&\quad - \left(1 + \frac{h}{3} + \frac{h^2}{12}\right) ht_i \\
&\quad + 1 + \frac{h}{2} - \frac{h^2}{6} - \frac{h^3}{24}
\end{aligned}
$$

From these equations we have, Taylor of order two

$$w_0 = 0.5$$

$$w_{i+1} = w_i + h\left[\left(1 + \frac{h}{2}\right)(w_i - t_i^2 - 2t_i + 1) - ht_i\right]$$

and Taylor of order 4

$$
\begin{aligned}
w_{i+1} &= w_i + h\Bigg[\left(1 + \frac{h}{2} + \frac{h^2}{6} + \frac{h^3}{24}\right)(w_i - t_i^2) \\
&\quad - \left(1 + \frac{h}{3} + \frac{h^2}{12}\right) ht_i + 1 + \frac{h}{2} - \frac{h^2}{6} - \frac{h^3}{24}\Bigg]
\end{aligned}
$$

The local truncation error for the 2nd order method is

$$\tau_{i+1}(h) = \frac{y_{i+1} - y_i}{h} - T^2(t_i, y_i) = \frac{h^2}{6} f^2(\xi_i, y(x_i))$$

where $\xi \in (t_i, t_{i+1})$.

In general if $y \in C^{n+1}[a, b]$

$$\tau_{i+1}(h) = \frac{h^n}{(n+1)!} f^n(\xi_i, y(\xi_i))\ O(h^n).$$

The issue is that for every differential equation a new method has be to derived.

3

RUNGE–KUTTA METHOD

The Runge-Kutta method (RK) method is closely related to the Taylor series expansions but no differentiation of f is necessary.
All RK methods will be written in the form

$$w_{n+1} = w_n + hF(t, w, h; f), \quad n \geq 0. \tag{12}$$

The truncation error for (12) is defined by

$$T_n(y) = y(t_{n+1}) - y(t_n) - hF(t_n, y(t_n), h; f)$$

where the error is written as $\tau_n(y)$

$$T_n = h\tau_n(y).$$

Rearranging we get

$$y(t_{n+1}) = y(t_n) - hF(t_n, y(t_n), h; f) + h\tau_n(y).$$

Theorem 3.0.1. *Suppose f(t,y) and all its partial derivatives of order less than or equal to n+1 are continuous on* $D = \{(t, y)|a \leq t \leq b, c \leq y \leq d\}$ *and let* $(t_0, y_0) \in D$ *for every* $(t, y) \in D$, $\exists\ \xi \in (t, t_0)$ *and* $\mu \in (y, y_0)$ *with*

$$f(t, y) = P_n(t, y) + R_n(t, y)$$

where

$$\begin{aligned}
P_n(t,y) \quad = \quad & f(t_0,y_0) + \left[(t-t_0)\frac{\partial f}{\partial t}(t_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(t_0,y_0)\right] \\
& + \left[\frac{(t-t_0)^2}{2}\frac{\partial^2 f}{\partial t^2}(t_0,y_0) + (y-y_0)(t-t_0)\frac{\partial^2 f}{\partial y \partial t}(t_0,y_0)\right. \\
& \left. + \frac{(y-y_0)^2}{2}\frac{\partial^2 f}{\partial y^2}(t_0,y_0)\right] \\
& + \ldots + \\
& + \left[\frac{1}{n!}\sum_{j=0}^{n}\left(\begin{array}{c} n \\ j \end{array}\right)(t-t_0)^{n-j}(y-y_0)^j\frac{\partial^n f}{\partial y^j \partial t^{n-j}}(t_0,y_0)\right]
\end{aligned}$$

and

$$R_n(t,y) = \left[\frac{1}{(n+1)!}\sum_{j=0}^{n+1}\left(\begin{array}{c} n+1 \\ j \end{array}\right)(t-t_0)^{n+1-j}(y-y_0)^j\frac{\partial^{n+1}f}{\partial y^j\partial t^{n+1-j}}(\xi,\mu)\right]$$

3.1 DERIVATION OF SECOND ORDER RUNGE KUTTA

Consider the explicit one-step method

$$\frac{w_{i+1}-w_i}{h} = F(f,t_i,w_i,h) \tag{13}$$

with

$$F(f,t,y,h) = a_0k_1 + a_1k_2, \tag{14}$$

$$F(f,t,y,h) = a_0f(t,y) + a_1f(t+\alpha_1, y+\beta_1), \tag{15}$$

where $a_0 + a_1 = 1$.

There is a free parameter in the derivation of the Runge Kutta method for this reason a_0 must be choosen

Deriving the second order Runge-Kutta method by using Theorem 3.0.1 to determine values for values a_1, α_1 and β_1 with the property that $a_1f(t+\alpha_1, y+\beta_1)$ approximates the second order Taylor

$$f(t,y) + \frac{h}{2}f^{'}(t,y)$$

with error no greater than than $O(h^2)$, the local truncation error for the Taylor method of order two.
Using

$$f^{'}(t,y) = \frac{\partial f}{\partial t}(y,t) + \frac{\partial f}{\partial y}(t,y).y^{'}(t),$$

the second order Taylor can be re-written as

$$f(t,y) + \frac{h}{2}\frac{\partial f}{\partial t}(y,t) + \frac{h}{2}\frac{\partial f}{\partial y}(t,y).f(t,y). \tag{16}$$

Expanding $a_1f(t+\alpha_1, y+\beta_1)$ in its Taylor polynomial of degree one about (t,y) gives

$$a_1f(t+\alpha_1,y+\beta_1) = a_1f(t,y) + a_1\alpha_1\frac{\partial f}{\partial t}(t,y) + a_1\beta_1\frac{\partial f}{\partial y} + a_1R_1(t+\alpha_1,y+\beta_1) \tag{17}$$

where

$$R_1(t+\alpha_1,y+\beta_1) = \frac{\alpha_1^2}{2}\frac{\partial^2 f}{\partial t^2}(\xi,\mu) + \alpha_1\beta_1\frac{\partial^2 f}{\partial t\partial y}(\xi,\mu) + \frac{\beta_1^2}{2}\frac{\partial^2 f}{\partial y^2}(\xi,\mu),$$

for some $\xi \in [t, t+\alpha_1]$ and $\mu \in [y, y+\beta_1]$.
Matching the coefficients and its derivatives in eqns (16) and (17) gives the equations

$$f(t,y) : a_1 = 1$$

$$\frac{\partial f}{\partial t}(t,y) : a_1\alpha_1 = \frac{h}{2}$$

and

$$\frac{\partial f}{\partial y}(t,y) : a_1\beta_1 = \frac{h}{2}f(t,y).$$

3.1.1 *Runge Kutta second order: Midpoint method*

Choosing $a_0 = 0$ gives the unique values $a_1 = 1$, $\alpha_1 = \frac{h}{2}$ and $\beta_1 = \frac{h}{2}f(t,y)$ so

$$T^2(t,y) = f(t+\frac{h}{2}, y+\frac{h}{2}f(t,y)) - R_1(t+\frac{h}{2}, y+\frac{h}{2}f(t,y))$$

and from

$$R_1(t+\frac{h}{2}, y+\frac{h}{2}f(t,y)) = \frac{h^2}{8}\frac{\partial^2 f}{\partial t^2}(\xi,\mu) + \frac{h^2}{4}\frac{\partial^2 f}{\partial t \partial y}(\xi,\mu) + \frac{h^2}{8}g(t,y)^2\frac{\partial^2 f}{\partial y^2}(\xi,\mu),$$

for some $\xi \in [t, t+\frac{h}{2}]$ and $\mu \in [y, y+\frac{h}{2}f(t,y)]$.
If all the second-order partial derivatives are bounded then

$$R_1(t+\frac{h}{2}, y+\frac{h}{2}f(t,y)) \sim O(h^2).$$

The Midpoint second order Runge-Kutta for the initial value problem

$$y' = f(t,y)$$

with the initial condition $y(t_0) = \alpha$ is given by

$$w_0 = \alpha,$$

$$w_{i+1} = w_i + hf(t_i + \frac{h}{2}, y_i + \frac{h}{2}f(t_i, w_i)),$$

with an error of order $O(h^2)$. The Figure 3.3.2 illustrates the solution to the $y^{'} = -xy$

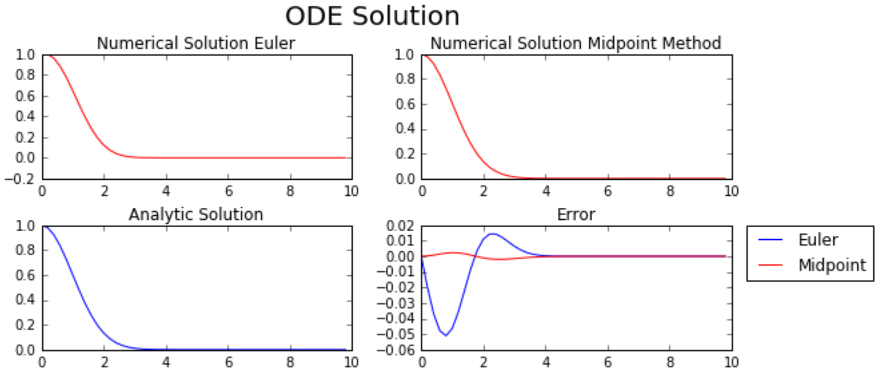


Figure 3.1.1: Python output: Illustrating upper bound $y' = -xy$ with the initial condition $y(0) = 1$

for each $i = 0, 1, ...N - 1$.

3.1.2 *2nd Order Runge Kutta $a_0 = 0.5$: Heun's method*

Choosing $a_0 = 0.5$ gives the unique values $a_1 = 0.5$, $\alpha_1 = h$ and $\beta_1 = hf(t, y)$ such that

$$T^2(t, y) = F(t, y) = 0.5f(t, y) + 0.5f(t + h, y + hf(t, y)) - R_1(t + h, y + hf(t, y))$$

and the error value from

$$R_1(t + h, y + hf(t, y)) = \frac{h^2}{2}\frac{\partial^2 f}{\partial t^2}(\xi, \mu) + h^2\frac{\partial^2 f}{\partial t \partial y}(\xi, \mu) + \frac{h^2}{2}f(t, y)^2\frac{\partial^2 f}{\partial y^2}(\xi, \mu),$$

for some $\xi \in [t, t + h]$ and $\mu \in [y, y + hf(t, y)]$.

Thus Heun's second order Runge-Kutta for the initial value problem

$$y' = f(t, y)$$

with the initial condition $y(t_0) = \alpha$ is given by

$$w_0 = \alpha,$$

$$w_{i+1} = w_i + \frac{h}{2}[f(t_i, w_i) + f(t_i + h, y_i + hf(t_i, w_i))]$$

with an error of order $O(h^2)$.
For ease of calculation this can be rewritten as:

$$k_1 = f(t_i, w_i),$$

$$k_2 = f(t_i + h, w_i + hk_1),$$

$$w_{i+1} = w_i + \frac{h}{2}[k_1 + k_2].$$

3.2 THIRD ORDER RUNGE KUTTA METHODS

Higher order methods are derived in a similar fashion. For the Third Order Runge Kutta methods

$$\frac{w_{i+1} - w_i}{h} = F(f, t_i, w_i, h), \tag{18}$$

with

$$F(f, t, w, h) = a_0 k_1 + a_1 k_2 + a_2 k_3, \tag{19}$$

where

$$a_0 + a_1 + a_2 = 1$$

and

$$k_1 = f(t_i, w_i),$$
$$k_2 = f(t_i + \alpha_1 h, t_i + \beta_{11} k_1),$$
$$k_3 = f(t_i + \alpha_2 h, t_i + \beta_{21} k_1 + \beta_{22} k_2)).$$

The values of a_0, a_1, a_2, α_1, α_2, β_{11},β_{21}, β_{22} are derived by group the Taylor expansion,

$$\begin{aligned} y_{i+1} &= y_i + hf(t_i, y_i) + \frac{h^2}{2}(f_t + f_y f)_{(t_i,y_i)} \\ &+\frac{h^3}{6}\left(f_{tt} + 2f_{ty}f + f_t f_y + f_{yy}f^2 + f_y f_y f\right)_{(t_i,y_i)} \\ &+O(h^4), \end{aligned}$$

with the 3rd order expand form:

$$y_{i+1} = y_i + ha_1 f(t_i, y_i) + ha_2(f + \alpha_1 h f_t + \beta_{11} h f_y f$$
$$+\frac{h^2}{2}(f_{tt}\alpha_1^2 + f_{yy}\beta_{11}^2 f^2 + 2f_{ty}\alpha_1\beta_{11}f) + O_2(h^3))$$
$$+ha_3(f + \alpha_2 h f_t + f_y\left(\beta_{21}hf + \beta_{22}h(f + \alpha_1 h f_t + \beta_{11}hf_y f + O_3(h^2))\right)$$
$$+\frac{1}{2}(f_{tt}(\alpha_2 h)^2 + f_{yy}h^2(\beta_{21}f + \beta_{22}(f + \underbrace{\alpha_1 h f_t + \beta_{11} h f_y f + O_4(h^2)}_{O_5(h)}))^2$$
$$+2f_{ty}\alpha_2 h^2(\beta_{21}f + \beta_{22}(f + \underbrace{\alpha_1 h f_t + \beta_{11} h f_y f + O_4(h^2)}_{O_5(h)}))))).$$

This results in 8 equations with 8 unknowns, but only 6 of these equations are independent. For this reason the are two free parameters to choose.

For example, we can choose that

$$\alpha_2 = 1, \beta_{11} = \frac{1}{2},$$

then we obtain the following difference equation.

$$w_{i+1} = w_i + \frac{h}{6}(k_1 + 4k_2 + k_3),$$

where

$$k_1 = f(t_i, w_i),$$
$$k_2 = f(t_n + 1/2h, w_n + 1/2hk_1),$$
$$k_3 = f(t_n + h, w_n - hk_1 + 2hk_2).$$

3.3 Runge Kutta fourth order

$$w_0 = \alpha,$$
$$k_1 = hf(t_i, w_i),$$
$$k_2 = hf(t_i + \frac{h}{2}, w_i + \frac{1}{2}k_1),$$
$$k_3 = hf(t_i + \frac{h}{2}, w_i + \frac{1}{2}k_2),$$
$$k_4 = hf(t_{i+1}, w_i + k_3),$$
$$w_{i+1} = w_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

Example 12

Example Midpoint method,

$$y^{'} = y - t^2 + 1, \quad 0 \leq t \leq 2, \quad y(0) = 0.5,$$

$$N = 10, \quad t_i = 0.2i, \quad h = 0.2,$$

$$w_0 = 0.5,$$

$$\begin{aligned} w_{i+1} &= w_i + 0.2f(t_i + \frac{0.2}{2}, w_i + \frac{0.2}{2}f(t_i, w_i)) \\ &= w_i + 0.2f(t_i + 0.1, w_i + 0.1(w_i - t_i^2 + 1)) \\ &= w_i + 0.2(w_i + 0.1(w_i - t_i^2 + 1) - (t_i + 0.1)^2 + 1) \end{aligned}$$

Example 13

Example Runge Kutta fourth order method

$$y^{'} = y - t^2 + 1, \quad 0 \leq t \leq 2, \quad y(0) = 0.5,$$

$$N = 10, \quad t_i = 0.2i, \quad h = 0.2,$$

$$w_0 = 0.5,$$

$$\begin{aligned}
k_1 &= h(w_i - t_i^2 + 1), \\
k_2 &= h(w_i + \frac{1}{2}k_1 - (t_i + \frac{h}{2})^2 + 1) \\
k_3 &= h(w_i + \frac{1}{2}k_2 - (t_i + \frac{h}{2})^2 + 1) \\
k_4 &= h(w_i + \frac{1}{2}k_3 - (t_i + h)^2 + 1) \\
w_{i+1} &= w_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)
\end{aligned}$$

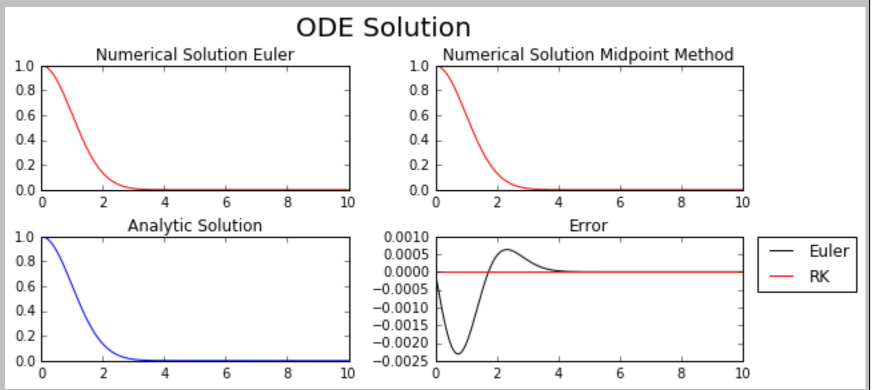


Figure 3.3.1: Python output: Illustrating upper bound $y' = -xy$ with the initial condition $y(0) = 1$

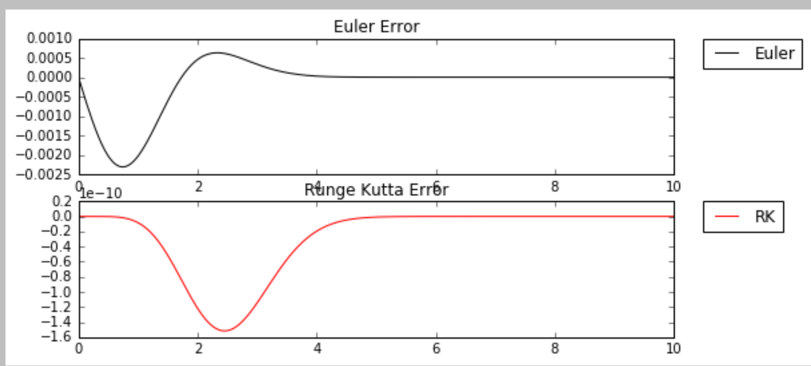


Figure 3.3.2: Python output: Illustrating upper bound $y' = -xy$ with the initial condition $y(0) = 1$

3.4 BUTCHER TABLEAU

Another way of representing a Runge Kutta method is called the Butcher tableau named after John C Butcher (31 March 1933).

$$y_{i+1} = y_i + h \sum_{n=1}^{s} a_n k_n,$$

where

$$\begin{aligned}
k_1 &= f(t_i, y_i), \\
k_2 &= f(t_i + \alpha_2 h, y_i + h(\beta_{21} k_1)), \\
k_3 &= f(t_i + \alpha_3 h, y_i + h(\beta_{31} k_1 + \beta_{32} k_2)), \\
&\vdots \\
k_s &= f(t_i + \alpha_s h, y_i + h(\beta_{s1} k_1 + \beta_{s2} k_2 + \cdots + \beta_{s,s-1} k_{s-1})).
\end{aligned}$$

These data are usually arranged in a mnemonic device, known as a Butcher tableau

$$\begin{array}{c|ccccc}
0 & & & & & \\
\alpha_2 & \beta_{21} & & & & \\
\alpha_3 & \beta_{31} & \beta_{31} & & & \\
\vdots & \vdots & \vdots & & & \\
\alpha_s & \beta_{s1} & \beta_{s2} & \cdots & \beta_{ss-1} & \\
\hline
 & a_1 & a_2 & \cdots & a_{s-1} & a_s
\end{array}$$

The method is consistent if

$$\sum_{j}^{s-1} \beta_{sj} = \alpha_s.$$

3.4.1 *Heun's Method*

The Butcher's Tableau for Heun's Method is:

$$\begin{array}{c|cc}
0 & & \\
1 & 1 & \\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}$$

3.4.2 *4th Order Runge Kutta*

The Butcher's Tableau for the 4th Order Runge Kutta is:

$$\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{2} & \frac{1}{2} & & & \\
\frac{1}{2} & 0 & \frac{1}{2} & & \\
1 & 0 & 0 & 1 & \\
\hline
 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6}
\end{array}$$

3.5 CONVERGENCE ANALYSIS

In order to obtain convergence of the general Runge Kutta we need to have the truncation $\tau_n(y) \to 0$ as $h \to 0$. Since,

$$\tau(y) = \frac{y(t_{n+1}) - y(t_n)}{h} - F(t_n, y(t_n), h; f),$$

we require,

$$F(x, y, h; f) \to y'(x) = f(x, y(x)).$$

More precisely define,

$$\delta(h) = \max_{a \le t \le b; -\infty < y < \infty} |f(t, y) - F(t, y, h; f)|,$$

and assume,

$$\delta(h) \to 0, \quad \text{as } h \to 0. \tag{20}$$

This is called the consistency condition for the RK method.
We will also need a Lipschitz Condition on F:

$$|F(t, y, h; f) - F(t, z, h; f)| \le |y - z|, \tag{21}$$

for all $a \le t \le b$ and $-\infty < y, z < \infty$ and small $h > 0$.

Example 14

Looking at the midpoint method

$$\begin{aligned}
|F(t, w, h; f) - F(t, z, h; f)| &= \Big| f(t + \frac{h}{2}, w + \frac{h}{2} f(t, w)) \\
&\quad - f(t + \frac{h}{2}, z + \frac{h}{2} f(t, z)) \Big| \\
&\le K \left| w - z + \frac{h}{2} [f(t, w) - f(t, z)] \right| \\
&\le K \left(1 + \frac{h}{2} K\right) |w - z|
\end{aligned}$$

Theorem 3.5.1. *Assume that the Runge Kutta method satisfies the Lipschitz Condition. Then for the initial value problems*

$$y' = f(x, y),$$

$$y(x_0) = y_0.$$

The numerical solution $\{w_n\}$ satisfies

$$\max_{a \le x \le b} |y(x_n) - w_n| \le e^{(b-a)L} |y_0 - w_0| + \left[\frac{e^{(b-a)L} - 1}{L}\right] \tau(h)$$

where

$$\tau(h) = \max_{a \leq x \leq b} |\tau_n(h)|,$$

If the consistency condition

$$\delta(h) \to 0 \text{ as } h \to 0,$$

where

$$\delta(h) = \max_{a \leq x \leq b} |f(x,y) - F(x,y;h;f)|.$$

Proof. Subtracting

$$w_{n+1} = w_n + hF(t_n, w_n, h; f),$$

and

$$y(t_{n+1}) = y(t_n) + hF(t_n, y(t_n), h; f) + h\tau_n(h),$$

we obtain

$$e_{n+1} = e_n + h[F(t_n, w_n, h; f) - F(t_n, w_n, h; f)] + h\tau_n(h),$$

in which $e_n = y(t_n) - w_n$. Apply the Lipschitz Condition L and the truncation error we obtain

$$|e_{n+1}| \leq (1 + hL)|e_n| + h\tau_n(h).$$

This nicely leads to the result.
In most cases it is known by direct computation that $\tau(h) \to 0$ as $h \to 0$ an in that case convergence of $\{w_n\}$ and $y(t_n)$ is immediately proved.
But all we need to know is that (20) is satisfied . To see this we write

$$\begin{array}{rcl} h\tau_n & = & y(t_{n+1}) - y(t_n) - hF(t_n, y(t_n), h; f), \\ & = & hy^{'}(t_n) + \frac{h^2}{2}y^{''}(\xi_n) - hF(t_n, y(t_n), h; f), \\ h|\tau_n| & \leq & h\delta(h) + \frac{h^2}{2}|y^{''}|. \\ |\tau_n| & \leq & \delta(h) + \frac{h}{2}|y^{''}|. \end{array}$$

Thus $\tau(h) \to 0$ as $h \to 0$ □

From this we have

Corollary 3.5.2. *If the RK method has a truncation error $\tau(h) = O(h^{m+1})$ then the rate of convergence of $\{w_n\}$ to $Y(t)$ is $O(h^m)$.*

3.6 THE CHOICE OF METHOD AND STEP-SIZE

An interesting question is since Runge-Kutta method is 4th order but requires 4 steps and Euler only required 3 is it more beneficial to use a smaller h than a higher order method?
But this does lead us to the question of how do we define our h to

maximize the solution we have.
An ideal difference-equation method

$$w_{i+1} = w_i + h\phi(t_i, w_i, h) \quad i = 0, .., n-1$$

for approximating the solution $y(t)$ to the Initial Value Problem $y' = f(t, y)$ would have the property that given a tolerance $\varepsilon > 0$ the minimal number of mesh points would be used to ensure that the global error $|y(t_i) - w_i|$ would not exceed ε for any $i = 0, ..., N$.
We do this by finding an appropriate choice of mesh points. Although we cannot generally determine the global error of a method there is a close relation between local truncation and global error. By using methods of differing order we can predict the local truncation error and using this prediction choose a step size that will keep global error in check.
Suppose we have two techniques

1. An nth order Taylor method of the form

$$y(t_{i+1}) = y(t_i) + h\phi(t_i, y(t_i), h_i) + O(h^{n+1})$$

producing approximations

$$w_0 = \alpha$$

$$w_{i+1} = w_i + h\phi(t_i, w_i, h_i)$$

with local truncation $\tau_{i+1} = O(h^n)$.

2. An (n+1)st order Taylor of the form

$$y(t_{i+1}) = y(t_i) + h\psi(t_i, y(t_i), h_i) + O(h^{n+2})$$

producing approximations

$$v_0 = \alpha$$

$$v_{i+1} = v_i + h\psi(t_i, v_i, h_i)$$

with local truncation $v_{i+1} = O(h^{n+1})$.

We first make the assumption that $w_i \approx y(t_i) \approx v_i$ and choose a fixed step size to generate w_{i+1} and v_{i+1} to approximate $y(t_{i+1})$. Then

$$\begin{aligned}
\tau_{i+1} &= \frac{y(t_{i+1}) - y(t_i)}{h} - \phi(t_i, y(t_i), h) \\
&= \frac{y(t_{i+1}) - w_i}{h} - \phi(t_i, w_i, h) \\
&= \frac{y(t_{i+1}) - (w_i + h\phi(t_i, w_i, h))}{h} \\
&= \frac{y(t_{i+1}) - w_{i+1}}{h}
\end{aligned}$$

Similarly

$$Y_{i+1} = \frac{y(t_{i+1}) - v_{i+1}}{h}$$

As a consequence

$$\begin{aligned}
\tau_{i+1} &= \frac{y(t_{i+1}) - w_{i+1}}{h} \\
&= \frac{(y(t_{i+1}) - v_{i+1}) + (v_{i+1} - w_{i+1})}{h} \\
&= Y_{i+1}(h) + \frac{(v_{i+1} - w_{i+1})}{h}.
\end{aligned}$$

but $\tau_{i+1}(h)$ is $O(h^n)$ and $Y_{i+1}(h)$ is $O(h^{n+1})$ so the significant factor of $\tau_{i+1}(h)$ must come from $\frac{(v_{i+1}-w_{i+1})}{h}$. This gives us an easily computed approximation of $O(h^n)$ method,

$$\tau_{i+1} \approx \frac{(v_{i+1} - w_{i+1})}{h}.$$

The object is not to estimate the local truncation error but to adjust step size to keep it within a specified bound. To do this we assume that since $\tau_{i+1}(h)$ is $O(h^n)$ a number K independent of h exists with,

$$\tau_{i+1}(h) \approx Kh^n.$$

Then the local truncation error produced by applying the nth order method with a new step size qh can be estimated using the original approximations w_{i+1} and v_{i+1}

$$\tau_{i+1}(qh) \approx K(qh)^n \approx q^n \tau_{i+1}(h) \approx \frac{q^n}{h}(v_{i+1} - w_{i+1}),$$

to bound $\tau_{i+1}(qh)$ by ε we choose q such that

$$\frac{q^n}{h}|v_{i+1} - w_{i+1}| \approx \tau_{i+1}(qh) \leq \varepsilon,$$

which leads to

$$q \leq \left(\frac{\varepsilon h}{|v_{i+1} - w_{i+1}|}\right)^{\frac{1}{n}},$$

which can be used to control the error.

3.7 PROBLEM SHEET 2

1. Apply the Taylor method to approximate the solution of initial value problem

$$y' = ty + ty^2, \quad (0 \leq t \leq 2), \quad y(0) = \frac{1}{2}$$

using $N = 4$ steps.

2. Apply the Midpoint Method to approximate the solution of the given initial value problems using the indicated number of time steps. Compare the approximate solution with the given exact solution

   a) $y' = t - y, \ \ (0 \leq t \leq 4)$,
   with the initial condition $y(0) = 1$,
   $N = 4$, with the exact solution $y(t) = 2e^{-t} + t - 1$.

   b) $y' = y - t, \ \ (0 \leq t \leq 2)$,
   with the initial condition $y(0) = 2$,
   $N = 4$, with the exact solution $y(t) = e^{t} + t + 1$.

3. Apply the 4th Order Runge Kutta Method to approximate the solution of the given initial value problems using the indicated number of time steps. Compare the approximate solution with the given exact solution

   a) $y' = t - y, \ \ (0 \leq t \leq 4)$,
   with the initial condition $y(0) = 1$,
   $N = 4$, with the exact solution $y(t) = 2e^{-t} + t - 1$.

   b) $y' = y - t, \ \ (0 \leq t \leq 2)$
   with the initial condition $y(0) = 2$,
   $N = 4$, with the exact solution $y(t) = e^{t} + t + 1$.

4. Derive the difference equation for the Midpoint Runge Kutta method

$$w_{n+1} = w_n + k_2,$$
$$k_1 = hf(t_n, w_n),$$
$$k_2 = hf(t_n + \frac{1}{2}h, w_n + \frac{1}{2}k_1)$$

for dolving the ordinary differential equation

$$\frac{dy}{dt} = f(t, y),$$

$$y(t_0) = y_0,$$

by using a formula of the form

$$w_{n+1} = w_n + ak_1 + bk_2,$$

where k_1 is defined as above,

$$k_2 = hf(t_n + \alpha h, w_n + \beta k_1),$$

and a, b, α and β are constants are deteremined. Prove that $a + b = 1$ and $b\alpha = b\beta = \frac{1}{2}$ and choose appropriate values to give the Midpoint Runge Kutta method.

4

MULTI-STEP METHODS

Methods using the approximation at more than one previous point to determine the approx at the next point are called multi-step methods.

Definition An m-step multi-step method for solving the Initial Value Problem

$$y' = f(t,y) \quad a \leq t \leq b \quad y(a) = \alpha$$

is on whose difference equation for finding the approximation w_{i+1} at the mesh points t_{i+1} can be represented by the following equation, when m is an integer greater than 1,

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + ... + a_0w_{i+1-m}$$

$$+h[b_mf(t_{i+1},w_{i+1}) + b_{m-1}f(t_i,w_i) + ... + b_0f(t_{i+1-m},w_{i+1-m})] \quad (22)$$

for $i = m-1, m, ..., N-1$ where $h = \frac{b-a}{N}$ the $a_0, a_1, ..., a_m - 1$ and $b_0, b_1, ..., b_m$ are constants, and the starting values

$$w_0 = \alpha, \;\; w_1 = \alpha_1, \;\; w_2 = \alpha_2, \; ... \; w_{m-1} = \alpha_{m-1}$$

are specified.
When $b_m = 0$ the method is called **explicit** or open since (22) then gives w_{i+1} explicitly in terms of previously determined approximations.

When $b_m \neq 0$ the method is called **implicit** or closed since w_{i+1} occurs on both sides of (22). ⋄

Example 15

Fourth order Adams-Bashforth

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad w_2 = \alpha_2 \quad w_3 = \alpha_3$$

$$
\begin{aligned}
w_{i+1} &= w_i + \frac{h}{24}[55f(t_i, w_i) - 59f(t_{i-1}, w_{i-1}) \\
&\quad +37f(t_{i-2}, w_{i-2}) - 9f(t_{i-3}, w_{i-3})]
\end{aligned}
$$

For each $i = 3, 4, ..., N-1$ define an explicit four step method known as the fourth order Adams-Bashforth technique.
The equation

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad w_2 = \alpha_2$$

$$
\begin{aligned}
w_{i+1} &= w_i + \frac{h}{24}[9f(t_{i+1}, w_{i+1}) + 19f(t_i, w_i) \\
&\quad -5f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})]
\end{aligned}
$$

For each $i = 2, 4, ..., N-1$ define an implicit three step method known as the fourth order Adams-Moulton technique.

For the previous methods we need to generate α_1, α_2 and α_3 by using a one step method.

4.1 Derivation of a explicit multistep method

4.1.1 *General Derivation of a explicit method Adams-Bashforth*

$$y(t_{i+1}) - y(t_i) = \int_{t_i}^{t_{i+1}} y^{'}(t)dt = \int_{t_i}^{t_{i+1}} f(t, y(t))dt$$

Consequently

$$y(t_{i+1}) = y(t_i) + \int_{t_i}^{t_{i+1}} f(t, y(t))dt$$

Since we cannot integrate $f(t, y(t))$ without knowing $y(t)$ the solution to the problem we instead integrate an interpolating poly. $P(t)$ to $f(t, y(t))$ that is determined by some of the previous obtained data points $(t_0, w_0), (t_1, w_1), ..., (t_i, w_i)$. When we assume in addition that $y(t_i) \approx w_i$

$$y(t_{i+1}) \approx w_i + \int_{t_i}^{t_{i+1}} P(t)dt$$

We use Newton back-substitution to derive an Adams-Bashforth explicit m-step technique, we form the backward difference poly $P_{m-1}(t)$ through $(t_i, f(t_i)), (t_{i-1}, f(t_{i-1})), ..., (t_{i+1-m}, f(t_{i+1-m}))$

$$f(t, y(t)) = P_{m-1}(t) + f^m(\xi, y(\xi))\frac{(t-t_i)...(t-t_{i+1-m})}{m!} \tag{23}$$

$$P_{m-1}(t) = \Sigma_{j=1}^{m} L_{m-1,j}(t)\nabla^j f(t_{i+1-j}, y(t_{i+1-j})) \tag{24}$$

where

$$\nabla f(t_i, y(t_i)) = f(t_i, y(t_i)) - f(t_{i-1}, y(t_{i-1})),$$
$$\nabla^2 f(t_i, y(t_i)) = \nabla f(t_i, y(t_i)) - \nabla f(t_{i-1}, y(t_{i-1}))$$
$$= f(t_i, y(t_i)) - 2f(t_{i-1}, y(t_{i-1})) + f(t_{i-2}, y(t_{i-2})).$$

Derivation of a explicit two-step method Adams Bashforth

To derive two step Adams-Bashforth technique

$$\begin{aligned}
\int_{t_i}^{t_{i+1}} f(t,y)dt &= \int_{t_i}^{t_{i+1}} [f(t_i, y(t_i)) + \frac{(t-t_i)}{h}\nabla f(t_i, y(t_i)) + error]dt \\
y_{i+1} - y_i &= [tf(t_i, y(t_i)) + \frac{t(\frac{t}{2}-t_i)}{h}\nabla f(t_i, y(t_i))]_{t_i}^{t_{i+1}} + Error \\
y_{i+1} &= y(t_i) + (t_{i+1}-t_i)f(t_i, y(t_i)) \\
&\quad + \frac{\frac{t_{i+1}}{2} - t_{i+1}t_i + \frac{t_i^2}{2} - t_i^2}{h}\nabla(f(t_i, y(t_i)) + Error \\
&= y(t_i) + hf(t_i, y(t_i)) \\
&\quad + \frac{(t_{i+1}-t_i)^2}{2h}(f(t_i, y(t_i)) - f(t_{i-1}, y(t_{i-1}))) + Error \\
&= y(t_i) + hf(t_i, y(t_i)) \\
&\quad + \frac{1}{2}(f(t_i, y(t_i)) - f(t_{i-1}, y(t_{i-1}))) + Error \\
&= y(t_i) + \frac{h}{2}[3f(t_i, y(t_i)) - f(t_{i-1}, y(t_{i-1})) + Error]
\end{aligned}$$

The two step Adams-Bashforth is $w_0 = \alpha_0$ and $w_1 = \alpha_1$ with

$$w_{i+1} = w_i + \frac{h}{2}[3w_i - w_{i-1}] \quad for \;\; i = 1, .., N-1$$

The local truncation error is

$$\tau_{i+1}(h) = \frac{y(t_i+1) - y(t_i)}{h} - \frac{1}{2}[3f(t_i, y(t_i)) - f(t_{i-1}, y(t_{i-1}))]$$

$$\tau_{i+1}(h) = \frac{Error}{h}$$

$$\begin{aligned}
Error &= \int_{t_i}^{t_{i+1}} \frac{(t-t_i)(t-t_{i-1})}{(t_{i+1}-t_i)(t_{i+1}-t_{i-1})} h^3 f^2(\mu_i, y(\mu_i))]dt \\
&= \frac{5}{12}h^3 f^2(\mu_i, y(\mu_i))
\end{aligned}$$

$$\tau_{i+1}(h) = \frac{\frac{5}{12}h^3 f^2(\mu_i, y(\mu_i))}{h}$$

The local truncation error for the two step Adams-Bashforth methods is of order 2

$$\tau_{i+1}(h) = O(h^2)$$

General Derivation of a explicit method Adams-Bashforth (cont.)

Definition The Lagrange polynomial $L_{m-1,j}(t)$ has a degree of $m-1$ and is associated with the interpolation point t_j in the sense

$$L_{m-1,j}(t) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$$L_{m-1,j}(t) = \frac{(t-t_0)...(t-t_{m-1})}{(t_j-t_0)...(t_j-t_{m-1})} = \prod_{k=0,k\neq j}^{m-1} \frac{t-t_k}{t_j-t_k} \tag{25}$$

Introducing the variable $t = t_k + sh$ with $dt = hds$ into $L_{m-1}(t)$

$$L_{m-1,j}(t) = \prod_{k=0,k\neq j}^{m-1} \frac{t_i+sh-t_k}{(i+1)h} = (-1)^{(m-1)} \begin{pmatrix} -s \\ (m-1) \end{pmatrix} \tag{26}$$

$$\begin{aligned} \int_{t_i}^{t_{i+1}} f(t,y(t))dt = & \quad \textstyle\int_{t_i}^{t_{i+1}} \sum_{k=0}^{m-1} \begin{pmatrix} -s \\ k \end{pmatrix} \nabla^k f(t_i,y(t_i))dt \\ & + \textstyle\int_{t_i}^{t_{i+1}} f^m(\xi,y(\xi)) \frac{(t-t_i)...(t-t_{i+1-m})}{m!} dt \\ = & \quad \textstyle\sum_{k=0}^{m-1} \nabla^k f(t_i,y(t_i)) h (-1)^k \int_0^1 \begin{pmatrix} -s \\ k \end{pmatrix} ds \\ & + \textstyle\frac{h^{m+1}}{m!} \int_0^1 s(s+1)...(s+m-1) f^m(\xi,y(\xi))ds \end{aligned}$$

The integrals $(-1)^k \int_0^1 \begin{pmatrix} -s \\ k \end{pmatrix} ds$ for various values of k are computed as such,

Example 16

Example $k = 2$

$$\begin{aligned} (-1)^2 \int_0^1 \begin{pmatrix} -s \\ 2 \end{pmatrix} ds &= \int_0^1 \frac{-s(-s-1)}{1.2} ds \\ &= \frac{1}{2} \int_0^1 s^2 + sds \\ &= \frac{1}{2} \left[\frac{s^3}{3} + \frac{s^2}{2} \right]_0^1 = \frac{5}{12} \end{aligned}$$

k	0	1	2	3	4	...
$(-1)^k \int_0^1 \begin{pmatrix} -s \\ k \end{pmatrix} ds$	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	.	..

Table 1: Table of Adams-Bashforth coefficients

As a consequence

$$\int_{t_i}^{t_{i+1}} f(t,y(t))dt = h\left[f(t_i,y(t_i)) + \frac{1}{2}\nabla f(t_i,y(t_i)) + \frac{5}{12}\nabla^2 f(t_i,y(t_i)) + ...\right]$$

$$+\frac{h^{m+1}}{m!}\int_0^1 s(s+1)...(s+m-1)f^m(\xi,y(\xi))ds$$

Since $s(s+1)...(s+m-1)$ does not change sign on $[0,1]$ it can be stated that for some μ_i where $t_{i+1-m} < \mu_i < t_{i+1}$ the error term becomes

$$\frac{h^{m+1}}{m!}\int_0^1 s(s+1)...(s+m-1)f^m(\xi,y(\xi))ds$$

$$\frac{h^{m+1}}{m!}f^m(\mu,y(\mu))\int_0^1 s(s+1)...(s+m-1)ds$$

Since $y(t_{i+1}) - y(t_i) = \int_{t_i}^{t_{i+1}} f(s,y(s))ds$ this can be written as

$$y(t_{i+1}) = y(t_i) + h\left[f(t_i,y(t_i)) + \frac{1}{2}\nabla f(t_i,y(t_i)) + \frac{5}{12}\nabla^2 f(t_i,y(t_i)) + ...\right]$$

Example 17

To derive the two step Adams-Bashforth method

$$\begin{aligned} y(t_{i+1}) &\approx y(t_i) + h[f(t_i,y(t_i)) + \frac{1}{2}(\nabla f(t_i,y(t_i)))] \\ &= y(t_i) + h[f(t_i,y(t_i)) + \frac{1}{2}(f(t_i,y(t_i)) - f(t_{i-1},y(t_{i-1})))] \\ &= y(t_i) + \frac{h}{2}[3f(t_i,y(t_i)) - f(t_{i-1},y(t_{i-1}))] \end{aligned}$$

The two step Adams-Bashforth is $w_0 = \alpha_0$ and $w_1 = \alpha_1$ with

$$w_{i+1} = w_i + \frac{h}{2}[3w_i - w_{i-1}] \quad for \;\; i = 1,..,N-1$$

Definition If $y(t)$ is a solution of the Initial Value Problem

$$y' = f(t,y), \quad a \le t \le b \;\; y(a) = \alpha$$

and

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + ... + a_0 w_{i+1-m}$$

$$+h[b_m f(t_{i+1}, w_{i+1}) + b_{m-1} f(t_i, w_i) + ... + b_0 f(t_{i+1-m}, w_{i+1-m})]$$

is the (i+1)th step in a multi-step method, the local truncation error at this step is

$$\tau_{i+1}(h) = \frac{y(t_{i+1}) - a_{m-1}y(t_i) - ... - a_0 y(t_{i+1-m})}{h}$$

$$-[b_m f(t_{i+1}, y(t_{i+1})) + b_{m-1} f(t_i, y(t_i)) + ... + b_0 f(t_{i+1-m}, y(t_{i+1-m}))]$$

for each $i = m-1, ..., N-1$.

Example 18

Truncation error for the two step Adams-Bashforth method is

$$h^3 f^2(\mu_i, y(\mu_i))(-1)^2 \int_0^1 \begin{pmatrix} -s \\ 2 \end{pmatrix} ds = \frac{5h^3}{12} f^2(\mu_i, y(\mu_i))$$

using the fact that $f^2(\mu_i, y(\mu_i)) = y^3(\mu_i)$

$$\tau_{i+1}(h) = \frac{y(t_{i+1}) - y(t_i)}{h} - \frac{1}{2}[3f(t_i, y(t_i)) - f(t_{i-1}, y(t_{i-1}))]$$

$$= \frac{1}{h}\left[\frac{5h^3}{12} f^2(\mu_i, y(\mu_i))\right] = \frac{5}{12} h^2 y^3(\mu_i)$$

4.1.2 *Adams-Bashforth three step method*

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad w_2 = \alpha_2$$

$$w_{i+1} = w_i + \frac{h}{12}[23f(t_i, w_i) - 16f(t_{i-1}, w_{i-1}) + 5f(t_{i-2}, w_{i-2})]$$

where i=2,3,...,N-1
The local truncation error is of order 3

$$\tau_{i+1}(h) = \frac{3}{8} h^3 y^4(\mu_i)$$

$\mu_i \in (t_{i-2}, t_{i+1})$

4.1.3 *Adams-Bashforth four step method*

$$w_0 = \alpha, \quad w_1 = \alpha_1, \quad w_2 = \alpha_2, \quad w_3 = \alpha_3,$$

$$w_{i+1} = w_i + \frac{h}{24}[55f(t_i, w_i) - 59f(t_{i-1}, w_{i-1}) + 37f(t_{i-2}, w_{i-2}) - 9f(t_{i-3}, w_{i-3})],$$

where $i = 3, ..., N-1$.
The local truncation error is of order 4

$$\tau_{i+1}(h) = \frac{251}{720}h^4 y^5(\mu_i),$$

$\mu_i \in (t_{i-3}, t_{i+1})$.

Example 19

$$y' = y - t^2 + 1, \quad 0 \leq t \leq 2, \quad y(0) = 0.5,$$

Adams-Bashforth two step $w_0 = \alpha_0$ and $w_1 = \alpha_1$ with

$$w_{i+1} = w_i + \frac{h}{2}[3f(t_i, w_i) - f(t_{i-1}, w_{i-1})], \quad for, \ i = 1, .., N-1,$$

truncation error

$$\tau_{i+1}(h) = \frac{5}{12}h^2 y^3(\mu_i), \qquad \mu_i \in (t_{i-1}, t_{i+1}).$$

1. Calculate α_0 and α_1
   From the initial condition we have $w_0 = 0.5$
   To calculate w_1 we use the modified Euler method.

$$\begin{aligned} w_0 &= \alpha \\ w_{i+1} &= w_i \frac{h}{2}[f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))] \end{aligned}$$

   We only need this to calculate w_1

$$\begin{aligned} w_0 &= 0.5 \\ w_1 &= w_0 + \frac{h}{2}[f(t_0, w_0) + f(t_1, w_0 + hf(t_0, w_0))] \\ w_1 &= w_0 + \frac{0.2}{2}[w_0 - t_0^2 + 1 + w_0 + h(w_0 - t_0^2 + 1) - t_1^2 + 1] \\ &= 0.5 + \frac{0.2}{2}[0.5 - 0 + 1 + 0.5 + 0.2(1.5) - (0.2)^2 + 1] \\ &= 0.826 \end{aligned}$$

   we now have $\alpha_1 = w_1 = 0.826$

2. Calculate w_i for $i = 2, ...N$

$$\begin{aligned} w_2 &= w_1 + \frac{h}{2}[3f(t_1, w_1) - f(t_0, w_0)] \\ &= w_1 + \frac{h}{2}[3(w_1 - t_1^2 + 1) - (w_0 - t_0^2 + 1)] \\ &= 0.826 + \frac{0.2}{2}[3(0.826 - 0.2^2 + 1) - (0.5 - 0^2 + 1)] \\ &= 0.8858 \\ &. \\ &. \\ &. \\ w_{i+1} &= w_i + \frac{0.2}{2}[3(w_i - t_i^2 + 1) - (w_{i-1} - t_{i-1}^2 + 1)] \end{aligned}$$

this method can be generalised for all Adams-Bashforth.

4.2 Derivation of the implicit multi-step method

4.2.0.1 *Derivation of an implicit one-step method Adams Moulton*

To derive one step Adams-Moulton technique

$$\begin{aligned} \int_{t_i}^{t_{i+1}} f(t,y)dt &= \int_{t_i}^{t_{i+1}} [f(t_{i+1}, y(t_{i+1})) + \frac{(t - t_{i+1})}{h} \nabla f(t_{i+1}, y(t_{i+1})) + error]dt \\ y_{i+1} - y_i &= [tf(t_{i+1}, y(t_{i+1})) \\ &\quad + \frac{t(\frac{t}{2} - t_{i+1})}{h} \nabla f(t_{i+1}, y(t_{i+1}))]_{t_i}^{t_{i+1}} + Error \\ y_{i+1} &= y(t_i) + (t_{i+1} - t_i) f(t_{i+1}, y(t_{i+1})) \\ &\quad + \frac{\frac{t_{i+1}^2}{2} - t_{i+1}^2 + t_i t_{i+1} - \frac{t_i^2}{2}}{h} \nabla(f(t_{i+1}, y(t_{i+1})) \\ &\quad + Error \\ &= y(t_i) + hf(t_{i+1}, y(t_{i+1})) \\ &\quad + \frac{-(t_{i+1} - t_i)^2}{2h}(f(t_{i+1}, y(t_{i+1})) - f(t_i, y(t_i))) \\ &\quad + Error \\ &= y(t_i) + hf(t_{i+1}, y(t_{i+1})) \\ &\quad - \frac{h}{2}(f(t_{i+1}, y(t_{i+1})) - f(t_i, y(t_i))) + Error \\ &= y(t_i) + \frac{h}{2}[f(t_{i+1}, y(t_{i+1})) + f(t_{i-1}, y(t_{i-1}))] + Error \end{aligned}$$

The two step Adams-Moulton is $w_0 = \alpha_0$ and $w_1 = \alpha_1$ with

$$w_{i+1} = w_i + \frac{h}{2}[w_{i+1} + w_i] \quad for \;\; i = 0, .., N - 1$$

The local truncation error is

$$\tau_{i+1}(h) = \frac{y(t_i+1) - y(t_i)}{h} - \frac{1}{2}[f(t_{i+1}, y(t_{i+1})) + f(t_i, y(t_i))]$$

$$\tau_{i+1}(h) = \frac{Error}{h}$$

$$\begin{aligned} Error &= \int_{t_i}^{t_{i+1}} \frac{(t-t_{i+1})(t-t_i)}{(t_i-t_{i+1})(t_i-t_{i-1})} h^3 f^2(\mu_i, y(\mu_i))]dt \\ &= \frac{1}{12} h^3 f^2(\mu_i, y(\mu_i)) \end{aligned}$$

$$\tau_{i+1}(h) = \frac{\frac{1}{12} h^3 f^2(\mu_i, y(\mu_i))}{h}$$

The local truncation error for the one step Adams-Moulton methods is of order 2

$$\tau_{i+1}(h) = O(h^2)$$

derivation of the implicit multi-step method (cont)

As before

$$\begin{aligned} y(t_{i+1}) - y(t_i) &= \int_{-1}^{0} y'(t_{i+1} + sh) ds \\ &= \sum_{k=0}^{m-1} \nabla^k f(t_{i+1}, y(t_{i+1})) h (-1)^k \int_{-1}^{0} \left(\begin{array}{c} -s \\ k \end{array} \right) ds \\ &+ \frac{h^{m+1}}{m!} \int_{-1}^{0} s(s+1)...(s+m-1) f^m(\xi, y(\xi)) ds \end{aligned}$$

Example 20

For k=3 we have

$$\begin{aligned} (-1)^3 \int_{-1}^{0} \left(\begin{array}{c} -s \\ k \end{array} \right) ds &= \int_{-1}^{0} \frac{-s(-s-1)(-s-2)}{1.2.3} ds \\ &= \frac{1}{6} \left[\frac{s^4}{4} + s^3 + s^2 \right]_{-1}^{0} = -\frac{1}{24} \end{aligned}$$

The general form of the Adams-Moulton method is

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + h[f(t_{i+1}, y(t_{i+1})) - \frac{1}{2} \nabla f(t_{i+1}, y(t_{i+1})) - \frac{1}{12} \nabla^2 f(t_{i+1}, y(t_{i+1})) - ...] \\ &+ \frac{h^{m+1}}{m!} \int_{-1}^{0} s(s+1)...(s+m-1) f^m(\xi, y(\xi)) ds \end{aligned}$$

Adams-Moulton two step method

$$w_0 = \alpha \quad w_1 = \alpha_1$$

$$w_{i+1} = w_i + \frac{h}{12}[5f(t_{i+1}, w_{i+1}) + 8f(t_i, w_i) - f(t_{i-1}, w_{i-1})]$$

where i=2,3,...,N-1
The local truncation error is

$$\tau_{i+1}(h) = -\frac{1}{24}h^3 y^4(\mu_i)$$

$\mu_i \in (t_{i-1}, t_{i+1})$

Adams-Moulton three step method

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad w_2 = \alpha_2$$

$$w_{i+1} = w_i + \frac{h}{24}[9f(t_{i+1}, w_{i+1}) + 19f(t_i, w_i) - 5f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})]$$

where i=3,...,N-1
The local truncation error is

$$\tau_{i+1}(h) = -\frac{19}{720}h^4 y^5(\mu_i)$$

$\mu_i \in (t_{i-2}, t_{i+1})$

Example 21

$$y' = y - t^2 + 1 \quad 0 \leq t \leq 2 \quad y(0) = 0.5$$

Adams-Moulton two step $w_0 = \alpha_0$ and $w_1 = \alpha_1$ with

$$w_{i+1} = w_i + \frac{h}{12}[5f(t_{i+1}, w_{i+1}) + 8f(t_i, w_i) - f(t_{i-1}, w_{i-1})]$$

$$for \quad i = 2, .., N-1$$

truncation error $\tau_{i+1}(h) = -\frac{1}{24}h^3 y^4(\mu_i)$ $\mu_i \in (t_{i-1}, t_{i+1})$.

1. Calculate α_0 and α_1
   From the initial condition we have $w_0 = 0.5$
   To calculate w_1 we use the modified Euler method.
   we now have $\alpha_1 = w_1 = 0.826$

2. Calculate w_i for $i = 2, ...N$

$$\begin{aligned}
w_2 &= w_1 + \frac{h}{12}[5f(t_2, w_2) \\
&\quad +8f(t_1, w_1) - f(t_0, w_0)] \\
&= w_1 + \frac{h}{12}[5(w_2 + t_2^2 + 1) \\
&\quad +8(w_1 + t_1^2 + 1) - (w_0 + t_0^2 + 1)] \\
&= w_1 + \frac{h}{12}[5(w_2 + t_2^2 + 1) \\
&\quad +8(w_1 + t_1^2 + 1) - (w_0 + t_0^2 + 1)] \\
&. \\
&. \\
&. \\
(1 - \frac{5h}{12})w_{i+1} &= \frac{h}{12}[[8 + \frac{12}{h}]w_i - 5t_{i+1}^2 \\
&\quad -8t_i^2 + t_{i-1}^2 + 12]
\end{aligned}$$

This, of course can be generalised.
The only unfortunate aspect of the implicit method is that you must convert it into an explicit method, this is not always possible. eg $y^{'} = e^y$.

4.3 TABLE OF ADAM'S METHODS

Order	Formula	LTE
1	$y_{n+1} = y_n + hf_n$	$\frac{h^2}{2}y''(\eta)$
2	$y_{n+1} = y_n + \frac{h}{2}[3f_n - f_{n-1}]$	$\frac{5h^3}{12}y'''(\eta)$
3	$y_{n+1} = y_n + \frac{h}{12}[23f_n - 16f_{n-1} + 5f_{n-2}]$	$\frac{3h^4}{8}y^{(4)}(\eta)$
4	$y_{n+1} = y_n + \frac{h}{24}[55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}]$	$\frac{251h^5}{720}y^{(5)}(\eta)$

Table 2: Adams-Bashforth formulas of different order. LTE stands for local truncation error.

Order	Formula	LTE
0	$y_{n+1} = y_n + hf_{n+1}$	$-\frac{h^2}{2}y''(\eta)$
1	$y_{n+1} = y_n + \frac{h}{2}[f_{n+1} + f_n]$	$-\frac{h^3}{12}y'''(\eta)$
2	$y_{n+1} = y_n + \frac{h}{12}[5f_{n+1} + 8f_n - f_{n-1}]$	$-\frac{h^4}{24}y^{(4)}(\eta)$
3	$y_{n+1} = y_n + \frac{h}{24}[9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2}]$	$-\frac{19h^5}{720}y^{(5)}(\eta)$

Table 3: Adams-Moulton formulas of different order. LTE stands for local truncation error.

4.4 PREDICTOR-CORRECTOR METHOD

In practice implicit methods are not used as above. They are used to improve approximations obtained by explicit methods. The combination of the two is called predictor-corrector method.

Example 22

Consider the following forth order method for solving an initial-value problem.

1. Calculate w_0, w_1, w_2, w_3 for the four step Adams-Bashforth method, to do this we use a 4th order one step method, eg Runge Kutta.
2. Calculate an approximation w_4 to $y(t_4)$ using the Adams-Bashforth method as the predictor.

$$w_4^0 = w_3 + \frac{h}{24}[55f(t_3, w_3) - 59f(t_2, w_2) + 37f(t_1, w_1) - 9f(t_0, w_0)]$$

3. This approximation is improved by inserting w_4^0 in the RHS of the three step Adams-Moulton and using it as a corrector

$$w_4^1 = w_3 + \frac{h}{24}[9f(t_4, w_4^0) + 19f(t_3, w_3) - 5f(t_2, w_2) + f(t_1, w_1)]$$

   The only new function evaluation is $f(t_4, w_4^0)$.

4. w_4^1 is the approximation of $y(t_4)$.
5. Repeat steps 2-4 for calculating the approximation of $y(t_5)$.
6. Repeat til $y(t_n)$.

Improved approximations to $y(t_{i+1})$ can be obtained by integrating the Adams Moulton formula

$$w_{i+1}^{k+1} = w_i + \frac{h}{24}[9f(t_{i+1}, w_{i+1}^k) + 19f(t_i, w_i) - 5f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})]$$

w_{i+1}^{k+1} converges to the approximation of the implicit method rather than the solution $y(t_{i+1})$.
A more effective method is to reduce step-size if improved accuracy is needed.

4.5 IMPROVED STEP-SIZE MULTI-STEP METHOD

As the predictor corrector technique produces two approximations of each step it is a natural candidate for error-control. (see previous section)

Example 23

The Adams-Bashforth 4-step method

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad w_2 = \alpha_2 \quad w_3 = \alpha_3$$

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + \frac{h}{24}[55f(t_i, y(t_i)) - 59f(t_{i-1}, y(t_{i-1})) \\ &\quad +37f(t_{i-2}, y(t_{i-2})) - 9f(t_{i-3}, y(t_{i-3}))] \\ &\quad +\frac{251}{720}h^5 y^5(\mu_i) \end{aligned}$$

$\mu_i \in (t_{i-3}, t_{i+1})$ the truncation error is

$$\frac{y(t_{i+1}) - w^0_{i+1}}{h} = \frac{251}{720}h^4 y^5(\mu_i) \tag{27}$$

Similarly for Adams-Moulton three step method

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + \frac{h}{24}[9f(t_{i+1}, y(t_{i+1})) + 19f(t_i, y(t_i)) \\ &\quad -5f(t_{i-1}, y(t_{i-1})) + f(t_{i-2}, y(t_{i-2}))] \\ &\quad -\frac{19}{720}h^5 y^5(\xi_i) \end{aligned}$$

$\xi_i \in (t_{i-2}, t_{i+1})$ the truncation error is

$$\frac{y(t_{i+1}) - w_{i+1}}{h} = -\frac{19}{720}h^4 y^5(\xi_i) \tag{28}$$

We make the assumption that for small h

$$y^5(\xi_i) \approx y^5(\mu_i)$$

Subtracting (27) from (28) we have

$$\frac{w_{i+1} - w^0_{i+1}}{h} = \frac{h^4}{720}[251y^5(\mu_i) + 19y^5(\xi_i)] \approx \frac{3}{8}h^4 y^5(\xi_i)$$

$$y^5(\xi_i) \approx \frac{8}{3h^5}(w_{i+1} - w^0_{i+1}) \tag{29}$$

Using this result to eliminate $h^4 y^5(\xi_i)$ from (28)

$$|\tau_{i+1}(h)| = \frac{|y(t_{i+1}) - w_{i+1}|}{h} \approx \frac{19h^4}{720}\frac{8}{3h^5}(|w_{i+1} - w^0_{i+1}|)$$

$$= \frac{19|w_{i+1} - w^0_{i+1}|}{270h}$$

Now consider a new step size qh generating new approximations $\hat{w}_0, \hat{w}_1, .., \hat{w}_i$ The objective is to choose a q so that the local truncation is bounded by a tol ε

$$\frac{|y(t_i+qh)-\hat{w}_{i+1}|}{qh} = \frac{19h^4}{720}|y^5(\nu)|q^4 \approx \frac{19h^4}{720}\frac{8}{3h^5}(|w_{i+1}-w^0_{i+1}|)q^4$$

and we need to choose a q so that

$$\frac{|y(t_i+qh)-\hat{w}_{i+1}|}{qh} \approx \frac{19q^4|w_{i+1}-w^0_{i+1}|}{270h} < \varepsilon$$

that is, we choose so that

$$q < \left(\frac{270}{19}\frac{h\varepsilon}{|w_{i+1}-w^0_{i+1}|}\right)^{\frac{1}{4}} \approx 2\left(\frac{h\varepsilon}{|w_{i+1}-w^0_{i+1}|}\right)^{\frac{1}{4}}$$

q is normally chosen as

$$q = 1.5\left(\frac{h\varepsilon}{|w_{i+1}-w^0_{i+1}|}\right)^{\frac{1}{4}}$$

With this knowledge we can change step sizes and control out error.

4.6 PROBLEM SHEET 3

1. Apply the 3-step Adams-Bashforth to approximate the solution of the given initial value problems using the indicated number of time steps. Compare the approximate solution with the given exact solution

   a) $y' = t - y, \quad (0 \le t \le 4)$
   with the initial condition $y(0) = 1$,
   $N = 4$, $y(t) = 2e^{-t} + t - 1$

   b) $y' = y - t, \quad (0 \le t \le 2)$
   with the initial condition $y(0) = 2$,
   $N = 4$, $y(t) = e^{t} + t + 1$

2. Apply the 2-step Adams-Moulton Method to approximate the solution of the given initial value problems using the indicated number of time steps. Compare the approximate solution with the given exact solution

   a) $y' = t - y, \quad (0 \le t \le 4)$
   with the initial condition $y(0) = 1$,
   $N = 4$, $y(t) = 2e^{-t} + t - 1$

   b) $y' = y - t, \quad (0 \le t \le 2)$
   with the initial condition $y(0) = 2$,
   $N = 4$, $y(t) = e^{t} + t + 1$

3. Derive the difference equation for the 1-step Adams-Bashforth method:

$$w_{n+1} = w_n + hf(t_n, w_n),$$

   with the local truncation error

$$\tau_{n+1}(h) = \frac{h}{2} y^2(\mu_n)$$

   where $\mu_n \in (t_n, t_{n+1})$.

4. Derive the difference equation for the 2-step Adams-Bashforth method:

$$w_{n+1} = w_n + (\frac{3}{2} hf(t_n, w_n) - \frac{1}{2} hf(t_{n-1}, w_{n-1})),$$

   with the local truncation error

$$\tau_{n+1}(h) = \frac{5h^2}{12} y^3(\mu_n)$$

where $\mu_n \in (t_{n-1}, t_{n+1})$.

5. Derive the difference equation for the 3-step Adams-Bashforth method:

$$w_{n+1} = w_n + (\frac{23}{12}hf(t_n, w_n) - \frac{4}{3}hf(t_{n-1}, w_{n-1}) + \frac{5}{12}hf(t_{n-2}, w_{n-2})),$$

with the local truncation error

$$\tau_{n+1}(h) = \frac{9h^3}{24}y^4(\mu_n)$$

where $\mu_n \in (t_{n-2}, t_{n+1})$.

6. Derive the difference equation for the 0-step Adams-Moulton method:

$$w_{n+1} = w_n + hf(t_{n+1}, w_{n+1}),$$

with the local truncation error

$$\tau_{n+1}(h) = -\frac{h}{2}y^2(\mu_n)$$

where $\mu_n \in (t_{n-2}, t_{n+1})$.

7. Derive the difference equation for the 1-step Adams-Moulton method:

$$w_{n+1} = w_n + \frac{1}{2}hf(t_{n+1}, w_{n+1}) + \frac{1}{2}hf(t_n, w_n),$$

with the local truncation error

$$\tau_{n+1}(h) = -\frac{h^2}{12}y^3(\mu_n)$$

where $\mu_n \in (t_n, t_{n+1})$.

8. Derive the difference equation for the 2-step Adams-Moulton method:

$$w_{n+1} = w_n + \frac{5}{12}hf(t_{n+1}, w_{n+1}) + \frac{8}{12}hf(t_n, w_n) - \frac{1}{12}hf(t_{n-1}, w_{n-1}),$$

with the local truncation error

$$\tau_{n+1}(h) = -\frac{h^3}{24}y^4(\mu_n)$$

where $\mu_n \in (t_{n-1}, t_{n+1})$.

9. Derive the difference equation for the 3-step Adams-Moulton method:

$$w_{n+1} = w_n + \frac{9}{24}hf(t_{n+1}, w_{n+1}) + \frac{19}{24}hf(t_n, w_n) - \frac{5}{24}hf(t_{n-1}, w_{n-1}) + \frac{1}{24}hf(t_{n-2}, w_{n-2}),$$

with the local truncation error

$$\tau_{n+1}(h) = -\frac{h^4}{720} y^5(\mu_n)$$

where $\mu_n \in (t_{n-2}, t_{n+1})$.

5

CONSISTENCY, CONVERGENCE AND STABILITY

5.1 ONE STEP METHODS

Stability is why some methods give satisfactory results and some do not.

Definition A one-step method with local truncation error $\tau_i(h)$ at the ith step is said to be **consistent** with the differential equation it approximates if

$$\lim_{h\to 0}(\max_{1\leq i\leq N}|\tau_i(h)|) = 0$$

where

$$\tau_i(h) = \frac{y_{i+1}-y_i}{h} - F(t_i, y_i, h, f)$$

As $h \to 0$ does $F(t_i, y_i, h, f) \to f(t, y)$.

Definition A one step method difference equation is said to be **convergent** with respect to the differential equation and w_i, the approximation obtained from the difference method at the ith step.

$$\max_{h\to 0}\max_{1\leq i\leq N}|y(t_i) - w_i| = 0$$

For Euler's method we have

$$\max_{1\leq i\leq N}|w_i - y(t_i)| \leq \frac{Mh}{2L}|e^{L(b-a)} - 1|$$

so Euler's method is convergent wrt to a differential equation.

Theorem 5.1.1. *Suppose the initial value problem*

$$y' = f(t, y) \quad a \leq t \leq b \;\; y(a) = \alpha$$

is approximated by a one step difference method in the form

$$\begin{aligned} w_0 &= \alpha \\ w_{i+1} &= w_i + hF(t_i, w_i : h) \end{aligned}$$

Suppose also that a number $h_0 > 0$ exists and that $F(t_i, w_i : h)$ is continuous and satisfies a Lipschitz Condition in the variable w with Lipschitz constant L on the set

$$D = \{(t, w, h) | a \leq t \leq b, -\infty < w < \infty, 0 \leq h \leq h_0\}$$

Then

1. *The method is stable;*
2. *The difference method is convergent if and only if it is consistent - that is iff*
$$F(t_i, w_i : 0) = f(t, y) \text{ for all } a \leq t \leq b$$
3. *If a function τ exists and for each $i = 1, 2, .., N$, the local truncation error $\tau_i(h)$ satisfies $|\tau_i(h)| \leq \tau(h)$ whenever $0 \leq h \leq h_0$, then*
$$|y(t_i) - w_i| \leq \frac{\tau(h)}{L} e^{L(t_i - a)}.$$

Example 24

Consider the modified Euler method given by

$$w_0 = \alpha$$

$$w_{i+1} = w_i + \frac{h}{2}[f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))]$$

Verify that this method satisfies the theorem. For this method

$$F(t, w : h) = \frac{1}{2} f(t, w) + \frac{1}{2} f(t + h, w + hf(t, w))$$

If f satisfies the Lipschitz Condition on $\{(t, w) | a \leq t \leq b, -\infty < w < \infty\}$ in the variable w with constant L, then

$$\begin{aligned} F(t, w : h) - F(t, \hat{w} : h) &= \frac{1}{2} f(t, w) + \frac{1}{2} f(t + h, w + hf(t, w)) \\ &\quad - \frac{1}{2} f(t, \hat{w}) - \frac{1}{2} f(t + h, \hat{w} + hf(t, \hat{w})) \end{aligned}$$

the Lipschitz Condition on f leads to

$$\begin{aligned}
|F(t,w:h)-F(t,\hat{w}:h)| &\leq \frac{1}{2}L|w-\hat{w}| \\
&\quad +\frac{1}{2}L|w+hf(t,w)-\hat{w}-hf(t,\hat{w})| \\
&\leq L|w-\hat{w}|+\frac{1}{2}L|hf(t,w)-hf(t,\hat{w})| \\
&\leq L|w-\hat{w}|+\frac{1}{2}hL^2|w-\hat{w}| \\
&\leq \left(L+\frac{1}{2}hL^2\right)|w-\hat{w}|
\end{aligned}$$

Therefore, F satisfies a Lipschitz Condition in w on the set

$$D=\{(t,w,h)|a\leq t\leq b,-\infty<w<\infty,0\leq h\leq h_0\}$$

for any $h_0>0$ with constant $L^{'}=\left(L+\frac{1}{2}hL^2\right)$
Finally, if f is continuous on $\{(t,w)|a\leq t\leq b,-\infty< w<\infty\}$, then F is continuous on

$$D=\{(t,w,h)|a\leq t\leq b,-\infty<w<\infty,0\leq h\leq h_0\}$$

so this implies that the method is stable. Letting $h=0$ we have

$$F(t,w:0)=\frac{1}{2}f(t,w)+\frac{1}{2}f(t+0,w+0f(t,w))=f(t,w)$$

so the consistency condition holds.
We know that the method is of order $O(h^2)$

5.2 MULTI-STEP METHODS

The general multi-step method for approximating the solution to the Initial Value Problem

$$y^{'}=f(t,y)\quad a\leq t\leq b\;\; y(a)=\alpha$$

can be written in the form

$$w_0=\alpha\quad w_1=\alpha_1\quad ...\quad w_{m-1}=\alpha_{m-1}$$

$$w_{i+1}=a_{m-1}w_i+a_{m-2}w_{i-1}+...+a_0w_{i+1-m}+hF(t_i,h,w_{i+1},...,w_{i+1-m}),$$

for each $i = m-1, ..., N-1$ where $a_0, a_1, .., a_{m-1}$ are constants.
The local truncation error for a multi-step method expressed in this form is

$$\tau_{i+1}(h) = \frac{y(t_{i+1}) - a_{m-1}y(t_i) - a_{m-2}y(t_{i-1}) + ... - a_0 y(t_{i+1-m})}{h}$$

$$+F(t_i, h, y(t_{i+1}), ..., y(t_{i+1-m}))$$

for each $i = m-1, ..., N-1$.

Definition A multi-step method is **consistent** if both

$$\lim_{h\to 0} |\tau_i(h)| = 0 \quad \text{for all } i = m, ..., N$$

$$\lim_{h\to 0} |\alpha_i - y(t_i)| = 0 \quad \text{for all } i = 0, .., m-1$$

Definition A multi-step method is **convergent** if the solution to the difference equation approaches the solution of the differential equation as the step size approaches zero.

$$\lim_{h\to 0} |w_i - y(t_i)| = 0$$

Theorem 5.2.1. *Suppose the Initial Value Problem*

$$y' = f(t, y) \quad a \leq t \leq b \quad y(a) = \alpha$$

is approximated by an Adams predictor-corrector method with an m-step Adams-Bashforth predictor equation

$$w_{i+1} = w_i + h[b_{m-1}f(t_i, w_i) + ... + b_0 f(t_{i+1-m}, w_{i+1-m})]$$

with local truncation error $\tau_{i+1}(h)$ and an (m-1)-step Adams-Moulton equation

$$w_{i+1} = w_i + h[\hat{b}_{m-1}f(t_{i+1}, w_{i+1}) + ... + \hat{b}_0 f(t_{i+2-m}, w_{i+2-m})]$$

with local truncation error $\hat{\tau}_{i+1}(h)$. In addition suppose that $f(t, y)$ and $f_y(t, y)$ are continuous on $= \{(t, y)|a \leq t \leq b, -\infty < y < \infty\}$ and that fy is bounded. Then the local truncation error $\sigma_{i+1}(h)$ of the predictor-corrector method is

$$\sigma_{i+1}(h) = \hat{\tau}_{i+1}(h) + h\tau_{i+1}(h)\hat{b}_{m-1}\frac{\partial f}{\partial y}(t_{i+1}, \theta_{i+1})$$

where $\theta_{i+1} \in [0, h\tau_{i+1}(h)$.
Moreover, there exists constants k_1 and k_2 such that

$$|w_i - y(t_i)| \leq \left[\max_{0\leq j\leq m-1} |w_j - y(t_j)| + k_1\sigma(h)\right] e^{k_2(t_i - a)}.$$

where $\sigma(h) = \max_{m \le i \le N} |\sigma_i(h)|$.

Definition Associated with the difference equation

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad ... \quad w_{m-1} = \alpha_{m-1}$$

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + ... + a_0 w_{i+1-m} + hF(t_i, h, w_{i+1}, ..., w_{i+1-m}),$$

is the characteristic equation given by

$$\lambda^m - a_{m-1}\lambda^{m-1} - a_{m-2}\lambda^{m-2} - ... - a_0 = 0$$

Definition Let $\lambda_1, ..., \lambda_m$ denote the roots of the that characteristic equation

$$\lambda^m - a_{m-1}\lambda^{m-1} - a_{m-2}\lambda^{m-2} - ... - a_0 = 0$$

associated with the multi-step difference method

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad ... \quad w_{m-1} = \alpha_{m-1}$$

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + ... + a_0 w_{i+1-m} + hF(t_i, h, w_{i+1}, ..., w_{i+1-m}),$$

If $|\lambda_i| \le 1$ for each $i = 1, ..., m$ and all roots with absolute value 1 are simple roots then the difference equation is said to satisfy the **root condition**.

Definition

1. Methods that satisfy the root condition and have $\lambda = 1$ as the only root of the characteristic equation of magnitude one are called **strongly stable**;
2. Methods that satisfy the root condition and have more than one distinct root with magnitude one are called **weakly stable**;
3. Methods that do not satisfy the root condition are called **unstable**.

Theorem 5.2.2. *A multi-step method of the form*

$$w_0 = \alpha \quad w_1 = \alpha_1 \quad ... \quad w_{m-1} = \alpha_{m-1}$$

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + ... + a_0 w_{i+1-m} + hF(t_i, h, w_{i+1}, ..., w_{i+1-m})$$

is stable iff it satisfies the root condition. Moreover if the difference method is consistent with the differential equation then the method is stable iff it is convergent.

Example 25

We have seen that the fourth order Adams-Bashforth method can be expressed as

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + ... + a_0 w_{i+1-m} + hF(t_i, h, w_{i+1}, w_i, ..., w_{i-3})$$

where

$$F(t_i, h, w_{i+1}, w_i, ..., w_{i-3}) =$$

$$\frac{1}{24}[55f(t_i, w_i) - 59f(t_{i-1}, w_{i-1}) + 37f(t_{i-2}, w_{i-2}) - 9f(t_{i-3}, w_{i-3})]$$

so $m = 4, a_0 = 0, a_1 = 0, a_2 = 0$ and $a_3 = 1$.
The characteristic equation is

$$\lambda^4 - \lambda^3 = \lambda^3(\lambda - 1) = 0$$

which has the roots $\lambda_1 = 1, \lambda_2 = 0, \lambda_3 = 0$ and $\lambda_4 = 0$.
It satisfies the root condition and is strongly stable.

Example 26

The explicit multi-step method given by

$$w_{i+1} = w_{i-3} + \frac{4h}{3}[2f(t_i, w_i) - f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})]$$

has a characteristic equation

$$\lambda^4 - 1 = 0$$

which has the roots $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = i$ and $\lambda_4 = -i$, the method satisfies the root condition, but is only weakly stable.

Example 27

The explicit multi-step method given by

$$w_{i+1} = aw_{i-3} + \frac{4h}{3}[2f(t_i, w_i) - f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})]$$

has a characteristic equation

$$\lambda^4 - a = 0$$

which has the roots $\lambda_1 = \sqrt[4]{a}, \lambda_2 = -\sqrt[4]{a}, \lambda_3 = i\sqrt[4]{a}$ and $\lambda_4 = -i\sqrt[4]{a}$, when $a > 1$ the method dose not satisfy the root condition, and hence is unstable.

Example 28

Solving the Initial Value Problem

$$y' = -0.5y^2 \quad y(0) = 1$$

Using a weakly stable method

$$w_{i+1} = w_{i-3} + \frac{4h}{3}[2w_i - w_{i-1} + w_{i-2}]$$

Using an two different unstable method

1.
$$w_{i+1} = 1.0001w_{i-3} + \frac{4h}{3}[2f(t_i, w_i) - f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})]$$

2.
$$w_{i+1} = 1.5w_{i-3} + \frac{4h}{3}[2f(t_i, w_i) - f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})]$$

$$\lambda^4 - a = 0$$

which has the roots $\lambda_1 = \sqrt[4]{a}, \lambda_2 = -\sqrt[4]{a}, \lambda_3 = i\sqrt[4]{a}$ and $\lambda_4 = -i\sqrt[4]{a}$, when $a > 1$ the method dose not satisfy the root condition, and hence is unstable.

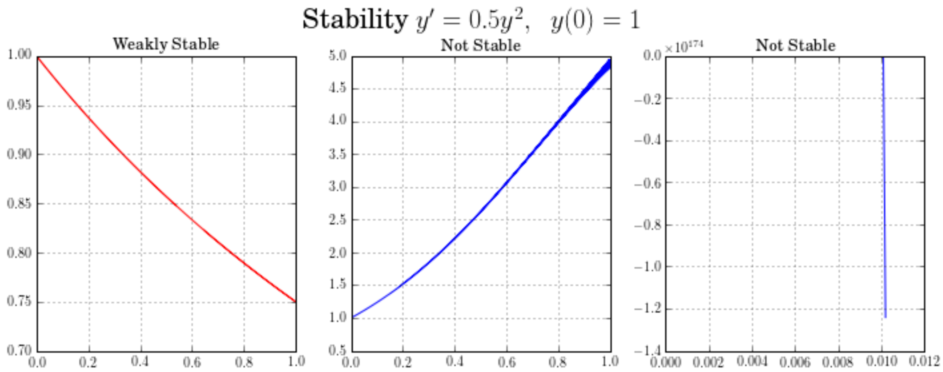


Figure 5.2.1: Python output: Left: Weakly stable solution, middle: unstable, right: very unstable

5.3 PROBLEM SHEET 4

1. Determine whether the 2-step Adams-Bashforth Method is consistent, stable and convergent

$$w_{n+1} = w_n + (\frac{3}{2}hf(t_n, w_n) - \frac{1}{2}hf(t_{n-1}, w_{n-1})),$$

2. Determine whether the 2-step Adams-Moulton Method is consistent, stable and convergent

$$w_{n+1} = w_n + \frac{5}{12}hf(t_{n+1}, w_{n+1}) + \frac{8}{12}hf(t_n, w_n) - \frac{1}{12}hf(t_{n-1}, w_{n-1}),$$

3. Determine whether the linear multistep following methods are consistent, stable and convergent

   a)

$$w_{n+1} = w_{n-1} + \frac{1}{3}h[f(t_{n+1}, w_{n+1}) + 4f(t_n, w_n) + f(t_{n-1}, w_{n-1})].$$

   b)

$$w_{n+1} = \frac{4}{3}w_n - \frac{1}{3}w_{n-1} + \frac{2}{3}h[f(t_{n+1}, w_{n+1})].$$

5.4 INITIAL VALUE PROBLEM REVIEW QUESTIONS

1. a) Derive the Euler approximation show it has a local truncation error of $O(h)$ of the Ordinary Differential Equation

$$y'(x) = f(x,y)$$

with initial condition

$$y(a) = \alpha.$$

[8 marks]

b) Suppose f is a continuous and satisfies a Lipschitz condition with constant L on $D = \{(t,y)|a \leq t \leq b, -\infty < y < \infty\}$ and that a constant M exists with the property that

$$|y''(t)| \leq M.$$

Let $y(t)$ denote the unique solution of the IVP

$$y' = f(t,y) \quad a \leq t \leq b \quad y(a) = \alpha$$

and $w_0, w_1, ..., w_N$ be the approx generated by the Euler method for some positive integer N. Then show for $i = 0, 1, ..., N$

$$|y(t_i) - w_i| \leq \frac{Mh}{2L}|e^{L(t_i-a)} - 1|$$

You may assume the two lemmas:
If s and t are positive real numbers $\{a_i\}_{i=0}^{N}$ is a sequence satisfying $a_0 \geq \frac{-t}{s}$ and $a_{i+1} \leq (1+s)a_i + t$ then

$$a_{i+1} \leq e^{(i+1)s}\left(a_0 + \frac{t}{s}\right) - \frac{t}{s}$$

For all $x \geq 0.1$ and any positive m we have

$$0 \leq (1+x)^m \leq e^{mx}$$

[17 marks]

c) Use Euler's method to estimate the solution of

$$y' = (1-x)y^2 - y; \quad y(0) = 1$$

at x=1, using $h = 0.25$.

[8 marks]

2. a) Derive the difference equation for the Midpoint Runge Kutta method

$$w_{n+1} = w_n + k_2$$
$$k_1 = hf(t_n, w_n)$$
$$k_2 = hf(t_n + \frac{1}{2}h, w_n + \frac{1}{2}k_1)$$

for dolving the ordinary differential equation

$$\frac{dy}{dt} = f(t, y)$$
$$y(t_0) = y_0$$

by using a formula of the form

$$w_{n+1} = w_n + ak_1 + bk_2$$

where k_1 is defined as above,

$$k_2 = hf(t_n + \alpha h, w_n + \beta k_1)$$

and a, b, α and β are constants are deteremined. Prove that $a + b = 1$ and $b\alpha = b\beta = \frac{1}{2}$ and choose appropriate values to give the Midpoint Runge Kutta method.

[18 marks]

b) Show that the midpoint Runge Kutta method is stable.

[5 marks]

c) Use the Runge Kutta method to approximate the solutions to the following initial value problem

$$y^{'} = 1 + (t - y)^2, \;\; 2 \leq t \leq 3, \;\; y(2) = 1$$

with $h = 0.2$ with the exact solution $y(t) = t + \frac{1}{1-t}$.

[10 marks]

3. a) Derive the two step Adams-Bashforth method:

$$w_{n+1} = w_n + (\frac{3}{2}hf(t_n, w_n) - \frac{1}{2}hf(t_{n-1}, w_{n-1})),$$

and the local truncation error

$$\tau_{n+1}(h) = -\frac{5h^2}{12}y^3(\mu_n)$$

[18 marks]

b) Apply the two step Adams-Bashforth method to approximate the soluion of the initial value problem:

$$y' = ty - y, \quad (0 \leq t \leq 2) \quad y(0) = 1$$

. Using $N = 4$ steps, given that $y_1 = 0.6872$.

[15 marks]

4. a) Derive the Adams-Moulton two step method and its truncation error which is of the form

$$w_0 = \alpha_0 \quad w_1 = \alpha_1$$

$$w_{n+1} = w_n + \frac{h}{12}[5f(t_{n+1}, w_{n+1}) + 8f(t_n, w_{n1}) - f(t_{n-2}, w_{n-2})]$$

and the local truncation error

$$\tau_{n+1}(h) = -\frac{h^3}{24} y^4(\mu_n)$$

[23 marks]

b) Define the terms strongly stable, weakly stable and unstable with respect to the characteristic equation.

[5 marks]

c) Show that the Adams-Bashforth two step method is stongly stable.

[5 marks]

5. a) Given the initial value problem:

$$y' = f(t, y), \quad y(t_0) = y_0$$

and a numerical method which generates a numerical solution $(w_n)_{n=0}^{N}$, explain what it means for the method to be convergent.

[5 marks]

b) Using the 2-step Adams-Bashforth method:

$$w_{n+1} = w_n + \frac{3}{2} h f(t_n, w_n) - \frac{1}{2} h f(t_{n-1}, w_{n-1})$$

as a predictor, and the 2-step Adams-Moulton method:

$$w_{n+1} = w_n + \frac{h}{12}[5f(t_{n+1}, w_{n+1}) + 8f(t_n, w_{n1}) - f(t_{n-2}, w_{n-2})]$$

as a corrector, apply the 2-step Adams predicitor-corrector method to approximate the solution of the initial value problem

$$y' = ty^3 - y, \quad (0 \leq t \leq 2), \ y(0) = 1$$

using N=4 steps, given $y_1 = 0.5$.

[18 marks]

c) Using the predictor corrector define a bound for the error by controlling the step size.

[10 marks]

6. a) Given the Midpoint point (Runge- Kutta) method

$$w_0 = y_0$$

$$w_{i+1} = w_i + hf(x_i + \frac{h}{2}, w_i + \frac{h}{2}f(x_i, w_i))$$

Assume that the Runge Kutta method satisfies the Lipschitz condition. Then for the initial value problems

$$y^{'} = f(x, y)$$

$$y(x_0) = Y_0$$

Show that the numerical solution $\{w_n\}$ satisfies

$$\max_{a \leq x \leq b} |y(x_n) - w_n| \leq e^{(b-a)L}|y_0 - w_0| + \left[\frac{e^{(b-a)L} - 1}{L}\right] \tau(h)$$

where

$$\tau(h) = \max_{a \leq x \leq b} |\tau_n(y)|$$

If the consistency condition

$$\delta(h) \to 0 \text{ as } h \to 0$$

where

$$\delta(h) = \max_{a \leq x \leq b} |f(x, y) - F(x, y; h; f)|$$

is satisfied then the numerical solution w_n converges to $Y(x_n)$.

[18 marks]

b) Consider the differential equation

$$y^{'} - y + x - 2 = 0, \ 0 \leq x \leq 1, \ y(0) = 0.$$

Apply the midpoint method to approximate the solution at $y(0.4)$ using $h = 0.2$

[11 marks]

c) How would you improve on this result.

[4 marks]

Part II

NUMERICAL SOLUTIONS TO BOUNDARY VALUE PROBLEMS

6

BOUNDARY VALUE PROBLEMS

6.1 SYSTEMS OF EQUATIONS

An m-th order system of equation of first order Initial Value Problem can be expressed in the form

$$\begin{aligned} \tfrac{du_1}{dt} &= f_1(t, u_1, ..., u_m) \\ \tfrac{du_2}{dt} &= f_2(t, u_1, ..., u_m) \\ &. \\ &. \\ &. \\ \tfrac{du_m}{dt} &= f_m(t, u_1, ..., u_m) \end{aligned} \tag{30}$$

for $a \leq t \leq b$ with the the initial conditions

$$\begin{aligned} u_1(a) &= \alpha_1 \\ u_2(a) &= \alpha_2 \\ &. \\ &. \\ &. \\ u_m(a) &= \alpha_m. \end{aligned} \tag{31}$$

This can also be written in vector from

$$\mathbf{u}' = \mathbf{f}(t, \mathbf{u})$$

with initial conditions

$$\mathbf{u}(\mathbf{a}) = \mathbf{ff}.$$

Definition The function $f(t, u_1, ..., u_m)$ defined on the set

$$D = \{(t, u_1, .., u_m) | a \leq t \leq b, -\infty < u_i < \infty, i = 1, .., m\}$$

is said to be a **Lipschitz Condition** on D in the variables $u_1, ..., u_m$ if a constant L, the Lipschitz Constant, exists with the property that

$$|f(t, u_1, ..., u_m) - f(t, z_1, z_2, ..., z_m)| \leq L \sum_{j=1}^{m} |u_j - z_j|$$

for all $(t, u_1, ..., u_m)$ and $(t, z_1, z_2, ..., z_m)$ in D.

Theorem 6.1.1. *Suppose*

$$D = \{(t, u_1, .., u_m) | a \leq t \leq b, -\infty < u_i < \infty, i = 1, .., m\}$$

is continuous on D and satisfy a Lipschitz Condition. The system of 1st order equations subject th the initial conditions, has a unique solution $u_1(t), u_2(t), ..., u_m(t)$ *for* $a \leq t \leq b$.

Example 29

Using Euler method on the system

$$\begin{aligned} u' &= u^2 - 2uv \quad u(0) = 1 \\ v' &= tu + u^2 sinv \quad v(0) = -1 \end{aligned}$$

for $0 \leq t \leq 0.5$ and $h = 0.05$ the general Euler difference system of equations is of the form

$$\begin{aligned} \hat{u}_{i+1} &= \hat{u}_i + hf(t_i, \hat{u}_i, \hat{v}_i) \\ \hat{v}_{i+1} &= \hat{v}_i + hg(t_i, \hat{u}_i, \hat{v}_i) \end{aligned}$$

Applied the the Initial Value Problem

$$\begin{aligned} \hat{u}_{i+1} &= \hat{u}_i + 0.05(\hat{u}_i^2 - 2\hat{u}_i\hat{v}_i) \\ \hat{v}_{i+1} &= \hat{v}_i + 0.05(t_i\hat{u}_i + \hat{u}_i^2 sin(\hat{v}_i)) \end{aligned}$$

We know for $i = 0$, $\hat{u}_0 = 1$ and $\hat{v}_0 = -1$ from the initial conditions.
For i=0 we have

$$\begin{aligned} \hat{u}_1 &= \hat{u}_0 + 0.05(\hat{u}_0^2 - 2\hat{u}_0\hat{v}_0) = 1.15 \\ \hat{v}_1 &= \hat{v}_0 + 0.05(t_0\hat{u}_0 + \hat{u}_0^2 sin(\hat{v}_0)) = -1.042 \end{aligned}$$

and so forth.

6.2 Higher order equations

Definition A general mth order initial value problem

$$y^{(m)}(t) = f(t, y, ..., y^{(m-1)}) \quad a \leq t \leq b$$

with initial conditions

$$y(a) = \alpha_1, y^{'}(a) = \alpha_2, ..., y^{(m-1)}(a) = \alpha_m$$

can be converted into a system of equations as in (30) and (31)
Let $u_1(t) = y(t), u_1(t) = y^1(t), ..., u_m(t) = y^{(m-1)}(t)$. This produces the first order system of equations

$$\begin{aligned}
\frac{du_1}{dt} &= \frac{dy}{dt} = u_2 \\
\frac{du_2}{dt} &= \frac{dy'}{dt} = u_3 \\
&\vdots \\
\frac{du_{m-1}}{dt} &= \frac{dy^{(m-2)}}{dt} = u_m \\
\frac{du_m}{dt} &= \frac{dy^{(m-1)}}{dt} f_m(t, y, ..., y^{(m-1)}) = f(t, u_1, ..., u_m)
\end{aligned}$$

with initial conditions

$$\begin{aligned}
u_1(a) &= y(a) = \alpha_1 \\
u_2(a) &= y'(a) = \alpha_2 \\
&\vdots \\
u_m(a) &= y^{(m-1)}(a) = \alpha_m
\end{aligned}$$

Example 30

$$y'' + 3y' + 2y = e^t$$

with initial conditions $y(0) = 1$ and $y'(0) = 2$ can be converted to the system

$$\begin{aligned}
u' &= v & u(0) &= 1 \\
v' &= e^t - 2u - 3v & v(0) &= 2
\end{aligned}$$

the difference Euler equation is of the form

$$\begin{aligned}
\hat{u}_{i+1} &= \hat{u}_i + hv(t_i, \hat{u}_i, \hat{v}_i) \\
\hat{v}_{i+1} &= \hat{v}_i + h(e^{t_i} - 2\hat{u}_i - 3\hat{v}_i)
\end{aligned}$$

6.3 Boundary Value Problems

Consider the second order differential equation

$$y'' = f(x, y, y') \tag{32}$$

defined on an interval $a \leq x \leq b$. Here f is a function of three variables and y is an unknown. The general solution to 32 contains two arbitrary constants so in order to determine it uniquely it is necessary to impose two additional conditions on y. When one of these is given at $x = a$ and the other at $x = b$ the problem is called a boundary value problem and associated conditions are called boundary condi-

tions.
The simplest type of boundary conditions are

$$y(a) = \alpha$$

$$y(b) = \beta$$

for a given numbers α and β. However more general conditions such as

$$\lambda_1 y(a) + \lambda_2 y^{'}(a) = \alpha_1$$
$$\mu_1 y(b) + \mu_2 y^{'}(b) = \alpha_2$$

for given numbers α_i, λ_i and μ_i (i=1,2) are sometimes imposed.
Unlike Initial Value Problem whose problems are uniquely solvable boundary value problem can have no solution or many.

Example 31

The differential equation

$$y^{''} + y = 0$$

$$y_1(x) = y(x) \quad y_2(x) = y^{'}(x)$$
$$y_1^{'} = y_2$$
$$y_2^{'} = -y_1$$

It has the general solution

$$w(x) = C_1 sin(x) + C_2 cos(x)$$

where C_1, C_2 are constants.
The special solution $w(x) = sin(x)$ is the only solution that satisfies

$$w(0) = 0 \quad w(\frac{\pi}{2}) = 1$$

All functions of the form $w(x) = C_1 sin(x)$ where C_1 is an arbitrary constant, satisfies

$$w(0) = 0 \quad w(\pi) = 0$$

while there is no solution for the boundary conditions

$$w(0) = 0 \quad w(\pi) = 1.$$

◇

While we cannot state that all boundary value problem are unique we can say a few things.

6.4 Some theorems about boundary value problem

Writing the general linear subset Boundary Value Problem

$$\begin{aligned} & y'' = p(x)y' + q(x)y + g(x) \quad a < x < b \\ & A \begin{pmatrix} y(a) \\ y'(a) \end{pmatrix} + B \begin{pmatrix} y(b) \\ y'(b) \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \end{aligned} \tag{33}$$

The homogeneous problem is the case in which $g(x)$ and $\gamma_1 = \gamma_2 = 0$.

Theorem 6.4.1. *The non-homogeneous problem (33) has a unique solution $y(x)$ on $[a, b]$ for each set of given $\{g(x), \gamma_1, \gamma_2\}$ if and only if the homogeneous problem has only the trivial solutions $y(x) = 0$.*

For conditions under which the homogeneous problem (33) has only the zero solution we consider the following subset of problem

$$\begin{aligned} & y'' = p(x)y' + q(x)y + g(x) \quad a < x < b \\ & a_0 y(a) - a_1 y'(a) = \gamma_1 \\ & b_0 y(b) + b_1 y'(b) = \gamma_2 \end{aligned} \tag{34}$$

Assume the following conditions

$$\begin{aligned} & q(x) > 0 \quad a \leq x \leq b \\ & a_0, a_1 \geq 0 \quad b_0, b_1 \geq 0 \end{aligned} \tag{35}$$

$|a_1| + |a_0| \neq 0, |b_1| + |b_0| \neq 0, |a_0| + |b_0| \neq 0$ Then the homogeneous problem for (34) has only the zero solution therefore the theorem is applicable and the non-homogeneous problem has a unique solution for each set of data $\{g(x), \gamma_1, \gamma_2\}$.
The theory for a non-linear problem is far more complicated than that of a linear problem. Looking at the class of problems

$$\begin{aligned} & y'' = f(x, y, y') \quad a < x < b \\ & a_0 y(a) - a_1 y'(a) = \gamma_1 \\ & b_0 y(b) + b_1 y'(b) = \gamma_2 \end{aligned} \tag{36}$$

The function f is assumed to satisfy the following Lipschitz Condition

$$\begin{aligned} & |f(x, u_1, v_1) - f(x, u_2, v_2)| \leq K_1 |u_1 - u_2| \\ & |f(x, u_1, v_1) - f(x, u_2, v_2)| \leq K_2 |v_1 - v_2| \end{aligned} \tag{37}$$

for all points in the region

$$R = \{(x, u, v) | a \leq x \leq b, -\infty < u, v < \infty\}$$

Theorem 6.4.2. *The problem (36) assumes $f(x,u,v)$ is continuous on the region R and it satisfies the Lipschitz condition (37). In addition assume that f, on R, satisfies*

$$\frac{\partial f(x,u,v)}{\partial u} > 0 \quad \left|\frac{\partial f(x,u,v)}{\partial v}\right| \leq M$$

for some constant $M > 0$ for the boundary conditions of 36 assume that $|a_1| + |a_0| \neq 0, |b_1| + |b_0| \neq 0, |a_0| + |b_0| \neq 0$. The boundary value problem has a unique solution.

Example 32

The boundary value problem boundary value problem

$$y^{''} + e^{-xy} + sin(y^{'}) = 0 \quad 1 < x < 2$$

with $y(1) = y(2) = 0$, has

$$f(x,y,y^{'}) = -e^{-xy} - sin(y^{'})$$

Since

$$\frac{\partial f(x,y,y^{'})}{\partial y} = xe^{xy} > 0$$

and

$$\left|\frac{\partial f(x,y,y^{'})}{\partial y^{'}}\right| = |-cos(y^{'}) \leq 1$$

this problem has a unique solution. ⋄

6.5 SHOOTING METHODS

The principal of the shooting method is to change our original boundary value problem boundary value problem into 2 Initial Value Problem.

6.5.1 *Linear Shooting method*

Looking at problem class (34). We break this down into two Initial Value Problem.

$$\begin{aligned} y_1^{''} &= p(x)y_1^{'} + q(x)y_1 + r(x), \quad y_1(a) = \alpha, \quad y_1^{'}(a) = 0 \\ y_2^{''} &= p(x)y_2^{'} + q(x)y_2, \quad y_2(a) = 0, \quad y_2^{'}(a) = 1 \end{aligned} \tag{38}$$

combining these results together to get the unique solution

$$y(x) = y_1(x) + \frac{\beta - y_1(b)}{y_2(b)} y_2(x) \tag{39}$$

provided that $y_2(b) \neq 0$.

Example 33

$$y'' = 2y' + 3y - 6$$

with boundary conditions

$$y(0) = 3$$

$$y(1) = e^3 + 2$$

The exact solution is

$$y = e^{3x} + 2$$

breaking this boundary value problem into two Initial Value Problem's

$$y_1'' = 2y_1' + 3y_1 - 6 \quad y_1(a) = 3, \quad y_1'(a) = 0 \tag{40}$$

$$y_2'' = 2y_2' + 3y_2 \quad y_2(a) = 0, \quad y_2'(a) = 1 \tag{41}$$

Discretising (40)

$$y_1 = u_1 \quad y_1' = u_2$$

$$u_1' = u_2 \quad u_1(a) = 3$$

$$u_2' = 2u_2 + 3u_1 - 6 \quad u_2(a) = 0$$

using the Euler method we have the two difference equations

$$u_{1i+1} = u_{1i} + hu_{2i}$$

$$u_{2i+1} = u_{2i} + h(2u_{2i} + 3u_{1i} - 6)$$

Discretising (41)

$$y_2 = w_1 \quad y_2' = w_2$$

$$w_1' = w_2 \quad w_1(a) = 0$$

$$w_2' = 2w_2 + 3w_1 \quad w_2(a) = 1$$

using the Euler method we have the two difference equations

$$w_{1i+1} = w_{1i} + hw_{2i}$$

$$w_{2i+1} = w_{2i} + h(2w_{2i} + 3w_{1i})$$

combining all these to get our solution

$$y_i = u_{1i} + \frac{\beta - u_1(b)}{w_1(b)} w_{1i}$$

It can be said

$$|y_i - y(x_i)| \leq Kh^n \left|1 + \frac{w_{1i}}{u_{1i}}\right|$$

$O(h^n)$ is the order of the method.

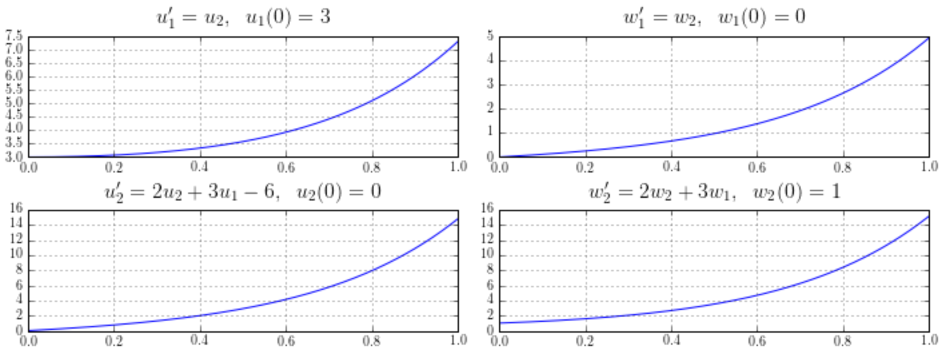


Figure 6.5.1: Python output: Shooting Method

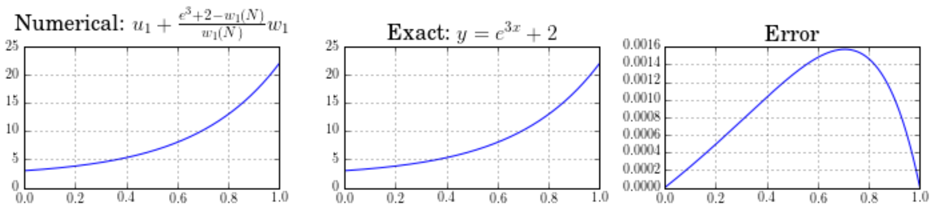


Figure 6.5.2: Python output: Shooting Method error

6.5.2 *The Shooting method for non-linear equations*

Example 34

$$y'' = -2yy' \quad y(0) = 0 \quad y(1) = 1$$

The corresponding initial value problem is

$$y'' = -2yy' \quad y(0) = 0 \quad y'(0) = \lambda \tag{42}$$

Which reduces to the first order system, letting $y_1 = y$ and $y_2 = y'$.

$$y_1' = y_2 \quad y_1(0) = 0$$

$$y_2' = -2y_1y_2 \quad y_2'(0) = \lambda$$

Taking $\lambda = 1$ and $\lambda = 2$ as the first and second guess of $y'(0)$. (42) depends on two variable x and λ. $\diamond$

How to choose λ?
Our goal is to choose λ such that.

$$F(\lambda) = y(b, \lambda) - \beta = 0$$

We use Newton's method to generate the sequence λ_k with only the initial approx λ_0. The iteration has the form

$$\lambda_k = \lambda_{k-1} - \frac{y(b, \lambda_{k-1}) - \beta}{\frac{dy}{d\lambda}(b, \lambda_{k-1})}$$

and requires knowledge of $\frac{dy}{d\lambda}(b, \lambda_{k-1})$. This present a difficulty since an explicit representation for $y(b, \lambda)$ is unknown we only know $y(b, \lambda_0)$, $y(b, \lambda_1), ..., y(b, \lambda_{k-1})$.
Rewriting our Initial Value Problem we have it so that it depends on both x and λ.

$$y''(x, \lambda) = f(x, y(x, \lambda), y'(x, \lambda)) \quad a \leq x \leq b$$

$$y(a, \lambda) = \alpha \quad y'(a, \lambda) = \lambda$$

differentiating with respect to λ and let $z(x, \lambda)$ denote $\frac{\partial y}{\partial \lambda}(x, \lambda)$ we have

$$\frac{\partial}{\partial \lambda}(y'') = \frac{\partial f}{\partial \lambda} = \frac{\partial f}{\partial y}\frac{\partial y}{\partial \lambda} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \lambda}$$

Now

$$\frac{\partial y'}{\partial \lambda} = \frac{\partial}{\partial \lambda}\frac{\partial y}{\partial x} = \frac{\partial}{\partial x}\left(\frac{\partial y}{\partial \lambda}\right) = \frac{\partial z}{\partial x} = z'$$

we have

$$z''(x, \lambda) = \frac{\partial f}{\partial y} z(x, \lambda) + \frac{\partial f}{\partial y'} z'(x, \lambda)$$

for $a \leq x \leq b$ and the boundary conditions

$$z(a, \lambda) = 0, \quad z'(a, \lambda) = 1$$

Now we have

$$\lambda_k = \lambda_{k-1} - \frac{y(b, \lambda_{k-1}) - \beta}{z(b, \lambda_{k-1})}$$

We can solve the original non-linear subset Boundary Value Problem by solving the 2 Initial Value Problem's.

Example 35

(cont.)

$$\frac{\partial f}{\partial y} = -2y' \quad \frac{\partial f}{\partial y'} = -2y$$

We now have the two Initial Value Problem's

$$y'' = -2yy' \quad y(0) = 0 \;\; y'(0) = \lambda$$

$$\begin{aligned} z'' &= \frac{\partial f}{\partial y} z(x,\lambda) + \frac{\partial f}{\partial y'} z'(x,\lambda) \\ &= -2y'z - 2yz' \quad z(0) = 0 \;\; z'(0) = 1 \end{aligned}$$

Discretising we let $y_1 = y$, $y_2 = y'$, $y_3 = z$ and $y_4 = z'$.

$$\begin{aligned} y_1' = y_2 \quad &: \quad y_1(0) = 0 \\ y_2' = -2y_1y_2 \quad &: \quad y_2(0) = \lambda_k \\ z_1' = z_2 \quad &: \quad z_1(0) = 0 \\ z_2' = -2z_1y_2 - 2y_1z_2 \quad &: \quad y_2(0) = 1 \end{aligned}$$

with

$$\lambda_k = \lambda_{k-1} - \frac{y_1(b) - \beta}{y_3(b)}$$

Then solve using a one step method. ◇

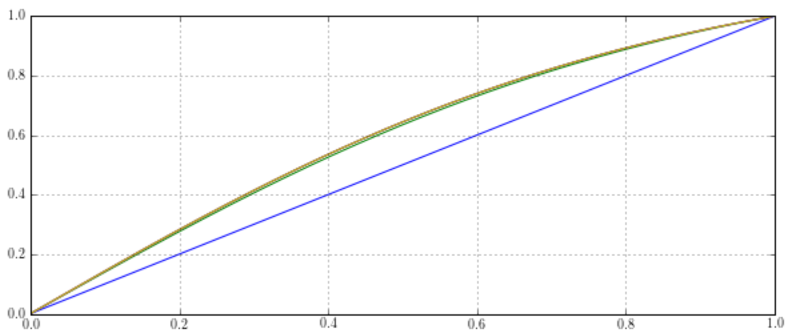


Figure 6.5.3: Python output: Nonlinear Shooting Method

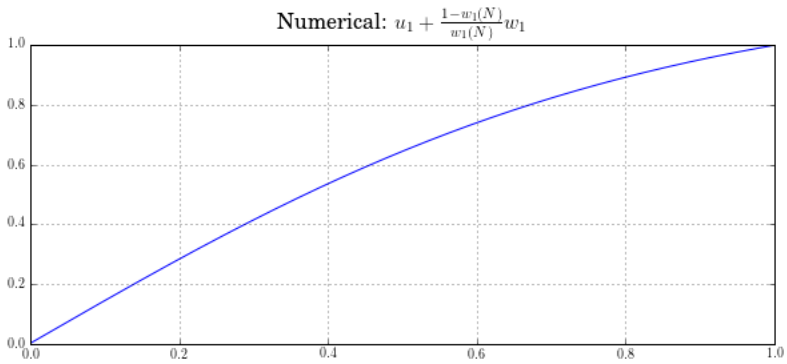


Figure 6.5.4: Python output: Nonlinear Shooting Method result

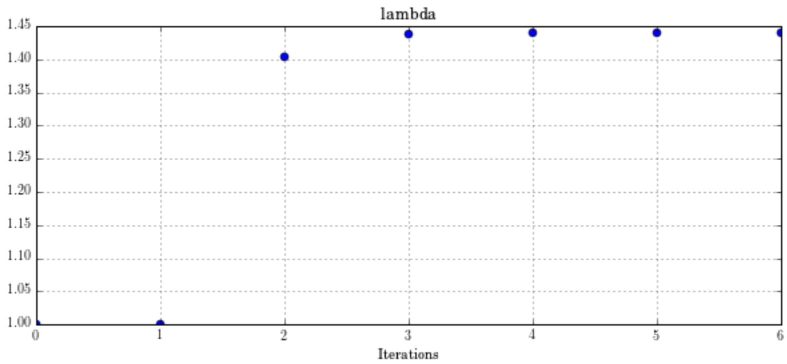


Figure 6.5.5: Python output: Nonlinear Shooting Method λ

6.6 Finite Difference Method

Each finite difference operator can be derived from Taylor expansion. Once again looking at a linear second order differential equation

$$y'' = p(x)y' + q(x)y + r(x)$$

on $[a, b]$ subject to boundary conditions

$$y(a) = \alpha \quad y(b) = \beta$$

As with all cases we divide the area into even spaced mesh points

$$x_0 = a, \quad x_N = b \quad x_i = x_0 + ih \quad h = \frac{b-a}{N}$$

We now replace the derivatives $y'(x)$ and $y''(x)$ with the centered difference approximations

$$y'(x) = \frac{1}{2h}(y(x_{i+1}) - y(x_{i-1})) - \frac{h^2}{12}y^3(\xi_i)$$

$$y''(x) = \frac{1}{h^2}(y(x_{i+1}) - 2y(x_i) + y(x_{i-1})) - \frac{h^2}{6}y^4(\mu_i)$$

for some $x_{i-1} \leq \xi_i \mu_i \leq x_{i+1}$ for i=1,...,N-1.
We now have the equation

$$\frac{1}{h^2}(y(x_{i+1}) - 2y(x_i) + y(x_{i-1})) = p(x_i)\frac{1}{2h}(y(x_{i+1}) - y(x_{i-1})) + q(x_i)y(x_i) + r(x_i)$$

This is rearranged such that we have all the unknown together,

$$\left(1 + \frac{hp(x_i)}{2}\right) y(x_{i-1}) - (2 + h^2 q(x_i))y(x_i) + \left(1 - \frac{hp(x_i)}{2}\right) y(x_{i+1}) = h^2 r(x_i)$$

for $i = 1, .., N-1$.
Since the values of $p(x_i), q(x_i)$ and $r(x_i)$ are known it represents a linear algebraic equation involving $y(x_{i-1}), y(x_i), y(x_{i+1})$.
This produces a system of $N-1$ linear equations with $N-1$ unknowns $y(x_1), ..., y(x_{N-1})$.
The first equation corresponding to $i = 1$ simplifies to

$$-(2 + h^2 q(x_1))y(x_1) + \left(1 - \frac{hp(x_1)}{2}\right) y(x_2) = h^2 r(x_1) - \left(1 + \frac{hp(x_1)}{2}\right) \alpha$$

because of the boundary condition $y(a) = \alpha$, and for $i = N-1$

$$\left(1 + \frac{hp(x_{N-1})}{2}\right) y(x_{N-2}) - (2 + h^2 q(x_{N-1}))y(x_{N-1}) = h^2 r(x_{N-1}) - \left(1 - \frac{hp(x_{N-1})}{2}\right) \beta$$

because $y(b) = \beta$.
The values of y_i, $(i = 1, ..., N-1)$ can therefore be found by solving the tridiagonal system

$$A\mathbf{y} = \mathbf{b}$$

where

$$A =$$

$$\begin{bmatrix} -(2 + h^2 q(x_1)) & \left(1 - \frac{hp(x_1)}{2}\right) & 0 & . & 0 \\ \left(1 + \frac{hp(x_2)}{2}\right) & -(2 + h^2 q(x_2)) & \left(1 - \frac{hp(x_2)}{2}\right) & 0 & . \\ 0 & . & . & 0 & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & 0 & \left(1 + \frac{hp(x_{N-2})}{2}\right) & -(2 + h^2 q(x_{N-2})) & \left(1 - \frac{hp(x_{N-2})}{2}\right) \\ . & 0 & 0 & \left(1 + \frac{hp(x_{N-1})}{2}\right) & -(2 + h^2 q(x_{N-1})) \end{bmatrix}$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_{N-2} \\ y_{N-1} \end{pmatrix} b = \begin{pmatrix} h^2 r_1 - \left(1 + \frac{hp_1}{2}\right)\alpha \\ h^2 r_2 \\ . \\ . \\ h^2 r_{N-2} \\ h^2 r_{N-1} - \left(1 - \frac{hp_1}{2}\right)\beta \end{pmatrix}$$

Example 36

Looking at the simple case

$$\frac{d^2y}{dx^2} = 4y, \quad y(0) = 1.1752, \quad y(1) = 10.0179.$$

Our difference equation is

$$\frac{1}{h^2}(y(x_{i+1}) - 2y(x_i) + y(x_{i-1})) = 4y(x_i) \quad i = 1, .., N-1$$

dividing $[0, 1]$ into 4 subintervals we have $h = \frac{1-0}{4}$

$$x_i = x_0 + ih = 0 + i(0.25)$$

In this simple example $q(x) = 4$, $p(x) = 0$ and $r(x) = 0$. Rearranging the equation we have

$$\frac{1}{h^2}(y(x_{i+1})) - \left(\frac{2}{h^2} + 4\right) y(x_i) + \frac{1}{h^2}(y(x_{i-1})) = 0$$

multiplying across by h^2

$$y(x_{i+1}) - (2 + 4h^2)y(x_i) + (y(x_{i-1})) = 0$$

with the boundary conditions $y(x_0) = 1.1752$ and $y(x_4) = 10.0179$. Our equations are of the form

$$\begin{aligned} y(x_2) - 2.25y(x_1) &= -1.1752 \\ y(x_3) - 2.25y(x_2) + y(x_1) &= 0 \\ -2.25y(x_3) + y(x_2) &= -10.0179 \end{aligned}$$

Putting this into matrix form

$$\begin{pmatrix} -2.25 & 1 & 0 \\ 1 & -2.25 & 1 \\ 0 & 1 & -2.25 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1.1752 \\ 0 \\ -10.0179 \end{pmatrix}$$

x	y	Exact $sinh(2x+1)$
0	1.1752	1.1752
0.25	2.1467	2.1293
0.5	3.6549	3.6269
0.75	6.0768	6.0502
1.0	10.0179	10.0179

◇

Example 37

Looking at a more involved boundary value problem

$$y'' = xy' - 3y + e^x \quad y(0) = 1 \;\; y(1) = 2$$

Let N=5 then $h = \frac{1-0}{5} = 0.2$. The difference equation is of the form

$$\frac{1}{h^2}(y(x_{i+1}) - 2y(x_i) + y(x_{i-1})) = x_i\frac{1}{2h}(y(x_{i+1}) - y(x_{i-1})) - 3y(x_i) + e^{x_i}$$

Re arranging and putting $h = 0.2$

$$(1 + \frac{0.2(x_i)}{2})y(x_{i-1}) - (1.88)y(x_i) + (1 - \frac{0.2(x_i)}{2})y(x_{i+1}) = 0.04e^{x_i}$$

In matrix form this is

$$\begin{pmatrix} -1.88 & 0.98 & 0 & 0 \\ 1.04 & -1.88 & 0.96 & 0 \\ 0 & 1.06 & -1.88 & 0.94 \\ 0 & 0 & 1.08 & -1.88 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0.04e^{0.2} - 1.02 \\ 0.04e^{0.4} \\ 0.04e^{0.6} \\ 0.04e^{0.8} - 1.84 \end{pmatrix}$$

$y_1 = 1.4651$, $y_2 = 1.8196$, $y_3 = 2.0283$ and $y_4 = 2.1023$.

SOLVING A TRI-DIAGONAL SYSTEM

To solve a tri-diagonal system we can use the method discussed in the approximation theory.

Part III

NUMERICAL SOLUTIONS TO PARTIAL DIFFERENTIAL EQUATIONS

7

PARTIAL DIFFERENTIAL EQUATIONS

7.1 Introduction

Partial Differential Equations (PDE), occur frequently in maths, natural science and engineering.
PDE's are problems involving rates of change of functions of several variables.
The following involve 2 independent variables:

$$\begin{aligned} -\nabla^2 u = -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} &= f(x,y) \quad \text{Poisson Eqn} \\ \frac{\partial u}{\partial t} + v\frac{\partial u}{\partial x} &= 0 \quad \text{Advection Eqn} \\ \frac{\partial u}{\partial t} - D\frac{\partial^2 u}{\partial x^2} &= 0 \quad \text{Heat Eqn} \\ \frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} &= 0 \quad \text{Wave Equation} \end{aligned}$$

Here v, D, c are real positive constants. In these cases x, y are the space coordinates and t, x are often viewed as time and space coordinates, respectively.
These are only examples and do not cover all cases. In real occurrences PDE's usually have 3 or 4 variables.

7.2 PDE Classification

PDE's in two independent variables x and y have the form

$$\Phi\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial^2 u}{\partial x^2}, ...\right) = 0$$

where the symbol Φ stands for some functional relationship.
As we saw with BVP this is too general a case so we must define new classes of the general PDE.

Definition The order of a PDE is the order of the highest derivative that appears.
ie Poisson is 2nd order, Advection eqn is 1st order. $\circ$

Most of the mathematical theory of PDE's concerns linear equations of first or second order.
After order and linearity (linear or non-linear), the most important classification scheme for PDE's involves geometry.
Introducing the ideas with an example:

Example 38

$$\alpha(t,x)\frac{\partial u}{\partial t} + \beta\frac{\partial u}{\partial x} = \gamma(t,x) \tag{43}$$

A solution $u(t,x)$ to this PDE defines a surface $\{t,x,u(t,x)\}$ lying over some region of the (t,x)-plane.
Consider any smooth path in the (t,x)-plane lying below the solution $\{t,x,u(t,x)\}$.
Such a path has a parameterization $(t(s),x(s))$ where the parameter s measures progress along the path.
What is the rate of change $\frac{du}{ds}$ of the solution as we travel along the path $(t(s),x(s))$.
The chain rule provides the answer

$$\frac{dt}{ds}\frac{\partial u}{\partial t} + \frac{dx}{ds}\frac{\partial u}{\partial x} = \frac{du}{ds} \tag{44}$$

Equation (44) holds for an arbitrary smooth path in the (t,x)-plane. Restricting attention to a specific family of paths leads to a useful observation: When

$$\frac{dt}{ds} = \alpha(t,x) \text{ and } \frac{dx}{ds} = \beta(t,x) \tag{45}$$

the simultaneous validity of (43) and (44) requires that

$$\frac{du}{ds} = \gamma(t,x). \tag{46}$$

Equation (46) defines a family of curves $(t(s),x(s))$ called characteristic curves in the plane (t,x).
Equation (46) is an ode called the characteristic equation that the solution must satisfy along only the characteristic curve.
Thus the original PDE collapses to an ODE along the characteristic curves. Characteristic curves are paths along which information about the solution to the PDE propagates from points where the initial value or boundary values are known. ⋄

Consider a second order PDE having the form

$$\alpha(x,y)\frac{\partial^2 u}{\partial x^2} + \beta(x,y)\frac{\partial^2 u}{\partial x \partial y} + \gamma(x,y)\frac{\partial^2 u}{\partial y^2} = \Psi(x,y,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}) \tag{47}$$

Along an arbitrary smooth curve $(x(s), y(s))$ in the (x, y)-plane, the gradient $\left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right)$ of the solution varies according to the chain rule:

$$\frac{dx}{ds}\frac{\partial^2 u}{\partial y \partial x} + \frac{dy}{ds}\frac{\partial^2 u}{\partial y \partial x} = \frac{d}{ds}\left(\frac{\partial u}{\partial x}\right)$$

$$\frac{dx}{ds}\frac{\partial^2 u}{\partial x \partial y} + \frac{dy}{ds}\frac{\partial^2 u}{\partial y^2} = \frac{d}{ds}\left(\frac{\partial u}{\partial y}\right)$$

if the solution $u(x, y)$ is continuously differentiable then these relationships together with the original PDE yield the following system:

$$\begin{pmatrix} \alpha & \beta & \gamma \\ \frac{dx}{ds} & \frac{dy}{ds} & 0 \\ 0 & \frac{dx}{ds} & \frac{dy}{ds} \end{pmatrix} \begin{pmatrix} \frac{\partial^2 u}{\partial x^2} \\ \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial y^2} \end{pmatrix} = \begin{pmatrix} \Psi \\ \frac{d}{ds}\left(\frac{\partial u}{\partial x}\right) \\ \frac{d}{ds}\left(\frac{\partial u}{\partial y}\right) \end{pmatrix} \tag{48}$$

By analogy with the first order case we determine the characteristic curves bu where the PDE is redundant with the chain rule. This occurs when the determinant of the matrix in (48) vanishes that is when

$$\alpha\left(\frac{dy}{ds}\right)^2 - \beta\left(\frac{dy}{ds}\right)\left(\frac{dx}{ds}\right) + \gamma\left(\frac{dx}{ds}\right)^2 = 0$$

eliminating the parameter s reduces this equation to the equivalent condition

$$\alpha\left(\frac{dy}{dx}\right)^2 - \beta\left(\frac{dy}{dx}\right) + \gamma = 0$$

Formally solving this quadratic for $\frac{dy}{dx}$, we find

$$\frac{dy}{dx} = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

This pair of ODE's determine the characteristic curves. From this equation we divide into 3 classes each defined with respect to $\beta^2 - 4\alpha\gamma$.

1. **HYPERBOLIC**
   $\beta^2 - 4\alpha\gamma > 0$ This gives two families of real characteristic curves.

2. **PARABOLIC**
   $\beta^2 - 4\alpha\gamma = 0$ This gives exactly one family of real characteristic curves.

3. **ELLIPTIC**
   $\beta^2 - 4\alpha\gamma < 0$ This gives no real characteristic equations.

Example 39

The wave equation

$$c^2 \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$$

now equating this with our formula for the characteristics we have

$$\frac{dt}{dx} = \frac{0 \pm \sqrt{0 + 4c^2}}{2} = \pm c$$

this implies that the characteristics are $x + ct = const$ and $x - ct = const$. This means that the effects travel along the characteristics.

Laplace equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

from this we have $-4(1)(1) < 0$ which implies it is elliptic.

This means that information at one point affects all other points.

Heat equation

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial u}{\partial t} = 0$$

from this we have $\beta^2 - 4\alpha\gamma = 0$ this implies that the equation is parabolic thus we have

$$\frac{\partial t}{\partial x} = 0$$

$\diamond$

We can also state that hyperbolic and parabolic are Boundary value problems and initial value problems. While, elliptic problems are boundary value problems.

7.3 DIFFERENCE OPERATORS

Through out this chapter we will use U to denote the exact solution and w to denote the numerical (approximate) solution.
1-D difference operators

$$D^+U_i = \frac{U_{i+1} - U_i}{h_{i+1}} \quad \text{Forward}$$
$$D^-U_i = \frac{U_i - U_{i-1}}{h_i} \quad \text{Backward}$$
$$D^0U_i = \frac{U_{i+1} - U_{i-1}}{x_{i+1} - x_{i-1}} \quad \text{Centered}$$

For 2-D Differences Schemes is similar when dealing with the x-direction we hold the y-direction constant and then dealing with the y-direction hold the x-direction constant.

$$D_x^+U_{ij} = \frac{U_{i+1j} - U_{ij}}{x_{i+1} - x_i} \quad \text{Forward in the x-direction}$$
$$D_y^+U_{ij} = \frac{U_{ij+1} - U_{ij}}{y_{i+1-y_i}} \quad \text{Forward in the y-direction}$$
$$D_x^-U_{ij} = \frac{U_{ij} - U_{i-1j}}{x_i - x_{i-1}} \quad \text{Backward in the x direction}$$
$$D_y^-U_{ij} = \frac{U_{ij} - U_{ij-1}}{y_i - y_{i-1}} \quad \text{Backward in the y direction}$$
$$D_x^0U_{ij} = \frac{U_{i+1j} - U_{i-1j}}{x_{i+1} - x_{i-1}} \quad \text{Centered in the x direction}$$
$$D_y^0U_{ij} = \frac{U_{ij+1} - U_{ij-1}}{y_{i+1} - y_{i-1}} \quad \text{Centered in the y direction}$$

Second derivatives

$$\delta_x^2U_{ij} = \frac{2}{x_{i+1} - x_{i-1}}\left(\frac{U_{i+1j} - U_{ij}}{x_{i+1} - x_i} - \frac{U_{ij} - U_{i-1j}}{x_i - x_{i-1}}\right) \quad \text{Centered in } x \text{ direction}$$
$$\delta_y^2U_{ij} = \frac{2}{y_{i+1} - y_{i-1}}\left(\frac{U_{ij+1} - U_{ij}}{y_{i+1} - y_i} - \frac{U_{ij} - U_{ij-1}}{y_i - y_{i-1}}\right) \quad \text{Centered in } y \text{ direction}$$

8

PARABOLIC EQUATIONS

We will look at the Heat equation as our sample parabolic equation.

$$\frac{\partial U}{\partial T} = K\frac{\partial^2 U}{\partial X^2} \text{ on } \Omega$$

and

$$U = g(x, y) \text{ on the boundary } \delta\Omega$$

this can be transformed without loss of generality by a non-dimensional transformation to

$$\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2} \tag{49}$$

with the domain

$$\Omega = \{(t, x) |\ 0 \leq t, 0 \leq x \leq 1\}.$$

8.1 EXAMPLE HEAT EQUATION

In this case we look at a rod of unit length with each end in ice.
The rod is heat insulated along its length so that temp changes occur through heat conduction along its length and heat transfer at its ends, where w denotes temp.
Given that the ends of the rod are kept in contact with ice and the initial temp distribution is non dimensional form is

1. $U = 2x$ for $0 \leq x \leq \frac{1}{2}$
2. $U = 2(1 - x)$ for $\frac{1}{2} \leq x \leq 1$

In other words we are seeking a numerical solution of

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

which satisfies

1. $U = 0$ at $x = 0$ and $x = 1$ for all $t > 0$ (the boundary condition)
2. $U = 2x$ for $0 \leq x \leq \frac{1}{2}$ for $t = 0$
   $U = 2(1 - x)$ for $\frac{1}{2} \leq x \leq 1$ for $t = 0$ (the initial condition).

Due to the initial conditions the problem is symmetric with respect to $x = 0.5$. To illustrate the implementation and limitations of the explicit, implicit and Crank-Nicholson methods we will numerically solve the Heat Equation of the rod for three different values of r:

Case 1 Let $h = \frac{1}{10}$ and $k = \frac{1}{1000}$ so that $r = \frac{k}{h^2} = \frac{1}{10}$;

Case 2 Let $h = \frac{1}{10}$ and $k = \frac{1}{200}$ so that $r = \frac{k}{h^2} = \frac{1}{2}$;

Case 3 Let $h = \frac{1}{10}$ and $k = \frac{1}{100}$ so that $r = \frac{k}{h^2} = 1$.

8.2 An explicit method for the heat eqn

The explicit Forwards Time Centered Space (FTCS) equation difference equation of the differential equation (49) is

$$\frac{w_{ij+1} - w_{ij}}{t_{j+1} - t_j} = \frac{w_{i+1j} - 2w_{ij} + w_{i-1j}}{h^2}, \tag{50}$$

$$\frac{w_{ij+1} - w_{ij}}{k} = \frac{w_{i+1j} - 2w_{ij} + w_{i-1j}}{h^2}$$

when approaching this we have divided up the area into two uniform meshes one in the x direction and the other in the t-direction. We define $t_j = jk$ where k is the step size in the time direction.
We define $x_i = ih$ where h is the step size in the space direction.
w_{ij} denotes the numerical approximation of U at (x_i, t_j).
Rearranging the equation we get

$$w_{ij+1} = rw_{i-1j} + (1 - 2r)w_{ij} + rw_{i+1j} \tag{51}$$

where $r = \frac{k}{h^2}$.
This gives the formula for the unknown term w_{ij+1} at the $(ij + 1)$ mesh points in terms of all x_i along the jth time row.
Hence we can calculate the unknown pivotal values of w along the first row $t = k$ or $j = 1$ in terms of the known boundary conditions.
This can be written in matrix form:

$$\mathbf{w}_{j+1} = A\mathbf{w}_j + \mathbf{b}_j$$

for which A is an $N - 1 \times N - 1$ matrixL

$$\begin{pmatrix} 1-2r & r & 0 & . & . & . & . \\ r & 1-2r & r & 0 & . & . & . \\ 0 & r & 1-2r & r & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & r & 1-2r & r \\ . & . & . & . & . & r & 1-2r \end{pmatrix}$$

where $r = \frac{k}{h_x^2} > 0$, $\mathbf{w}_j$ is

$$\begin{pmatrix} w_{1j} \\ w_{2j} \\ . \\ w_{N-2j} \\ w_{N-1j} \end{pmatrix}$$

and $\mathbf{b}_j$ is

$$\begin{pmatrix} rw_{0j} \\ 0 \\ . \\ 0 \\ rw_{Nj} \end{pmatrix}.$$

It is assumed that the boundary values w_{0j} and w_{Nj} are known for $j = 1, 2, ...$, and w_{i0} is the initial condition.

Example 40

8.2.0.1 *Explicit FTCS method for the Heat Equation with $r = \frac{1}{10}$*

Let $h = \frac{1}{5}$ and $k = \frac{1}{250}$ so that $r = \frac{k}{h^2} = \frac{1}{10}$ difference equation (51) becomes

$$w_{ij+1} = \frac{1}{10}(w_{i-1j} + 8w_{ij} + w_{i+1j}).$$

This can be written in matrix form

$$\begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ w_{3j+1} \\ w_{4j+1} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 & 0 & 0 \\ 0.1 & 0.8 & 0.1 & 0 \\ 0 & 0.1 & 0.8 & 0.1 \\ 0 & 0 & 0.1 & 0.8 \end{pmatrix} \begin{pmatrix} w_{1j} \\ w_{2j} \\ w_{3j} \\ w_{4j} \end{pmatrix} + 0.1 \begin{pmatrix} w_{0j} \\ 0 \\ 0 \\ w_{5j} \end{pmatrix}.$$

Figure 8.2.1 shows a graphical representation of the matrix.

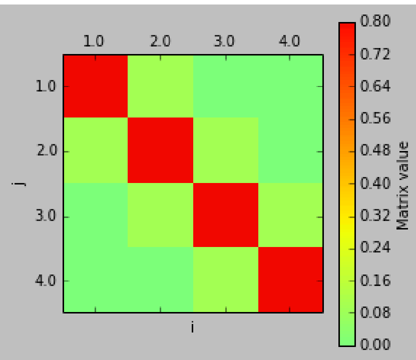


Figure 8.2.1: Graphical representation of the matrix A for $r = \frac{1}{10}$.

To solve for w_{21} we have

$$w_{21} = \frac{1}{10}(w_{10} + 8w_{20} + w_{30}) = \frac{1}{10}(0.4 + 8 \times 0.8 + 0.8) = 0.76.$$

Table 4 shows the initial condition and one time step for the Heat Equation.

j/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{250}$	0	0.4	0.76	0.76	0.4	0.0

Table 4: The explicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 1 time step.

Figure 8.2.2 shows the explicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 10 time steps each represented by a different line.

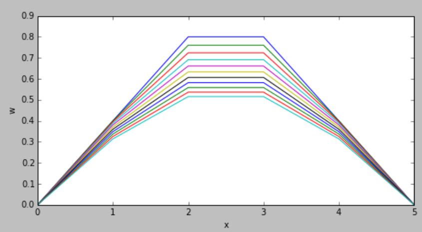


Figure 8.2.2: The explicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 10 time steps each represented by a different line.

Figure 8.2.3 shows the explicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 10 time steps as a colour plot.

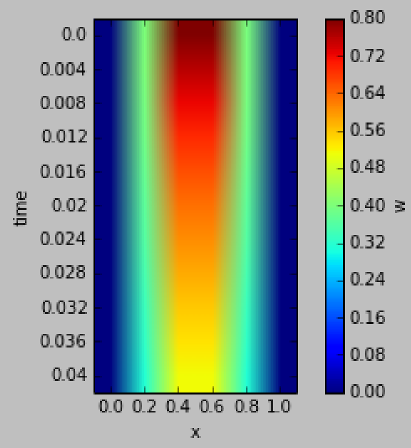


Figure 8.2.3: The colour plot of the explicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$.

The forward time and centered space numerical solution for the Heat Equation shown in Figures 8.2.3 and 8.2.2 tends to 0 in a monotonic fashion as time progresses for $r = \frac{1}{10}$.

Example 41

8.2.0.2 *Explicit FTCS method for the Heat Equation with $r = \frac{1}{2}$*

Let $h = \frac{1}{5}$ and $k = \frac{1}{50}$ so that $r = \frac{k}{h^2} = \frac{1}{2}$ difference equation (51) becomes

$$w_{ij+1} = \frac{1}{2}(w_{i-1j} + w_{i+1j}).$$

This can be written in matrix form

$$\begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ w_{3j+1} \\ w_{4j+1} \end{pmatrix} = \begin{pmatrix} 0.0 & 0.5 & 0 & 0 \\ 0.5 & 0.0 & 0.5 & 0 \\ 0 & 0.5 & 0.0 & 0.5 \\ 0 & 0 & 0.5 & 0.0 \end{pmatrix} \begin{pmatrix} w_{1j} \\ w_{2j} \\ w_{3j} \\ w_{4j} \end{pmatrix} + 0.5 \begin{pmatrix} w_{0j} \\ 0 \\ 0 \\ w_{Nj} \end{pmatrix}.$$

Figure 8.2.4 is a graphical representation of the matrix A.

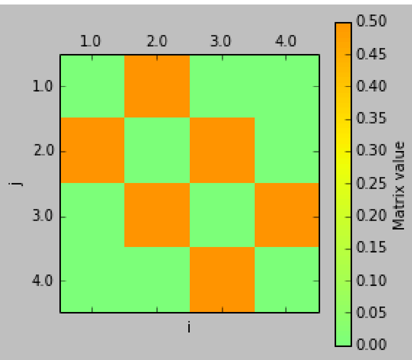


Figure 8.2.4: Graphical representation of the matrix A for $r = \frac{1}{2}$.

Table 8.2.4 shows the explicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 1 time step.

t/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{50}$	0	0.4	0.6	0.6	0.4	0.0

Table 5: The explicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 1 time step.

Figure 8.2.5 shows the explicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 10 time steps each represented by a different line.

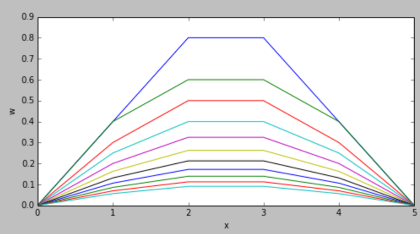


Figure 8.2.5: The explicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 10 time steps each represented by a different line.

Figure 8.2.6 shows the explicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 10 time steps as a colour plot.

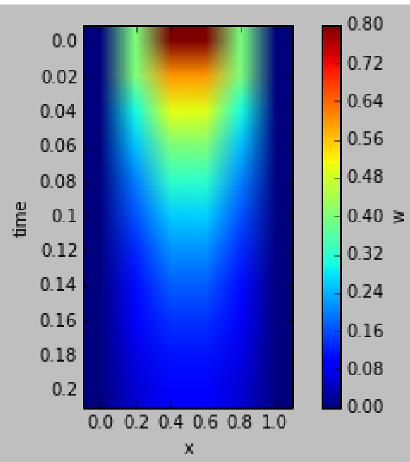


Figure 8.2.6: The colour plot of the explicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$.

The choice of $r = \frac{1}{2}$ gives an acceptable approximation to the solution of the Heat Equation as shown in Figures 8.2.5 and 8.2.6.

Example 42

8.2.0.3 *Explicit FTCS method for the Heat Equation with $r = 1$*

Let $h = \frac{1}{5}$ and $k = \frac{1}{25}$ so that $r = \frac{k}{h^2} = 1$ difference equation (51) becomes

$$w_{ij+1} = w_{i-1j} - w_{ij} + w_{i+1j}.$$

This can be written in matrix form

$$\begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ w_{3j+1} \\ w_{4j+1} \end{pmatrix} = \begin{pmatrix} -1.0 & 1.0 & 0 & 0 \\ 1.0 & -1.0 & 1.0 & 0 \\ 0 & 1.0 & -1.0 & 1.0 \\ 0 & 0 & 1.0 & -1.0 \end{pmatrix} \begin{pmatrix} w_{1j} \\ w_{2j} \\ w_{3j} \\ w_{4j} \end{pmatrix}$$

$$+ \begin{pmatrix} w_{0j} \\ 0 \\ 0 \\ w_{5j} \end{pmatrix}.$$

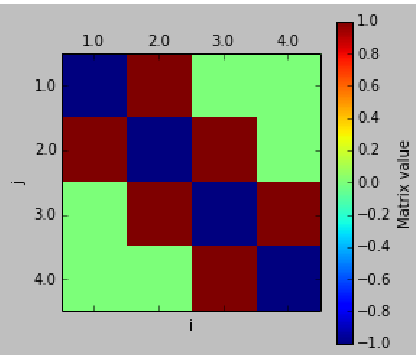


Figure 8.2.7: Graphical representation of the matrix A for $r = 1$.

t/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{25}$	0	0.4	0.4	0.4	0.4	0.0
$\frac{2}{25}$	0	0.0	0.4	0.4	0.0	0.0
$\frac{3}{25}$	0	0.4	0.0	0.0	0.4	0.0

Table 6: The explicit numerical solution w of the Heat Equation for $r = 1$ for 3 time steps.

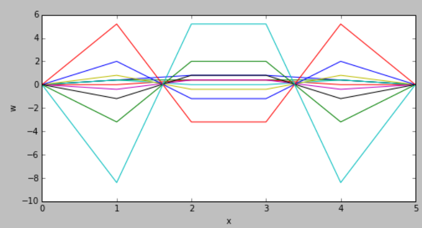


Figure 8.2.8: The explicit numerical solution w of the Heat Equation for $r = 1$ for 10 time steps each represented by a different line.

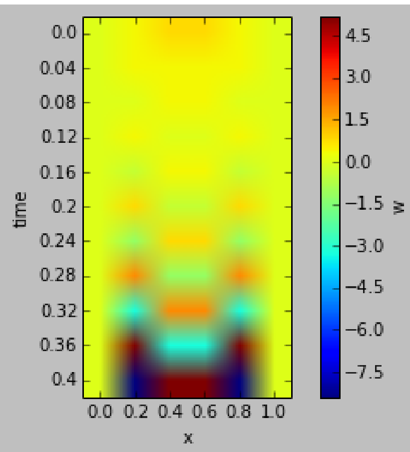


Figure 8.2.9: The colorplot of The explicit numerical solution w of the Heat Equation for $r = 1$.

Considered as a solution to the Heat Equation this is meaningless although it is the correct solution of the difference equation with respect to the initial conditions and the boundary conditions.

8.3 An implicit (BTCS) method for the Heat Equation

The implicit Backward Time Centered Space (BTCS) difference equation of the differential Heat equation (49) is

$$\frac{w_{ij+1} - w_{ij}}{k} = \frac{w_{i+1j+1} - 2w_{ij+1} + w_{i-1j+1}}{h^2} \tag{52}$$

when approaching this we have divided up the area into two uniform meshes one in the x direction and the other in the t-direction. We define $t_j = jk$ where k is the step size in the time direction.
We define $x_i = ih$ where h is the step size in the space direction.
w_{ij} denotes the numerical approximation of U at (x_i, t_j).
Rearranging the equation we get

$$-rw_{i-1j+1} + (1 + 2r)w_{ij+1} - rw_{i+1j+1} = w_{ij} \tag{53}$$

where $r = \frac{k}{h^2}$.
This gives the formula for the unknown term w_{ij+1} at the $(ij + 1)$ mesh points in terms of terms along the jth time row.
Hence we can calculate the unknown pivotal values of w along the first row $t = k$ or $j = 1$ in terms of the known boundary conditions.
This can be written in matrix from

$$A\mathbf{w}_{j+1} = \mathbf{w}_j + \mathbf{b}_{j+1}$$

for which A is

$$\left(\begin{array}{ccccccc} 1+2r & -r & 0 & . & . & . & . \\ -r & 1+2r & -r & 0 & . & . & . \\ 0 & -r & 1+2r & -r & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & -r & 1+2r & -r \\ . & . & . & . & . & -r & 1+2r \end{array}\right)$$

where $r = \frac{k}{h_x^2} > 0$, $\mathbf{w}_j$ is

$$\left(\begin{array}{c} w_{1j} \\ w_{2j} \\ . \\ w_{N-2j} \\ w_{N-1j} \end{array}\right)$$

and $\mathbf{b}_{j+1}$ is

$$\left(\begin{array}{c} rw_{0j+1} \\ 0 \\ . \\ 0 \\ rw_{Nj+1} \end{array}\right).$$

It is assumed that the boundary values w_{0j} and w_{Nj} are known for $j = 1, 2, ...,$ and w_{i0} is the initial condition.

8.3.1 *Example implicit (BTCS) for the Heat Equation*

In this case we look at a rod of unit length with each end in ice.
The rod is heat insulated along its length so that temperature changes occur through heat conduction along its length and heat transfer at its ends, where w denotes temperature.
Given that the ends of the rod are kept in contact with ice and the initial temperature distribution is non dimensional form is

1. $U = 2x$ for $0 \leq x \leq \frac{1}{2}$
2. $U = 2(1-x)$ for $\frac{1}{2} \leq x \leq 1$

In other words we are seeking a numerical solution of

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

which satisfies

1. $U = 0$ at $x = 0$ for all $t > 0$ (the boundary condition),
2. $U = 2x$ for $0 \leq x \leq \frac{1}{2}$ for $t = 0$,
   $U = 2(1-x)$ for $\frac{1}{2} \leq x \leq 1$ for $t = 0$ (the initial condition).

Example 43

8.3.1.1 *Implicit (BTCS) method for the Heat Equation* $r = \frac{1}{10}$

Let $h = \frac{1}{5}$ and $k = \frac{1}{250}$ so that $r = \frac{k}{h^2} = \frac{1}{10}$ difference equation (51) becomes

$$\frac{1}{10}(-w_{i-1j+1} + (12)w_{ij+1} - w_{i+1j+1}) = w_{ij}$$

This can be written in matrix form

$$\begin{pmatrix} 1.2 & -0.1 & 0 & 0 \\ -0.1 & 1.2 & -0.1 & 0 \\ 0 & -0.1 & 1.2 & -0.1 \\ 0 & 0 & -0.1 & 1.2 \end{pmatrix} \begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ w_{3j+1} \\ w_{4j+1} \end{pmatrix} = \begin{pmatrix} w_{1j} \\ w_{2j} \\ w_{3j} \\ w_{4j} \end{pmatrix} + 0.1 \begin{pmatrix} w_{0j+1} \\ 0 \\ 0 \\ w_{5j+1} \end{pmatrix}.$$

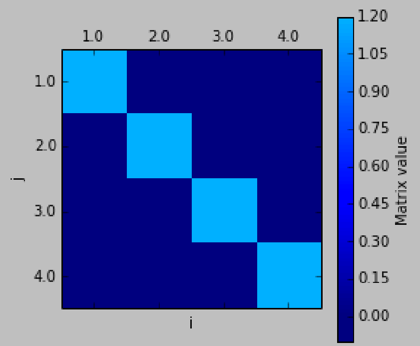


Figure 8.3.1: Graphical representation of the matrix A for $r = \frac{1}{10}$.

To solve we need to invert the matix, to get

$$\mathbf{w}_{j+1} = A^{-1}(\mathbf{w}_j + \mathbf{b}_j)$$

j/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{250}$	0.	0.39694656	0.76335878	0.76335878	0.39694656	0.

Table 7: The implicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 1 time step.

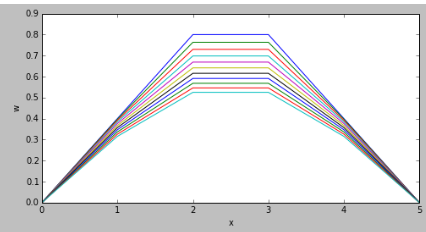


Figure 8.3.2: The implicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 10 time steps each represented by a different line.

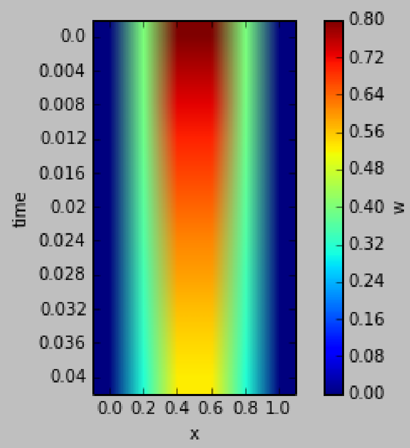


Figure 8.3.3: The colorplot of The implicit numerical solution w of the Heat Equation for $r = \frac{1}{10}$.

Example 44

8.3.1.2 *Implicit (BTCS) method for the Heat Equation for $r = \frac{1}{2}$*

Let $h = \frac{1}{5}$ and $k = \frac{1}{50}$ so that $r = \frac{k}{h^2} = \frac{1}{2}$ difference equation (51) becomes

$$\frac{1}{2}(-w_{i-1j+1} + (4)w_{ij+1} - w_{i+1j+1}) = w_{ij}$$

This can be written in matrix form

$$\begin{pmatrix} 2 & -0.5 & 0 & 0 \\ -0.5 & 2 & -0.5 & 0 \\ 0 & -0.5 & 2 & -0.5 \\ 0 & 0 & -0.5 & 2 \end{pmatrix} \begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ w_{3j+1} \\ w_{4j+1} \end{pmatrix} = \begin{pmatrix} w_{1j} \\ w_{2j} \\ w_{3j} \\ w_{4j} \end{pmatrix}.$$

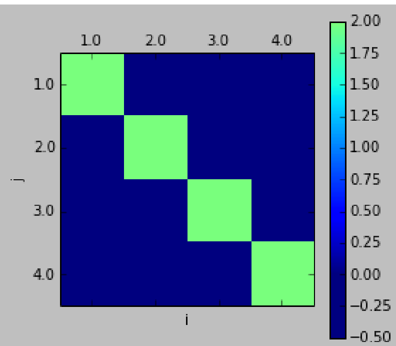


Figure 8.3.4: Graphical representation of the matrix A for $r = \frac{1}{2}$

t/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{50}$	0.	0.36363636	0.65454545	0.65454545	0.36363636	0.

Table 8: The implicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 1 time step

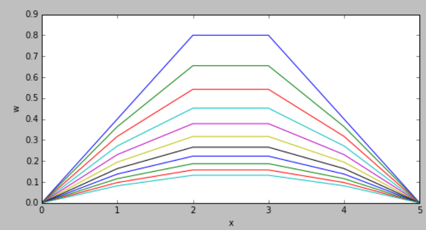


Figure 8.3.5: The implicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 10 time steps each represented by a different line

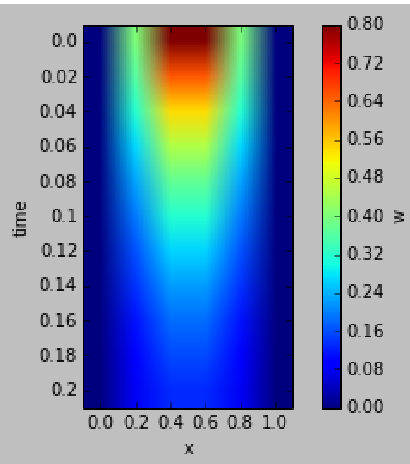


Figure 8.3.6: The colorplot of the implicit numerical solution w of the Heat Equation for $r = \frac{1}{2}$.

This method also gives an acceptable approximation to the solution of the PDE.

Example 45

8.3.1.3 *Implicit (BTCS) method for the Heat Equation for* $r = 1$

Let $h = \frac{1}{5}$ and $k = \frac{1}{25}$ so that $r = \frac{k}{h^2} = 1$ difference equation (51) becomes

$$(-w_{i-1j+1} + (3)w_{ij+1} - w_{i+1j+1}) = w_{ij}$$

This can be written in matrix form

$$\begin{pmatrix} 3 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 3 \end{pmatrix} \begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ w_{3j+1} \\ w_{4j+1} \end{pmatrix} = \begin{pmatrix} w_{1j} \\ w_{2j} \\ w_{3j} \\ w_{4j} \end{pmatrix} + \begin{pmatrix} w_{0j+1} \\ 0 \\ 0 \\ w_{5j+1} \end{pmatrix}.$$

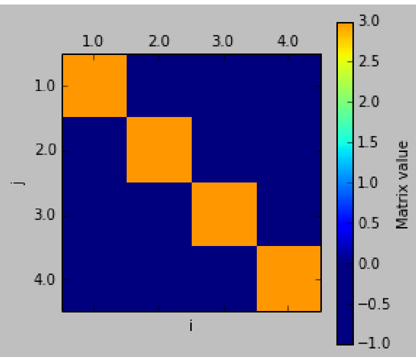


Figure 8.3.7: Graphical representation of the matrix A for $r = 1$

t/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{25}$	0.	0.32	0.56	0.56	0.32	0.
$\frac{2}{25}$	0.	0.24	0.4	0.4	0.24	0.
$\frac{3}{25}$	0.	0.176	0.288	0.288	0.176	0.

Table 9: The implicit numerical solution w of the Heat Equation for $r = 1$ for 3 time step

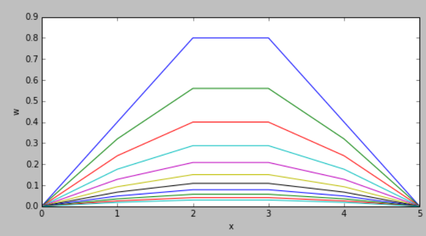


Figure 8.3.8: The implicit numerical solution w of the Heat Equation for $r = 1$ for 10 time steps each represented by a different line

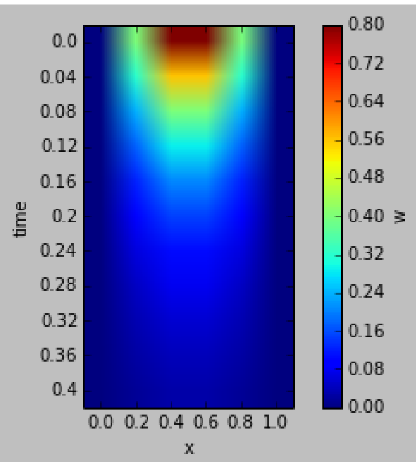


Figure 8.3.9: The colorplot of the implicit numerical solution w of the Heat Equation for $r = 1$.

8.4 Crank Nicholson Implicit Method

Since the implicit method requires that $k \leq \frac{1}{2}h^2$ a new method was needed which would work for all finite values of r.
They considered the partial differential equation as being satisfied at the midpoint $\{ih, (j+\frac{1}{2})k\}$ and replace $\frac{\delta^2 U}{\delta x^2}$ by the mean of its finite difference approximations at the jth and (j+1)th time levels. In other words they approximated the equation

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

by

$$\frac{w_{i,j+1} - w_{ij}}{k} = \frac{1}{2}\left\{\frac{w_{i+1j+1} - 2w_{ij+1} + w_{i-1j+1}}{h^2} + \frac{w_{i+1j} - 2w_{ij} + w_{i-1j}}{h^2}\right\}$$

giving

$$-rw_{i-1j+1} + (2+2r)w_{ij+1} - rw_{i+1j+1} = rw_{i-1j} + (2-2r)w_{ij} + rw_{i+1j} \quad (54)$$

with $r = \frac{k}{h^2}$.
In general the LHS contains 3 unknowns and the RHS 3 known pivotal values.
If there are N intervals mesh points along each row then for $j = 0$ and $i = 1, .., N$ it gives N simultaneous equations for N unknown pivotal values along the first row.
Which can be described in matrix form

$$B\mathbf{w}_{j+1} = C\mathbf{w}_j + \mathbf{b}_j$$

as

$$\begin{pmatrix} 2+2r & -r & 0 & . & . & . & . \\ -r & 2+2r & -r & 0 & . & . & . \\ 0 & -r & 2+2r & -r & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & -r & 2+2r & -r \\ . & . & . & . & . & -r & 2+2r \end{pmatrix} \begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ . \\ . \\ w_{N-2j+1} \\ w_{N-1j+1} \end{pmatrix}$$

$$= \begin{pmatrix} 2-2r & r & 0 & . & . & . & . \\ r & 2-2r & r & 0 & . & . & . \\ 0 & r & 2-2r & r & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & r & 2-2r & r \\ . & . & . & . & . & r & 2-2r \end{pmatrix} \begin{pmatrix} w_{1j} \\ w_{2j} \\ . \\ . \\ w_{N-2j} \\ w_{N-1j} \end{pmatrix} + \mathbf{b}_j + \mathbf{b}_{j+1}$$

where $\mathbf{b}_j$ and $\mathbf{b}_{j+1}$ are vectors of known boundary conditions.

$$\mathbf{b}_j = \begin{pmatrix} rw_{0j} \\ 0 \\ . \\ . \\ 0 \\ rw_{Nj} \end{pmatrix}, \quad \mathbf{b}_{j+1} = \begin{pmatrix} rw_{0j+1} \\ 0 \\ . \\ . \\ 0 \\ rw_{Nj+1} \end{pmatrix} \tag{55}$$

8.4.1 *Example Crank-Nicholson solution of the Heat Equation*

In this case we look at a rod of unit length with each end in ice.
The rod is heat insulated along its length so that temperature changes occur through heat conduction along its length and heat transfer at its ends, where w denotes temperature.

Simple case

Given that the ends of the rod are kept in contact with ice and the initial temperature distribution is non dimensional form is

1. $U = 2x$ for $0 \leq x \leq \frac{1}{2}$
2. $U = 2(1-x)$ for $\frac{1}{2} \leq x \leq 1$

In other words we are seeking a numerical solution of

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

which satisfies

1. $U = 0$ at $x = 0$ for all $t > 0$ (the boundary condition)
2. $U = 2x$ for $0 \leq x \leq \frac{1}{2}$ for $t = 0$ $U = 2(1-x)$ for $\frac{1}{2} \leq x \leq 1$ for $t = 0$ (the initial condition)

Example 46

8.4.1.1 *Crank-Nicholson method* $r = \frac{1}{10}$

Let $h = \frac{1}{5}$ and $k = \frac{1}{250}$ so that $r = \frac{k}{h^2} = \frac{1}{10}$ difference equation (54) becomes

$$-0.1w_{i-1,j+1} + 2.2w_{i,j+1} - 0.1w_{i+1,j+1} = 0.1w_{i-1,j} + 1.8w_{i,j} + 0.1w_{i+1,j}$$

Let $j = 0$

$$\begin{aligned}
&i = 1: \\
&-0.1w_{0,1} + 2.2w_{1,1} - 0.1w_{2,1} &=& \; 0.1w_{0,0} + 1.8w_{1,0} + 0.1w_{2,0} \\
&i = 2: \\
&-0.1w_{1,1} + 2.2w_{2,1} - 0.1w_{3,1} &=& \; 0.1w_{1,0} + 1.8w_{2,0j} + 0.1w_{3,0} \\
&i = 3: \\
&-0.1w_{2,1} + 2.2w_{3,1} - 0.1w_{4,1} &=& \; 0.1w_{2,0} + 1.8w_{3,0} + 0.1w_{4,0} \\
&i = 4: \\
&-0.1w_{3,1} + 2.2w_{4,1} - 0.1w_{5,1} &=& \; 0.1w_{3,0} + 1.8w_{4,0} + 0.1w_{5,0}
\end{aligned}$$

In matrix form

$$\begin{pmatrix} 2.2 & -0.1 & 0 & 0 \\ -0.1 & 2.2 & -0.1 & 0 \\ 0 & -0.1 & 2.2 & -0.1 \\ 0 & 0 & -0.1 & 2.2 \end{pmatrix} \begin{pmatrix} w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{4,1} \end{pmatrix}$$

$$= \begin{pmatrix} 1.8 & 0.1 & 0 & 0 \\ 0.1 & 1.8 & 0.1 & 0 \\ 0 & 0.1 & 1.8 & 0.1 \\ 0 & 0 & 0.1 & 1.8 \end{pmatrix} \begin{pmatrix} w_{1,0} \\ w_{2,0} \\ w_{3,0} \\ w_{4,0} \end{pmatrix} + 0.1 \begin{pmatrix} w_{0,1} + w_{0,0} \\ 0 \\ 0 \\ w_{5,1} + w_{5,0} \end{pmatrix}$$

To solve we need to invert the matix, to get

$$\mathbf{w}_{j+1} = A^{-1}(B\mathbf{w}_j + \mathbf{b}_{j+1} + \mathbf{b}_j)$$

j/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{250}$	0.	0.39826464	0.76182213	0.76182213	0.39826464	0.

Table 10: The Crank-Nicholson numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 1 time step

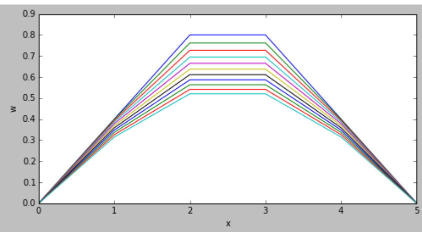


Figure 8.4.1: The Crank-Nicholson numerical solution w of the Heat Equation for $r = \frac{1}{10}$ for 10 time steps each represented by a different line

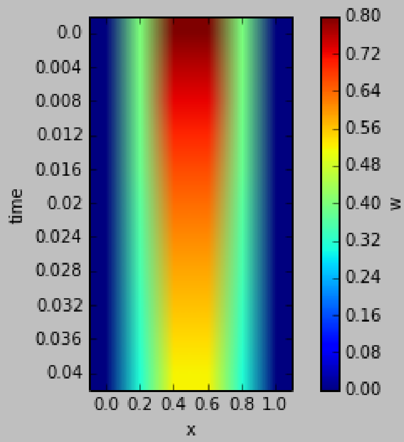


Figure 8.4.2: The colorplot of the Crank-Nicholson numerical solution w of the Heat Equation for $r = \frac{1}{10}$.

Example 47

8.4.1.2 *Crank-Nicholson method* $r = \frac{1}{2}$

Let $h = \frac{1}{5}$ and $k = \frac{1}{50}$ so that $r = \frac{k}{h^2} = \frac{1}{2}$ difference equation (54) becomes

$$-0.5w_{i-1,j+1} + 3w_{i,j+1} - 0.5w_{i+1,j+1} =$$

$$0.5w_{i-1,j} + 1w_{i,j} + 0.5w_{i+1,j}$$

Let $j = 0$

$$\begin{aligned}
&i = 1: \\
&-0.5w_{0,1} + 3w_{1,1} - 0.5w_{2,1} &=& \quad 0.5w_{0,0} + 1w_{1,0} + 0.5w_{2,0} \\
&i = 2: \\
&-0.5w_{1,1} + 3w_{2,1} - 0.5w_{3,1} &=& \quad 0.5w_{1,0} + 1w_{2,0j} + 0.5w_{3,0} \\
&i = 3: \\
&-0.5w_{2,1} + 3w_{3,1} - 0.5w_{4,1} &=& \quad 0.5w_{2,0} + 1w_{3,0} + 0.5w_{4,0} \\
&i = 4: \\
&-0.5w_{3,1} + 3w_{4,1} - 0.5w_{5,1} &=& \quad 0.5w_{3,0} + 1w_{4,0} + 0.5w_{5,0}
\end{aligned}$$

In matrix form

$$\begin{pmatrix} 3 & -0.5 & 0 & 0 \\ -0.5 & 3 & -0.5 & 0 \\ 0 & -0.5 & 3 & -0.5 \\ 0 & 0 & -0.5 & 3 \end{pmatrix} \begin{pmatrix} w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{4,1} \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0.5 & 1 \end{pmatrix} \begin{pmatrix} w_{1,0} \\ w_{2,0} \\ w_{3,0} \\ w_{4,0} \end{pmatrix} + 0.5 \begin{pmatrix} w_{0,1} + w_{0,0} \\ 0 \\ 0 \\ w_{5,1} + w_{5,0} \end{pmatrix}$$

t/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{50}$	0.	0.37241379	0.63448276	0.63448276	0.37241379	0.

Table 11: The Crank-Nicholson numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 1 time step

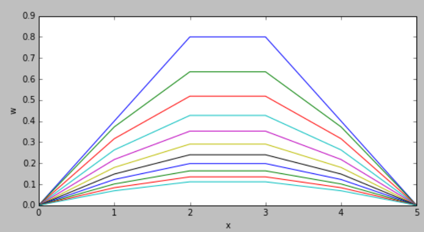


Figure 8.4.3: The Crank-Nicholson numerical solution w of the Heat Equation for $r = \frac{1}{2}$ for 10 time steps each represented by a different line

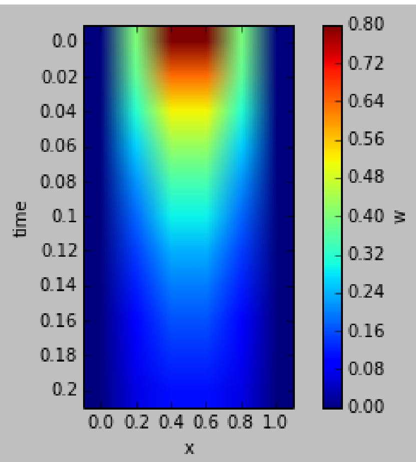


Figure 8.4.4: The colorplot of the Crank-Nicholson numerical solution w of the Heat Equation for $r = \frac{1}{2}$.

This method also gives an good approximation to the solution of the PDE.

Example 48

8.4.1.3 *Crank-Nicholson method for the Heat Equation with $r = 1$*

Let $h = \frac{1}{5}$ and $k = \frac{1}{25}$ so that $r = \frac{k}{h^2} = 1$ difference equation (54) becomes

$$-w_{i-1,j+1} + 4w_{i,j+1} - w_{i+1,j+1} = w_{i-1,j} + 0w_{i,j} + w_{i+1,j}$$

Let $j = 0$

$$\begin{array}{lcl} i = 1: & & \\ -w_{0,1} + 4w_{1,1} - w_{2,1} & = & w_{0,0} + 0w_{1,0} + w_{2,0} \\ i = 2: & & \\ -w_{1,1} + 4w_{2,1} - w_{3,1} & = & w_{1,0} + 0w_{2,0j} + w_{3,0} \\ i = 3: & & \\ -w_{2,1} + 4w_{3,1} - w_{4,1} & = & w_{2,0} + 0w_{3,0} + w_{4,0} \\ i = 4: & & \\ -w_{3,1} + 4w_{4,1} - w_{5,1} & = & w_{3,0} + 0w_{4,0} + w_{5,0} \end{array}$$

In matrix form

$$\begin{pmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{pmatrix} \begin{pmatrix} w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{4,1} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} w_{1,0} \\ w_{2,0} \\ w_{3,0} \\ w_{4,0} \end{pmatrix} + \begin{pmatrix} w_{0,1} + w_{0,0} \\ 0 \\ 0 \\ w_{5,1} + w_{5,0} \end{pmatrix}$$

t/x	0	0.2	0.4	0.6	0.8	1.0
0	0	0.4	0.8	0.8	0.4	0.0
$\frac{1}{25}$	0.	0.32727273	0.50909091	0.50909091	0.32727273	0.
$\frac{2}{25}$	0.	0.21487603	0.35041322	0.35041322	0.21487603	0.
$\frac{3}{25}$	0.	0.14695718	0.23741548	0.23741548	0.14695718	0.

Table 12: The Crank-Nicholson numerical solution w of the Heat Equation for $r = 1$ for 3 time step

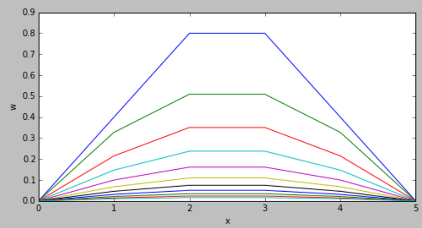


Figure 8.4.5: The Crank-Nicholson numerical solution w of the Heat Equation for $r = 1$ for 10 time steps each represented by a different line

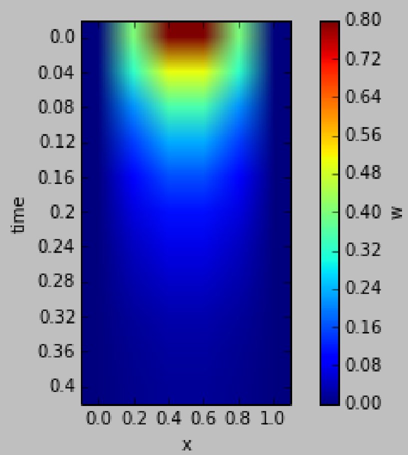


Figure 8.4.6: The colorplot of the Crank-Nicholson numerical solution w of the Heat Equation for $r = 1$.

8.5 THE THETA METHOD

The Theta Method is a generalization of the Crank-Nicholson method and expresses our partial differential equation as

$$\frac{w_{i,j+1}-w_{ij}}{k} = \left\{\theta\left(\frac{w_{i+1j+1}-2w_{ij+1}+w_{i-1j+1}}{h^2}\right)+(1-\theta)\left(\frac{w_{i+1j}-2w_{ij}+w_{i-1j}}{h^2}\right)\right\} \tag{56}$$

- when $\theta = 0$ we get the explicit scheme,
- when $\theta = \frac{1}{2}$ we get the Crank-Nicholson scheme,
- and $\theta = 1$ we get fully implicit backward finite difference method.

The equations are unconditionally valid for $\frac{1}{2} \leq \theta \leq 1$. For $0 \leq \theta < \frac{1}{2}$ we must have

$$r \leq \frac{1}{2(1-2\theta)}.$$

8.6 THE GENERAL MATRIX FORM

Let the solution domain of the PDE be the finite rectangle $0 \leq x \leq 1$ and $0 \leq t \leq T$ and subdivide it into a uniform rectangular mesh by the lines $x_i = ih$ for $i = 0$ to N and $t_j = jk$ for $j = 0$ to J it will be assumed that h is related to k by some relationship such as $k = rh$ or $k = rh^2$ with $r > 0$ and finite so that as $h \to 0$ as $k \to 0$.
Assume that the finite difference equation relating the mesh point values along the $(j+1)$th and jth row is

$$b_{i-1}w_{i-1j+1}+b_iw_{ij+1}+b_{i+1}w_{i+1j+1} = c_{i-1}w_{i-1j}+c_iw_{ij}+c_{i+1}w_{i+1j}$$

where the coefficients are constant. If the boundary values at $i = 0$ and N for $j > 0$ are known these $(N-1)$ equations for $i = 1$ to $N-1$ can be written in matrix form.

$$\begin{pmatrix} b_1 & b_2 & 0 & . & . & . & . \\ b_1 & b_2 & b_3 & 0 & . & . & . \\ 0 & b_2 & b_3 & b_4 & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & b_{N-3} & b_{N-2} & b_{N-1} \\ . & . & . & . & . & b_{N-2} & b_{N-1} \end{pmatrix} \begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ . \\ . \\ w_{N-2j+1} \\ w_{N-1j+1} \end{pmatrix}$$

$$= \begin{pmatrix} c_1 & c_2 & 0 & . & . & . & . \\ c_1 & c_2 & c_3 & 0 & . & . & . \\ 0 & c_2 & c_3 & c_4 & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & c_{N-3} & c_{N-2} & c_{N-1} \\ . & . & . & . & . & c_{N-2} & c_{N-1} \end{pmatrix} \begin{pmatrix} w_{1j} \\ w_{2j} \\ . \\ . \\ w_{N-2j} \\ w_{N-1j} \end{pmatrix} + \begin{pmatrix} c_0w_{0j}-b_0w_{0j+1} \\ 0 \\ . \\ . \\ 0 \\ c_Nw_{Nj}-b_Nw_{Nj+1} \end{pmatrix}$$

Which can be written as

$$B\mathbf{w}_{j+1} = C\mathbf{w}_j + \mathbf{d}_j$$

Where B and C are of order $(N-1)$ $\mathbf{w}_j$ denotes a column vector and $\mathbf{d}_j$ denotes a column vector of boundary values.
Hence

$$\mathbf{w}_{j+1} = B^{-1}C\mathbf{w}_j + B^{-1}\mathbf{d}_j.$$

Expressed in a more conventional manner as

$$\mathbf{w}_{j+1} = A\mathbf{w}_j + \mathbf{f}_j$$

Where $A = B^{-1}C$ and $\mathbf{f}_j = B^{-1}\mathbf{d}_j$.

8.7 Derivative Boundary Conditions

Boundary conditions expressed in terms of derivatives occur frequently.

8.7.1 *Example Derivative Boundary Conditions*

$$\frac{\partial U}{\partial x} = H(U - v_0) \quad at\ x = 0$$

where H is a positive constant and v_0 is the surrounding temperature.

How do we deal with this type of boundary condition?

1. By using forward difference for $\frac{\partial U}{\partial x}$, we have

$$\frac{w_{1j} - w_{0j}}{h_x} = H(w_{0j} - v_0)$$

where $h_x = x_1 - x_0$. This gives us one extra equation for the temp w_{ij}.

2. If we wish to represent $\frac{\partial U}{\partial x}$ more accurately at x=0, we use a central difference formula. It is necessary to introduce a fictitious temperature w_{-1j} at the external mesh points $(-h_x, jk)$. The temperature w_{-1j} is unknown and needs another equation. This is obtained by assuming that the heat conduction equation is satisfied at the end points. The unknown w_{-1j} can be eliminated between these equations.

Solve for the equation

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

satisfying the initial condition

$$U = 1 \text{ for } 0 \leq x \leq 1 \text{ when } t = 0$$

and the boundary conditions

$$\frac{\partial U}{\partial x} = U \text{ at } x = 0 \text{ for all t}$$

$$\frac{\partial U}{\partial x} = -U \text{ at } x = 1 \text{ for all t.}$$

8.7.1.1 *Example 1*

Using forward difference approximation for the derivative boundary condition and the explicit method to approximate the PDE.
Our difference equation is,

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

$$w_{ij+1} = w_{ij} + r(w_{i-1j} - 2w_{ij} + w_{i+1j}) \tag{57}$$

where $r = \frac{k}{h_x^2}$.
At i=1, (57) is,

$$w_{1j+1} = w_{1j} + r(w_{0j} - 2w_{1j} + w_{2j}) \tag{58}$$

The boundary condition at $x = 0$ is $\frac{\partial U}{\partial x} = U$ in terms of forward difference this is

$$\frac{w_{1j} - w_{0j}}{h_x} = w_{0j}$$

rearranging

$$w_{0j} = \frac{w_{1j}}{1 + h_x} \tag{59}$$

Using (61) and (58) to eliminate we get,

$$w_{1j+1} = \left(1 - 2r + \frac{r}{1 + h_x}\right) w_{1j} + r w_{2j}.$$

At $i = N - 1$, (57) is,

$$w_{N-1j+1} = w_{N-1j} + r(w_{N-2j} - 2w_{N-1j} + w_{Nj}) \tag{60}$$

The boundary condition at $x = 1$ is $\frac{\partial U}{\partial x} = U$ in terms of forward difference this is

$$\frac{w_{Nj} - w_{N-10}}{h_x} = w_{Nj}$$

rearranging

$$w_{Nj} = \frac{w_{N-1j}}{1 - h_x} \tag{61}$$

Using (61) and (60) to eliminate we get,

$$w_{N-1j+1} = rw_{N-2j} + \left(1 - 2r + \frac{r}{1-h_x}\right) w_{N-1j}.$$

Choose $h_s = \frac{1}{5}$ and $k = \frac{1}{100}$ such that $r = \frac{1}{4}$.
The equations become

$$w_{1j+1} = \frac{7}{24}w_{1j} + \frac{1}{4}w_{2j},$$

$$w_{ij+1} = \frac{1}{4}(w_{i-1j} + 2w_{ij} + w_{i+1j}) \quad i = 2, 3$$

and

$$w_{5j+1} = \frac{1}{4}w_{3j} + (\frac{13}{16}w_{4j})$$

In matrix form

$$\begin{pmatrix} w_{1j+1} \\ w_{2j+1} \\ w_{3j+1} \\ w_{4j+1} \end{pmatrix} = \begin{pmatrix} \frac{7}{24} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{13}{16} \end{pmatrix} \begin{pmatrix} w_{1j} \\ w_{2j} \\ w_{3j} \\ w_{4j} \end{pmatrix}.$$

with the boundaries given by

$$w_{0j+1} = \frac{10}{12}w_{1j+1},$$

$$w_{0j+1} = \frac{10}{8}w_{1j+1}.$$

8.7.1.2 *Example 2*

Using central difference approximation for the derivative boundary condition and the explicit method to approximate the PDE.
Our difference equation is as in (57).
At $i = 0$ we have

$$w_{0j+1} = w_{0j} + r(w_{-1j} - 2w_{0j} + w_{1j}) \tag{62}$$

The boundary condition at $x = 0$, in terms of central differences can be written as

$$\frac{w_{1j} - w_{-1j}}{2h_x} = w_{0j} \tag{63}$$

Using (63) and (62) to eliminate the fictitious term w_{-1j} we get,

$$w_{0j+1} = w_{0j} + 2r((-1 - h_x)w_{0j} + w_{1j})$$

8.7.1.3 *Example 3*

Using central difference approximation for the derivative boundary condition and the Crank-Nicholson method to approximate the PDE. The difference equation is,

$$\frac{w_{i,j+1} - w_{ij}}{k} = \frac{1}{2}\left\{\frac{w_{i+1j+1} - 2w_{ij+1} + w_{i-1j+1}}{h^2} + \frac{w_{i+1j} - 2w_{ij} + w_{i-1j}}{h^2}\right\}$$

giving

$$-rw_{i-1j+1} + (2+2r)w_{ij+1} - rw_{i+1j+1} = rw_{i-1j} + (2-2r)w_{ij} + rw_{i+1j} \tag{64}$$

with $r = \frac{k}{h^2}$.
The boundary condition at $x = 0$, in terms of central differences can be written as

$$\frac{w_{1j} - w_{-1j}}{2h_x} = w_{0j}$$

Rearranging we have

$$w_{-1j} = w_{1j} - 2h_x w_{0j} \tag{65}$$

and

$$w_{-1j+1} = w_{1j+1} - 2h_x w_{0j+1} \tag{66}$$

Let $j = 0$ and $i = 0$ the difference equation becomes

$$-rw_{-11} + (2+2r)w_{01} - rw_{11} = rw_{-10} + (2-2r)w_{00} + rw_{10} \tag{67}$$

Using, (65), (66) and (67) we can eliminate the fictious terms w_{-1j} and w_{-1j+1}.

8.8 LOCAL TRUNCATION ERROR AND CONSISTENCY

Let $F_{ij}(w)$ represent the difference equation approximating the PDE at the ijth point with exact solution w.
If w is replaced by U at the mesh points of the difference equation where U is the exact solution of the PDE the value of $F_{ij}(U)$ is the local truncation error T_{ij} in at the ij mesh pont.
Using Taylor expansions it is easy to express T_{ij} in terms of h_x and k and partial derivatives of U at (ih_x, jk).
Although U and its derivatives are generally unknown it is worthwhile because it provides a method for comparing the local accuracies of different difference schemes approximating the PDE.

Example 49

The local truncation error of the classical explicit difference approach to

$$\frac{\partial U}{\partial t} - \frac{\partial^2 U}{\partial x^2} = 0$$

with

$$F_{ij}(w) = \frac{w_{ij+1} - w_{ij}}{k} - \frac{w_{i+1j} - 2w_{ij} + w_{i-1j}}{h_x^2} = 0$$

is

$$T_{ij} = F_{ij}(U) = \frac{U_{ij+1} - U_{ij}}{k} - \frac{U_{i+1j} - 2U_{ij} + U_{i-1j}}{h_x^2}$$

By Taylors expansions we have

$$\begin{aligned}
U_{i+1j} &= U((i+1)h_x, jk) = U(x_i + h, t_j) \\
&= U_{ij} + h_x \left(\frac{\partial U}{\partial x}\right)_{ij} + \frac{h_x^2}{2}\left(\frac{\partial^2 U}{\partial x^2}\right)_{ij} + \frac{h_x^3}{6}\left(\frac{\partial^3 U}{\partial x^3}\right)_{ij} + ... \\
U_{i-1j} &= U((i-1)h_x, jk) = U(x_i - h, t_j) \\
&= U_{ij} - h_x \left(\frac{\partial U}{\partial x}\right)_{ij} + \frac{h_x^2}{2}\left(\frac{\partial^2 U}{\partial x^2}\right)_{ij} - \frac{h_x^3}{6}\left(\frac{\partial^3 U}{\partial x^3}\right)_{ij} + ... \\
U_{ij+1} &= U(ih_x, (j+1)k) = U(x_i, t_j + k) \\
&= U_{ij} + k \left(\frac{\partial U}{\partial t}\right)_{ij} + \frac{k^2}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{ij} + \frac{k^3}{6}\left(\frac{\partial^3 U}{\partial t^3}\right)_{ij} + ...
\end{aligned}$$

substitution into the expression for T_{ij} then gives

$$\begin{aligned}
T_{ij} &= \left(\frac{\partial U}{\partial t} - \frac{\partial^2 U}{\partial x^2}\right)_{ij} + \frac{k}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{ij} - \frac{h_x^2}{12}\left(\frac{\partial^4 U}{\partial x^4}\right)_{ij} \\
&\quad + \frac{k^2}{6}\left(\frac{\partial^3 U}{\partial t^3}\right)_{ij} - \frac{h_x^4}{360}\left(\frac{\partial^6 U}{\partial x^6}\right)_{ij} + ...
\end{aligned}$$

But U is the solution to the differential equation so

$$\left(\frac{\partial U}{\partial t} - \frac{\partial^2 U}{\partial x^2}\right)_{ij} = 0$$

the principal part of the local truncation error is

$$\frac{k}{2}\left(\frac{\partial^2 U}{\partial t^2}\right)_{ij} - \frac{h_x^2}{12}\left(\frac{\partial^4 U}{\partial x^4}\right)_{ij}.$$

Hence

$$T_{ij} = O(k) + O(h_x^2)$$

8.9 Consistency and Compatibility

It is sometimes possible to approximate a parabolic or hyperbolic equation with a finite difference scheme that is stable but which does not converge to the solution of differential equation as the mesh lengths tend to zero. Such a scheme is called inconsistent or incompatible.
This is useful when considering the theorem which states that is a linear finite difference equation is consistent with a properly posed linear IVP then stability guarantees convergence of w to U as the mesh lengths tend to zero.

Definition Let $L(U) = 0$ represent the PDE in the independent variables x and t with the exact solution U.
Let $F(w) = 0$ represent the approximate finite difference equation with exact solution w.
Let v be a continuous function of x and t with sufficient derivatives to enable $L(v)$ to be evaluated at the point (ih_x, jk). Then the truncation error $T_{ij}(v)$ at (ih_x, jk) is defined by

$$T_{ij}(v) = F_{ij}(v) - L(v_{ij})$$

If $T_{ij}(v) \to 0$ as $h \to 0$, $k \to 0$ the difference equation is said to be consistent or compatible with the with the PDE. ∘

Looking back at the previous example it follows that the classical explicit approximation to

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

is consistent with the difference equation.

8.10 Convergence and Stability

Definition By convergence we mean that the results of the method approach the analytical solution as k and h_x tends to zero. ∘

Definition By stability we mean that errors at one stage of the calculations do not cause increasingly large errors as the computations are continued. ∘

8.11 Stability by the Fourier Series Method (von Neumann's Method)

This method uses a Fourier series to express $w_{pq} = w(ph_x, qk)$ which is

$$w_{pq} = e^{i\beta x}\xi^q$$

where $\xi = e^{\alpha k}$ in this case i denotes the complex number $i = \sqrt{-1}$ and for values of β needed to satisfy the initial conditions. ξ is known as the amplification factor. The finite difference equation will be stable if $|w_{pq}|$ remains bounded for all q as $h \to 0, k \to 0$ and all β.
If the exact solution does not increase exponentially with time then a necessary and sufficient condition is that

$$|\xi| \leq 1$$

8.11.1 *Stability for the explicit FTCS Method*

Investigating the stability of the fully explicit difference equation

$$\frac{1}{k}(w_{pq+1} - w_{pq}) = \frac{1}{h_x^2}(w_{p-1q} - 2w_{pq} + w_{p+1q})$$

approximating $\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$ at (ph_x, qk). Substituting $w_{pq} = e^{i\beta x}\xi^q$ into the difference equation

$$e^{i\beta ph}\xi^{q+1} - e^{i\beta ph}\xi^q = r\{e^{i\beta(p-1)h}\xi^q - 2e^{i\beta ph}\xi^q + e^{i\beta(p+1)h}\xi^q\}$$

where $r = \frac{k}{h_x^2}$. Divide across by $e^{i\beta(p)h}\xi^q$ leads to

$$\begin{aligned}
\xi - 1 &= r(e^{i\beta(-1)h} - 2 + e^{i\beta h}) \\
&= 1 + r(2\cos(\beta h) - 2) \\
&= 1 - 4r(\sin^2(\beta\frac{h}{2})).
\end{aligned}$$

Hence

$$\left|1 - 4r(\sin^2(\beta\frac{h}{2}))\right| \leq 1,$$

for this to hold

$$4r(\sin^2(\beta\frac{h}{2})) \leq 2,$$

which means

$$r \leq \frac{1}{2}.$$

$0 < \xi \leq 1$ for $r < \frac{1}{2}$ and all β therefore the equation is conditionally stable.

8.11.2 *Stability for the implicit BTCS Method*

Investigating the stability of the fully implicit difference equation

$$\frac{1}{k}(w_{pq+1} - w_{pq}) = \frac{1}{h_x^2}(w_{p-1q+1} - 2w_{pq+1} + w_{p+1q+1})$$

approximating $\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$ at (ph_x, qk). Substituting $w_{pq} = e^{i\beta x}\xi^q$ into the difference equation

$$e^{i\beta ph}\xi^{q+1} - e^{i\beta ph}\xi^q = r\{e^{i\beta(p-1)h}\xi^{q+1} - 2e^{i\beta ph}\xi^{q+1} + e^{i\beta(p+1)h}\xi^{q+1}\}$$

where $r = \frac{k}{h_x^2}$. Divide across by $e^{i\beta(p)h}\xi^q$ leads to

$$\begin{aligned}\xi - 1 &= r\xi(e^{i\beta(-1)h} - 2 + e^{i\beta h})\\ &= r\xi(2\cos(\beta h) - 2)\\ &= -4r\xi(\sin^2(\beta\frac{h}{2}))\end{aligned}$$

Hence

$$\xi = \frac{1}{1 + 4r\sin^2(\frac{\beta h}{2})}$$

$0 < \xi \leq 1$ for all $r > 0$ and all β therefore the equation is unconditionally stable.

8.11.3 *Stability for the Crank Nicholson Method*

Investigating the stability of the fully implicit difference equation

$$\frac{1}{k}(w_{pq+1} - w_{pq}) = \frac{1}{2h_x^2}(w_{p-1q+1} - 2w_{pq+1} + w_{p+1q+1}) + \frac{1}{2h_x^2}(w_{p-1q} - 2w_{pq} + w_{p+1q})$$

approximating $\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$ at (ph_x, qk). Substituting $w_{pq} = e^{i\beta x}\xi^q$ into the difference equation

$$\begin{aligned}e^{i\beta ph}\xi^{q+1} - e^{i\beta ph}\xi^q &= \frac{r}{2}\{e^{i\beta(p-1)h}\xi^{q+1} - 2e^{i\beta ph}\xi^{q+1} + e^{i\beta(p+1)h}\xi^{q+1}\\ &\quad + e^{i\beta(p-1)h}\xi^q - 2e^{i\beta ph}\xi^q + e^{i\beta(p+1)h}\xi^q\}\end{aligned}$$

where $r = \frac{k}{h_x^2}$. Divide across by $e^{i\beta(p)h}\xi^q$ leads to

$$\begin{aligned}\xi - 1 &= \frac{r}{2}\xi(e^{-i\beta h} - 2 + e^{i\beta h}) + \frac{r}{2}\{e^{-i\beta h} - 2 + e^{i\beta h}\}\\ &= \frac{r}{2}\xi(2\cos(\beta h) - 2) + \frac{r}{2}(2\cos(\beta h) - 2)\\ &= -2r\xi(\sin^2(\beta\frac{h}{2})) - 2r(\sin^2(\beta\frac{h}{2}))\end{aligned}$$

Hence

$$\xi = \frac{1 - 2r\sin^2(\frac{\beta h}{2})}{1 + 2r\sin^2(\frac{\beta h}{2})}$$

$0 < \xi \leq 1$ for all $r > 0$ and all β therefore the equation is unconditionally stable.

8.12 Parabolic Equations Questions

8.12.1 *Explicit Equations*

1. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the explicit numerical scheme

$$w_{j,k+1} = rw_{j-1,k} + (1-2r)w_{j,k} + rw_{j+1,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \; 0 \leq t, 0 \leq x \leq 1\}.$$

[10 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \; 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0,t) = 1, \; u(1,t) = 1,$$

and initial condition

$$u(x,0) = 4x^2 - 4x + 1.$$

Taking $h = \frac{1}{4}$ in the x-direction and $k = \frac{1}{32}$ in the t-direction, set up and solve the corresponding systems of finite difference equations for one time step.

[18 marks]

c) For the explicit method what is the step-size requirement for h and k for the method to be stable.

[5 marks]

2. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the explicit numerical scheme

$$w_{j,k+1} = rw_{j-1,k} + (1-2r)w_{j,k} + rw_{j+1,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x)|\ 0 \leq t, 0 \leq x \leq 1\}.$$

[10 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x)|\ 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0,t) = 1,\ u(1,t) = 1,$$

and initial condition

$$u(x,0) = \begin{cases} 1-x & \text{for } 0 \leq t \leq \frac{1}{2} \\ x & \text{for } \frac{1}{2} \leq t \leq 1 \end{cases}.$$

Taking $h = \frac{1}{5}$ in the x-direction and $k = \frac{1}{250}$ in the t-direction, set up and solve the corresponding systems of finite difference equations for one time step.

[18 marks]

c) For the explicit method what is the step-size requirement for h and k for the method to be stable.

[5 marks]

3. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the explicit numerical scheme

$$w_{j,k+1} = rw_{j-1,k} + (1-2r)w_{j,k} + rw_{j+1,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \, 0 \leq t, 0 \leq x \leq 1\}.$$

[10 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \, 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0,t) = 0, \; u(1,t) = 0,$$

and initial condition

$$u(x,0) = 2\sin(2\pi x)$$

Taking $h = \frac{1}{6}$ in the x-direction and $k = \frac{1}{144}$ in the t-direction, set up and solve the corresponding systems of finite difference equations for one time step.

[18 marks]

c) For the explicit method what is the step-size requirement for h and k for the method to be stable.

[5 marks]

8.12.2 *Implicit Methods*

4. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the implicit numerical scheme

$$-rw_{j-1,k} + (1+2r)w_{j,k} - rw_{j+1,k} = w_{j,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \ 0 \leq t, 0 \leq x \leq 1\}.$$

[13 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \ 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0,t) = 1, \ u(1,t) = 1,$$

and initial condition

$$u(x,0) = 4x^2 - 4x + 1.$$

Taking $h = \frac{1}{4}$ in the x-direction and $k = \frac{1}{32}$ in the t-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations for one time step.

[20 marks]

5. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the implicit numerical scheme

$$-rw_{j-1,k} + (1+2r)w_{j,k} - rw_{j+1,k} = w_{j,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x)|\ 0 \leq t, 0 \leq x \leq 1\}.$$

[13 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x)|\ 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0,t) = 1,\ u(1,t) = 1,$$

and initial condition

$$u(x,0) = \begin{cases} 1-x & \text{for } 0 \leq t \leq \frac{1}{2} \\ x & \text{for } \frac{1}{2} \leq t \leq 1 \end{cases}.$$

Taking $h = \frac{1}{5}$ in the x-direction and $k = \frac{1}{250}$ in the t-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations for one time step.

[20 marks]

6. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the implicit numerical scheme

$$-rw_{j-1,k} + (1+2r)w_{j,k} - rw_{j+1,k} = w_{j,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x)|\ 0 \leq t, 0 \leq x \leq 1\}.$$

[13 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x)|\ 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0,t) = 0,\ u(1,t) = 0,$$

and initial condition

$$u(x,0) = 2\sin(2\pi x)$$

Taking $h = \frac{1}{6}$ in the x-direction and $k = \frac{1}{144}$ in the t-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations for one time step.

[20 marks]

8.12.3 *Crank Nicholson Methods*

7. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the Crank Nicholson numerical scheme

$$-rw_{j-1,k} + (2+2r)w_{j,k} - rw_{j+1,k} = rw_{j-1,k} + (2-2r)w_{j,k} + rw_{j+1,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \ 0 \leq t, 0 \leq x \leq 1\}.$$

[13 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t,x) | \ 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0,t) = 1, \ u(1,t) = 1,$$

and initial condition

$$u(x,0) = 4x^2 - 4x + 1.$$

Taking $h = \frac{1}{4}$ in the x-direction and $k = \frac{1}{32}$ in the t-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations for one time step.

[20 marks]

8. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the Crank Nicholson numerical scheme

$$-rw_{j-1,k} + (2+2r)w_{j,k} - rw_{j+1,k} = rw_{j-1,k} + (2-2r)w_{j,k} + rw_{j+1,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t, x) | \ 0 \leq t, 0 \leq x \leq 1\}.$$

[13 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t, x) | \ 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0, t) = 1, \ u(1, t) = 1,$$

and initial condition

$$u(x, 0) = \begin{cases} 1 - x & \text{for } 0 \leq t \leq \frac{1}{2} \\ x & \text{for } \frac{1}{2} \leq t \leq 1 \end{cases}.$$

Taking $h = \frac{1}{5}$ in the x-direction and $k = \frac{1}{250}$ in the t-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations for one time step.

[20 marks]

9. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0 + h) - 2f(x_0) + f(x_0 - h)}{h^2} + \mathcal{O}(h^2)$$

to derive the Crank Nicholson numerical scheme

$$-rw_{j-1,k} + (2 + 2r)w_{j,k} - rw_{j+1,k} = rw_{j-1,k} + (2 - 2r)w_{j,k} + rw_{j+1,k},$$

where $r = \frac{k}{h^2}$, k is the step in the time direction and h is the step in the x direction, for the Heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t, x) | \; 0 \leq t, 0 \leq x \leq 1\}.$$

[13 marks]

b) Consider the problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

on the rectangular domain

$$\Omega = \{(t, x) | \; 0 \leq t, 0 \leq x \leq 1\},$$

with the boundary conditions

$$u(0, t) = 0, \; u(1, t) = 0,$$

and initial condition

$$u(x, 0) = 2\sin(2\pi x)$$

Taking $h = \frac{1}{6}$ in the x-direction and $k = \frac{1}{144}$ in the t-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations for one time step.

[20 marks]

9

ELLIPTIC PDE'S

The Poisson equation is,

$$-\nabla^2 U(x,y) = f(x,y), \quad (x,y) \in \Omega = (0,1) \times (0,1), \tag{68}$$

where ∇ is the Laplacian,

$$\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial x^2},$$

with boundary conditions,

$$U(x,y) = g(x,y), \quad (x,y) \in \delta\Omega\text{-boundary.}$$

9.1 THE FIVE POINT APPROXIMATION OF THE LAPLACIAN

To numerically approxiamte the solution of the Poisson Equation 68 the unit square region $\bar{\Omega} = [0,1] \times [0,1] = \Omega \bigcup \partial\Omega$ must be discretised into a uniform grid.

$$\triangle = \{(x_i, y_j) \in [0,1] \times [0,1] : x_i = ih, y_j = jh\}$$

for $i = 0, 1, .., N$ and $i = 0, 1, .., N$, where N is a positive constant. The interior nodes of the grid are defined as:

$$\Omega_h = \{(x_i, y_j) \in \triangle : 1 \leq i, j \leq N-1\},$$

the boundary nodes are

$$\partial\Omega_h = \{(x_0, y_j), (x_N, y_j), (x_i, y_0), (x_i, y_N) : 1 \leq i, j \leq N-1\}.$$

The Poisson Equation 68 is discretised using δ_x^2 the central difference approximation of the second derivative in the x direction

$$\delta_x^2 = \frac{1}{h^2}(w_{i+1j} - 2w_{ij} + w_{i-1j}),$$

and δ_y^2 the central difference approximation of the second derivative in the y direction

$$\delta_y^2 = \frac{1}{h^2}(w_{ij+1} - 2w_{ij} + w_{ij-1}).$$

The gives the Poisson Difference Equation,

$$-\nabla_h w_{ij} = f_{ij} \quad (x_i, y_j) \in \Omega_h, \tag{69}$$
$$-(\delta_x^2 w_{ij} + \delta_y^2 w_{ij}) = f_{ij} \quad (x_i, y_j) \in \Omega_h, \tag{70}$$
$$w_{ij} = g_{ij} \quad (x_i, y_j) \in \partial\Omega_h, \tag{71}$$

where w_{ij} is the numerical approximation of U at x_i and y_j. Expanding the the Poisson Difference Equation 71 gives the five point method,

$$-(w_{i-1j} + w_{ij-1} - 4w_{ij} + w_{ij+1} + w_{i+1j}) = h^2 f_{ij}$$

for $i = 1, ..., N-1$ and $j = 1, ..., N-1$, which is depicted in Figure 9.1.1 on a 6 discrete grid.

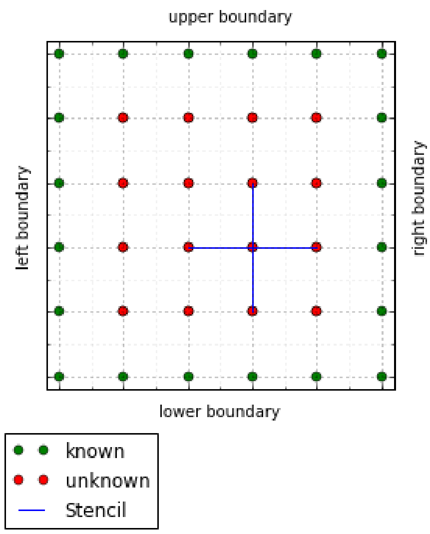


Figure 9.1.1: Graphical representation of the difference equation stencil

Unlike the Parabolic equation, the Elliptic equation cannot be estimated by holding one variable constant and then stepping in that direction. The approximation must be solved at all points at the same instant.

9.1.1 *Matrix representation of the five point scheme*

The five point scheme results in a system of $(N-1)^2$ equations for the $(N-1)^2$ unknowns. This is depicted in Figure 9.1.1 on a $6 \times 6 = 36$ where there is a grid of $4 \times 4 = 16$ unknowns (red) surround by the boundary of 20 known values. The general set of 4×4 equations of the Poisson difference equation on the 6×6 grid where

$$h = \frac{1}{6-1} = \frac{1}{5},$$

can be written as:

$$\begin{array}{l|lcl}
j=1 & & & \\
i=1 & w_{0,1} + w_{1,0} - 4w_{1,1} + w_{1,2} + w_{2,1} & = & \frac{1}{5}^2 f_{11} \\
i=2 & w_{1,1} + w_{2,0} - 4w_{2,1} + w_{2,2} + w_{3,1} & = & \frac{1}{5}^2 f_{21} \\
i=3 & w_{2,1} + w_{3,0} - 4w_{3,1} + w_{3,2} + w_{4,1} & = & \frac{1}{5}^2 f_{31} \\
i=4 & w_{3,1} + w_{4,0} - 4w_{4,1} + w_{4,2} + w_{5,1} & = & \frac{1}{5}^2 f_{41}
\end{array}$$

$$\begin{array}{l|lcl}
j=2 & & & \\
i=1 & w_{0,2} + w_{1,1} - 4w_{1,2} + w_{1,3} + w_{2,2} & = & \frac{1}{5}^2 f_{12} \\
i=2 & w_{1,2} + w_{2,1} - 4w_{2,2} + w_{2,3} + w_{3,2} & = & \frac{1}{5}^2 f_{22} \\
i=3 & w_{2,2} + w_{3,1} - 4w_{3,2} + w_{3,3} + w_{4,2} & = & \frac{1}{5}^2 f_{32} \\
i=4 & w_{3,2} + w_{4,1} - 4w_{4,2} + w_{4,3} + w_{5,2} & = & \frac{1}{5}^2 f_{42}
\end{array}$$

$$\begin{array}{l|lcl}
j=3 & & & \\
i=1 & w_{0,3} + w_{1,2} - 4w_{1,3} + w_{1,4} + w_{2,3} & = & \frac{1}{5}^2 f_{13} \\
i=2 & w_{1,3} + w_{2,2} - 4w_{2,3} + w_{2,4} + w_{3,3} & = & \frac{1}{5}^2 f_{23} \\
i=3 & w_{2,3} + w_{3,2} - 4w_{3,3} + w_{3,4} + w_{4,3} & = & \frac{1}{5}^2 f_{33} \\
i=4 & w_{3,3} + w_{4,2} - 4w_{4,3} + w_{4,4} + w_{5,3} & = & \frac{1}{5}^2 f_{43}
\end{array}$$

$$\begin{array}{l|lcl}
j=4 & & & \\
i=1 & w_{0,4} + w_{1,3} - 4w_{1,4} + w_{1,5} + w_{2,4} & = & \frac{1}{5}^2 f_{14} \\
i=2 & w_{1,4} + w_{2,3} - 4w_{2,4} + w_{2,5} + w_{3,4} & = & \frac{1}{5}^2 f_{24} \\
i=3 & w_{2,4} + w_{3,3} - 4w_{3,4} + w_{3,5} + w_{4,4} & = & \frac{1}{5}^2 f_{34} \\
i=4 & w_{3,4} + w_{4,3} - 4w_{4,4} + w_{4,5} + w_{5,4} & = & \frac{1}{5}^2 f_{44}.
\end{array}$$

This set of equations can be re-arranged by bringing the known boundary conditions $w_{0,j}$, $w_{5,j}$, $w_{i,0}$ and $w_{i,5}$, to the right hand side. This can be written as a 16×16 Matrix equation of the form:

$$
\left(\begin{array}{cccc|cccc|cccc|cccc}
-4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
1 & 0 & 0 & 0 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 0 & 0 & 0 & 1 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -4 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4
\end{array}\right)
\left(\begin{array}{c}
w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{4,1} \\ \hline
w_{1,2} \\ w_{2,2} \\ w_{3,2} \\ w_{4,2} \\ \hline
w_{1,3} \\ w_{2,3} \\ w_{3,3} \\ w_{4,3} \\ \hline
w_{1,4} \\ w_{2,4} \\ w_{3,4} \\ w_{4,4}
\end{array}\right)
= -\frac{1}{5}^2
\left(\begin{array}{c}
f_{1,1} \\ f_{2,1} \\ f_{3,1} \\ f_{4,1} \\ \hline
f_{1,2} \\ f_{2,2} \\ f_{3,2} \\ f_{4,2} \\ \hline
f_{1,3} \\ f_{2,3} \\ f_{3,3} \\ f_{4,3} \\ \hline
f_{1,4} \\ f_{2,4} \\ f_{3,4} \\ f_{4,4}
\end{array}\right)
+
\left(\begin{array}{c}
-w_{1,0} \\ -w_{2,0} \\ -w_{3,0} \\ -w_{4,0} \\ \hline
0 \\ 0 \\ 0 \\ 0 \\ \hline
0 \\ 0 \\ 0 \\ 0 \\ \hline
-w_{1,4} \\ -w_{2,4} \\ -w_{3,4} \\ -w_{4,4}
\end{array}\right)
+
\left(\begin{array}{c}
-w_{0,1} \\ 0 \\ 0 \\ -w_{5,1} \\ \hline
-w_{0,2} \\ 0 \\ 0 \\ -w_{5,2} \\ \hline
-w_{0,3} \\ 0 \\ 0 \\ -w_{5,3} \\ \hline
-w_{0,4} \\ 0 \\ 0 \\ -w_{5,4}
\end{array}\right)
$$

The horizontal and vertical lines are for display purposes to help indicated each set of the four sets of four equations.

9.1.2 *Generalised Matrix form of the discrete Poisson Equation*

The generalised form of this matrix of the system of equations for the parabolic case results in $(N-1)$ equations, that are written as an $(N-1)^2 \times (N-1)^2$ square matrix A and the $(N-1)^2 \times 1$ vectors $\mathbf{w}$, $\mathbf{r}$ and $\mathbf{b}$:

$$A\mathbf{w} = -h\mathbf{r} + \mathbf{b}.$$

The matrix can be written as following block tridiagonal structure (Figure SparseMatrix) :

$$\begin{pmatrix} T & I & 0 & 0 & . & . & . \\ I & T & I & 0 & 0 & . & . \\ 0 & . & . & . & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & 0 & I & T & I \\ . & . & . & . & 0 & I & T \end{pmatrix} \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ . \\ . \\ \mathbf{w}_{N-2} \\ \mathbf{w}_{N-1} \end{pmatrix} = -h^2 \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ . \\ . \\ \mathbf{f}_{N-2} \\ \mathbf{f}_{N-1} \end{pmatrix} + \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ . \\ . \\ \mathbf{b}_{N-2} \\ \mathbf{b}_{N-1} \end{pmatrix},$$

where I denotes an $(N-1) \times (N-1)$ identity matrix and T is an $(N-1) \times (N-1)$ tridiagonal matrix of the form:

$$T = \begin{pmatrix} -4 & 1 & 0 & 0 & . & . & . \\ 1 & -4 & 1 & 0 & 0 & . & . \\ 0 & . & . & . & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & 0 & 1 & -4 & 1 \\ . & . & . & . & 0 & 1 & -4 \end{pmatrix},$$

$\mathbf{w}_j$ is an $(N-1) \times 1$ vector of approximations w_{ij},

$$\mathbf{w}_j = \begin{pmatrix} w_{1j} \\ w_{2j} \\ . \\ . \\ w_{N-2j} \\ w_{N-1j} \end{pmatrix}$$

the vector $\mathbf{f}$ is made up of $(N-1)$ vectors of length $(N-1) \times 1$,

$$\mathbf{f}_j = \begin{pmatrix} f_{1j} \\ f_{2j} \\ . \\ . \\ f_{N-2j} \\ f_{N-1j} \end{pmatrix},$$

finally $\mathbf{b}$ is the vector of boundary conditions made up of two $(N-1)$ vectors of length $(N-1)\times 1$,

$$\mathbf{b}_j = \mathbf{b}_{left,right,j} + \mathbf{b}_{top,bottom,j} = -\begin{pmatrix} g_{0j} \\ 0 \\ . \\ . \\ 0 \\ g_{Nj} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 0 \\ 0 \end{pmatrix}$$

for $j = 2, .., N-2$, for $j = 1$ and $j = N-1$

$$\mathbf{b}_1 = -\begin{pmatrix} g_{10} \\ 0 \\ . \\ . \\ 0 \\ g_{1N} \end{pmatrix} - \begin{pmatrix} g_{10} \\ g_{20} \\ . \\ . \\ g_{N-20} \\ g_{1N} \end{pmatrix}, \ \mathbf{b}_{N-1} = -\begin{pmatrix} g_{0N-1} \\ 0 \\ . \\ . \\ 0 \\ g_{NN-1} \end{pmatrix} - \begin{pmatrix} g_{1N} \\ g_{2N} \\ . \\ . \\ g_{N-2N} \\ g_{N-1N} \end{pmatrix}.$$

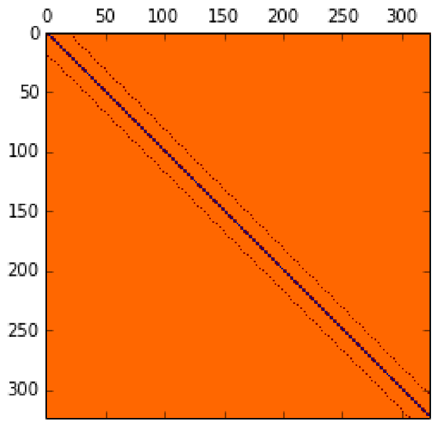


Figure 9.1.2: Graphical representation of the large sparse matrix A for the discrete solution of the Poisson Equation

The matrix has a unique solution. For sparse matrices of this form an iterative method is used as it would be to computationally expensive to compute the inverse.

9.2 Specific Examples

This section will work through three example problems:

1. Homogenous form of the Poisson Equation (Lapalacian),
2. Poisson Equation with zero boundary conditions,
3. Poisson Equation with non-zero boundary conditions.

9.2.1 *Example 1:Homogeneous equation with non-zero boundary*

Consider the Homogeneous Poisson Equation (also known as the Laplacian):

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x^2} = 0, \quad (x, y) \in \Omega = (0,1) \times (0,1),$$

with boundary conditions:
lower boundary,

$$u(x, 0) = \sin(2\pi x),$$

upper boundary,

$$u(x, 1) = \sin(2\pi x),$$

left boundary,

$$u(0, y) = 2\sin(2\pi y),$$

right boundary.

$$u(1, y) = 2\sin(2\pi y).$$

The general difference equation for the Laplacian is of the form

$$-(w_{i-1j} + w_{ij-1} - 4w_{ij} + w_{ij+1} + w_{i+1j}) = 0.$$

Here, $N = 4$, which gives the step-size,

$$h = \frac{1}{4},$$

and

$$x_i = i\frac{1}{4}, \quad y_j = j\frac{1}{4},$$

for $i = 0, 1, 2, 3, 4$ and $j = 0, 1, 2, 3, 4$. This gives the system of 3×3 equations:

$$\begin{array}{r|lcl}
j = 1 & & & \\
i = 1 & w_{0,1} + w_{1,0} - 4w_{1,1} + w_{1,2} + w_{2,1} & = & \frac{1}{4}^2 0 \\
i = 2 & w_{1,1} + w_{2,0} - 4w_{2,1} + w_{2,2} + w_{3,1} & = & \frac{1}{4}^2 0 \\
i = 3 & w_{2,1} + w_{3,0} - 4w_{3,1} + w_{3,2} + w_{4,1} & = & \frac{1}{4}^2 0
\end{array}$$

$$\begin{array}{r|lcl}
j = 2 & & & \\
i = 1 & w_{0,2} + w_{1,1} - 4w_{1,2} + w_{1,3} + w_{2,2} & = & \frac{1}{4}^2 0 \\
i = 2 & w_{1,2} + w_{2,1} - 4w_{2,2} + w_{2,3} + w_{3,2} & = & \frac{1}{4}^2 0 \\
i = 3 & w_{2,2} + w_{3,1} - 4w_{3,2} + w_{3,3} + w_{4,2} & = & \frac{1}{4}^2 0
\end{array}$$

$$\begin{array}{r|lcl}
j = 3 & & & \\
i = 1 & w_{0,3} + w_{1,2} - 4w_{1,3} + w_{1,4} + w_{2,3} & = & \frac{1}{4}^2 0 \\
i = 2 & w_{1,3} + w_{2,2} - 4w_{2,3} + w_{2,4} + w_{3,3} & = & \frac{1}{4}^2 0 \\
i = 3 & w_{2,3} + w_{3,2} - 4w_{3,3} + w_{3,4} + w_{4,3} & = & \frac{1}{4}^2 0.
\end{array}$$

This system is then rearranged by bringing the known boundary conditions to the right hand side, to give:

$$
\begin{array}{l|rcl}
j=1 & & & \\
i=1 & -4w_{1,1}+w_{1,2}+w_{2,1} & = & \frac{1}{4}^2 0 - w_{0,1} - w_{1,0} \\
i=2 & w_{1,1}-4w_{2,1}+w_{2,2}+w_{3,1} & = & \frac{1}{4}^2 0 - w_{2,0} \\
i=3 & w_{2,1}-4w_{3,1}+w_{3,2} & = & \frac{1}{4}^2 0 - w_{4,1} - w_{3,0}
\end{array}
$$

$$
\begin{array}{l|rcl}
j=2 & & & \\
i=1 & w_{1,1}-4w_{1,2}+w_{1,3}+w_{2,2} & = & \frac{1}{4}^2 0 - w_{0,2} \\
i=2 & w_{1,2}+w_{2,1}-4w_{2,2}+w_{2,3}+w_{3,2} & = & \frac{1}{4}^2 0 \\
i=3 & w_{2,2}+w_{3,1}-4w_{3,2}+w_{3,3} & = & \frac{1}{4}^2 0 - w_{4,2}
\end{array}
$$

$$
\begin{array}{l|rcl}
j=3 & & & \\
i=1 & w_{1,2}-4w_{1,3}+w_{2,3} & = & \frac{1}{4}^2 0 - w_{0,3} - w_{1,4} \\
i=2 & w_{1,3}+w_{2,2}-4w_{2,3}+w_{3,3} & = & \frac{1}{4}^2 0 - w_{2,4} \\
i=3 & w_{2,3}+w_{3,2}-4w_{3,3} & = & \frac{1}{4}^2 0 - w_{4,3} - w_{3,4}
\end{array}
$$

Given the discrete boundary conditions:

Left boundary	**Right boundary**
$x_0 = 0$	$x_4 = 1$
$u(0,y) = 2\sin(2\pi y)$	$u(1,y) = 2\sin(2\pi y)$
$w_{0,0} = 0$	$w_{4,0} = 0$
$w_{0,1} = 2\sin(2\pi y_1) = 2$	$w_{4,1} = 2\sin(2\pi y_1) = 2$
$w_{0,2} = 2\sin(2\pi y_2) = 0$	$w_{4,2} = 2\sin(2\pi y_2) = 0$
$w_{0,3} = 2\sin(2\pi y_3) = -2$	$w_{4,3} = 2\sin(2\pi y_3) = -2$
$w_{0,4} = 2\sin(2\pi y_4) = 0$	$w_{4,4} = 2\sin(2\pi y_4) = 0$

Lower boundary	**Upper boundary**
$y_0 = 0$	$y_4 = 1$
$u(x,0) = \sin(2\pi x)$	$u(x,1) = \sin(2\pi x)$
$w_{0,0} = 0$	$w_{0,4} = 0$
$w_{1,0} = \sin(2\pi x_1) = 1$	$w_{1,4} = \sin(2\pi x_1) = 1$
$w_{2,0} = \sin(2\pi x_2) = 0$	$w_{2,4} = \sin(2\pi x_2) = 0$
$w_{3,0} = \sin(2\pi x_3) = -1$	$w_{3,4} = \sin(2\pi x_3) = -1$
$w_{4,0} = \sin(2\pi x_4) = 0$	$w_{4,4} = \sin(2\pi x_4) = 0$

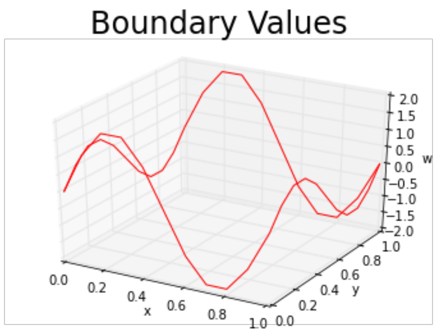


Figure 9.2.1: Sine Wave Boundary Conditions.

The system of equations are written in matrix form:

$$\left(\begin{array}{ccccccccc} -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 \end{array}\right) \left(\begin{array}{c} w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{1,2} \\ w_{2,2} \\ w_{3,2} \\ w_{1,3} \\ w_{2,3} \\ w_{3,3} \end{array}\right) = \left(\begin{array}{c} -w_{1,0} \\ -w_{2,0} \\ -w_{3,0} \\ 0 \\ 0 \\ 0 \\ -w_{1,4} \\ -w_{2,4} \\ -w_{3,4} \end{array}\right) + \left(\begin{array}{c} -w_{0,1} \\ 0 \\ -w_{4,1} \\ -w_{0,2} \\ 0 \\ -w_{4,2} \\ -w_{0,3} \\ 0 \\ -w_{4,3} \end{array}\right),$$

where the matrix is $3^2 \times 3^2$ which is graphically represented in Figure 9.2.2, where the colours indicated the values in each cell.

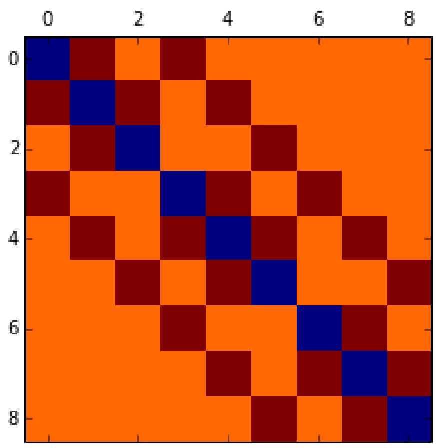


Figure 9.2.2: Graphical representation of the matrix

For the given boundary conditions the matrix equation is written as :

$$\begin{pmatrix} -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 \end{pmatrix} \begin{pmatrix} w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{1,2} \\ w_{2,2} \\ w_{3,2} \\ w_{1,3} \\ w_{2,3} \\ w_{3,3} \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ -2 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 2 \end{pmatrix}.$$

Figure 9.2.3 shows the approximate solution of the Laplacian Equation for the given boundary conditions and $h = \frac{1}{4}$.

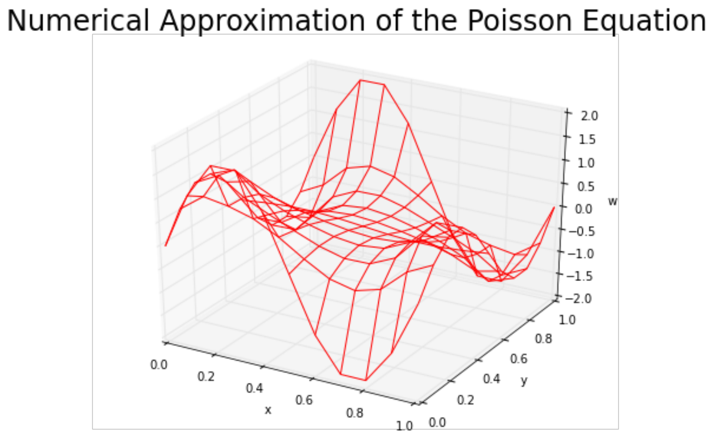


Figure 9.2.3: Numerical solution of the homogeneous differential equation

9.2.2 *Example 2: non-homogeneous equation with zero boundary*

Consider the Poission Equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x^2} = x^2 + y^2 \quad (x, y) \in \Omega = (0,1) \times (0,1)$$

with zero boundary conditions: Left boundary:

$$u(x, 0) = 0$$

Right boundary:

$$u(x, 1) = 0$$

Lower boundary:

$$u(0, y) = 0$$

Upper boundary:

$$u(1, y) = 0.$$

The difference equation is of the form:

$$-(w_{i-1j} + w_{ij-1} - 4w_{ij} + w_{ij+1} + w_{i+1j}) = h^2(x_i^2 + y_j^2).$$

Here, $N = 4$, which gives the step-size,

$$h = \frac{1}{4},$$

and

$$x_i = i\frac{1}{4}, \quad y_j = j\frac{1}{4},$$

for $i = 0, 1, 2, 3, 4$ and $j = 0, 1, 2, 3, 4$. This gives the system of 3×3 equations:

$$\begin{array}{c|rcl}
j = 1 & & & \\
i = 1 & w_{0,1} + w_{1,0} - 4w_{1,1} + w_{1,2} + w_{2,1} & = & \frac{1}{4}^2(x_1^2 + y_1^2) \\
i = 2 & w_{1,1} + w_{2,0} - 4w_{2,1} + w_{2,2} + w_{3,1} & = & \frac{1}{4}^2(x_2^2 + y_1^2) \\
i = 3 & w_{2,1} + w_{3,0} - 4w_{3,1} + w_{3,2} + w_{4,1} & = & \frac{1}{4}^2(x_3^2 + y_1^2)
\end{array}$$

$$\begin{array}{c|rcl}
j = 2 & & & \\
i = 1 & w_{0,2} + w_{1,1} - 4w_{1,2} + w_{1,3} + w_{2,2} & = & \frac{1}{4}^2(x_1^2 + y_2^2) \\
i = 2 & w_{1,2} + w_{2,1} - 4w_{2,2} + w_{2,3} + w_{3,2} & = & \frac{1}{4}^2(x_2^2 + y_2^2) \\
i = 3 & w_{2,2} + w_{3,1} - 4w_{3,2} + w_{3,3} + w_{4,2} & = & \frac{1}{4}^2(x_3^2 + y_2^2)
\end{array}$$

$$\begin{array}{c|rcl}
j = 3 & & & \\
i = 1 & w_{0,3} + w_{1,2} - 4w_{1,3} + w_{1,4} + w_{2,3} & = & \frac{1}{4}^2(x_1^2 + y_3^2) \\
i = 2 & w_{1,3} + w_{2,2} - 4w_{2,3} + w_{2,4} + w_{3,3} & = & \frac{1}{4}^2(x_2^2 + y_3^2) \\
i = 3 & w_{2,3} + w_{3,2} - 4w_{3,3} + w_{3,4} + w_{4,3} & = & \frac{1}{4}^2(x_3^2 + y_3^2)
\end{array}$$

This system is then rearranged by bringing the known boundary conditions to the right hand side, to give:

$$\begin{array}{c|rcl}
j = 1 & & & \\
i = 1 & -4w_{1,1} + w_{1,2} + w_{2,1} & = & \frac{1}{4}^2(x_1^2 + y_1^2) - w_{0,1} - w_{1,0} \\
i = 2 & w_{1,1} - 4w_{2,1} + w_{2,2} + w_{3,1} & = & \frac{1}{4}^2(x_2^2 + y_1^2) - w_{2,0} \\
i = 3 & w_{2,1} - 4w_{3,1} + w_{3,2} & = & \frac{1}{4}^2(x_3^2 + y_1^2) - w_{4,1} - w_{3,0}
\end{array}$$

$$\begin{array}{c|rcl}
j = 2 & & & \\
i = 1 & w_{1,1} - 4w_{1,2} + w_{1,3} + w_{2,2} & = & \frac{1}{4}^2(x_1^2 + y_2^2) - w_{0,2} \\
i = 2 & w_{1,2} + w_{2,1} - 4w_{2,2} + w_{2,3} + w_{3,2} & = & \frac{1}{4}^2(x_2^2 + y_2^2) \\
i = 3 & w_{2,2} + w_{3,1} - 4w_{3,2} + w_{3,3} & = & \frac{1}{4}^2(x_3^2 + y_2^2) - w_{4,2}
\end{array}$$

$$
\begin{array}{c|rcl}
j=3 & & & \\
i=1 & w_{1,2}-4w_{1,3}+w_{2,3} & = & \frac{1}{4}^2(x_1^2+y_3^2)-w_{0,3}-w_{1,4} \\
i=2 & w_{1,3}+w_{2,2}-4w_{2,3}+w_{3,3} & = & \frac{1}{4}^2(x_2^2+y_3^2)-w_{2,4} \\
i=3 & w_{2,3}+w_{3,2}-4w_{3,3} & = & \frac{1}{4}^2(x_3^2+y_3^2)-w_{4,3}-w_{3,4}.
\end{array}
$$

Given the zero boundary conditions

Lower Boundary	**Upper Boundary**
$x_0 = 0$	$x_4 = 1$
$u(0, y) = 0$	$u(1, y) = 0$
$w_{0,0} = 0$	$w_{4,0} = 0$
$w_{0,1} = 0$	$w_{4,1} = 0$
$w_{0,2} = 0$	$w_{4,2} = 0$
$w_{0,3} = 0$	$w_{4,3} = 0$
$w_{0,4} = 0$	$w_{4,4} = 0$

Left Boundary	**Right Boundary**
$y_0 = 0$	$y_4 = 1$
$u(x, 0) = 0$	$u(x, 1) = 0$
$w_{0,0} = 0$	$w_{0,4} = 0$
$w_{1,0} = 0$	$w_{1,4} = 0$
$w_{2,0} = 0$	$w_{2,4} = 0$
$w_{3,0} = 0$	$w_{3,4} = 0$
$w_{4,0} = 0$	$w_{4,4} = 0$

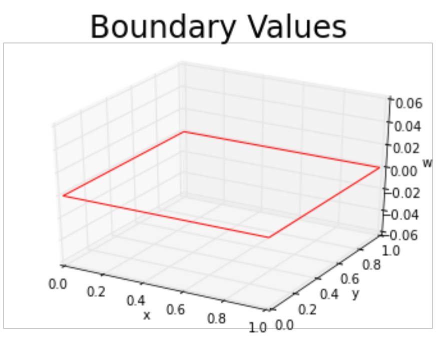


Figure 9.2.4: Sine Wave Boundary Conditions.

The system of equations can be written in matrix form:

$$
\begin{pmatrix}
-4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4
\end{pmatrix}
\begin{pmatrix}
w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{1,2} \\ w_{2,2} \\ w_{3,2} \\ w_{1,3} \\ w_{2,3} \\ w_{3,3}
\end{pmatrix}
= h^2
\begin{pmatrix}
(x_1^2+y_1^2) \\ (x_2^2+y_1^2) \\ (x_3^2+y_1^2) \\ (x_1^2+y_2^2) \\ (x_2^2+y_2^2) \\ (x_3^2+y_2^2) \\ (x_1^2+y_3^2) \\ (x_2^2+y_3^2) \\ (x_3^2+y_3^2)
\end{pmatrix}.
$$

Substituting values into the right hand side gives the specific matric form:

$$\begin{pmatrix} -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 \end{pmatrix} \begin{pmatrix} w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{1,2} \\ w_{2,2} \\ w_{3,2} \\ w_{1,3} \\ w_{2,3} \\ w_{3,3} \end{pmatrix} = \begin{pmatrix} 0.0078125 \\ 0.01953125 \\ 0.0390625 \\ 0.01953125 \\ 0.0312 \\ 0.05078125 \\ 0.0390625 \\ 0.05078125 \\ 0.0703125 \end{pmatrix}.$$

Figure 9.2.5 shows the numerical solution of the Poisson Equation with zero boundary conditions.

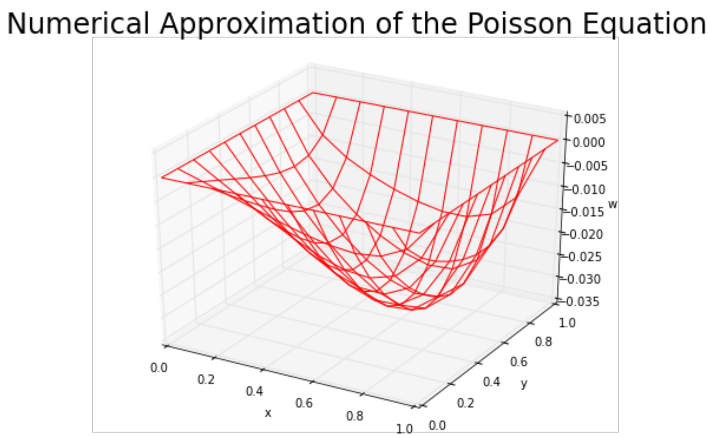


Figure 9.2.5: Numerical solution of the differential equation with zero boundary conditions

9.2.3 *Example 3: Inhomogeneous equation with non-zero boundary*

Consider the Poisson Equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x^2} = xy, \quad (x, y) \in \Omega = (0,1) \times (0,1),$$

with boundary conditions
Right Boundary

$$u(x, 0) = -x^2 + x$$

Left Boundary

$$u(x, 1) = x^2 - x$$

Lower Boundary

$$u(0, y) = -y^2 + y$$

Upper Boundary

$$u(1,y) = -y^2 + y.$$

The five point difference equation is of the form

$$-(w_{i-1j} + w_{ij-1} - 4w_{ij} + w_{ij+1} + w_{i+1j}) = h^2(x_i y_j).$$

Here, $N = 4$, which gives the step-size,

$$h = \frac{1}{4},$$

and

$$x_i = i\frac{1}{4}, \quad y_j = j\frac{1}{4},$$

for $i = 0, 1, 2, 3, 4$ and $j = 0, 1, 2, 3, 4$. This gives the system of 3×3 equations:

$$\begin{array}{c|rcl} i=1 & w_{0,1} + w_{1,0} - 4w_{1,1} + w_{1,2} + w_{2,1} & = & \frac{1}{4}^2(x_1y_1) \\ i=2 & w_{1,1} + w_{2,0} - 4w_{2,1} + w_{2,2} + w_{3,1} & = & \frac{1}{4}^2(x_2y_1) \\ i=3 & w_{2,1} + w_{3,0} - 4w_{3,1} + w_{3,2} + w_{4,1} & = & \frac{1}{4}^2(x_3y_1) \end{array}$$

$$\begin{array}{c|rcl} j=2 & & & \\ i=1 & w_{0,2} + w_{1,1} - 4w_{1,2} + w_{1,3} + w_{2,2} & = & \frac{1}{4}^2(x_1y_2) \\ i=2 & w_{1,2} + w_{2,1} - 4w_{2,2} + w_{2,3} + w_{3,2} & = & \frac{1}{4}^2(x_2y_2) \\ i=3 & w_{2,2} + w_{3,1} - 4w_{3,2} + w_{3,3} + w_{4,2} & = & \frac{1}{4}^2(x_3y_2) \end{array}$$

$$\begin{array}{c|rcl} j=3 & & & \\ i=1 & w_{0,3} + w_{1,2} - 4w_{1,3} + w_{1,4} + w_{2,3} & = & \frac{1}{4}^2(x_1y_3) \\ i=2 & w_{1,3} + w_{2,2} - 4w_{2,3} + w_{2,4} + w_{3,3} & = & \frac{1}{4}^2(x_2y_3) \\ i=3 & w_{2,3} + w_{3,2} - 4w_{3,3} + w_{3,4} + w_{4,3} & = & \frac{1}{4}^2(x_3y_3). \end{array}$$

Re-arranging the system such that the known values are on the right hand side:

$$\begin{array}{c|rcl} j=1 & & & \\ i=1 & -4w_{1,1} + w_{1,2} + w_{2,1} & = & \frac{1}{4}^2(x_1y_1) - w_{0,1} - w_{1,0} \\ i=2 & w_{1,1} - 4w_{2,1} + w_{2,2} + w_{3,1} & = & \frac{1}{4}^2(x_2y_1) - w_{2,0} \\ i=3 & w_{2,1} - 4w_{3,1} + w_{3,2} & = & \frac{1}{4}^2(x_3y_1) - w_{4,1} - w_{3,0} \end{array}$$

$$\begin{array}{c|rcl} j=2 & & & \\ i=1 & w_{1,1} - 4w_{1,2} + w_{1,3} + w_{2,2} & = & \frac{1}{4}^2(x_1y_2) - w_{0,2} \\ i=2 & w_{1,2} + w_{2,1} - 4w_{2,2} + w_{2,3} + w_{3,2} & = & \frac{1}{4}^2(x_2y_2) \\ i=3 & w_{2,2} + w_{3,1} - 4w_{3,2} + w_{3,3} & = & \frac{1}{4}^2(x_3y_2) - w_{4,2} \end{array}$$

$$\begin{array}{l|rcl}
j=3 & & & \\
i=1 & w_{1,2}-4w_{1,3}+w_{2,3} & = & \frac{1}{4}^2(x_1y_3)-w_{0,3}-w_{1,4} \\
i=2 & w_{1,3}+w_{2,2}-4w_{2,3}+w_{3,3} & = & \frac{1}{4}^2(x_2y_3)-w_{2,4} \\
i=3 & w_{2,3}+w_{3,2}-4w_{3,3} & = & \frac{1}{4}^2(x_3y_3)-w_{4,3}-w_{3,4}.
\end{array}$$

The discrete boundary conditions are

Left boundary	**Right boundary**
$x_0=0$	$x_4=1$
$u(0,y)=-y^2+y$	$u(1,y)=-y^2+y$
$w_{0,0}=0$	$w_{4,0}=0$
$w_{0,1}=-y_1^2+y_1=\frac{3}{16}$	$w_{4,1}=-y_1^2+y_1=\frac{1}{16}$
$w_{0,2}=-y_2^2+y_2=\frac{1}{4}$	$w_{4,2}=-y_2^2+y_2=\frac{1}{4}$
$w_{0,3}=-y_3^2+y_3=\frac{3}{16}$	$w_{4,3}=-y_3^2+y_3=\frac{3}{16}$
$w_{0,4}=-y_4^2+y_4=0$	$w_{4,4}=-y_4^2+y_4=0$

Lower boundary	**Upper boundary**
$y_0=0$	$y_4=1$
$u(x,0)=-x^2+x$	$u(x,1)=x^2-x$
$w_{0,0}=0$	$w_{0,4}=0$
$w_{1,0}=-x_1^2+x_1=\frac{3}{16}$	$w_{1,4}=x_1^2-x_1=-\frac{3}{16}$
$w_{2,0}=-x_2^2+x_2=\frac{1}{4}$	$w_{2,4}=x_2^2-x_2=-\frac{1}{4}$
$w_{3,0}=-x_3^2+x_3=\frac{3}{16}$	$w_{3,4}=x_3^2-x_3=-\frac{3}{16}$
$w_{4,0}=0$	$w_{4,4}=0$

The system of equations can be written in 9×9 Matrix form:

$$\begin{pmatrix}
-4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4
\end{pmatrix}
\begin{pmatrix}
w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{1,2} \\ w_{2,2} \\ w_{3,2} \\ w_{1,3} \\ w_{2,3} \\ w_{3,3}
\end{pmatrix} =$$

$$h^2\begin{pmatrix}
(x_1y_1) \\ (x_2y_1) \\ (x_3y_1) \\ (x_1y_2) \\ (x_2y_2) \\ (x_3y_2) \\ (x_1y_3) \\ (x_2y_3) \\ (x_3y_3)
\end{pmatrix} + \begin{pmatrix}
-w_{1,0} \\ -w_{2,0} \\ -w_{3,0} \\ 0 \\ 0 \\ 0 \\ -w_{1,4} \\ -w_{2,4} \\ -w_{3,4}
\end{pmatrix} + \begin{pmatrix}
-w_{0,1} \\ 0 \\ -w_{4,1} \\ -w_{0,2} \\ 0 \\ -w_{4,2} \\ -w_{0,3} \\ 0 \\ -w_{4,3}
\end{pmatrix},$$

inputting the specific boundary values and the right hand side of the equation gives:

$$\begin{pmatrix} -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 \end{pmatrix} \begin{pmatrix} w_{1,1} \\ w_{2,1} \\ w_{3,1} \\ w_{1,2} \\ w_{2,2} \\ w_{3,2} \\ w_{1,3} \\ w_{2,3} \\ w_{3,3} \end{pmatrix} =$$

$$\left(\frac{1}{4}\right)^2 \begin{pmatrix} 0.0625 \\ 0.125 \\ 0.1875 \\ 0.125 \\ 0.25 \\ 0.375 \\ 0.1875 \\ 0.375 \\ 0.5626 \end{pmatrix} + \begin{pmatrix} -\frac{3}{16} \\ -\frac{1}{4} \\ -\frac{3}{16} \\ 0 \\ 0 \\ 0 \\ \frac{3}{16} \\ \frac{1}{4} \\ \frac{3}{16} \end{pmatrix} + \begin{pmatrix} -\frac{3}{16} \\ 0 \\ -\frac{3}{16} \\ -\frac{1}{4} \\ 0 \\ -\frac{1}{4} \\ -\frac{3}{16} \\ 0 \\ -\frac{3}{16} \end{pmatrix}.$$

Figure 9.2.6 shows the numerical solution of the Poisson Equation with non-zero boundary conditions.

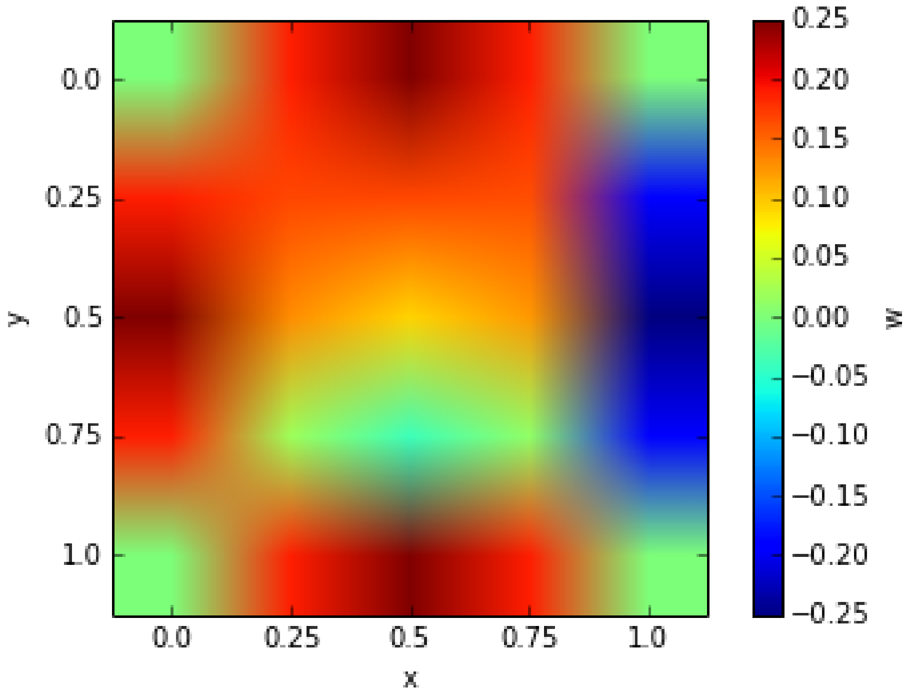


Figure 9.2.6: Numerical solution of the differential equation with non-zero boundary conditions

9.3 CONSISTENCY AND CONVERGENCE

We now ask how well the grid function determined by the five point scheme approximates the exact solution of the Poisson problem.

Definition Let L_h denote the finite difference approximation associated with the grid Ω_h having the mesh size h, to a partial differential operator L defined on a simply connected, open set $\Omega \subset R^2$. For a given function $\varphi \in C^\infty(\Omega)$, the truncation error of L_h is

$$\tau_h(\mathbf{x}) = (L - L_h)\varphi(x)$$

The approximation L_h is consistent with L if

$$\lim_{h\to 0} \tau_h(x) = 0,$$

for all $\mathbf{x} \in D$ and all $\varphi \in C^\infty(\Omega)$. The approximation is consistent to order p if $\tau_h(\mathbf{x}) = O(h^p)$. $\circ$

While we have seen this definition a few times it is always interesting how the terms are denoted and expressed but the ideas are always the same.

Proposition 9.3.1. *The five-point difference analog $-\nabla_h^2$ is consistent to order 2 with $-\nabla^2$.*

Proof. Pick $\varphi \in C^\infty(D)$, and let $(x, y) \in \Omega$ be a point such that $(x \pm h, y), (x, y \pm h) \in \Omega \bigcup \partial\Omega$. By the Taylor Theorem

$$\varphi(x \pm h, y) \quad = \quad \varphi(x, y) \pm h\frac{\partial \varphi}{\partial x}(x, y) + \frac{h^2}{2!}\frac{\partial^2 \varphi}{\partial x^2}(x, y) \pm \frac{h^3}{3!}\frac{\partial^3 \varphi}{\partial x^3}(x, y) + \frac{h^4}{4!}\frac{\partial^4 \varphi}{\partial x^4}(\zeta^\pm, y)$$

where $\zeta^\pm \in (x - h, x + h)$. Adding this pair of equation together and rearranging , we get

$$\frac{1}{h^2}[\varphi(x+h, y) - 2\varphi(x, y) + \varphi(x-h, y)] - \frac{\partial^2 \varphi}{\partial x^2}(x, y) = \frac{h^2}{4!}\left[\frac{\partial^4 \varphi}{\partial x^4}(\zeta^+, y) + \frac{\partial^4 \varphi}{\partial x^4}(\zeta^-, y)\right]$$

By the intermediate value theorem

$$\left[\frac{\partial^4 \varphi}{\partial x^4}(\zeta^+, y) + \frac{\partial^4 \varphi}{\partial x^4}(\zeta^-, y)\right] = 2\frac{\partial^4 \varphi}{\partial x^4}(\zeta, y),$$

for some $\zeta \in (x - h, x + h)$. Therefore,

$$\delta_x^2(x, y) = \frac{\partial^2 \varphi}{\partial x^2}(x, y) + \frac{h^2}{2!}\frac{\partial^4 \varphi}{\partial x^4}(\zeta, y)$$

Similar reasoning shows that

$$\delta_y^2(x, y) = \frac{\partial^2 \varphi}{\partial y^2}(x, y) + \frac{h^2}{2!}\frac{\partial^4 \varphi}{\partial y^4}(x, \eta)$$

for some $\eta \in (y-h, y+h)$. We conclude that $\tau_h(x,y) = (\nabla - \nabla_h)\varphi(x,y) = O(h^2)$. • □

Consistency does not guarantee that the solution to the difference equations approximates the exact solution to the PDE.

Definition Let $L_h w(\mathbf{x}_j) = f(\mathbf{x}_j)$ be a finite difference approximation, defined on a grid mesh size h, to a PDE $LU(\mathbf{x}) = f(\mathbf{x})$ on a simply connected set $D \subset R^n$. Assume that $w(x,y) = U(x,y)$ at all points (x,y) on the boundary $\partial\Omega$. The finite difference scheme converges (or is convergent) if

$$\max_j |U(\mathbf{x}_j) - w(\mathbf{x}_j)| \to 0 \text{ as } h \to 0.$$

◦

For the five point scheme there is a direct connection between consistency and convergence. Underlying this connection is an argument based on the following principle:

Theorem 9.3.2. *(DISCRETE MAXIMUM PRINCIPLE). If $\nabla_h^2 V_{ij} \geq 0$ for all points $(x_i, y_j) \in \Omega_h$, then*

$$\max_{(x_i,y_j)\in\Omega_h} V_{ij} \leq \max_{(x_i,y_j)\in\partial\Omega_h} V_{ij}$$

If $\nabla_h^2 V_{ij} \leq 0$ for all points $(x_i, y_j) \in \Omega_h$, then

$$\min_{(x_i,y_j)\in\Omega_h} V_{ij} \geq \min_{(x_i,y_j)\in\partial\Omega_h} V_{ij}$$

In other words, a grid function V for which $\nabla_h^2 V$ is nonnegative on Ω_h attains its maximum on the boundary $\partial\Omega_h$ of the grid. Similarly, if $\nabla_h^2 V$ is nonpositive on Ω_h, then V attains its minimum on the boundary $\partial\Omega_h$.

Proof. The proof is by contradiction. We argue for the case $\nabla_h^2 V_{ij} \geq 0$, reasoning for the case $\nabla_h^2 V_{ij} \leq 0$ begin similar.
Assume that V attains its maximum value M at an interior grid point (x_I, y_J) and that $\max_{(x_i,y_j)\in\partial\Omega_h} V_{ij} < M$. The hypothesis $\nabla_h^2 V_{ij} \geq 0$ implies that

$$V_{IJ} \leq \frac{1}{4}(V_{I+1J} + V_{I-1J} + V_{IJ+1} + V_{IJ-1})$$

This cannot hold unless

$$V_{I+1J} = V_{I-1J} = V_{IJ+1} = V_{IJ-1} = M.$$

If any of the corresponding grid points $(x_{I+1}, y_L), (x_{J-1}, y_L), (x_I, y_{L+1}), (x_I, y_{L-1})$ lies in $\partial\Omega_h$, then we have reached the desired contradiction.
Otherwise, we continue arguing in this way until we conclude that

$V_{I+iJ+j} = M$ for some point $(x_{I+iJ+j}) \in \partial\Omega$, which again gives a contradiction. • □

This leads to interesting results

Proposition 9.3.3. *1. The zero grid function (for which $U_{ij} = 0$ for all $(x_i, y_j) \in \Omega_h \bigcup \partial\Omega_h$ is the only solution to the finite difference problem*

$$\nabla_h^2 U_{ij} = 0 \textit{ for } (x_i, y_j) \in \Omega_h,$$

$$U_{ij} = 0 \textit{ for } (x_i, y_j) \in \partial\Omega_h.$$

2. For prescribed grid functions f_{ij} and g_{ij}, there exists a unique solution to the problem

$$\nabla_h^2 U_{ij} = f_{ij} \textit{ for } (x_i, y_j) \in \Omega_h,$$

$$U_{ij} = g_{ij} \textit{ for } (x_i, y_j) \in \partial\Omega_h.$$

Definition For any grid function $V : \Omega_h \bigcup \partial\Omega_h \to R$,

$$||V||_\Omega = \max_{(x_i,y_j)\in\Omega_h} |V_{ij}|,$$

$$||V||_{\partial\Omega} = \max_{(x_i,y_j)\in\partial\Omega_h} |V_{ij}|.$$

∘

Lemma 9.3.4. *If the grid function $V : \Omega_h \bigcup \partial\Omega_h \to R$ satisfies the boundary condition $V_{ij} = 0$ for $(x_i, y_j) \in \partial\Omega_h$, then*

$$||V_|||_\Omega \leq \frac{1}{8}||\nabla_h^2 V||_\Omega$$

Proof. Let $\nu = ||\nabla_h^2 V||_\Omega$. Clearly for all points $(x_i, y_j) \in \Omega_h$,

$$-\nu \leq \nabla_h^2 V_{ij} \leq \nu \tag{72}$$

Now we define $W : \Omega_h \bigcup \partial\Omega_h \to R$ by setting $W_{ij} = \frac{1}{4}[(x_i - \frac{1}{2})^2 + (y_j - \frac{1}{2})^2]$, which is nonnegative. Also $\nabla_h^2 W_{ij} = 1$ and that $||W||_{\partial\Omega} = \frac{1}{8}$. The inequality (72) implies that, for all points $(x_i, y_j) \in \Omega_h$,

$$\nabla_h^2(V_{ij} + \nu W_{ij}) \geq 0$$

$$\nabla_h^2(V_{ij} - \nu W_{ij}) \leq 0$$

By the discrete minimum principle and the fact that V vanishes on $\partial\Omega_h$

$$V_{ij} \leq V_{ij} + \nu W_{ij} \leq \nu||W||_{\partial\Omega}$$

$$V_{ij} \geq V_{ij} - \nu W_{ij} \geq -\nu||W||_{\partial\Omega}$$

Since $||W||_{\partial\Omega} = \frac{1}{8}$

$$||V_|||_\Omega \leq \frac{1}{8}\nu = \frac{1}{8}||\nabla_h^2 V||_\Omega$$

• □

Finally we prove that the five point scheme for the Poisson equation is convergent.

Theorem 9.3.5. *Let U be a solution to the Poisson equation and let w be the grid function that satisfies the discrete analog*

$$-\nabla_h^2 w_{ij} = f_{ij} \quad \text{for } (x_i, y_j) \in \Omega_h,$$

$$w_{ij} = g_{ij} \quad \text{for } (x_i, y_j) \in \partial\Omega_h.$$

Then there exists a positive constant K such that

$$||U - w||_\Omega \leq KMh^2$$

where

$$M = \left\{ \left|\left| \frac{\partial^4 U}{\partial x^4} \right|\right|_\infty, \left|\left| \frac{\partial^4 U}{\partial x^3 \partial y} \right|\right|_\infty, ..., \left|\left| \frac{\partial^4 U}{\partial y^4} \right|\right|_\infty \right\}$$

The statement of the theorem assumes that $U \in C^4(\bar{\Omega})$. This assumption holds if f and g are smooth enough.

Proof. Following from the proof of the Proposition we have

$$(\nabla_h^2 - \nabla^2) U_{ij} = \frac{h^2}{12} \left[\frac{\partial^4 U}{\partial x^4}(\zeta_i, y_j) + \frac{\partial^4 U}{\partial y^4}(x_i, \eta_j) \right]$$

for some $\zeta \in (x_{i-1}, x_{i+1})$ and $\eta_j \in (y_{j-1}, y_{j+1})$. Therefore,

$$-\nabla_h^2 U_{ij} = f_{ij} - \frac{h^2}{12} \left[\frac{\partial^4 U}{\partial x^4}(\zeta_i, y_j) + \frac{\partial^4 U}{\partial y^4}(x_i, \eta_j) \right].$$

If we subtract from this the identity equation $-\nabla_h^2 w_{ij} = f_{ij}$ and note that $U - w$ vanishes on $\partial\Omega_h$, we find that

$$\nabla_h^2 (U_{ij} - w_{ij}) = \frac{h^2}{12} \left[\frac{\partial^4 U}{\partial x^4}(\zeta_i, y_j) + \frac{\partial^4 U}{\partial y^4}(x_i, \eta_j) \right].$$

It follows that

$$||U - w||_\Omega \leq \frac{1}{8} ||\nabla_h^2 (U - w)||_\Omega \leq KMh^2$$

□

9.4 ELLIPTIC EQUATIONS QUESTIONS

1. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the explicit numerical scheme

$$w_{j-1,k} + w_{j+1,k} + w_{j,k-1} + w_{j,k+1} - 4w_{j,k} = h^2 f_{j,k}$$

for the Elliptic equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y)$$

on the rectangular domain

$$\Omega = \{(x,y) |\ a \leq x \leq b, c \leq y \leq d\}.$$

[10 marks]

b) Consider the problem

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

on the rectangular domain

$$\Omega = \{(x,y) |\ 0 \leq x \leq 1, 0 \leq y \leq 1\},$$

with the boundary conditions

$$u(x,0) = 4x^2 - 4x + 1,\ u(x,1) = 4x^2 - 4x + 1,$$

$$u(0,y) = 4y^2 - 4y + 1,\ u(1,y) = 4y^2 - 4y + 1.$$

Taking $N = 4$ steps in the x-direction and $M = 4$ steps in the y-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations.

[23 marks]

2. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the explicit numerical scheme

$$w_{j-1,k} + w_{j+1,k} + w_{j,k-1} + w_{j,k+1} - 4w_{j,k} = h^2 f_{j,k}$$

for the Elliptic equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y)$$

on the rectangular domain

$$\Omega = \{(x,y) |\ a \leq x \leq b, c \leq y \leq d\}.$$

[10 marks]

b) Consider the problem

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = xy + x^2$$

on the rectangular domain

$$\Omega = \{(x,y) |\ 0 \leq x \leq 1, 0 \leq y \leq 1\},$$

with the boundary conditions

$$u(x,0) = 0,\ u(x,1) = 0,$$

$$u(0,y) = 0,\ u(1,y) = 0.$$

Taking $N = 4$ steps in the x-direction and $M = 4$ steps in the y-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations.

[23 marks]

3. a) Use the central difference formula for the second derivative

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \mathcal{O}(h^2)$$

to derive the explicit numerical scheme

$$w_{j-1,k} + w_{j+1,k} + w_{j,k-1} + w_{j,k+1} - 4w_{j,k} = h^2 f_{j,k}$$

for the Elliptic equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y)$$

on the rectangular domain

$$\Omega = \{(x,y)|\ a \leq x \leq b, c \leq y \leq d\}.$$

[10 marks]

b) Consider the problem

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = y^2$$

on the rectangular domain

$$\Omega = \{(x,y)|\ 0 \leq x \leq 1, 0 \leq y \leq 1\},$$

with the boundary conditions

$$u(x,0) = x,\ u(x,1) = x,$$

$$u(0,y) = 0,\ u(1,y) = 1.$$

Taking $N = 4$ steps in the x-direction and $M = 4$ steps in the y-direction, set up and write in matrix form (but do not solve) the corresponding systems of finite difference equations.

[23 marks]

10

HYPERBOLIC EQUATIONS

First-order scalar equation

$$\begin{array}{ll} \frac{\partial U}{\partial t} = -a\frac{\partial U}{\partial x} & x \in R \;\; t > 0 \\ U(x,0) = U_0(x) & x \in R \end{array} \tag{73}$$

where a is a positive real number. Its solution is given by

$$U(x,t) = U_0(x - at) \quad t \geq 0$$

and represents a traveling wave with velocity a. The curves $(x(t), t)$ in the plane (x, t) are the characteristic curves. They are the straight lines $x(t) = x_0 + at$, $t > 0$. The solution of (73) remains constant along them.
For the more general problem

$$\begin{array}{ll} \frac{\partial U}{\partial t} + a\frac{\partial U}{\partial x} + a_0 = f & x \in R \;\; t > 0 \\ U(x,0) = U_0(x) & x \in R \end{array} \tag{74}$$

where a, a_0 and f are given functions of the variables (x, t), the characteristic curves are still defined as before. In this case the solutions of (74) satisfy along the characteristics the following differential equation

$$\frac{du}{dt} = f - a_0 u \text{ on } (x(t), t)$$

10.1 THE WAVE EQUATION

Consider the second-order hyperbolic equation

$$\frac{\partial^2 U}{\partial t^2} - \gamma\frac{\partial^2 U}{\partial x^2} = f \quad x \in (\alpha, \beta), \;\; t > 0 \tag{75}$$

with initial data

$$U(x,0) = u_0(x) \text{ and } \frac{\partial U}{\partial t}(x,0) = v_0(x), \;\; x \in (\alpha, \beta)$$

and boundary data

$$U(\alpha, t) = 0 \text{ and } U(\beta, t) = 0, \;\; t > 0$$

In this case, U may represent the transverse displacement of an elastic vibrating string of length $\beta - \alpha$, fixed at the endpoints and γ is a coefficient depending on the specific mass of the string and its tension. The spring is subject to a vertical force of density f.

The functions $u_0(x)$ and $v_0(x)$ denote respectively the initial displacement and initial velocity of the string.
The change of variables

$$\omega_1 = \frac{\partial U}{\partial x}, \quad \omega_2 = \frac{\partial U}{\partial t}$$

transforms (75) into

$$\frac{\partial \hat{\omega}}{\partial t} + A\frac{\partial \hat{\omega}}{\partial x} = \mathbf{0}$$

where

$$\hat{\omega} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$$

Since the initial conditions are $\omega_1(x,0) = u_0'(x)$ and $\omega_2(x,0) = v_0(x)$.

Aside

Notice that replacing $\frac{\partial^2 u}{\partial t^2}$ by t^2, $\frac{\partial^2 u}{\partial x^2}$ by x^2 and f by 1, the wave equation becomes

$$t^2 - \gamma^2 x^2 = 1$$

which represents an hyperbola in (x,t) plane. Proceeding analogously in the case of the heat equation we end up with

$$t - x^2 = 1$$

which represents a parabola in the (x,t) plane. Finally, for the Poisson equation we get

$$x^2 + y^2 = 1$$

which represents an ellipse in the (x,y) plane.
Due to the geometric interpretation above, the corresponding differential operators are classified as hyperbolic, parabolic and elliptic.

10.2 Finite Difference Method for Hyperbolic Equations

As always we discretise the domain by space-time finite difference. To this aim, the half-plane $\{(x,t) : -\infty < x < \infty, t > 0\}$ is discretised by choosing a spatial grid size Δx, a temporal step Δt and the grid points (x_j, t^n) as follows

$$x_j = j\Delta x \quad j \in Z, \quad t^n = n\Delta t \quad n \in N$$

and let

$$\lambda = \frac{\Delta t}{\Delta x}.$$

10.2.1 *Discretisation of the scalar equation*

Here are some explicit methods

- Forward Euler/centered method:

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n),$$

- Lax-Friedrichs method,

$$u_j^{n+1} = \frac{u_{j+1}^n + u_{j-1}^n}{2} - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n),$$

- Lax-Wendroff method,

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n) + \frac{\lambda^2}{2}a^2(u_{j+1}^n - 2u_j^n + u_{j-1}^n),$$

- Upwind method,

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}(u_{j+1}^n - u_{j-1}^n) + \frac{\lambda}{2}|a|(u_{j+1}^n - 2u_j^n + u_{j-1}^n).$$

The last three methods can be obtained from the forward Euler/-centered method by adding a term proportional to a numerical approximation of a second derivative term so that they can be written in the equivalent form

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2}a(u_{j+1}^n - u_{j-1}^n) + \frac{1}{2}k\frac{(u_{j+1}^n - 2u_j^n + u_{j-1}^n)}{(\Delta x)^2}$$

where k is an artificial viscosity term.
An example of an implicit method is the backward Euler/ centered scheme

$$u_j^{n+1} + \frac{\lambda}{2}a(u_{j+1}^{n+1} - u_{j-1}^{j+1}) = u_j^n$$

10.3 Analysis of the Finite Difference Methods

10.4 Consistency

A numerical method is convergent if

$$\lim_{\Delta t, \Delta x \to 0} \max_{j,n} |U(x_j, t^n) - w_j^n|$$

The local truncation error at x_j, t^n is defined as

$$\tau_j^n = L(U_j^n)$$

the truncation error is

$$\tau(\Delta t, \Delta x) = \max_{j,n} |\tau_j^n|$$

When $\tau(\Delta t, \Delta x)$ goes to zero as Δt and Δx tend to zero independently is said to be consistent.

10.5 Stability

10.6 Courant Freidrich Lewy condition

A method is said to be stable if, for any T here exist a constant $C_T > 0$ and δ_0 such that

$$||\mathbf{u}^n||_\Delta \leq C_T ||\mathbf{u}^0||_\Delta$$

for any n such that $n\Delta t \leq T$ and for any $\Delta t, \Delta x$ such that $0 < \Delta t \leq \delta_0, 0 < \Delta x \leq \delta_0$. We have denoted by $||.||_\Delta$ a suitable discrete norm.

Forward Euler/centered

$$u_j^{n+1} = u_j^n - \frac{\lambda}{2} a(u_{j+1}^n - u_{j-1}^n)$$

Truncation error

$$O(\Delta t, (\Delta x)^2)$$

For an explicit method to be stable we need

$$|a\lambda| = \left| a\frac{\delta t}{\delta x} \right| \leq 1$$

this is known as the Courant Freidrich Lewy condition.
Using Von Neumann stability analysis we can show that the method is stable under the Courant Freidrich Lewy condition.

10.6.1 *von Neumann stability for the Forward Euler*

$$u_j^n = e^{i\beta j \Delta x} \xi^n$$

where

$$\xi = e^{\alpha \Delta t}$$

It is sufficient to show

$$|\xi| \leq 1$$

$$\xi^{n+1} e^{i\beta(j)\Delta x} = \xi^n e^{i\beta(j)\Delta x} + \frac{\lambda}{2} a(\xi^n e^{i\beta(j+1)\Delta x} - \xi^n e^{i\beta(j-1)\Delta x})$$

$$\xi = 1 - \frac{\lambda}{2} a(e^{i\beta\Delta x} - e^{-i\beta\Delta x})$$
$$\xi = 1 - i\frac{\lambda}{2} a(2sin(\beta\Delta x))$$
$$\xi = 1 - i\lambda a(sin(\beta\Delta x))$$
$$|\xi| = \sqrt{1 + (\lambda a(sin(\beta\Delta x)))^2}$$

Hence
$$\xi > 1$$
therefore the method is unstable for the Courant Freidrich Lewy.

10.6.2 *von Neumann stability for the Lax-Friedrich*

$$\xi^{n+1} e^{i\beta(j)\Delta x} = \frac{\xi^n e^{i\beta(j+1)\Delta x} + \xi^n e^{i\beta(j-1)\Delta x}}{2} + \frac{\lambda}{2} a(\xi^n e^{i\beta(j+1)\Delta x} - \xi^n e^{i\beta(j-1)\Delta x})$$
$$\xi = \frac{e^{i\beta\Delta x} - e^{-i\beta\Delta x}}{2} + \frac{\lambda}{2} a(e^{i\beta\Delta x} - e^{-i\beta\Delta x})$$
$$\xi = \frac{1+\lambda a}{2} e^{i\beta\Delta x} + \frac{1-\lambda a}{2} e^{-i\beta\Delta x}$$
$$\xi = \cos(\beta\Delta x) + i\lambda a \sin(\beta\Delta x)$$
$$|\xi|^2 \le (\cos(\beta\Delta x))^2 + (a\lambda)^2 (\sin(\beta\Delta x))^2$$

Hence
$$\xi < 1$$
for $a\lambda \le 1$.

Example 50

These methods can be applied to the Burgers equation
$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0$$
as there it is a non-trivial non-linear hyperbolic equation. Taking initial condition
$$u(x,0) = u_0(x) = \begin{cases} 1, & x_0 \le 0, \\ 1-x, & 0 \le x_0 \le 1, \\ 0, & x_0 \ge 1. \end{cases}$$
the characteristic line issuing from the point $(x_0, 0)$ is given by
$$x(t) = x_0 + tu_0(x_0) = \begin{cases} x_0 + t, & x_0 \le 0 \\ x_0 + t(1-x_0), & 0 \le x_0 \le 1, \\ x_0 & x_0, \ge 1. \end{cases}$$

11

VARIATIONAL METHODS

Variational methods are based on the fact that the solutions of some Boundary Value Problems,

$$\begin{aligned} &-(p(x)u'(x))' + q(x)u(x) = g(x, u(x)) \\ &u(a) = \alpha, \quad u(b) = \beta, \end{aligned} \tag{76}$$

under the assumptions that,

$$\begin{aligned} &p \in C^1[a,b], && p(x) \geq p_0 > 0, \\ &q \in C^1[a,b], && q(x) \geq 0, \\ &g \in C^1([a,b] \times R), && g_u(x,u) \leq \lambda_0 \end{aligned} \tag{77}$$

then if $u(x)$ is the solution of (76), it can be written in the form $y(x) = u(x) - l(x)$ with

$$l(x) = \alpha\frac{b-x}{b-a} + \beta\frac{a-x}{a-b}, \ \ l(a) = \alpha, \ \ l(b) = \beta,$$

and y is the solution of a boundary value problem

$$\begin{aligned} &-(p(x)y'(x))' + q(x)y(x) = f(x), \\ &y(a), = 0 \quad y(b) = 0, \end{aligned} \tag{78}$$

with zero boundary values. Without loss of generality we can just consider problems of the form (78), is known as the:

Classical Problem (D)

$$-(p(x)u'(x))' + q(x)u(x) = f(x),$$

$$u(a) = 0, \quad u(b) = 0.$$

The assumptions on the Classical Problem can be relaxed such that $f \in L_2([0,1])$, such that

$$u(x) \in D_L = \{u \in C^2[a,b] \mid u(a) = 0, u(b) = 0\}.$$

Convolving the Classical Problem (D) with the function $v(x)$ gives the problem

$$\int_a^b [-(p(x)u'(x))' + q(x)u(x)]v(x)dx = \int_a^b f(x)v(x)dx,$$

where $v \in D_L$. Integrating by parts gives the simplified problem gives the:

Weak Form Problem (w)

$$\int_a^b [p(x)u^{'}(x)v^{'}(x) + q(x)u(x)v(x)]dx = \int_a^b f(x)v(x)dx.$$

It is sufficient to solve the *Weak Form (W)* of the Classical Problem (D). From the Weak Form we have two definitions:

1. **Definition (Bilinear Form)**

$$a(u,v) = \int_a^b [p(x)u^{'}(x)v^{'}(x) + q(x)u(x)v(x)]dx;$$

2. $$(f,v) = \int_a^b f(x)v(x)dx,$$

   where $f \in L_2([a,b])$.

From these definitions the *Weak Form* of the ODE problem (D) is then given by

$$a(u,v) = (f,v),$$

where $u \in D_L$ is the solution to the Classical Problem.
The Weak Form of the problem can be equivalently written in a *Variational or Minimisation* form of the problem is given by,

Variational/Minimization form (M):

$$F(v) = \frac{1}{2}a(v,v) - (f,v).$$

where $f \in L_2([a,b])$. This gives the problem

$$F(u) \leq F(v), \quad \text{all } v \in D_L$$

such that the function u that minimizes F over D_L.

Theorem 11.0.1. *We have the following relationships between the solutions to the three problems* ***Classical Problem (D), Weak Form (W)*** *and* ***Minimization Form (M)****.*

1. *If the function u solves* ***Classical Problem (D)****, then u solves* ***Weak Form (W)****.*

2. *The function u solves* ***Weak Form (W)*** *if and only if u solves* ***Minimization Form (M)****.*

3. *If $f \in C([0,1])$ and $u \in C^2([0,1])$ solves* ***Weak Form (W)****, then u solves* ***Classical Problem (D)****.*

Proof. 1. Let u be the solution to **Classical Problem (D)**; then u solves **Weak Form (W)** is obvious, since the Weak Form (W) derives directly from **Classical Problem (D)**.

2. a) Show **Weak Form (W)** $\Rightarrow$ **Minimization Form (M)**.
   Let u solve **Weak Form (W)**, and define $v(x) = u(x) + z(x)$, $u, z \in D_L$. By linearity

$$\begin{aligned} F(v) &= \tfrac{1}{2}a(u+z, u+z) - (f, u+z) \\ &= F(u) + \tfrac{1}{2}a(z,z) + a(u,z) - (f,z) \\ &= F(u) + \tfrac{1}{2}a(z,z) \end{aligned}$$

   which implies that $F(v) \geq F(u)$, and therefore u solves **Minimization Form (M)**.

   b) Show **Weak Form (W)** $\Leftarrow$ **Minimization Form (M)**.
   Let u solve **Minimization Form (M)** and choose $\varepsilon \in R$, $v \in D_L$. Then $F(u) \leq F(u + \varepsilon v)$, since $u + \varepsilon v \in D_L$. Now $F(u + \varepsilon v)$ is a quadratic form in ε and its minimum occurs at $\varepsilon = 0$ ie

$$\begin{aligned} 0 &= \tfrac{dF(u+\varepsilon v)}{d\varepsilon}\big|_{\varepsilon=0} \\ &= a(u,v) - (f,v), \end{aligned}$$

   it follows that u solves the **Weak Form (W)**.

3. Is immediate.

□

11.1 Ritz -Galerkin Method

This is a classical approach which we exploit to fined "discrete" approximation to the problem **Weak Form (W) / Minimization Form (M)**. We look for a solution u_S in a finite dimensional subspace S of D_L such that u_S is an approximation to the solution of the continuous problem,

$$u_S = u_1\phi_1 + u_2\phi_2 + ... + u_n\phi_n.$$

Discrete Weak Form (W_S):

Find $u_S \in S = span\{\phi_1, \phi_2, ..., \phi_n\}, \ \ n < \infty$ such that

$$a(u_S, v) = (f, v),$$

$$u \approx u_S = u_1\phi_1 + u_2\phi_2 + ... + u_n\phi_n.$$

Similarly the

Discrete Variational/Minimization form (M_S):

Find $u_S \in S = \text{span}\{\phi_1, \phi_2, ..., \phi_n\}, \ \ n < \infty$ that satisfies

$$F(u_S) \leq F(v) \qquad \text{all } v \in S,$$

where

$$F(v) = \frac{1}{2}a(v,v) - (f,v).$$

$v \in D_L$.

Theorem 11.1.1. *Given $f \in L_2([0,1])$, then (W_S) has a unique solution.*

Proof. We write $u_S = \sum_1^n u_j\phi_j(x)$ and look for constants u_j, $j = 1, ..., n$ to solve the discrete problem. We define

$$A = \{A_{ij}\} = \{a(\phi_i, \phi_j)\} = \int_a^b [p(x)\phi_i'\phi_j' + q(x)\phi_i\phi_j]dx$$

and

$$\bar{F} = \{F_j\} = \{(f, \phi_j)\} = \{\int_a^b f\phi_i dx\}$$

Then we require

$$a(u_S, v) = a(\sum_1^n u_j\phi_j(x), v) = (f, v) \text{ all } v \in S$$

Hence, for each basis function $\phi_i \in S$ we must have,

$$a(u_S, \phi_i) = a(\sum_1^n u_j\phi_j(x), \phi_i) = (f, \phi_i) \text{ all } i = 1, ..., n \in S$$

this gives the matrix,

$$\begin{bmatrix} a(\phi_1,\phi_1) & ... & a(\phi_n,\phi_1) \\ . & . & . \\ . & . & . \\ . & . & . \\ a(\phi_1,\phi_n) & ... & a(\phi_n,\phi_n) \end{bmatrix} \begin{bmatrix} u_1 \\ . \\ . \\ . \\ u_n \end{bmatrix} = \begin{bmatrix} (f,\phi_1) \\ . \\ . \\ . \\ (f,\phi_n) \end{bmatrix}$$

which can be written as,

$$A\bar{u} = \bar{F}$$

Hence u_S is found by the solution to a matrix equation. We now show existence/uniqueness of the solution to the algebraic problem. We show by contradiction that A is full-rank ie that the only solution to $A\bar{u} = 0$ is $\bar{u} = 0$.
Suppose that there exists a vector $\bar{v} = \{v_j\} \neq 0$ such that $A\bar{v} = 0$ and construct $v(x) = \sum_1^n v_j\phi_j \in S$. Then

$$\begin{aligned} A\bar{v} = 0 \quad &\Leftrightarrow \quad \textstyle\sum_j a(\phi_j, \phi_k)v_j = a(v, \phi_k) = 0 \text{ all } k \\ &\Leftrightarrow \quad \textstyle\sum_k a(v, \phi_k)v_j = a(v, \sum v_k\phi_k) = a(v, v) = 0 \\ &\Leftrightarrow \quad v = 0 \end{aligned}$$

Therefore a contradiction. □

Classically, in the Ritz-Galerkin method, the basis functions are chosen to be continuous functions over the entire interval $[a, b]$, for example, $\{\sin(mx), \cos(mx)\}$ give us trigonometric polynomial approximations to the solutions of the ODEs.

11.2 FINITE ELEMENT

We choose the basis functions $\{\phi_i\}_1^n$ to be piecewise polynomials with compact support. In the simplest case ϕ_i is linear. We divide the region in to n intervals or "elements",

$$a = x_0 < x_1 < \dots < x_n = b$$

and let E_i denote the element $[x_{i-1}, x_i]$, $h_i = x_i - x_{i-1}$.

Definition Let $S^h \subset D$ be the space of functions such that $v(x) \in [0, 1]$, $v(x)$ is linear on E_i and $v(a) = v(b) = 0$ ie

$$S^h = \{v(x) : \text{ piecewise linear on } [0, 1], v(a) = v(b) = 0\}$$

The basis functions $\phi_i(x)$ for S^h are defined such that $\phi_i(x)$ is linear on E_i, E_{i+1} and $\phi_i(x_j) = \delta_{ij}$.

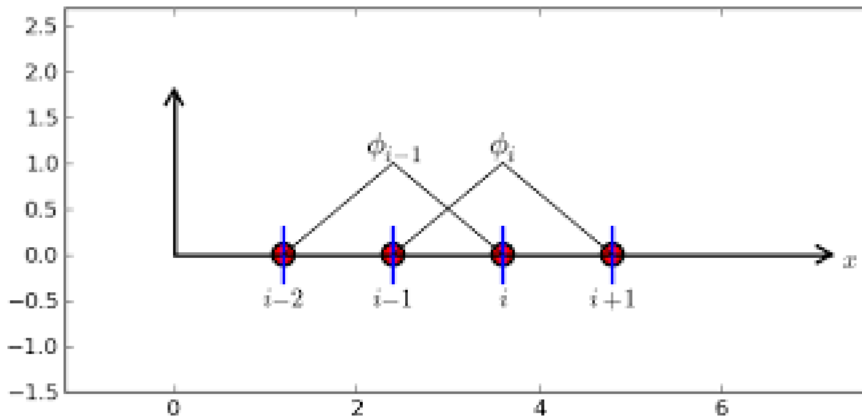


Figure 11.2.1: Hat functions ϕ_i form a basis for the space S^h

We now show that the hat functions ϕ_i form a basis for the space S^h (Figure 11.2.1).

Lemma 11.2.1. *The set of functions $\{\phi_i\}_i^n$ is a basis for the space S^h.*

Proof. We show first that the set $\{\phi_i\}_1^n$ is linearly independent. If $\sum_1^n c_i\phi_i(x) = 0$ for all $x \in [a, b]$, then taking $x = x_j$, implies $c_j = 0$ for each value of j, and hence the functions are independent.
To show $S^h = \text{span}\{\phi_i\}$, we only need to show that

$$v(x) = v_I = \sum v_j\phi_j, \text{ all } v(x) \in S^h$$

This is proved by construction. Since $(v - v_I)$ is linear on $[x_{i-1}, x_i]$ and $v - v_I = 0$ at all points x_j, it follows that $v = v_I$ on E_i. □

We now consider the matrix $A\hat{u} = \hat{F}$ in the case where the basis functions are chosen to be the "hat functions". In this case the elements of A can be found We have

$$\phi_i = 0, \phi_i' = 0, \text{ for } x \notin [x_{i-1}, x_{i+1}) = E_i \bigcup E_{i+1},$$

where

$$\phi_i = \frac{x - x_{i-1}}{x_i - x_{i-1}} = \frac{1}{h_i}(x - x_{i-1}), \quad \phi_i' = \frac{1}{h_i}, \text{ on } E_i.$$

and

$$\phi_i = \frac{x_{i+1} - x}{x_{i+1} - x_i} = \frac{1}{h_{i+1}}(x_{i+1} - x), \quad \phi_i' = \frac{-1}{h_{i+1}}, \text{ on } E_{i+1}.$$

Therefore we have the elements of the matrix A

$$\begin{aligned} A_{i,i} &= \int_{x_{i-1}}^{x_i} \tfrac{1}{h_i^2} p(x)dx + \int_{x_i}^{x_{i+1}} \tfrac{1}{h_{i+1}^2} p(x)dx \\ &+ \int_{x_{i-1}}^{x_i} \tfrac{1}{h_i^2}(x - x_{i-1})^2 q(x)dx + \int_{x_i}^{x_{i+1}} \tfrac{1}{h_{i+1}^2}(x_{i+1} - x)^2 q(x)dx, \end{aligned}$$

$$A_{i,i+1} = \int_{x_i}^{x_{i+1}} \tfrac{-1}{h_{i+1}^2} p(x)dx + \int_{x_i}^{x_{i+1}} \tfrac{1}{h_{i+1}^2}(x_{i+1} - x)(x - x_i)q(x)dx,$$

$$A_{i,i-1} = \int_{x_{i-1}}^{x_i} \tfrac{-1}{h_i^2} p(x)dx + \int_{x_i}^{x_{i+1}} \tfrac{1}{h_i^2}(x_i - x)(x - x_{i-1})q(x)dx,$$

and

$$F_i = \int_{x_{i-1}}^{x_i} \tfrac{1}{h_i}(x - x_{i-1})f(x)dx + \int_{x_i}^{x_{i+1}} \tfrac{1}{h_{i+1}}(x_{i+1} - x)f(x)dx.$$

11.2.1 *Error bounds of Finite Element methods*

Lemma 11.2.2. *Assume u_S solves $(\mathbf{W_S})$. Then*

$$a(u - u_S, w) = 0, \textit{ for all } x \in S$$

Proof. Given that

$$a(u_S, w) = (f, w),$$

and

$$a(u, w) = (f, w),$$

for all $w \in S$. Since a is bilinear, taking the differences gives

$$a(u - u_S, w) = 0.$$

□

The error bounds we are interested in will be in term of the energy norm,

$$||v||_E = [a(v, v)]^{\frac{1}{2}}$$

for all $v \in D_L$. The function satisfies the properties:

$$||\alpha v||_E = \alpha||a||_E, \quad ||v+z||_E \leq ||v||_E + ||z||_E$$

Theorem 11.2.3. *To show u_S is the best fit we show that*

$$||u - u_S||_E = \min_{v \in S} ||u - v||_E$$

Proof. By the Cauchy -Schwarz Lemma, we have $|a(u,v)| \leq ||u||_E||v||_E$. Let $w = u_S - v \in S$. Using the previous lemma we obtain

$$\begin{aligned} ||u - u_S||_E^2 &= a(u-u_s, u-u_s) \\ &\leq a(u-u_s, u-u_s) + a(u-u_s, w) \\ &\leq a(u-u_s, u-u_s+w) = a(u-u_s, u-v) \\ &\leq ||u-u_s||_E||u-v||_E. \end{aligned}$$

If $||u - u_S||_E = 0$, then the theorem holds. Otherwise

$$\min ||u-v||_E \leq ||u-u_S|| \leq \min ||u-v||_E,$$

the result follows. □

Theorem 11.2.4. *Error bounds*

$$||u - u_S||_E \leq Ch||u^{''}||_\infty$$

where C is a constant.

Proof. First from the previous theorem we have that

$$||u - u_S||_E = \min_{v \in S} ||u - v||_E \leq ||u - u_I||_E$$

We look for a bound on $||u - u_I||_E$, where

$$u_I(x) = \sum_j \bar{u}_j \phi_j, \quad \bar{u}_j = u(x_j).$$

We assume that

$$u_S(x) = \sum_j u_j \phi_j$$

where $\mathbf{u} = \{u_j\}$ solves $A\mathbf{u} = \mathbf{F}$. We define $e = u - u_I$. Since $u_I \in S$ implies that u_I is piecewise linear, then $u_I^{''} = 0$. Therefore $e^{''} = u^{''}$. Looking at the subinterval $[x_i, x_{i+1}]$ The Schwarz inequality yields the estimate

$$\begin{aligned} (e)^2 &\leq \int_{x_i}^{x} 1^2 d\xi \int_{x_i}^{x} (e'(\xi))^2 d\xi \\ &\leq (x - x_i) \int_{x_i}^{x} (e'(\xi))^2 d\xi \\ &\leq h_i \int_{x_i}^{x_{i+1}} (e'(\xi))^2 d\xi \end{aligned}$$

and thus

$$||e||_\infty^2 \leq h_i \int_{x_i}^{x_{i+1}} (e'(\xi))^2 d\xi \leq h_i^2 ||e'||_\infty^2$$

Similarly,

$$\begin{aligned}
(e')^2 &\leq \int_{x_i}^{x} 1^2 d\xi \int_{x_i}^{x} (e''(\xi))^2 d\xi \\
&\leq (x - x_i) \int_{x_i}^{x} (e''(\xi))^2 d\xi \\
&\leq h_i \int_{x_i}^{x_{i+1}} (e''(\xi))^2 d\xi
\end{aligned}$$

and thus

$$||e'||_\infty^2 \leq h_i \int_{x_i}^{x_{i+1}} (e''(\xi))^2 d\xi \leq h_i^2 ||e''||_\infty^2$$

Finally we also have

$$\begin{aligned}
a(e,e) = \; & \int_{x_i}^{x_{i+1}} (p(x)[e']^2 + q(x)[e(x)]^2) dx \\
& \leq ||p||_\infty \int_{x_i}^{x_{i+1}} [e']^2 + ||q||_\infty \int_{x_i}^{x_{i+1}} [e(x)]^2 dx \\
& \leq ||p||_\infty h_i^2 ||e''||_\infty^2 + ||q||_\infty h_i^2 ||e'||_\infty^2 \\
& \leq ||p||_\infty h_i^2 ||e''||_\infty^2 + ||q||_\infty h_i^4 ||e'||_\infty^2 \\
& \leq C h_i^2 ||u''||_\infty^2
\end{aligned}$$

$$||u - u_S||_E = \min_{v \in S} ||u - v||_E \leq ||u - u_I||_E \leq Ch||u''||_\infty$$

where $h = max\{h_i\}$. □

PROBLEM SHEET

1. a) State the 3 classes and conditions of 2nd order Partial Differential Equations defined by the characteristic curves.

   b) Given the non-dimensional form of the heat equation

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

   supply sample boundary conditions to specify this problem.
   Write a fully implicit scheme to solve this partial differential equation.

   c) Derive the local truncation error for the fully implicit method, for the heat equation.

   d) Show that the method is unconditionally stable using von Neumann's method.

2. a) State the 3 classes and conditions of 2nd order Partial Differential Equations defined by the characteristic curves.

   b) Given the non-dimensional form of the heat equation

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

   supply sample boundary conditions to specify this problem.
   Write an explicit scheme to solve this partial differential equation.

   c) Derive the local truncation error for the explicit method, for the heat equation.

   d) Show that the method is consistent, convergent and stable for $\frac{k}{h^2} < \frac{1}{2}$, where k is the step-size in the t direction and h is the step-size in the x direction.

3. a) State the 3 classes and conditions of 2nd order Partial Differential Equations defined by the characteristic curves.

   b) Given the non-dimensional form of the heat equation

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

supply sample boundary conditions to specify this problem.
Write an the Crank-Nicholson method to solve this partial differential equation.

c) Derive the local truncation error for the Crank-Nicholson method, for the heat equation.

d) Show that the method is unconditionally stable using von Neumann's method.

4. a) Approximate the Poisson equation

$$-\nabla^2 U(x,y) = f(x,y) \qquad (x,y) \in \Omega = (0,1) \times (0,1)$$

with boundary conditions

$$U(x,y) = g(x,y) \qquad (x,y) \in \delta\Omega - boundary$$

using the five point method. Sketch how the finite difference scheme may be rewritten in the form $Ax = b$, where A is a sparse $N^2 \times N^2$ matrix, b is an N^2 component matrix and x is an N^2 component vector of unknowns. (Assume your 2d discretised grid contains N components in the x and y direction).

b) Prove (DISCRETE MAXIMUM PRINCIPLE). if $\nabla_h^2 V_{ij} \geq 0$ for all points $(x_i, y_j) \in \Omega_h$, then

$$\max_{(x_i,y_j)\in\Omega_h} V_{ij} \leq \max_{(x_i,y_j)\in\partial\Omega_h} V_{ij}$$

If $\nabla_h^2 V_{ij} \leq 0$ for all points $(x_i, y_j) \in \Omega_h$, then

$$\min_{(x_i,y_j)\in\Omega_h} V_{ij} \geq \min_{(x_i,y_j)\in\partial\Omega_h} V_{ij}$$

c) Hence prove:
Let U be a solution to the Poisson equation and let w be the grid function that satisfies the discrete analog

$$-\nabla_h^2 w_{ij} = f_{ij} \quad \text{for } (x_i, y_j) \in \Omega_h,$$

$$w_{ij} = g_{ij} \quad \text{for } (x_i, y_j) \in \partial\Omega_h.$$

Then there exists a positive constant K such that

$$||U - w||_\Omega \leq KMh^2$$

where

$$M = \left\{ \left|\left| \frac{\partial^4 U}{\partial x^4} \right|\right|_\infty, \left|\left| \frac{\partial^4 U}{\partial x^3 \partial y} \right|\right|_\infty, , ..., \left|\left| \frac{\partial^4 U}{\partial y^4} \right|\right|_\infty \right\}$$

You may assume:
Lemma
If the grid function $V : \Omega_h \bigcup \partial\Omega_h \rightarrow R$ satisfies the boundary condition $V_{ij} = 0$ for $(x_i, y_j) \in \partial\Omega_h$, then

$$||V||_\Omega \leq \frac{1}{8} ||\nabla_h^2 V||_\Omega$$

5. a) For a finite difference scheme approximating a partial differential equation of the form

$$\frac{\partial U}{\partial t} = -a\frac{\partial U}{\partial x} + f(x,t), \quad x \in R, \quad t > 0$$

$$U(x,0) = U_0(x), \quad x \in R$$

define what is meant by:

   i. convergence,
   ii. consistency,
   iii. stability.

   b) Describe the forward Euler/centered difference method for the transport equation and derive the local truncation error.

   c) Define the Courant Friedrichs Lewy condition and state how it is related to stability.

   d) Show that the method is stable under the Courant Friedrichs Lewy condition using Von Neumann analysis, you may assume $f(x,t) = 0$.

6. Consider the second order differential equation

$$\frac{d^2 u}{dx^2} + u = x$$

with boundary conditions

$$u(0) = 0 \qquad u(1) = 0$$

   a) Show that the solution $u(x)$ of this equation satisfies the weak form

$$\int_0^1 dx \left(-\frac{du}{dx}\frac{dv}{dx} + uv - xv \right) = 0$$

for all $v(x)$ which are sufficiently smooth and which satisfy

$$v(0) = 0 \qquad v(1) = 0$$

b) By splitting the interval $x \in [0,1]$ into N equal elements of size h, where $Nh = 1$, one can define nodes x_i and finite element shape functions as follows

$$x_i = ih$$

$$\phi_i(x) = \begin{cases} 0 & 0 \le x \le x_{i-1} \\ \frac{x-x_{i-1}}{h} & x_{i-1} \le x \le x_i \\ \frac{x_{i+1}-x}{h} & x_i \le x \le x_{i+1} \\ 0 & x_{i+1} \le x \le 1 \end{cases}$$

A finite element approximation to the differential equation is obtained by approximating $u(x)$ and $v(x)$ with linear combinations of these finite element shape functions, ϕ_i, where

$$u_n = \sum_{i=1}^{N-1} \alpha_i \phi_i(x)$$

$$v_n = \sum_{j=1}^{N-1} \beta_j \phi_j(x)$$

Show that the equation which results from this approximation has the form

$$K\alpha = F$$

where K is an $N-1 \times N-1$ sparse matrix, F is an $N-1$ component vector and α is an $N-1$ component vector of unknown co-efficient α_i.

c) What structure does the matrix K have? Evaluate the first component of the main diagonal of K.

BIBLIOGRAPHY

[1] Ian Jaques and Colin Judd, *Numerical Anaylsis*, Chapman and Hall Ltd, 1987

[2] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag 1980

[3] Walter Gautschi, *Numerical Analysis an Introduction*, Birkhauser 1997

[4] Curtis F. Gerald, Patrick O. Wheatley, *Applied Numerical Analysis*, Addison-Wesley 1994

[5] Richard L. Burden, J. Douglas Faires, *Numerical Analysis*, Brooks/Cole 1997

[6] Myron B Allen III, Eli L. Isaacson *Numerical Analysis for Applied Science* John Wiley and Sons, Inc. 1997

[7] G D Smith *Numerical Solution of Partial Differential Equations: Finite Difference Method* Oxford 1992

[8] Atkinson, Han *Elementary Numerical Analysis*

[9] Alfio Quarteroni, Riccardo Sacco, Fausto Saleri*Numerical Mathematics*

[10] John H Mathews, Kurtis Fink *Numerical Method using MATLAB*

[11] Brian Bradie, *A Friendly Introduction to Numerical Analysis*

INDEX