

ANSWERS OF THE TEST NUMERICAL METHODS FOR  
DIFFERENTIAL EQUATIONS (WI3097 TU)

Tuesday 1 February 2005, 14:00-17:00

1. (a) Use the transformation:

$$\begin{aligned}y_1 &= y, \\y_2 &= y',\end{aligned}$$

This implies that

$$\begin{aligned}y_1' &= y' = y_2, \\y_2' &= y'' = -qy_1 - py_2 + \cos t,\end{aligned}$$

So the matrix  $\mathbf{A}$  and vector  $\mathbf{g}$  are:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix}; \quad \mathbf{g}(t) = \begin{pmatrix} 0 \\ \cos t \end{pmatrix}.$$

Characteristic equation:  $\lambda^2 + p\lambda + q = 0$ .  $\lambda_{1,2} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$ .

- (b) Replace  $f(t, y)$  by  $\lambda y$  in the RK<sub>4</sub> formulas:

$$\begin{aligned}k_1 &= h\lambda w_n \\k_2 &= h\lambda(w_n + \frac{1}{2}k_1) = h\lambda(1 + \frac{1}{2}h\lambda)w_n \\k_3 &= h\lambda(w_n + \frac{1}{2}k_2) = h\lambda(1 + \frac{1}{2}h\lambda(1 + \frac{1}{2}h\lambda))w_n \\k_4 &= h\lambda(w_n + k_3) = h\lambda(1 + h\lambda(1 + \frac{1}{2}h\lambda(1 + \frac{1}{2}h\lambda)))w_n\end{aligned}$$

Substitution of these expressions into:

$$w_{n+1} = w_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

and collecting like powers of  $h\lambda$  yields:

$$w_{n+1} = [1 + h\lambda + \frac{1}{2}(h\lambda)^2 + \frac{1}{6}(h\lambda)^3 + \frac{1}{24}(h\lambda)^4]w_n.$$

The amplification factor is therefore:

$$Q(h\lambda) = 1 + h\lambda + \frac{1}{2}(h\lambda)^2 + \frac{1}{6}(h\lambda)^3 + \frac{1}{24}(h\lambda)^4.$$

(c) The local truncation error is defined as

$$\tau_{n+1} = \frac{y(t_{n+1}) - \overline{w_{n+1}}}{h}, \quad (1)$$

where  $\overline{w_{n+1}}$  is the numerical solution at  $t_{n+1}$ , obtained by starting from the exact value  $y(t_n)$  in stead of  $w_n$ . Repeating the derivation under (b), with  $w_n$  replaced by  $y(t_n)$ , gives:

$$\overline{w_{n+1}} = Q(h\lambda)y(t_n).$$

Using furthermore  $y(t_{n+1}) = e^{h\lambda}y(t_n)$  in (1) it follows that

$$\tau_{n+1} = \frac{e^{h\lambda} - Q(h\lambda)}{h}y(t_n).$$

Canceling the first five terms of the expansion of  $e^{h\lambda}$  against  $Q(h\lambda)$ , the required order of magnitude of  $\tau_{n+1}$  follows.

(d) Substitution of the values of p and q into the matrix **A** yields the eigenvalues  $\lambda_{1,2} = -1000 \pm 3i$ . From the given drawing of the stability region the following can be inferred. Because the imaginary part is much smaller than the real part, an approximate stability condition can be obtained by simply neglecting the imaginary part. Then  $h \leq 2.8/1000$  follows as the stability condition.

(e)

$$y'' + py' + qy = \cos t, \quad y(0) = y_0, \quad y'(0) = y'_0. \quad (2)$$

The general solution of (2) is the sum of a homogeneous part, governed by the eigenvalues, and the so called particular solution, which is some linear combination of  $\sin t$  and  $\cos t$ . Both exponentials in the homogeneous part are damped very rapidly because of the (in absolute value) large real part of the eigenvalues and, after a short time of order  $10^{-3}$ , the solution becomes practically equal to the 'smooth' particular solution (time scale of order 1). The smooth solution can be integrated accurately by RK<sub>4</sub> with a 'large' stepsize: a step size of 0.1, let us say, would give an error of order  $10^{-4}$  which is sufficient for most engineering purposes. However stability, governed by the eigenvalues, requires that the stepsize be restricted (see part (d)) to 0.0028. So the stability requirement forces us to choose a stepsize yielding an unnecessarily accurate solution, which is inefficient.

The Crank-Nicolson method, on the other hand, is stable for all stepsizes. So the stepsize is restricted by accuracy requirements only. Crank-Nicolson has a global error of order  $h^2$  such that a good accuracy may be expected for stepsizes of about 0.01, larger than the restriction 0.0028. An efficiency gain may be obtained in spite of the extra work connected with the implicitness of the method.

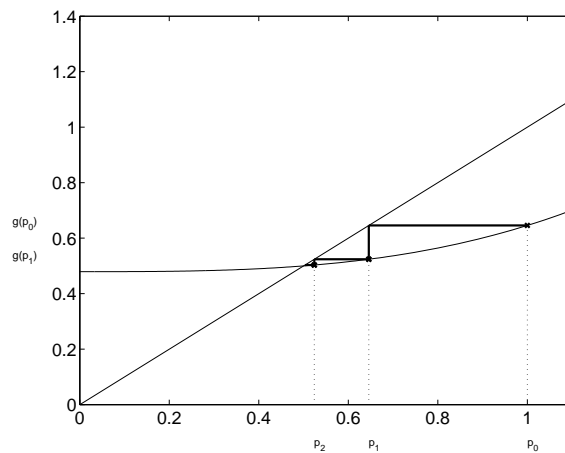
- (f) For the parameter values of part (d) and a right hand side  $\cos(1000t)$  in (2) the homogeneous and the particular part of the solution have the same time scales (viz. order  $10^{-3}$ ). Now much smaller stepsizes than 0.1 are required for RK<sub>4</sub> to compute the short scaled solution accurately and accuracy will most probably restrict the stepsize more severely than stability. There will be no need for better stability properties in this case and RK<sub>4</sub> is preferred because it is explicit.
2. (a) A fixed point  $p$  satisfies the equation  $p = g(p)$ . Substitution gives:  $p = \frac{p^3}{6} + \frac{23}{48}$ . Rewriting this expression gives:

$$\begin{aligned} -\frac{p^3}{6} + p - \frac{23}{48} &= 0 \\ -p^3 + 6p - \frac{23}{8} &= 0 \\ -p^3 + 6p - 2\frac{7}{8} &= 0 \\ f(p) &= 0 \end{aligned}$$

The fixed point iteration is defined by:  $p_{i+1} = g(p_i)$ . Starting with  $p_0 = 1$  one obtains:

$$\begin{aligned} p_1 &= 0.6458 \\ p_2 &= 0.5241 \\ p_3 &= 0.5032 \end{aligned}$$

- (b) The fixed point iteration is illustrated in the next figure.



(c) For the convergence two conditions should be satisfied:

- $g(p) \in [0, 1]$  for all  $p \in [0, 1]$ .
- $|g'(p)| \leq k < 1$  for all  $p \in [0, 1]$ .

Since  $g(p) = \frac{p^3}{6} + \frac{23}{48}$  it follows that  $g'(p) = \frac{p^2}{2}$ . Note that  $g'(p) \geq 0$  for all  $p \in [0, 1]$ . This implies that  $\frac{23}{48} = g(0) \leq g(p) \leq g(1) = \frac{31}{48}$  for all  $p \in [0, 1]$ , so the first condition holds. For the second condition we note that  $|g'(p)| = \frac{p^2}{2} \leq \frac{1}{2} = k < 1$  for all  $p \in [0, 1]$ , so the second condition is also satisfied, which implies that the fixed point iteration is convergent for all  $p_0 \in [0, 1]$ .

(d) Graphically the Newton-Raphson method is given in Figure 1. The tangent in

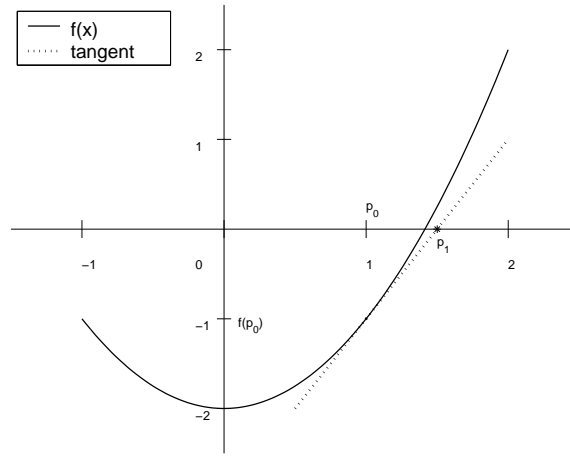


Figure 1: The Newton-Raphson method

$(p_0, f(p_0))$  is given by:

$$l(x) = f(p_0) + (x - p_0)f'(p_0)$$

Taking  $l(p_1) = 0$  leads to

$$f(p_0) + (p_1 - p_0)f'(p_0) = 0$$

Rewriting gives  $p_1 = p_0 - \frac{f(p_0)}{f'(p_0)}$ .

(e) Starting with  $p_0 = 1$  we note that

$$\begin{aligned} f(p) &= -p^3 + 6p - 2\frac{7}{8}, & f(1) &= -1 + 6 - 2\frac{7}{8} = 2\frac{1}{8} \\ f'(p) &= -3p^2 + 6, & f'(1) &= 3 \end{aligned}$$

Substituting this into the formula gives  $p_1 = 1 - \frac{2\frac{1}{8}}{3} = 0.2917$ .

(f) There are two answers possible:

- Note that the Newton-Raphson method can be written as a fixed point iteration  $g_{NR}(p) = p - \frac{f(p)}{f'(p)}$ . The convergence speed is determined by  $g'$  in the fixed point. Note that  $g'(0.5) = \frac{1}{8}$  and  $g'_{NR}(0.5) = 0$ , so the Newton-Raphson method converges faster than the fixed point method given in (a).
- One can also prove that the convergence of the Newton-Raphson method is quadratically, whereas the convergence of a fixed point method is linearly. The proof for Newton-Raphson runs as follows:

$$0 = f(p) = f(p_n) + (p - p_n)f'(p_n) + \frac{(p - p_n)^2}{2}f''(\xi_n), \xi_n \in (p_n, p).$$

Due to the definition one has

$$0 = f(p_n) + (p_{n+1} - p_n)f'(p_n).$$

Subtraction yields

$$p_{n+1} - p = (p_n - p)^2 \frac{f''(\xi)}{2f'(p_n)}.$$

which implies quadratic convergence.