Final Exam Solutions 550.386, May 4, 2007

1. Below is the array of finite-differences that is produced in applying the Newton algorithm to find the interpolating polynomial for the six points $(x, y) = (1, 7), (2, -4), (3, 3), (4, -2), (5, 11), (6, 4)$:

x	y	$1st$	$2nd$	$3rd$	$4th$	$5th$
1.0000	7.0000	0	0	0	0	0
2.0000	-4.0000	-11.0000	0	0	0	0
3.0000	3.0000	7.0000	9.0000	0	0	0
4.0000	-2.0000	-5.0000	-6.0000	-5.0000	0	0
5.0000	11.0000	13.0000	9.0000	5.0000	2.5000	0
6.0000	4.0000	-7.0000	-10.0000	-6.3333	-2.8333	-1.0667

The last five columns contain the 1st, 2nd, 3rd, 4th and 5th differences, respectively.

(a) Write down the interpolating polynomial $P_4(x)$ for the five points $(x, y) = (1, 7)$, $(2, -4)$, $(3, 3)$, $(4, -2)$, $(5, 11)$.

(b) Write down the interpolating polynomial $P_5(x)$ for all six points $(x, y) = (1, 7)$, $(2, -4)$, $(3, 3)$, $(4, -2)$, $(5, 11)$, and $(6, 4)$. If a seventh point is added with $x = 7$, how will the interpolating polynomial $P_6(x)$ change?

(c) Assume that the points are obtained from a smooth function f as $(x, y) = (x, f(x))$. Write down the error in approximating $f(x)$ at the point x with the polynomial $P_6(x)$, using an expression which involves a suitable finite-difference of f.

(d) Continuing (c), write down a formula for the error in approximating the derivative $f'(x)$ at the point x with the polynomial $P_6'(x)$, using again finite-differences of f.

Solution:

(a) Let $\psi_i(x) = (x - 1)(x - 2) \cdots (x - i)$ for $i = 1, ..., 6$. Then,

$$P_4(x) = 7 - 11\psi_1(x) + 9\psi_2(x) - 5\psi_3(x) + 2.5\psi_4(x).$$

(b) With the same notations,

$$P_5(x) = 7 - 11\psi_1(x) + 9\psi_2(x) - 5\psi_3(x) + 2.5\psi_4(x) - 1.0\bar{6}\psi_5(x).$$

If a seventh point is added at $x = 7$, then

$$P_6(x) = 7 - 11\psi_1(x) + 9\psi_2(x) - 5\psi_3(x) + 2.5\psi_4(x) - 1.0\bar{6}\psi_5(x) + f[1, 2, 3, 4, 5, 6, 7]\psi_6(x),$$

where $f[1,2,3,4,5,6,7]$ is the finite difference constructed with the new point.

(c) The error is

$$f(x) - P_5(x) = f[1,2,3,4,5,6,x]\psi_6(x).$$

(d) The error in the derivative is

$$f'(x) - P_5'(x) = f[1,2,3,4,5,6,x,x]\psi_6(x) + f[1,2,3,4,5,6,x]\psi_6'(x).$$

2. (a) Consider three times t_{n-1}, $t_n = t_{n-1} + h$, $t_{n+1} = t_n + h$ for some $h > 0$ and three vector values $\mathbf{y}_{n-1}$, $\mathbf{y}_n, \mathbf{y}_{n+1}$. Write down the Lagrange form of the interpolating polynomial $\mathbf{P}_2(t)$ that passes through the points $(t_i, \mathbf{y}_i)$ for $i = n-1, n, n+1$.

(b) Calculate the derivative $\mathbf{P}_2'(t)$.

(c) Use part (b) to construct the 2-point Backward Differentiation Formula (BDF2)

$$\mathbf{P}_2'(t_{n+1}) \doteq \mathbf{f}(t_{n+1}, \mathbf{y}_{n+1}).$$

Write this in the standard form

$$\mathbf{y}_{n+1} = \alpha_0 \mathbf{y}_n + \alpha_1 \mathbf{y}_{n-1} + \beta h \mathbf{f}(t_{n+1}, \mathbf{y}_{n+1}).$$

Solution:

(a)

$$\begin{aligned}
\mathbf{P}_2(t) &= \mathbf{y}_{n+1}\frac{(t-t_n)(t-t_{n-1})}{(t_{n+1}-t_n)(t_{n+1}-t_{n-1})} + \mathbf{y}_n\frac{(t-t_{n+1})(t-t_{n-1})}{(t_n-t_{n+1})(t_n-t_{n-1})} \\
&\qquad + \mathbf{y}_{n-1}\frac{(t-t_{n+1})(t-t_n)}{(t_{n-1}-t_{n+1})(t_{n-1}-t_n)} \\
&= \frac{1}{2h^2}\mathbf{y}_{n+1}(t-t_n)(t-t_{n-1}) - \frac{1}{h^2}\mathbf{y}_n(t-t_{n+1})(t-t_{n-1}) \\
&\qquad + \frac{1}{2h^2}\mathbf{y}_{n-1}(t-t_{n+1})(t-t_n).
\end{aligned}$$

(b)

$$\begin{aligned}
\mathbf{P}_2'(t) &= \frac{1}{2h^2}\mathbf{y}_{n+1}[(t-t_n)+(t-t_{n-1})] - \frac{1}{h^2}\mathbf{y}_n[(t-t_{n+1})+(t-t_{n-1})] \\
&\qquad + \frac{1}{2h^2}\mathbf{y}_{n-1}[(t-t_{n+1})+(t-t_n)].
\end{aligned}$$

(c)

$$\begin{aligned}
\mathbf{P}_2'(t_{n+1}) &= \frac{1}{2h^2}\mathbf{y}_{n+1}[h+2h] - \frac{1}{h^2}\mathbf{y}_n[0+2h] + \frac{1}{2h^2}\mathbf{y}_{n-1}[0+h] \\
&= \frac{1}{h}[(3/2)\mathbf{y}_{n+1} - 2\mathbf{y}_n + (1/2)\mathbf{y}_{n-1}] \\
&\doteq \mathbf{f}(t_{n+1}, \mathbf{y}_{n+1})
\end{aligned}$$

gives

$$\mathbf{y}_{n+1} = (4/3)\mathbf{y}_n - (1/3)\mathbf{y}_{n-1} + (2/3)h\mathbf{f}(t_{n+1}, \mathbf{y}_{n+1}).$$

3. Consider a general 4th-order Runge-Kutta scheme

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\mathbf{F}(t_n, \mathbf{y}_n; \mathbf{f}, h)$$

with local truncation error $\mathbf{T}_n(\mathbf{y}) = \boldsymbol{\varphi}(t_n)h^5 + O(h^6)$.

(a) If $\mathbf{y}(t)$ is the exact solution of the ODE $\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y})$ and if $\mathbf{y}(t; h)$ is the approximate solution with stepsize h, there is an asymptotic error formula of the form

$$\mathbf{y}(t) - \mathbf{y}(t; h) = \boldsymbol{\delta}(t)h^4 + O(h^5)$$

for $h \to 0$. What ODE does the coefficient $\boldsymbol{\delta}(t)$ satisfy?

(b) Use the above asymptotic error formula for approximate solutions $\mathbf{y}(t; h)$ and $\mathbf{y}(t; 2h)$ to derive an error estimate involving only these two approximate solutions. (*Hint:* eliminate the exact solution!)

(c) Apply the error estimate in (b) to approximate the error in calculating $y(1) = e$ by solving $\dot{y} = y$, $y(0) = 1$ with the classical 4th-order Runge-Kutta method for $N = 20$ steps, by comparing with the solution for $N = 10$ steps. Determine the true error and the approximation to the error.

(d) Use part (c) and extrapolation to develop an improved estimate for $y(1) = e$. Determine the error in the extrapolated estimate.

Solution:

(a) For $\mathbf{J}(t, \mathbf{y}(t))$ the Jacobian matrix, the coefficient satisfies

$$\dot{\boldsymbol{\delta}}(t) = \mathbf{J}(t, \mathbf{y}(t))\boldsymbol{\delta}(t) + \boldsymbol{\varphi}(t).$$

The initial condition is $\boldsymbol{\delta}(t_0) = \boldsymbol{\delta}_0$ if $\mathbf{y}(t_0) - \mathbf{y}_0 = \boldsymbol{\delta}_0 h^4 + O(h^5)$.

(b) Since

$$\mathbf{y}(t) - \mathbf{y}(t; 2h) = 16\boldsymbol{\delta}(t)h^4 + O(h^5),$$

one finds by subtracting that

$$\mathbf{y}(t; h) - \mathbf{y}(t; 2h) = 15\boldsymbol{\delta}(t)h^4 + O(h^5).$$

Thus,

$$(1/15)[\mathbf{y}(t; h) - \mathbf{y}(t; 2h)] = \boldsymbol{\delta}(t)h^4 + O(h^5) \doteq \mathbf{y}(t) - \mathbf{y}(t; h).$$

(c) In `MATLAB` we find

```
> [t2,y2]=rk4(f,[0 1],1,10,1);
> [t,y]=rk4(f,[0 1],1,20,1);
> exp(1)-y(21)
> ans =1.358027104103599e-07
> (y(21)-y2(11))/15
> ans =1.299014112869183e-07
```

(d) Again in `MATLAB` we find

```
> y(21)+(y(21)-y2(11))/15
> ans =2.718281822557747
> exp(1)-y(21)-(y(21)-y2(11))/15
> ans =5.901299123441573e-09
```

4. The optimal second-order Runge-Kutta scheme has the form

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\left[\frac{1}{4}\mathbf{k}_1 + \frac{3}{4}\mathbf{k}_2\right]$$

with

$$\mathbf{k}_1 = \mathbf{f}(t_n, \mathbf{y}_n), \quad \mathbf{k}_2 = \mathbf{f}\left(t_n + \frac{2}{3}h, \mathbf{y}_n + \frac{2}{3}h\mathbf{k}_1\right).$$

This method has the smallest error coefficient of all 2nd-order RK algorithms.

(a) Calculate the characteristic equation of this 1-step method and find its root(s).

(b) Is this method weakly stable? Relatively stable? Explain your answers.

(c) Find the stability threshold of this method. Is the scheme A-stable?

Solution:

(a) For the model problem with $f(t_n, y_n) = \lambda y_n$ one finds that

$$y_{n+1} = y_n + h\lambda\left[\frac{1}{4} + \frac{3}{4}\left(1 + \frac{2}{3}h\lambda\right)\right] y_n = \left[1 + h\lambda + \frac{1}{2}(h\lambda)^2\right] y_n$$

Thus, the characteristic equation is linear, or $r - r_0(z) = 0$, with the root

$$r_0(z) = 1 + z + (1/2)z^2.$$

(b) The method is weakly stable, as shown by Theorem 6.9 in Atkinson. Weak stability is the property that numerical solutions are continuous in initial data. One could also appeal to the root condition, since $r_0(0) = 1$ is simple. Finally, all 1-step methods are relatively stable, since there is only one root and thus no "parasitic solutions" due to secondary roots.

(c) The parabola $p(x) = 1 + x + (1/2)x^2 = (1/2) + (1/2)(x+1)^2$ is equal to 1 at $x = -2$ or 0 and achieves its minimum of 1/2 at $x = -1$. Thus, $|p(x)|$ first exceeds 1 at $x = -2$ along the negative x-axis and -2 is the stability threshold. The method is *not* A-stable, since not even the entire negative real axis is within its stability region.

5. Consider the following 2-step numerical scheme:

$$\mathbf{y}_{n+1} = \frac{3}{2}\mathbf{y}_{n-1} - \frac{1}{2}\mathbf{y}_n + \frac{5}{2}h\mathbf{f}(t_n, \mathbf{y}_n).$$

(a) Is this method consistent?

(b) Find the characteristic equation of the method and its two roots $r_0(z)$, $r_1(z)$.

(c) Is the method weakly stable? Convergent? When can this method be used?

Solution:

(a) This method has $a_0 = -1/2$, $a_1 = 3/2$ and $b_1 = 5/2$. The method is consistent since the two conditions

$$\sum_{j=0}^{p} a_j = (-1/2) + (3/2) = 1$$

and

$$\sum_{j=-1}^{p} b_j = 5/2 = 1 + 0(-1/2) + 1(3/2) = 1 + \sum_{j=0}^{p} j a_j$$

are both satisfied.

(b) The characteristic equation is obtained from

$$r^{n+1} = \frac{3}{2}r^{n-1} - \frac{1}{2}r^n + \frac{5}{2}(h\lambda)r^n,$$

or, with $z = h\lambda$,

$$r^2 + \frac{1}{2}(1 - 5z)r - \frac{3}{2} = 0.$$

Using the quadratic formula, the roots are

$$r_0(z) = \frac{-(1-5z) + 5\sqrt{1 - (2/5)z + z^2}}{4}, \quad r_1(z) = \frac{-(1-5z) - 5\sqrt{1 - (2/5)z + z^2}}{4}.$$

(c) For $z = 0$, the roots are $r_0(0) = 1$ and $r_1(0) = -3/2$. Since the second root lies outside the unit circle, the root conditon is violated. Thus, the method is *not* weakly stable. Since convergence is equivalent to weak stability for consistent schemes, this method is also *not* convergent. This method can *never* be used!

6. Consider the following initial-value problem

$$\dot{y} = -e^t y + 1 - e^{-t}, \quad y(0) = 1$$

for the time interval $0 < t < 8$. The exact solution is $y(t) = e^{-t}$.

(a) Consider solving this problem with Heun's method for a fixed time-step h. Using the stability threshold of this method, determine the largest step h and, thus, the smallest number of steps N so that the Heun integration remains absolutely stable for $0 < t < 8$.

(b) Apply the Heun method with $N = 10^4$ and report the results. Explain your observations using part (a).

(c) Repeat (a) for the trapezoidal method and (b) for the trapezoidal method with $N = 100$.

(d) Which of the following `MATLAB` integrators, `ode45`, `ode113` and `ode15s`, give convergent results for the above initial-value problem in the limit $TOL \to 0$? State which of these methods you would use for this problem and explain why.

Solution:

(a) Since the Jacobian of the equation is

$$J(t, y) = -e^t$$

and since the stability threshold of the Heun method is -2,

$$he^t < 2$$

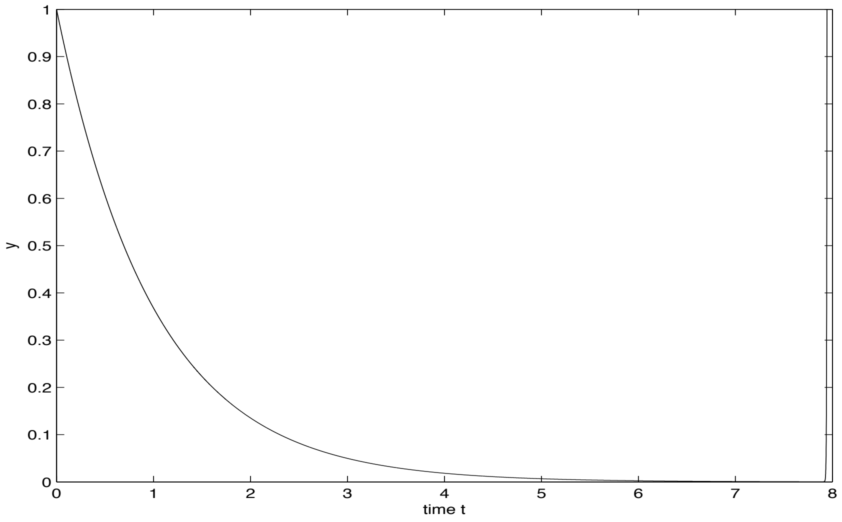
is required for absolute stability up to time t. For stability up to $t = 8$, Heun method must thus have $h < 2e^{-8} = 6.7093 \times 10^{-4}$ and $N = 8/h > 1.1924 \times 10^4$.

(b) In the plot below we see that the Heun method with $N = 10^4$ yields a good result until shortly before $t = 8$ but then blows up. The time at which absolute stability is lost for $N = 10^4$ is

$$t_* = \ln(10^4/4) \doteq 7.824.$$

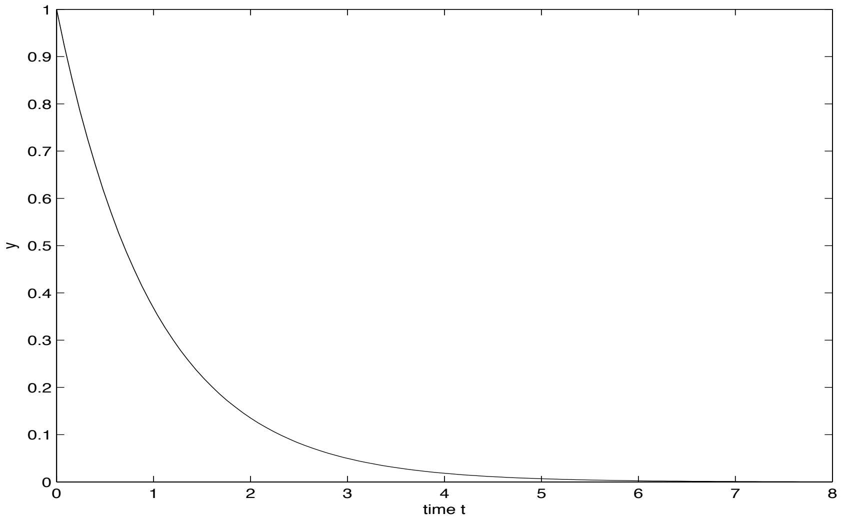
This time agrees well with the observed time of blow up.

Heun method with N=10,000



(c) The trapezoidal method is A-stable and thus its stability threshold is $-\infty$! There is no restriction placed on the size of h by stability considerations. In the plot below we see that the trapezoidal method yields an accurate solution over the whole time range with $N = 100$.

Trapezoidal method with N=100



(d) All of the integrators `ode45`, `ode113` and `ode15s` give convergent results for the above initial-value problem in the limit $h \rightarrow 0$. However, only `ode15s` is based upon implicit methods (BDF's or NDF's) that have the entire negative real axis in their domain of stability. Thus, `ode15s` is much more efficient for a **stiff** problem like the present one and it should be employed here.