HGCMCE—Exam 2007—Solutions and Feedback

1 (*a*) Truncation errors occur when a limiting process is stopped after a finite number of steps; for example, the approximation

$$\sin x \approx x - \tfrac{1}{6}x^3,$$

obtained by truncating the Taylor series for $\sin x$, has truncation error of $-\frac{1}{120}x^5 + \ldots$.

Rounding errors occur when arithmetic values are not represented exactly in a computer or calculator; for example, when $\frac{1}{3}$ or $\sqrt{2}$ are rounded to 12dp. [A form of truncation error, but less subject to analysis!]

Cancellation errors occur when the result of an addition or subtraction is much smaller than the addends/subtrahends, so that the result has fewer significant figures. For example, when working out $(f(x+h) - f(x))/h$ in an attempt at numerical differentiation, if h is taken small, then there is severe cancellation error, whereas if h is large there is severe truncation error; either way, the results are not good.

Bookwork; 2 marks each part [6]

(*b*) *(i)* A graph suggests three real roots, around 2, $-\frac{1}{3}$ and -2 [or $-\frac{3}{2}$]. Newton–Raphson gives the iteration

$$x \to x - \frac{x^3 - 3x - 1}{3x^2 - 3}.$$

A starting value of $2 \to \frac{17}{9} \to 1.8794\ldots \to 1.87938524 \to \ldots$ [same to 8dp]; similarly, $-\frac{1}{3} \to -0.34722\ldots \to -0.347296353\ldots \to \ldots$ [same to 7dp]; and $-2 \to -\frac{5}{3} \to -1.5486111\ldots \to -1.5323901\ldots \to -1.5320889\ldots \to \ldots$ [same to 6dp]. So, to 5dp, the roots are **1.87939**, **−0.34730** and **−1.53209**. [Check: sum of roots is zero.]

Similar to coursework and examples. 5 marks for appropriate method and application, 2 for each numerical root. Other methods than NR acceptable, of course. [11]

(ii) Graphs of $y = x^2$ and $y = \sin x$ show the root $x = 0$ and a second root near $x = 1$ The iteration $x \to \sqrt{\sin x}$ works well: $1 \to 0.9173\ldots \to 0.89105\ldots \to 0.88189\ldots \to \ldots \to 0.8767262\ldots$. The roots are **0** and **0.87673** [to 5dp].

Similar to coursework and examples. NR or other methods OK, but more work. 2 marks for each root, 4 for a suitable method and application. [8]

Comments: Mostly rather well done. A significant minority of students, despite drawing correct graphs, failed to find all three roots in *(i)* and/or the zero root in *(ii)*. If your graph shows three roots, perhaps there are three roots! A few students also had great difficulty describing the different sorts of error.

2 (*a*) The error term in the composite form of Simpson's Rule is proportional to [a power of the strip width and] the fourth derivative of the integrand at some point of the interval. So the composite form works well in cases where that derivative is 'well-behaved' [eg continuous], and especially if its average value is small; and badly if it is 'ill-behaved' [not defined, discontinuous, tending to infinity, and so on] at any point of the interval, including the end-points.

For example, consider $I = \int_0^1 \sqrt{\sin x}\, dx$. This is ill-behaved near $x = 0$, where the integrand is approximately $\sqrt{x}$, which is not itself too bad but has a derivative that tends to infinity. Two approaches that may be used are:
(i) A substitution, such as [in this case] $u^2 = x$, so that $2u\, du = dx$ and $I = \int_0^1 \sqrt{\sin u^2}\, 2u\, du$, where the integrand is now approximately $2u^2$ for small u, and is well behaved for $0 \leqslant u \leqslant 1$.
(ii) Subtracting off the bad behaviour. In the present case, for small x, $\sqrt{\sin x} \approx \sqrt{x}$, so $I = \int_0^1 (\sqrt{\sin x} - \sqrt{x} + \sqrt{x})\, dx$, and now $\sqrt{\sin x} - \sqrt{x}$ is [relatively] well-behaved for small x [enough for numerical integration], while $\int_0^1 \sqrt{x}\, dx = \frac{2}{3}$ can be evaluated analytically.

Bookwork: 3 marks each for the comments and for the two examples. [9]

(*b*) Following the hint,

$$\int_0^1 \log \sin x\, dx = [x \log \sin x]_0^1 - \int_0^1 x\, d/dx(\log \sin x)\, dx = \log \sin 1 - \int_0^1 x \cos x / \sin x\, dx,$$

in which the integrand $x \cot x$ is now well-behaved near $x = 0$.

Using Simpson's Rule with two strips, we have $ends = f(0) + f(1) \approx 1{+}0.642093 = 1.642092$ [working to 6dp throughout]; $evens = 0$, $odds = f(\frac{1}{2}) \approx 0.915244$, and so $I_2 = \frac{1}{2}(ends + 4odds + 2evens) / 3 \approx 0.883845$.

With four strips, $ends \approx 1.642092$, $evens = old\,evens + old\,odds \approx 0.915244$, and $odds = f(\frac{1}{4}) + f(\frac{3}{4}) \approx 0.979079{+}0.805070 = 1.784149$, and so $I_4 \approx \frac{1}{4}(1.642092 + 4{\times}1.784149 + 2{\times}0.915244)/3 \approx 0.884098$. As I_2 and I_4 differ by [only] 0.000253, our error estimate is $\varepsilon \approx -0.000253/15 \approx -0.000015$, so we can already expect I_4 to be correct to 4dp, with a 'corrected' value of $I_4 - \varepsilon \approx 0.884113$, and so

$$\int_0^1 \log \sin x\, dx = \log \sin 1 - \int_0^1 x \cot x\, dx \approx -0.172604 - 0.884113 = -1.056717,$$

or **−1.0567** to 4dp.

Familiar techniques. Maple gives −1.056720206. *A little more work, but unnecessary, to go to* I_8 *for more accuracy.* 4 marks for initial analysis, 4 for knowing and being able to apply SR, 2 each for I_2, I_4, the answer, and error control. [16]

Comments: Also mostly well done. A few students assumed that $f(0) = 0$, despite the evidence when they looked at other values of f, which rather spoiled convergence. In an exam, you might be expected to go to I_8, but I_{16}, let alone I_{32}, would be very unfair in an exam, so those who did just that really should have stopped and thought instead.

3 (*a*) The corresponding tableau is

	4	0	1	1	2	… (1)
	3	1	3	1	1	… (2)
	0	1	2	0	4	… (3)
	3	2	4	1	3	… (4)
$(2)-\frac{3}{4}\times(1)$	0	1	9/4	1/4	–1/2	… (5)
$(4)-\frac{3}{4}\times(1)$	0	2	13/4	1/4	3/2	… (6)
$(3)-(5)$	0	0	–1/4	–1/4	9/2	… (7)
$(6)-2\times(5)$	0	0	–5/4	–1/4	5/2	… (8)
$(8)-5\times(7)$	0	0	0	1	–20	… (9)
$-4\times(7)-(9)$	0	0	1	0	2	… (10)
$(5)-\frac{9}{4}(10)-\frac{1}{4}(9)$	0	1	0	0	0	… (11)
$\frac{1}{4}(1)-\frac{1}{4}(10)-\frac{1}{4}(9)$	1	0	0	0	5	… (12)

From equations (12), (11), (10) and (9), in that order, we read off $x = 5$, $y = 0$, $z = 2$ and $w = -20$, or equivalently $\boldsymbol{x} = \mathbf{(5, 0, 2, -20)^T}$.

Standard method on numerical example. 4 marks for method, 4 for initial tableau, 4 for accuracy. [No excuse for wrong answers!] [12]

Equations (1), (5), (7) and (9), in that order, constitute a lower triangular matrix obtained from A only by subtracting multiples of one row from another; so its determinant is the same as $\det A$, and is the product of its diagonal elements, $4\times1\times(-\frac{1}{4})\times1 = \mathbf{-1}$. *Standard method.* [3]

To find A^{-1}, write an identity matrix I to the right of the vertical line in place of $\boldsymbol{b}$. Then reduce as above; equations (12), (11), (10) and (9), in that order, then constitute I to the left of the line, and the corresponding right-hand sides constitute the rows of A^{-1}. *Bookwork.* [3]

(*b*) Partial pivoting means using the equation with the largest coefficient [in absolute value] for the variable to be eliminated in preference to the first [remaining] equation. This means that in eliminations of the form $equation(j) = equation(j)-\lambda\times equation(i)$, as the above tableau, the multiplier λ is always such that $|\lambda| \leqslant 1$, so that rounding errors in the equations do not grow unreasonably. If we allow $|\lambda| > 1$, then in systems with perhaps 100+ equations, rounding errors may grow exponentially, making the results meaningless.

The given equations are equivalent to $\varepsilon x+y = 1$, $x+y = 2$. If, following the tableau method as above, we use the first equation to eliminate x from the second, we have $y-y/\varepsilon = 2-1/\varepsilon$, so $y \approx 1$, and the value of $x = (1-y)/\varepsilon$ from the first equation is subject to severe cancellation error. If on the other hand we use the second equation to eliminate x, then we have $y-\varepsilon y = 1-2\varepsilon$, so again $y \approx 1$, but now x is found from the second equation, $x = 2-y$, and can be obtained to the same accuracy as y.

Bookwork, but numerical example only discussed in general terms. 2 marks for definition, 2 for reason and 3 for analysis of given example. [7]

Comments: Finding determinants was rather hit-or-miss. You do need to take care if you swap or scale equations. Partial pivoting defeated most of you; you were OK with describing what and why, but not with the actual example. The solution is ‘obviously’ $x \approx y \approx 1$, so the whole point was to track what happened to the errors with and without pivoting. Far too many of you just solved the equations, found the same ‘exact’ solution either way, and let that stand without comment.

4 (*a*) In order to take a 'step', the assumption is that we know y, and hence $f(x, y)$, at [for the three-point formula] three equally-spaced values of x, the current one, x_0 and two previous ones x_{-1}, x_{-2}, where in general a subscript i indicates values at $x_i = x_0 + ih$, with h the step length.

The Adams–Bashforth formula is used to extrapolate, using a quadratic approximation to f based on the last three values, to the next value of y, given that $y_1 = \int_{x_0}^{x_1} f(x, y)\,dx$. This gives a *predicted* value, y_1^*, for y_1, and hence for $f_1^* = f(x_1, y_1^*)$.

The predicted values are then typically used in the Adams–Moulton formula, which interpolates a cubic approximation to f based on the known and predicted values to estimate a *corrected* value of y_1. As interpolation is usually more accurate than extrapolation, this is usually an improvement, after which the step is complete, we can assume y known as far as x_1, and the start the whole process again to take the next step as far as is thought desirable.

If the corrected value is very different from the predicted, we could iterate the Adams–Moulton formula; but this is usually rather a sign that h is too large.

Bookwork. Rather less wordy version of the above acceptable! [6]

Pro: Fast [only two function evaluations per step, compared with Runge–Kutta four or five, even if more accurate 4- or 5-point formulas are used] and efficient [retains and uses information from previous steps]; predictor-corrector process gives a degree of error knowledge/control. *Con*: Must be 'seeded' with initial values, beyond the initial conditions, obtained by some other method [not *necessarily* hard if, eg, a Taylor series valid near the initial x is known]; relatively difficult to change step length adaptively.

Bookwork. [4]

(*b*) The results are as follows:

x	y	f	AB	y_{pred}	f_{pred}	AM	y_{corr}
–0.2	1.020201	–0.204040					
0.0	1.000000	0.000000					
0.2	1.020201	0.204040	0.306060	1.081413	0.432565	0.315242	1.083250
0.4	1.083250	0.433300	0.558438	1.194937	0.716962	0.569382	1.197126
0.6	1.197126	0.718276	0.883979	1.373921	1.099137	0.899042	1.376935
0.8	1.376935	1.101548	1.334140	1.643763	1.643763	1.356883	1.648311
1.0	1.648311						

So we estimate $\mathbf{y(1) \approx 1.6483}$. [For comparison, $\sqrt{e} \approx 1.648721$.] In the table, $f = xy$, AB is the extrapolated mean f [eg, in the line for $x = 0.4$, it is $(23\times0.433\ldots - 16\times0.204\ldots + 5\times0.000)/12 = 0.558\ldots$]; y_{pred} is the predicted next y [eg $1.083\ldots + 0.2\times0.558\ldots \approx 1.195$], followed by the corresponding predicted f; then the column AM is the interpolated mean f obtained from the same preceding values of f and the just-obtained predicted f, and y_{corr} is the corrected y, carried over to the next line of the table. [All calculations in the table shown to 6dp but carried out to full calculator accuracy.]

Standard technique on numerical problem. 5 marks for implementing the method, 2 marks per 'interesting' line for general numerical accuracy, 2 for the final result. Potential bonus for discussion of errors. [15]

Comments: The bookwork was reasonably well done. But there were only a very few good solutions to the numerical part. Even though the formulas were there in front of you, several students used the Modified Euler method instead. Others made numerical slips that really should have been picked up during the calculation—the question gave you a hint about what the 'exact' solution was, but even without that you should stop and think if a table seems to be 'blowing up'.

5 (*a*) Replacing the second derivatives in the wave equation by the given approximation, and replacing ϕ by its approximation u, we have

$$\frac{\partial^2\phi}{\partial t^2} \approx (\phi(x,t+k)-2\phi(x,t)+\phi(x,t-k))/k^2 \approx (u_{i,j+1}-2u_{i,j}+u_{i,j-1})/k^2$$
$$= c^2\frac{\partial^2\phi}{\partial x^2} \approx c^2(\phi(x+h,t)-2\phi(x,t)+\phi(x-h,t))/h^2 \approx c^2(u_{i+1,j}-2u_{i,j}+u_{i-1,j})/h^2,$$

whence

$$u_{i,j+1} \approx 2u_{i,j}-u_{i,j-1}+(kc/h)^2(u_{i-1,j}-2u_{i,j}+u_{i+1,j}),$$

Bookwork. [7]

(*b*) When $j=0$, we find $u_{i,1}$ depends on $u_{i,-1}$, which is unknown. However,

$$\phi(x,k) = \phi(x,0)+k\frac{\partial\phi}{\partial t}+\tfrac{1}{2}k^2\frac{\partial^2\phi}{\partial t^2}+\ldots = \phi(x,0)+k\frac{\partial\phi}{\partial t}+\tfrac{1}{2}k^2c^2\frac{\partial^2\phi}{\partial x^2}+\ldots,$$

from the Taylor series, and so

$$u_{i,1} \approx u_{i,0}+k\dot{u}_{i,0}+\tfrac{1}{2}k^2c^2(u_{i+1,0}-2u_{i,0}+u_{i-1,0}),$$

where everything on the RHS in known from the initial conditions.

Bookwork. 2 marks for the reason, 4 for the equation. [6]

(*c*) We have

$$u_{i,j+1} \approx 2u_{i,j}-u_{i,j-1}+(4/\pi)^2(u_{i+1,j}-2u_{i,j}+u_{i-1,j}) \quad [\text{for } j \geqslant 1,\ -1 \leqslant i \leqslant 1];$$
$$u_{i,1} \approx u_{i,0}+\tfrac{1}{2}(4/\pi)^2(u_{i+1,0}-2u_{i,0}+u_{i-1,0}) \quad [\text{as } \dot{u}_{i,0}=0];$$

with initial/boundary conditions $u_{\pm 2,j}=0$, $u_{\pm 1,j}=1/\sqrt{2}$, $u_{0,0}=1$.

So, successively, we have [all calculations to 5dp]:

$$u_{0,1} \approx 1+\tfrac{1}{2}(4/\pi)^2(2/\sqrt{2}-2) \approx 0.52518,$$
$$u_{\pm 1,1} \approx 1/\sqrt{2}+\tfrac{1}{2}(4/\pi)^2(0-2/\sqrt{2}+1) \approx 0.37136,$$
$$u_{0,2} \approx 1.05036-1+(4/\pi)^2(0.74272-1.05036) \approx -0.44837,$$
$$u_{\pm 1,2} \approx 0.74272-1/\sqrt{2}+(4/\pi)^2(0.52518-0.74272) \approx -0.31705,$$
$$u_{0,3} \approx -0.89674-0.52518+(4/\pi)^2(-0.63410+0.89674) \approx -0.99614, \text{ and}$$
$$u_{\pm 1,3} \approx -0.63410-0.37136+(4/\pi)^2(-0.44837+0.63410) \approx -0.70437.$$

So, in this approximation,

$$\boldsymbol{\phi(0,3) \approx -0.99614},\ \boldsymbol{\phi(\pm\tfrac{1}{4}\pi,3) \approx -0.70437}, \text{ and } \boldsymbol{\phi(\pm\tfrac{1}{2}\pi,3) = 0}.$$

Numerical example. 3 marks for equations, 3 for method, 4 for numerical accuracy. [10]

For stability, require $kc \leqslant h$; this is breached, but only mildly, by $k=1$, $c=1$, $h=\tfrac{1}{4}\pi$.

Bookwork. [The exact solution would show $\phi(0,3) = \cos 3 \approx -0.98999$, so the approximate solution shows increasing oscillations.] [2]

Comments: Again, the bookwork was well done, but the numbers proved elusive. The discussion in lectures showed that we had to 'progress' through the solution for increasing values of t; there is very little to store, and no point at all trying to solve the equations algebraically. If you find it hard to keep track of the results, then draw a grid of suitable size, and annotate the grid points with the values of $u_{i,j}$ as you find them.