Numerical Analysis PhD Qualifying Exam
University of Vermont, Spring 2010

Instructions: *Four problems must be completed, and one problem must be attempted. At least two problems from 1-4 and at least two problems from 4-7 must be completed.* **Note** *that Question 4 can count towards either section, but not both. To have attempted a problem, you must correctly outline the main idea of the solution and begin the calculation, but need not have finished. You have three hours to complete the exam.*

---

1. This question concerns number representation and errors. Normalized floating point numbers can be represented by $\pm 1.b_1 b_2 b_3 \ldots b_N \times 2^{\pm p}$ where b_i is either 0 or 1, N is the number of bits in the mantissa, p is an M-bit binary exponent, and two additional bits are used to store the signs. Assume we are using a machine for which $N = 23$ and $M = 7$. *Note* that you may not need all of the information above to solve the problem.
(a): Let $x = 2^{16} + 2^{-8} + 2^{-9} + 2^{-10}$ and let x^* be the machine number closest to x on the machine above. What is $|x - x^*|$?

   (b): The **Theorem on Loss of Precision** states that if x and y are positive normalized floating point binary machine numbers such that $x > y$ and

$$2^{-q} \leq 1 - \frac{y}{x} \leq 2^{-p}$$

   then at most q and at least p significant binary bits are lost in the subtraction $x - y$. The theorem is useful in estimating the likelihood of catastrophic cancellation, which is common when performing modular arithmetic.
Use the theorem to show that if $x > \pi \cdot 2^{25}$ is a number represented *exactly* on the machine described above, then $z \equiv x \pmod{2\pi} = x - 2k\pi$ can be computed with *no* significant digits. The value of z is required to compute $\cos x$, for example. *Hint*: Use $y = 2k\pi$ and solve for p.

2. This question concerns root finding. To avoid computing the derivative at each step in Newton's method, it has been proposed to replace $f'(x_n)$ by $f'(x_0)$. Define the error at step n to be $e_n = x_n - r$ where the function f has a single root at the point r, i.e. $f(r) = f(x_n - e_n) = 0$. Derive the rate of convergence for this method by finding the relationship between e_{n+1} and e_n. *Hint*: You will need a Taylor remainder at one point.

3. **(a)**: Compute a singular value decomposition $A = USV^\top = \sum_{i=1}^{2} s_i \vec{u}_i \vec{v}_i^\top$ of the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 3/2 \end{bmatrix}$$

   It is advised that you keep your entries as fractions to avoid nasty numbers.

   (b): What is the best rank-1 approximation of the matrix A?

   (c): What is the condition number of the matrix A and what does this say about the number of significant digits d in the solution to $A\vec{x} = \vec{b}$? *Note* that a good explanation of your reasoning is more important than an exact answer for d.

4. **(a)**: Determine the coefficients of an implicit, one-step, ODE method of the form

$$x(t+h) = ax(t) + bx'(t) + cx'(t+h)$$

so that it is exact for polynomials of as high a degree as possible. Begin by letting LHS $= x(h)$ and RHS $= ax(0) + bx'(0) + cx'(h)$ and fill in the missing entries in the table below. The first row and column have been filled in for you.

$x(t)$	$x'(t)$	LHS	RHS
1	0	1	a
t			
t^2			
...	...	...	...

(b): Once you have obtained the coefficients a, b, c in part **(a)**, use Taylor Series to find the order of the local truncation error term.

5. The Dahlquist method

$$Y_{n+1} - 2Y_n + Y_{n-1} = \frac{h^2}{4}(f_{n+1} + 2f_n + f_{n-1}), \qquad \text{where} \quad f_n \equiv f(x_n, Y_n), \text{ etc.} \tag{1}$$

can be used to solve the initial-value problem

$$y'' = f(x, y), \qquad y(x_0) = y_0, \quad y'(x_0) = y_0'. \tag{2}$$

(a) Show that method (1) has the global error of order 2 when applied to (2).

(b) Show that this method is stable for any value of $h\omega$ when applied to the oscillator equation

$$y'' = -\omega^2 y, \qquad \omega > 0. \tag{3}$$

6. **(a)** Propose a 2nd-order accurate discretization of the equation

$$\big(p(x)\, u_x\big)_x = q(x)u + r(x). \tag{1}$$

(b) Use this discretization to set up a linear system for the boundary-value problem given by Eq. (1) and by the boundary conditions

$$u_x(0) = \alpha, \qquad u(1) = \beta, \tag{2}$$

where α, β are some given constants. Use $h = 1/3$ and write out each equation in the linear system in question. Make sure to use the 2nd-order accurate approximation for the Neumann boundary conditions.

7. Consider a unidirectional wave equation

$$u_t = c\, u_x, \qquad -\infty < x < \infty \tag{1}$$

where c =const.

(a) Use the von Neumann analysis to determine under what condition on the ratio

$$\mu = \frac{c\kappa}{h}$$

the scheme

$$\frac{U_m^{n+1} - U_m^n}{\kappa} = c\, \frac{U_{m+1}^n - U_m^n}{h}\,, \tag{2}$$

approximating (1), is stable.

(b) Similarly, show that the scheme

$$\frac{U_m^{n+1} - U_m^n}{\kappa} = c\, \frac{U_{m+1}^n - U_{m-1}^n}{2h} \tag{3}$$

is unstable for any μ.

(c) Note that for a Fourier harmonic $u = \exp[i\beta x]$, the right-hand sides of (2) and (3) equal λu for some λ. (Of course, this λ is different for (2) and (3).) Use this fact to interpret your results in parts (a) and (b) in light of the stability of a certain numerical method for ODEs.