MASTER'S COMPREHENSIVE EXAM *NUMERICAL ANALYSIS*

AUGUST 1995

INSTRUCTIONS: Calculator required for some problems. Show at least eight (8) figures where numerical results occur.

1. Given t–bit binary floating point numbers a, b, c define $x = a * b * c$, $y = (a \tilde{*} b) \tilde{*} c$, $z = a \tilde{*} (b \tilde{*} c)$ barring over- or under-flow. Here, the usual machine floating point operations $x \tilde{\omega} y$ replace exact arithmetic operations $x \omega y$, with $\omega \in \{+, \ -, \ *, \ /\}$. Estimate the distance $\big| |x-y| - |y \tilde{-} z| \big|$ of the true rounding error $|x-y|$ from the machine–computable $|y \tilde{-} z|$, using only general properties of t–bit binary floating point arithmetic.

2.a) For a positive integer n and real numbers $x_i, \ y_i, \ i = 0, 1, \ldots, n, \ \ x_0 < x_1 < \ldots < x_n$, state Newton's formula for the interpolation polynomial $p_n(x) = \sum_{k=1}^{n} a_k x^k$ of degree at most n such that $p_n(x_i) = y_i, \ i = 0, 1, \ldots, n$. Recall that Newton's interpolation formula is based on divided differences.

b) State formulas for the error $f(x) - p_n(x)$ given data $y_i = f(x_i)$, $i = 0, 1, \ldots, n$.

c) For the special data $y_i = x_i^{n+1}, \ i = 0, 1, \ldots, n$, prove $p_n(0) = (-1)^n x_0 x_1 \cdots x_n$.

(CONTINUED ON NEXT PAGE • • •)

3. Given is the integration formula, valid for arbitrary functions $f \in \mathbf{C}^k[-1,1]$,

$$\int_{-1}^{1} f(x)dx = \frac{7}{15}f(-1) + \frac{16}{15}f(0) + \frac{7}{15}f(1) + \frac{1}{15}f'(-1) - \frac{1}{15}f'(1) + c_k f^{(k)}(\tilde{x}),$$

$-1 < \tilde{x} < 1$. Determine the numerical values of the order $k \in \mathbf{N}$ in the error term as well as the error constant c_k.

HINT: Test the formula for monomials $f(x) = x^m,\ m = 0, 1, \ldots$.

4. The matrix equation $A\mathbf{x} = \mathbf{b}$, with $A = \begin{bmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 0.217 \\ 0.254 \end{bmatrix}$, has the column vector $\mathbf{x} = [1, -1]^T$ as the exact solution. Thus, of the "approximate solutions" $\mathbf{y} = [0.999, -1.001]^T$ and $\mathbf{z} = [0.341, -0.087]^T$ the former would be considered to be more accurate (by far) as $\|\mathbf{x} - \mathbf{y}\|_\infty < \|\mathbf{x} - \mathbf{z}\|_\infty$.

a) For the residuals, however, $\|A\mathbf{y} - \mathbf{b}\|_\infty > \|A\mathbf{z} - \mathbf{b}\|_\infty$. Show this is so.

b) Give a quantitative explanation for these seemingly contradictory facts.

5.a) State a definition of convergence of order q ($q = 1$: linear, $q = 2$: quadratic, $q = 3$: cubic,... convergence) of an iteration of the form $x_n = G(x_n),\ n = 0, 1, \ldots$. Here, $G(x)$ is a suitable real function differentiable on a real interval.

b) Given an equation $f(x) = 0$, where again $f(x)$ is a suitable differentiable function with $f'(x) > 0$, define $G(x) = x - f'(x)^{-1}f(x)$. Under what conditions is the iteration in part **a)** at least quadratically ($q \geq 2$) convergent? Prove your answer is correct.