Math 405/607E 2015: Numerical Solution of Differential Equations
Final Exam, December 2015

Family Name: ______________________ Given Name: ______________________

Signature: ______________________ Student Number: ______________________

Course: 405 or 607E: ______________________

Question:	1	2	3	4	5	6	7	8	9	Total
Points:	15	10	15	5	7	0	8	6	20	86
Score:										

Instructions

- You have 150 minutes to write this exam.
- This exam contains 6 pages including this cover page. Ensure you have all the pages.
- Some parts of this exam are multiple choice: for these you do not need to show any calculations unless specifically asked.
- For other problems, please write your answers on an exam booklet.
- For short answer questions, write solutions to the questions in the space provided below each question. If you require more space, label clearly where you have written your solution.
- No aids are allowed. No notes, books, calculators etc.

15 marks 1. (a) True False The Trapezoidal rule and Simpson's rule are examples of numerical differentiation methods.

(b) Which of the following is often referred to as the "fundamental theorem of finite differences"?

A. Stability and consistency imply convergence.

B. Stability and convergence imply consistency.

C. Consistency and convergence imply stability.

(c) $u' = \lambda u$ with $u(0) = 1$. This is the:

A. LeVeque test problem

B. Dalquist test problem

C. Nyquist test problem

D. Macdonald test problem

(d) Consider the test problem $u' = \lambda u$ with $u(0) = 1$ and λ real. Which of the following is true?

A. When $\lambda < 0$ any stable numerical time-stepping method will give a solution which decays toward 0, for any choices of time-step k.

B. When $\lambda < 0$, a consistent numerical time-stepping method will give a solution which will decay toward zero.

C. When $\lambda < 0$, a stable numerical time-stepping method will give a solution which decays toward zero, for any λk contained in the absolute stability region.

(e) Consider the ODE IVP $u' = f(u)$ with $u(0) = \eta$. Assuming $v^n \approx u(t_n)$, which of the following is the backward Euler time-stepping method?

A. $v^{n+1} = v^n + kf(v^n)$ with $v^0 = \eta$

B. $v^n = v^{n+1} + kf(v^{n+1})$ with $v^0 = \eta$

C. $v^{n+1} = v^n + kf(v^{n+1})$ with $v^0 = \eta$

D. $v^{n+1} = v^n + \frac{k}{2}f(v^n) + \frac{k}{2}f(v^{n+1})$ with $v^0 = \eta$

(f) True False An advantage of the backward Euler method is increased stability due to its large absolute stability region. It can usually take larger time-steps k compared to the forward Euler method.

(g) True False A disadvantage of the backward Euler method is increased computational cost compared to the forward Euler method. For the same size of time-step k, the backward Euler method will use more computing resources (e.g., CPU time, memory).

(h) Which of the following is true?

A. Consistent numerical methods guarantee that the numerical solution is close to the exact solution, at least in the limit as $h, k \to 0$. This is because the local truncation error is small.

B. Consistent numerical methods approximate the correct equation in the limit as $h, k \to 0$ because the local truncation error limits to 0.

C. Consistent numerical methods are always stable so they approximate the exact solution, at least in the limit as $h, k \to 0$.

(i) Which of the following is true?

A. For smooth problems, both forward Euler and backward Euler have global errors of $O(k)$: they are both *first order* methods.

B. Both forward Euler and backward Euler have global errors of $O(k)$ for any ODE problem.

C. Because it is more stable, the backward Euler method is a more accurate method than the forward Euler method.

(j) True False A disadvantage of linear multistep methods is that they require multiple starting values whereas one-step methods require only the initial condition.

(k) A *backward stable algorithm* gives the/a/an ____________________ solution to

the/a/an ____________________ problem. [Fill in the blanks]

(l) What can be said about *stiffness*?

A. Boing-boing!

B. A stiff problem is one where implicit methods work better (e.g., compute the solution faster).

C. A stiff problem tends to involve both fast and slow time-scales.

D. One should refer to *stiff problems* (rather than stiff differential equation) because a problem could be stiff only for certain initial conditions or particular domains.

E. All of the above are true (especially A.)

(m) Suppose the condition number of a matrix A is $\kappa(A) = 10^7$. If we solve $Ax = b$ numerical (for some given b) roughly how accurate would you expect the answer to be? (Assume the computation is done on a machine implementing standard IEEE 754 floating point).

Answer: ______________.

(n) F.T.N.A.?

A. For The Next Adventure!

B. Fried (and) Tasty (but) Not Animals

C. Fundamental Theorem of Numerical Analysis

D. All of the above!

Solutions for the following problems should be written in exam booklets.

[10 marks] 2. (a) Derive the *real part of the absolute stability region* for the Backward Euler method. Sketch the result.

(b) Suppose $y' = 10y$ and $y(0) = 100$. If we select a time-step of $k = 1/2$, is the backward Euler method (absolutely) stable? Sketch the exact solution and numerical solution (with $k = 1/2$) on the same axes.

(c) Consider the ODE problem $u' = u^2$ with initial condition $u(0) = -6$. It can be shown (you don't have to) that this has a unique solution for $t > -\frac{1}{6}$.

i. Suppose we use stepsize $k = 1$. Compute one step using the Backward Euler method. What goes wrong?

ii. Can you think of how a practical algorithm (for general nonlinear RHS $f(u)$) might deal with this issue?

[15 marks] 3. (a) Suppose a is a positive real number. Consider the PDE problem $u_t + au_x = 0$ for $x \in [0, 1]$ and $t > 0$ with initial condition $u(0, x) = g(x)$ and boundary conditions $u(t, 0) = 1$ and $u(t, 1) = 0$. Give a first-order discretization of this problem using backward differences in space and Forward Euler in time.

(b) Starting with an ansatz of $u = \exp(i\xi jh)$, start a *von Neumann stability analysis* of the scheme above to find the growth factor $G(\xi)$. You may find it helpful to use the *Courant number* $\nu = \frac{ka}{h}$ where k is the time-step and h is the spatial step.

(c) If $a(x)$ is a specified function with $a(x) > 0$, modify your discretization (it should still be 1st-order accurate but no justification is required).

(d) Now if $-5 \le a(x) \le 5$, use *upwinding* to modify your discretization to construct a stable scheme (again, you do not need to justify this result). What time-step restriction do you expect?

[5 marks] 4. State the cardinal polynomial of degree n.

Suppose we wish to find a polynomial p_n of degree at most n such that $p_n(x_i) = f_i$ for data f_i at distinct $x_i, i = 0, 1, \ldots, n$. We proved in lectures and on the midterm that a solution exists. Prove that it is unique.

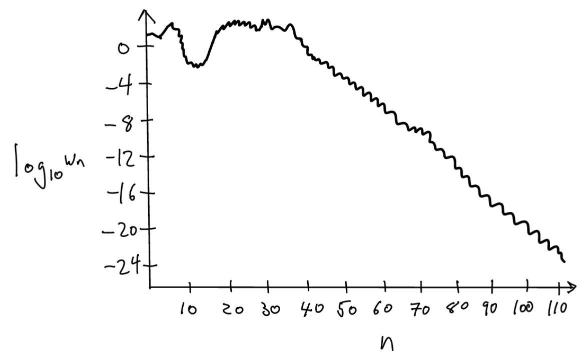
[7 marks] 5. (a) Consider $u_t = u_{xx} + u_{yy} + f(x, y, u)$ on the domain $(x, y) \in [0, 1] \times [0, 2]$ with initial condition $u(0, x, y) = g(x, y)$ and zero Dirichlet boundary conditions. Here $f(x, y, u)$ is a given nonlinear function.

For the interior nodes (away from the boundary), give a finite difference discretization of this problem in space (x, y) (a *semidiscretization*, continuous in time, discrete in space, also known as the *method of lines*).

(b) For the previous semidiscretization, suggest an appropriate time discretization which avoids the restriction of $k = O(h^2)$ and also avoids using Newton's method (or other iterative solvers.)

4 (bonus) 6. The *Chebfun* software project attempts to represent *functions* to machine precision accuracy (that is, with error approximately 2.2×10^{-16}) using a truncated expansion in Chebyshev polynomials: $u(x) \approx \sum_{n=0}^{M} w_n T_n(x)$ where w_n are the coefficients of the expansion. On the domain of interest, the Chebyshev polynomials have maximum value $|T_n(x)|_\infty = 1$.

Consider a function with expansion coefficients w_n as shown:



(a) How many terms M should Chebfun keep in its expansion?

(b) Do you think it is possible to design an algorithm to determine M? If so, describe briefly (one sentence). If not, describe how one might trick Chebfun (suggestion: draw a carefully labelled diagram).

8 marks 7. (a) What sort of matrix problem is solved by the *QR algorithm*? What is the output of the algorithm? State the most basic version of the algorithm (e.g., without shifts or preprocessing of the matrix).

(b) Suppose we have the SVD of a nonsingular square $n \times n$ matrix A. Describe the steps needed to efficiently solve $Ax = b$ using that result. What is operation count (in big Oh notation) for each step?

6 marks 8. (a) What is computed by the following pseudocode? (i.e., what is the meaning of "v"?)

```
i = sqrt(-1)
u = <vector of length n, consisting of samples of a function>
k = <vector of wave numbers>
uhat = fft(u)
vhat = -i * k.^3 * uhat
v = ifft(vhat)
```

(b) For each of the last three lines, what is the cost (in Big Oh notation or otherwise)? If helpful, you may assume n is a highly composite number (for example, that it factors into powers of 2 and 3).

(c) Suppose the "u" consists of n equispaced samples of a smooth periodic function. How might you expect the accuracy of "v" to scale with n?

20 marks 9. (a) What is an *orthogonal matrix*, Q? If the Euclidean length of a column vector $x \in \mathbb{R}^n$ is $\|x\| = \sqrt{x^T x}$, show that $\|Qx\| = \|x\|$.

(b) Suppose A is an invertible $n \times n$ matrix and $b \in \mathbb{R}^n$ is a vector. Given a factorization $A = QR$ where Q is orthogonal and R is upper triangular, explain how to solve $Ax = b$ for $x \in \mathbb{R}^n$ using $O(n^2)$ floating point operations.

(c) i. Define a *Givens rotation matrix* $J(i, j, \theta)$. Show explicitly that any such matrix is orthogonal.

ii. Given any vector $x \in \mathbb{R}^n$ with $n \geq 2$ and distinct integers i, j, with $1 \leq i, j \leq n$, show that θ can be so chosen that $y = J(i, j, \theta)x$ has $y_j = 0$ and $y_k = x_k$ for $k \neq i, j$. What is y_i, explicitly in terms of the entries of x?

(d) We adopt the common convention that, for an $n \times n$ matrix B, $J(i, j)B$ should be interpreted as $J(i, j, \theta)B$ where θ is obtained by applying the procedure described above with x equal to the i^{th} column of B.

i. Consider a matrix Q_1 defined by a product of Givens rotation matrices:

$$\underbrace{J(1, n)J(1, n-1) \ldots J(1, 3)J(1, 2)}_{Q_1} A,$$

where $A = [a \,|\, b]$ is an $n \times 2$ matrix with columns $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^n$. What is $Q_1 a$? (this notation means matrix Q_1 times vector a)

ii. Define a matrix Q_2, in terms of one or more Givens rotation matrices, such that $B = Q_2 Q_1 A$ has the following properties:

- $B(:, 1)$, the first column of B is unchanged from that of $Q_1 A$ and;
- $B(:, 2)$, the second column of B has $B(i, 2) = 0$ for $i \geq 3$.

iii. Now suppose that A is a $n \times 2$ matrix with orthonormal columns. What can you say about the first *row* of $Q_1 A$? Prove it.