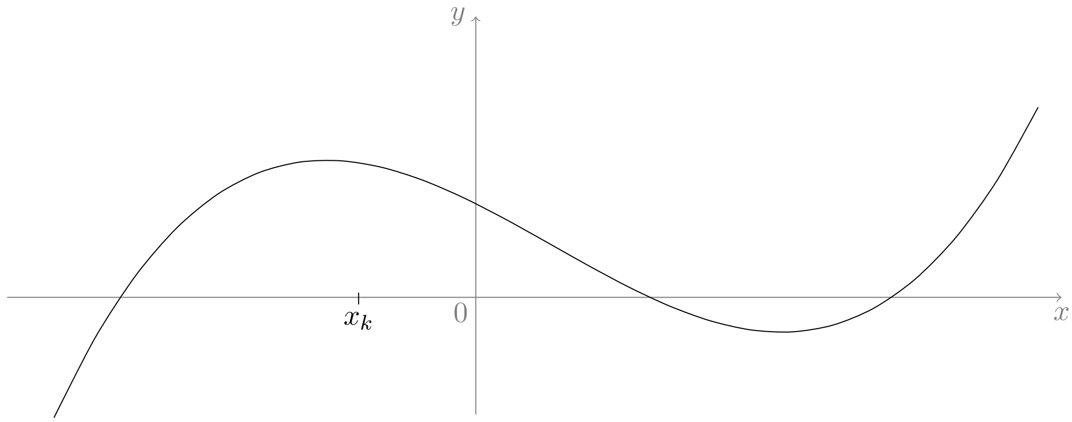
Name: ______________________________

Math 310 Numerical Analysis (Bueler) 10 December 2019

Final Exam

In class. No book or electronics. 1/2 sheet of notes allowed.
120 minutes maximum. 165 points total.

1. (a) [5 points] A differentiable function $f(x)$ and an iterate x_k are shown on the axes below. Sketch, with appropriate labeling, how **Newton's method** determines the next iterate x_{k+1}.



(b) [5 points] Write **Newton's method** as a formula for determining the next iterate x_{k+1} from the previous iterate x_k:

$x_{k+1} =$

(c) [5 points] Write the **secant method** as a formula for determining the next iterate x_{k+1} from the previous two iterates x_k and x_{k-1}:

$x_{k+1} =$

2. (a) [10 points] State **Taylor's theorem with remainder**. *Carefully state the hypotheses* and *the conclusion of the theorem.*

(b) [10 points] Assume that the basepoint is $a = 1$ and that the function is $f(x) = \cos(x/3)$. How close is the degree 4 Taylor polynomial $P_4(x)$ to the function $f(x)$ on the interval $[0, 2]$? (*Note that you are not asked to compute $P_4(x)$; no points will be given for that.*) Answer by completing the following with a concrete upper bound. (*You may leave your concrete expression unsimplified.*)

$$|f(x) - P_4(x)| \leq$$

3. [15 points] Table 10.3, printed on the last page, includes the order of accuracy $O(h^2)$ for the **composite trapezoid rule**. However, we proved that if $f \in C^2[a,b]$, and if $h = (b-a)/n$ and $x_i = a + ih$ for $i = 0, 1, \dots, n$, then there is $\xi \in [a,b]$ so that

$$\int_a^b f(x)\,dx = \frac{h}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)\right] - \frac{1}{12}(b-a)h^2 f''(\xi).$$

Suppose we want to use this rule to compute $\int_0^1 e^{-x^2}\,dx$ to an accuracy of 10^{-8}. What n is needed? (*Note you are not asked to apply the rule. You may leave your concrete expression for* n *unsimplified.*)

4. [10 points] Write two to four complete sentences to explain the ideas of **Clenshaw-Curtis quadrature (integration)** on the interval $[-1,1]$. (*Hint. The ideas are not that different from those of the Newton-Cotes formulas. Restate the basic ideas and then state the new ones, including a formula.*)

5. (a) [10 points] Suppose $f(x) = e^x$. Completely set up, *but do not solve*, the **Vandermonde** linear system to find the degree 2 polynomial $p(x) = c_0 + c_1 x + c_2 x^2$ which interpolates $f(x)$ at the points $x_0 = -1, x_1 = 0, x_2 = 1$.

(b) [10 points] For the same $f(x)$ and interpolation points as in part **(a)**, write down **Lagrange's form** of the polynomial $p(x)$. (*Do not simplify.*)

6. An actual proposed 8 bit version of the IEEE 754 standard for floating point representations, called *minifloat*, has this set up:

s	e_1	e_2	e_3	e_4	b_1	b_2	b_3

representing the number

$$x = (-1)^s \, (1.b_1 b_2 b_3)_2 \; 2^{(e_1 e_2 e_3 e_4)_2 + 2_{10}}$$

This scheme is designed to have the (surprising) property that all representable numbers are integers. (*The "$+2_{10}$" in the exponent is* not *a misprint.*) Note the usual exception cases:

- exponent bits $(0000)_2$ define the number zero or subnormal numbers
- exponent bits $(1111)_2$ define the other exceptions: $\pm\infty$ and NaN (*. . . ignore the details*)

(a) [10 points] What is the **largest number** that this system can represent? (*State the number in decimal notation and show the bits. There is no need to simplify the number.*)

(b) [10 points] What is the **smallest *normal* positive number** that this system can represent? (*State the number in decimal notation and show the bits. Please simplify the number.*)

Extra Credit. [3 points] How do you represent 1 in the system? What is the value of "machine epsilon" in this system? Indeed, how do you define "machine epsilon"? Write complete sentences. (*Hint. Think subnormal.*)

7. Consider the ODE IVP

$$y' = t - 2y, \qquad y(0) = 2.$$

(a) [10 points] Do two steps of the Euler method with step size $h = \frac{1}{2}$. This computes approximations of $y(0.5)$ and $y(1)$.

(b) [10 points] The trapezoid method, an implicit rule, is

$$y_{k+1} = y_k + \frac{h}{2}\left(f(t_k, y_k) + f(t_{k+1}, y_{k+1})\right).$$

Do two steps of this method, again with step size $h = \frac{1}{2}$. This computes new approximations of $y(0.5)$ and $y(1)$. (*Hint. This requires easy algebra.*)

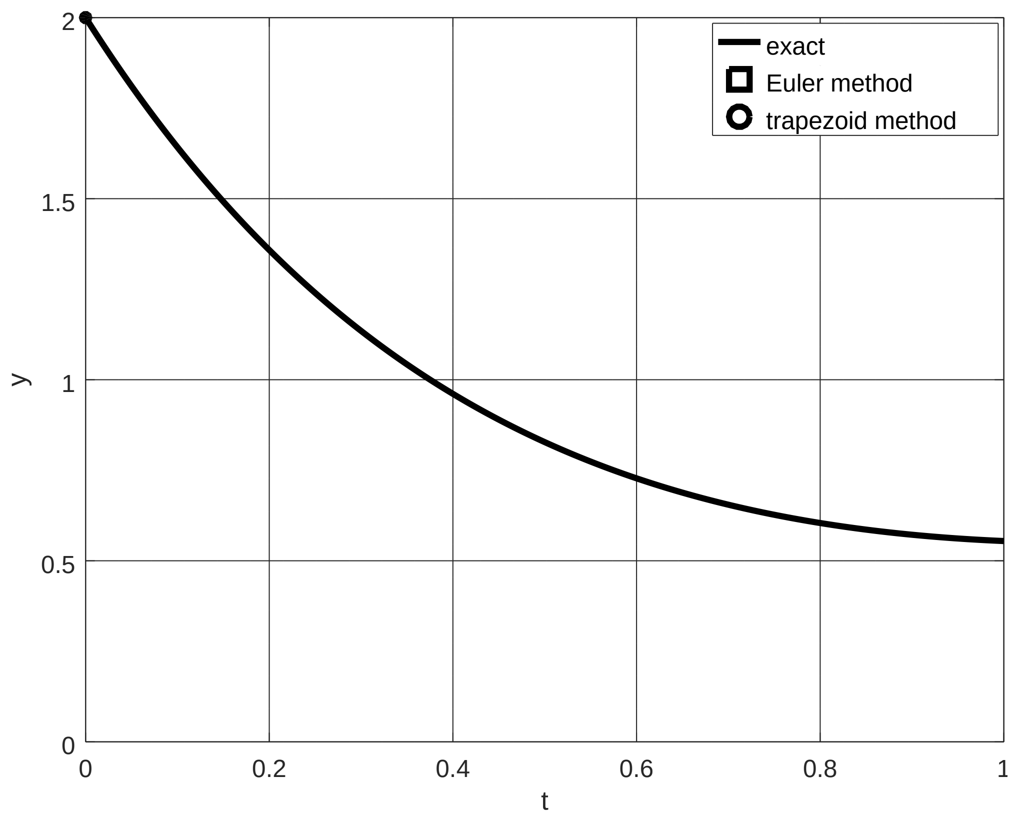
[continuation of problem 7]

(c) [5 points] The exact solution to this ODE IVP is

$$y(t) = \frac{t}{2} - \frac{1}{4} + \frac{9}{4}e^{-2t}.$$

Verify this.

(d) [5 points] The graph below already shows the exact solution. Using the symbols shown in the legend, add your results from parts **(a)** and **(b)** to this graph.



8. [15 points] The midpoint method for the ODE IVP $y' = f(t,y)$, $y(t_0) = y_0$ is the pair of formulas

$$y_{k+1/2} = y_k + \frac{h}{2} f(t_k, y_k)$$
$$y_{k+1} = y_k + h f(t_{k+1/2}, y_{k+1/2})$$

Write a MATLAB algorithm which does n steps of the midpoint method to approximately solve the ODE IVP on the interval $t_0 \le t \le t_f$.

Note that the time interval is described by a pair of numbers `tspan = [t0 tf]` *and that* $h = (t_f - t_0)/n$*. You may assume that the ODE is scalar, so* y_k *in the above formulas is a single number, not a column vector. Also, the returned variables should be suitable for plotting with* `plot(tt,yy)`*. Finally, do not worry about adding comments.*

```
function [tt,yy] = midpoint(f,tspan,y0,n)
```

10. [10 points] Solve the following system of linear equations by Gauss elimination with partial pivoting and back substitution. Show your steps, that is, indicate your row operations.

$$\begin{aligned} x_1 + x_2 &= 3 \\ 4x_1 - 3x_2 &= -2. \end{aligned}$$

11. [10 points] Suppose you have n linear equations in n unknowns,

$$A\mathbf{x} = \mathbf{b}.$$

The algorithm used to solve such linear systems has two stages:

(1) Gauss elimination with partial pivoting, and
(2) back substitution.

Approximately how many floating-point operations occur in these stages? Which stage is more costly? Answer quantitatively in a complete sentence or two. (*You do not need to count operations exactly, or prove your claims either.*)

TABLE 10.3
Quadrature formulas and their errors.

Method	Approximation to $\int_a^b f(x)\,dx$	Error
Trapezoid rule	$\frac{b-a}{2}[f(a)+f(b)]$	$-\frac{1}{12}(b-a)^3 f''(\eta),\ \eta \in [a,b]$
Simpson's rule	$\frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right]$	$\frac{1}{2880}(b-a)^5 f^{(4)}(\xi),\ \xi \in [a,b]$
Composite trapezoid rule	$\frac{h}{2}[f_0+2f_1+\ldots+2f_{n-1}+f_n]$	$O(h^2)$
Composite Simpson's rule	$\frac{h}{6}[f_0+4f_{1/2}+2f_1+\ldots+2f_{n-1}+4f_{n-1/2}+f_n]$	$O(h^4)$

[BLANK SPACE FOR SCRATCH WORK]