Assignment I

$\Big($**Taylor theorem for $f(x)$:** If the function f possesses continuous derivatives of orders $0, 1, 2, \cdots, (n+1)$ in closed interval $I = [a, b]$, then for any c and x in I,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + E_{n+1},$$

where the error term E_{n+1} can be given in the form

$$E_{n+1} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1},$$

where ξ is a point that lies between c and x and depends on both.$\Big)$

1. Determine the machine representation (write the answer in base 16) in single precision on a 32-bit word-length computer for the decimal numbers (i) 64.015625 (ii) -8×2^{-24}. Also, find the decimal numbers that have the machine representations (iii) $[3F27E520]_{16}$ (iv) $[CB187ABC]_{16}$

2. Given the exact value $a = 21.1456$ and approximate value $\hat{a} = 21.1523$. Find the absolute error, relative error and number of significant digits. Do the same calculation when $\hat{a} = 21.1462$.

3. (Loss of precision theorem) Let x and y be normalized floating-point machine numbers, where $x > y > 0$. If $2^{-p} \leq (1 - y/x) \leq 2^{-q}$ for some positive integers p and q, then show that at most p and at least q significant binary bits are lost in the subtraction $x - y$.

   Using this theorem, show that in the direct calculation of $e^x - e^{-3x}$, at most one significant bit will be lost whenever $|x| > \frac{1}{4}\ln 2$.

4. Consider the definite integral
$$I_n = \int_0^1 \frac{x^n}{5+x}dx,$$
   from which we obtain $I_n + 5I_{n-1} = 1/n$. To find I_n for $n = 0, 1, 2, \cdots, 20$, we may find I_0 first and then use the recurrence relation in forward direction. Alternatively, we may find I_{20} first and then use the recurrence relation in backward direction. Explain which strategy is better and why.

5. Suppose a calculator can handle maximum power of 10^{20}. How would you calculate $\sqrt{a^2 + b^2}$ where $a = 10^{15}$ and $b = 10^{14}$.

6. For each function below explain why a naive construction will be susceptible to significant rounding error (for x near certain value), and explain how to avoid this error.

   (a) $f(x) = \exp(x) - \exp(-x)$

   (b) $f(x) = 1 - \cos(x)$

(c) $f(x) = \sqrt{1+x^2} - \sqrt{1-x^2}$

(d) $f(x) = (\ln x - \sin \pi x)\,(1-x)^{-1}$

(e) $f(x) = (\cos(\pi + x) - \cos \pi)\, x^{-1}$

(f) $f(x) = (e^{1+x} - e^{1-x})\,(2x)^{-1}$

7. Solve the equation $x^2 - 40x + 1 = 0$ using the quadratic formula. Use five-digit decimal arithmetic to find the numerical values for the roots of the equation. Identify any loss-of -significance error that you encounter. Devise a alternative equivalent formula and find the numerical values for the roots using five-digit decimal arithmetic.

8. How to calculate $f(x) = (\cos x - e^{-x})/\sin x$ correctly near $x = 0$. Determine $f(0.005)$ correctly to ten decimal places.

9. Show that the condition numbers of a given (differentiable) function and its inverse function (assuming it exits and differentiable) are reciprocal of the other. Hence, either both are well conditioned or only one of them is well conditioned.

10. For small x, show that $(1+x)^2$ can sometimes be more accurately computed from $(x+2)x+1$. (Assume that the errors are due to arithmetic operations only)

11. Approximately how many terms are needed in the power series (about $x = 0$) to compute $\cos x$ for $|x| < 1/2$ accurate to 10 decimal places?

12. Given $\ln 10 \approx 2.3026$, how to calculate $\log_{10} 2$ without using table or calculator?

13. If $a_1 > a_2 > \cdots > a_n > 0$, in what order should $\mathrm{fl}(\sum_{i=1}^{n} a_i)$ be calculated to minimize the effect of rounding.

14. Show that arithmetic operations of addition, multiplication and division are backward stable.

15. Show that $x(y+z)$ is backward stable.

16. Show that the condition number for the calculation of a root of $x^2 - 2px + 1 = 0$ $(p \geq 1)$ becomes high for p close to 1.

17. Describe how $p(x) = 2(x+1) + 3(x+1)^5 - 6(x+1)^8 + 9(x+1)^{11}$ can be efficiently calculated.