Math 471 - Introduction to Numerical Methods - Fall 2019

Assignment # 1.
Due: Thursday, September 12, 2019.

1. **Conversion of bases:** Do the following calculations by hand.

   (a) Convert to base 10: $(1110101.101)_2$

   (b) Convert to base 2: $(2019)_{10}$

2. **Floating point numbers:** The floating point representation of a real number is $x = \pm(0.a_1 a_2 \ldots a_n)_\beta \cdot \beta^e$, where $a_1 \neq 0$, $-M \leq e \leq M$. Suppose that $\beta = 2$, $n = 5$, $M = 4$.

   (a) Find the smallest (positive) and largest floating point numbers that can be represented. Give the answers in decimal form.

   (b) Find the floating point number in this system which is closest to π.

3. **Loss of Significance due to cancellation error:** Near certain values of x, each of the following functions cannot be accurately computed as written. Identify these values of x (e.g. near $x = 0$ or large positive x) and propose a reformulation of the function (e.g., using Taylor series, rationalization, trigonometric identities, etc.) to remedy the problem.

   (a) $f(x) = 1 + \cos x$ **(b)** $f(x) = \sqrt{x^2+1} - \sqrt{x^2+4}$
   (c) $f(x) = \ln x - \ln(1/x)$ **(d)** $f(x) = x - \sin x$
   (e) $f(x) = 1 - 2\sin^2 x$ **(f)** $f(x) = \ln x - 1$

4. **Finite Differences:**

   a) Show that
$$f''(x) = \frac{f(x-h) - 2f(x) + f(x+h)}{h^2} + O(h^2)$$
   for sufficiently small h by expanding $f(x \pm h)$ out to at least fourth order using Taylor polynomials.

   b) Consider the function $f(x) = e^x$, $x_0 = 0$. Approximate the derivative $f'(x_0)$ by computing $Df = \frac{f(x_0+h) - f(x_0-h)}{2h}$ for a sequence of increasingly small h values, for example, $h = 5^{-n}$, $n = 1, \ldots, (20?30?)$. Plot the error against h. Make a table with the following information: h, Df, $E = |f'(0) - Df|$, E/h, E/h^2, E/h^3. You will need to keep enough significant digits here to be able to make meaningful observations. Describe what you observe. Can you tell from the results what is the order of accuracy of the approximation? Do you see anything unexpected?

5. The following algorithm

   step 1: $x_0 := x$; $j := 0$

   step 2: *while* $x_j \neq 0$, *do*

   $a_j :=$ remainder of integer divide $x_j/2$

   $x_{j+1} :=$ quotient of integer divide $x_j/2$

   $j := j + 1$

   end while

   can be used to convert a positive decimal integer x to its binary equivalent,

$$x = (a_n a_{n-1} \cdots a_1 a_0)_2.$$

   Implement the algorithm (write a computer program) and apply it to convert the following integers to their binary equivalents.

   (a) 471 (b) 2019

   (Two Matlab library functions *rem*, *mod* and *floor* might be helpful if you use Matlab. Try **help rem**, **help mod** and **help floor** to see how to use them.)

   (Please attach your code for completeness. No code, no credit!))