Computational Physics Problem Set 2

John Ferrante

September 17, 2024

1 Problem 1

100.98763 will be represented as a string of 32 bits with three portions representing sign, exponent, and mantissa. Using the function from lecture 1, these are found to be:

sign: 0
exponent: [1, 0, 0, 0, 0, 1, 0, 1]
mantissa: [1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1]

Using the formula given in lecture to determine the corresponding decimal yields the exact value of 100.987632751464847960 5 which differs from 100.98763 by 2.7514648479609605e-06

2 Problem 2

Smallest n for which 1 + n $\neq$ 1: 32-bit:0.0000001 64-bit:0.000000000000001
Highest number before overflow: 32-bit: 1e+38 64-bit: 1e+308
Lowest number before underflow: 32-bit: 1e-45 64-bit: 1e-323

3 Problem 2.9

(for L=100) M= -1.7418198158396654

4 Problem 3.7

I had difficulty using the sum function to iterate over the different c values.

5 Problem 4.2

5.1 A.

-9.999894245993346e-07 -999999.999999

5.2 B.

-1.000000000001e-06 -1000010.5755125057

The two different methods give different answers, one of which in each case is not correct. This can be explained by a floating point error, which occurs for one of the roots when -b $\pm\sqrt{b^2 - 4ac}$ is in the numerator and the other when it is in the denominator.