Assignment Set 5, MATH-UA0252-1

Olof Widlund

April 6, 2015

The following assignments are due on Friday April 17 at 12:00 noon. No homework will be accepted after that time. You should give me your homework in class or put it under my office door (WWH612). Do not use my mailbox in the WWH lobby.

You should turn in your MATLAB/OCTAVE programs and printouts which shows how they work in different cases. In addition, it is important that you comment on what you can learn about the performance of the methods. Always use format long e in MATLAB. Note that in MATLAB, all numbers are represented with approximately 16 decimals; we will discuss computer arithmetic during the first week of Apri.

This is all about floating point arithmetic, which will be subject of the lectures on April 7 and 9 and is the subject of the handout marked Chapter 2, The Real Numbers.

1. Problems 3.1, 3.4, 3.5, 3.8, and 3.10 of the handout.

2. Problems 4.1 and 4.2 of the handout.

3. Using MATLAB, which uses IEEE double precision, find the first power of 10, which is not represented exactly in that floating point system.

   What about powers of 10^{-1}?

   Using the table of double precision numbers, find an explanation of your experimental findings.

4. Find three floating point numbers, a, b, and c, such that $a + b + c$ is computed with a very large relative error. Do this problem using both paper and pencil and MATLAB.

5. Consider the following MATLAB program:

```
f(1)=1;
f(2)=1;
for i=2:(n-1)
```

```
f(i+1) = f(i) + f(i-1);
end
g(n)=f(n);
g(n-1)=f(n-1);
for i=(n-1):-1:2
g(i-1)=g(i+1) - g(i);
end
```

For any input n, a positive integer, the program first computes the n first Fibonacci numbers in the normal order and it then computes the same sequence of numbers in reverse order.

For small enough n there is no round-off errors at all. Explain why.

For very large values of n, the numbers will overflow. Find out how large n must be for this to happen and relate your finding to what you can learn from the tables of IEEE floating point numbers which is part of the handout.

Finally, for values of n in between, the $f(i)$ will differ very much from the $g(i)$. Get some experience with this and provide as full an explanation as you can for what is happening. In particular, explain why breakdown first happens at a particular value of n.

Does this failure indicates a serious problem with the arithmetic or can it be explained by saying that the problem is very sensitive to relatively small perturbations?