

Please write your family and given names and **underline** your family name on the front page of your paper.

1. [15 points] Assume there are two computer systems:

(i) One that uses fixed-point decimal arithmetic with 2 digits after the decimal point and 3 digits before the decimal point (i.e. ddd.dd). The digit required for the sign is separate and you don't need to worry about it.

(ii) Another that uses floating-point decimal arithmetic with 3 digits mantissa (after the decimal point) and 1 digit for the exponent (i.e.  $.ddd \times 10^d$ ). The digits required for the sign of the mantissa and the sign of the exponent are separate and you don't need to worry about them.

Compute the results of the following five arithmetic operations on each computer system. Assume that all computer operations return the correctly rounded result (i.e. the number closest to the correct answer in the representation being used). Indicate if rounding, overflow or underflow occurs.

$$(260 + 1.27) + 0.1 \quad 260 + (1.27 + 0.1) \quad 3.08 \times 0.1 \quad 3.08 \times 0.01 \quad -900 - 100$$

What is the range of representable numbers (i.e. the range of numbers for which **no** overflow happens) in each of the two systems?

What are the ranges of numbers for which underflow happens in each of the two systems? (Assume that, in the fixed-point system, the first decimal digit can be zero, while, in the floating-point system, only normalized mantissae are allowed.)

2.

(a) [10 points] Find the condition number of  $f(x) = \frac{1}{1+2x} - \frac{1-x}{1+x}$ , and study for what values of  $x$  in  $\mathbf{R}$  the function  $f(x)$  is ill-conditioned. (You may need to use de l'Hospital's rule.)

(b) [10 points] Consider the (numerical) stability of the computation of the expression  $\frac{1}{1+2x} - \frac{1-x}{1+x}$  (as it is given) for some  $x$  close to 0, either positive or negative. Explain what problems the computation of the expression  $\frac{1}{1+2x} - \frac{1-x}{1+x}$  may give rise to. Propose a mathematically equivalent expression that is likely to be more stable for  $x$  close to 0, and explain.

(c) [10 points] Write a matlab script that, for  $i = 1, 2, 3, \dots, 15$ , and  $x = -10^{-i}$ , computes  $f(x)$  as is given in (a) and as you proposed in (b). Also compute the relative error assuming the expression you proposed in (b) is the exact result. Comment on the results.

The script should look like the following:

```
for i = 1:15
    x = -10^(-i);
    f = -(1-x)/(1+x) + 1/(1+2*x);
    ff = ?; % fill-in your proposed expression
    fprintf('%10.2e %22.15e %22.15e %12.4e\n', x, f, ff, (f-ff)/ff);
end
```

Do not alter the output format.

Note:

In (b), we are interested in the stability of an expression that computes  $f(x)$  and not in the conditioning of  $f(x)$ .

Submit a hard-copy of your code and output, together with any explanations and comments. Also submit your code electronically through the cdf system.

3.

(a) [3 points] Give the Taylor's series for the function  $f(x) = \cos x$  about the point 0, writing explicitly all terms of the Taylor polynomial  $t_4(x)$  of degree 4 and the remainder  $R_6$ . Note that there are only 3 non-zero terms in  $t_4(x)$ . Indicate  $t_4(x)$  and  $R_6$ . Note that the remainder involves  $x$  and an unknown  $c$ . Indicate the (smallest) interval where  $c$  lies.

(b) [5 points] Using  $t_4(x)$  approximate  $\cos 1$ . Indicate the approximate value in 5 significant decimals. Using  $R_6$ , give an upper bound (as sharp as you can) for the (absolute value of the) error of the approximation to  $\cos 1$ . Explain how you got the bound.

(c) [5 points] Give the Taylor's series for  $f(x) = \cos x$  about the point 0, in a form so that the  $(2n)$ th term (for a general  $n$ ) of the Taylor polynomial  $t_{2n}(x)$  of degree  $2n$  and the remainder  $R_{2n+2}$  are shown explicitly. Using  $R_{2n+2}$ , give, in terms of  $n$  and  $x$ , an upper bound (as sharp as you can) for the (absolute value of the) error in the approximation to  $\cos x$

arising from  $t_{2n}(x)$ .

- (d) [12 points] Assume you have an accurate way of calculating the Taylor's series in (c), up to any  $n$ . How would you use the Taylor's series in (c) to obtain an efficient and accurate approximation to  $\cos x$  for
- (i)  $x = 2$ ?
  - (ii)  $x = 4$ ?
  - (iii)  $x = 6$ ?
  - (iv) some large  $x$  (e.g.  $x = 63$ )?
  - (v) some negative  $x$ ?

Generalize for any  $x$  and explain. This generalization can be given as pseudo-code of an if-then-else statement.

What is the maximum number of terms (or what is the maximum  $n$ ) that should be used to keep the absolute value of the remainder  $|R_{2n+2}|$  below  $10^{-16}$  for any  $x$ ? Explain.

- (e) [10 points] In the form of pseudo-code (a for-loop), give a way of calculating the Taylor's series in (c), without using powers (exponentiations) and without using factorials. (You can use additions, subtractions, multiplications and divisions.)

4. [20 points] A strictly column diagonally dominant matrix is a matrix  $A \in \mathbb{R}^{n \times n}$  for which  $|a_{jj}| > \sum_{i=1, i \neq j}^n |a_{ij}|$  for all  $j = 1, \dots, n$ .

Let  $A$  be a strictly column diagonally dominant matrix. Show that  $A$  is nonsingular, using a proof technique that employs Gauss elimination.

Hint: Starting from a generic column diagonally dominant matrix, apply one step of Gauss elimination, then continue working on appropriate submatrices. and/or appropriate submatrices.