COMPUTER SCIENCE TRIPOS Part IA – 2018 – Paper 1

6 Numerical Methods (DJG)

(a) A programmer plans to replace single-precision floating-point arithmetic in a subroutine with a fixed-point implementation that is guaranteed to have at least the same range and precision. Roughly how many bits must the fixed-point representation have? [5 marks]

(b) The designers of a new computer architecture provide an instruction that uses the fixed-point implementation of Part (a) to sum long lists of single-precision floating-point numbers. This implements the rounding and re-normalisation only once at the end of the operation.

What are the benefits of such an instruction compared with folding the standard two-argument addition operator over the list? [4 marks]

(c) A tri-diagonal square matrix has all entries zero except for the leading diagonal and the two diagonals on either side of the leading diagonal.

(i) What is the execution cost, in terms of the number of operations, of Gaussian Elimination without pivoting, both for 'school method' and for L/U decomposition? [3 marks]

(ii) Without any pivoting, the L/U decomposition of a tri-diagonal matrix results in L having ones on its leading diagonal and another partial diagonal of non-zero coefficients and U also being largely all zeros. Write out the equations or pseudocode that determine L and U. What is the cost of solving a tri-diagonal set of simultaneous equations? [5 marks]

(iii) What happens to the tri-diagonal structure when pivoting is used? [3 marks]