1995 Paper 3 Question 10

Numerical Analysis I

For Single Precision in the *IEEE* binary floating point standard (*IEEE* 754) the *precision* is defined as 24, and the *exponent* requires 8 bits of storage. With reference to *IEEE* Single Precision, explain the terms *exponent*, *significand*, *precision*, *sign bit*, *normalised number*, *denormal number*, *hidden bit*. [7 marks]

How many bits are required to store a Single Precision number? How is the exponent stored? What is the value of the hidden bit for (a) normalised, and (b) denormal numbers? [3 marks]

Define the terms *absolute error*, *relative error*, *machine epsilon*. [3 marks]

Suppose $x^* = 4.0027$, $y^* = 4.0047$ are numbers represented on a computer with *machine epsilon* $= 0.5 \times 10^{-4}$. Estimate roughly (i) the maximum *absolute error*, and (ii) maximum *relative error* in evaluating $(y - x)$.

Use your results to explain the term *loss of significance*. [7 marks]