1993 Paper 3 Question 10

Numerical Analysis I

In the IEEE binary standard (*IEEE* 754) what do the parameters p (precision), e_{min} and e_{max} specify? How is the value of an exponent e stored? [3 marks]

Explain the terms *normalised number*, *denormal number*, *hidden bit* and *NaN*. [4 marks]

In terms of the stored bit-pattern, how can each of the following be recognised: ± 0, $\pm\infty$, denormal number, NaN. [4 marks]

Suppose for some special-purpose hardware that a floating-point implementation is to be provided using only one byte for each representable number. Suppose also that, as far as possible, the principles of IEEE binary arithmetic are to be adhered to. If a sign bit of 0 represents a positive number, $p = 4$, $e_{min} = -2$ and $e_{max} = 3$ what should the following bit patterns represent?

00000000 00000001 00110000 11110000 11110001 [5 marks]

Consider the evaluation under IEEE arithmetic of the functions

$$(a) \quad \frac{x^2+2}{x^2+1} \qquad (b) \quad \ln(x^2+2)$$

where x and the function values are representable numbers, but x^2 is not. Show how you would formulate the evaluation of (a) and (b) to avoid this problem. [4 marks]