Numerical Analysis MATH50003 (2023–24) Problem Sheet 3

Problem 1 With 8-bit unsigned integers, what is the result for the following computations:

$$127 \oplus_{256} 200, \quad 2 \otimes_{256} 127, \quad 2 \otimes_{256} 128, \quad 0 \ominus_{256} 1$$

Problem 2(a) With 8-bit signed integers, what are the bits for the following: $10, 120, -10$.

Problem 2(b) With 8-bit signed integers, what is the result for the following computations:

$$127 \oplus_{256}^{\mathrm{s}} 200, \quad 2 \otimes_{256}^{\mathrm{s}} 127, \quad 2 \otimes_{256}^{\mathrm{s}} 128, \quad 0 \ominus_{256}^{\mathrm{s}} 1$$

Problem 3 What is π to 5 binary places? Hint: recall that $\pi \approx 3.14$.

Problem 4 What are the single precision $F_{32} = F_{127,8,23}$ floating point representations for the following:
$$2, \quad 31, \quad 32, \quad 23/4, \quad (23/4) \times 2^{100}$$

Problem 5 Let $m(y) = \min\{x \in F_{32} : x > y\}$ be the smallest single precision number greater than y. What is $m(2) - 2$ and $m(1024) - 1024$?