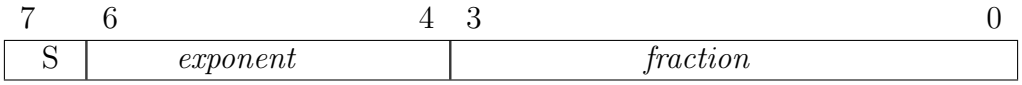
Question 1: (20 points)

NVIDIA has defined a "tiny-precision" floating pointing format for use in its Graphics Processing Units (GPUs). A floating-point number is represented in this format in 8 bits as follows: the most significant bit is the sign bit, next there are 3 bits used for the exponent, and 4 bits for the fraction. This format is illustrated below:



The exponent is expressed in excess-8 format (also known as a bias representation). Given the binary representation above, the decimal value of the number represented can be computed by the following expression:

$$N = \begin{cases} (-1)^S \times 0.0 & \text{if } exponent = 0 \text{ and } fraction = 0 \\ (-1)^S \times 0.fraction \times 2^{-2} & \text{if } exponent = 0 \text{ and } fraction \neq 0 \\ (-1)^S \times 1.fraction \times 2^{exponent-3} & \text{if } 0 < exponent < 7 \\ (-1)^S \times \infty & \text{if } exponent = 7 \text{ and } fraction = 0 \\ NaN & \text{if } exponent = 7 \text{ and } fraction \neq 0 \end{cases}$$

a. (8 points) Give the bit pattern for the representation of the number -1.25_{10} in this notation.

b. (8 points) What is the largest positive number that can be represented in this format? Give both the binary pattern in the representation and the decimal value.

c. (4 points) What is the smallest (closest to zero) positive number that can be represented in this format? Give both the bit pattern and the decimal value.