1.7 Write the number 0.06298828125 in the following forms (in part (*c*), follow the IEEE-754 standard): (*a*) Binary form. (*b*) Base 2 floating point representation. (*c*) 32-bit single-precision string.

Solution

(*a*) The largest power of 2 that can be divided into 0.06298828125 is $2^{-4} = 0.0625$. Next subtract: $0.06298828125 - 2^{-4} = 0.00048828125$. Now the largest power of 2 that can be divided into 0.000048828125 is $2^{-11} = 0.00048828125$. Thus, the number 0.06298828125 in binary form is 0.00010000001.

(*b*) Using part (*a*), the binary floating point representation of 0.06298828125 is:

$$\frac{0.06298828125}{2^{-4}} \times 2^{-4} = \frac{0.533203125}{0.5} \times 2^{-4} = 1.0078125 \times 2^{-4}$$

(*c*) According to the IEEE-754 standard, 0.06298828125 in single precision form is as follows:

- Since the number is positive, the first bit is 0
- From part (*b*), the exponent is -4. Adding a bias of 127, the value of the exponent that must be stored is -4+127 = 123. The number 123 in binary form is:

  $$\begin{aligned} 123 &= 1 \times 2^6 + 1 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 1 \times 2^1 + 1 \times 2^0 \\ &= 64 + 32 + 16 + 8 + 2 + 1 \end{aligned}$$

  Thus the number 123 in binary form is 1111011. In single precision, 8 bits can be used to store the exponent so that 123 is stored as 1111011 without the need for rounding or chopping.
- Next, the mantissa 0.0078125 is converted to binary form: 1×2^{-7} or, 0.0000001
- Since 23 bits are allocated for the mantissa, the binary number stored is 00000010000000000000000

Thus, the number 0.533203125 in single precision is stored as: |0|01111011|00000010000000000000000|.