

3.9. Expresar en formato binario de coma flotante de 32 bits según el estándar IEEE754

a) $-1023 \cdot 10^{-24}$

Según la expresión del estándar: $N = (-1)^S \cdot 2^{E-127} \cdot (1, m)$

- Paso 1: **Signo** negativo \Rightarrow el primer dígito en el formato IEEE754 será un **1**.
- Paso 2: **Mantisa**:
 1. Todo número se puede aproximar a una potencia de 2. (Se trabaja con el número sin signo a partir de aquí)

Entonces:

$$1023 \cdot 10^{-24} = 2^x$$

2. Aplicando logaritmos a ambos lados de la igualdad:

$$\log(1023 \cdot 10^{-24}) = \log(2^x)$$

3. El logaritmo del producto es la suma de los logaritmos:

$$\log(1023) + \log(10^{-24}) = \log(2^x)$$

4. El logaritmo de un número elevado a un exponente es el exponente multiplicado por el logaritmo del número.

$$\log(1023) + (-24) \cdot \log(10) = x \cdot \log(2)$$

5. $\log(10) = 1$ y despejamos la x :

$$x = \frac{\log(1023) - 24}{\log(2)} = \frac{3,0098 - 24}{0,30} = -69,72$$

6. El valor obtenido para x se aproxima al número entero inmediatamente menor. En este caso, $-69,72$ se aproxima a -70 , de donde tenemos que, para la igualdad del paso 1, se tiene que:

$$1023 \cdot 10^{-24} = 2^{-69,72}$$

Y aproximando: $1023 \cdot 10^{-24} \approx 2^{-70}$

7. Para que la aproximación vuelva a ser una igualdad, aplicamos un factor de corrección que va a ser nuestra mantisa

$$1023 \cdot 10^{-24} = M \cdot 2^{-70}$$

De donde $c = 1,20774522799391763402752$

8. Este valor es la mantisa que estamos buscando (el 1 obtenido en la parte entera ya está normalizado y no se tiene en cuenta.), pero hay que

convertirla a binario. Utilizamos el método para convertir partes fraccionarias visto en el tema 2

0,207745 X 2 = 0,41549	0, 18272 x 2 = 0,36544
0,41549 x 2 = 0,83098	0, 36544 x 2 = 0,73088
0,80398 x 2 = 1,66196	0, 73088 x 2 = 1,46176
0,66196 x 2 = 1,32392	0, 46176 x 2 = 0,92352
0,32392 x 2 = 0,64784	
0,64784 x 2 = 1,29568	
0,29568 x 2 = 0,59136	
0,59136 x 2 = 1,18272	

(Seguimos hasta donde queramos en función de la precisión que busquemos)

9. Por lo tanto, la mantisa buscada es:

$$m = 0011010$$

▪ Paso 3: **Exponente:**

1. Según el paso 7: $1023 \cdot 10^{-24} = M \cdot 2^{-70}$, con $M = (1,m)$, según acabamos de calcular.

Según la expresión del estándar: $N = (-1)^S \cdot 2^{E-127} \cdot (1,m)$.

De estas dos expresiones, se deduce que:

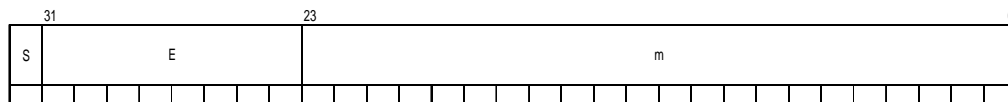
$$2^{e-127} = 2^{-70}$$

De donde $e = 127 - 70 = 57$

2. Este exponente en binario:

$$57)_{10} = 111001)_2$$

▪ Paso Final: **Escribirlo según el formato IEEE254 de 32 bits**



1 0 0 1 1 1 0 0 1 0 0 1 1 0 1 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0

En hexadecimal:

1001 1100 1001 1010 1001 0100 0000 0000
 9 C 9 A 9 6 0 0

b) $78,545 \cdot 10^{-16}$

Según la expresión del estándar: $N = (-1)^S \cdot 2^{E-127} \cdot (1,m)$

- Paso 1: **Signo** positivo \Rightarrow el primer dígito en el formato IEEE754 será un **0**.
- Paso 2: **Mantisa**:
 1. Todo número se puede aproximar a una potencia de 2. (Se trabaja con el número sin signo a partir de aquí)

Entonces:

$$78,545 \cdot 10^{-16} = 2^x$$

2. Aplicando logaritmos a ambos lados de la igualdad:

$$\log(78,545 \cdot 10^{-16}) = \log(2^x)$$

3. El logaritmo del producto es la suma de los logaritmos:

$$\log(78,545) + \log(10^{-16}) = \log(2^x)$$

4. El logaritmo de un número elevado a un exponente es el exponente multiplicado por el logaritmo del número.

$$\log(78,545) + (-16) \cdot \log(10) = x \cdot \log(2)$$

5. $\log(10) = 1$ y despejamos la x :

$$x = \frac{\log(78,545) - 16}{\log(2)} = \frac{1,895118 - 16}{0,30} = -46,855$$

6. El valor obtenido para x se aproxima al número entero inmediatamente menor. En este caso, $-46,855$ se aproxima a -47 , de donde tenemos que, para la igualdad del paso 1, se tiene que:

$$78,545 \cdot 10^{-16} = 2^{-46,855}$$

Y aproximando: $78,545 \cdot 10^{-16} \approx 2^{-47}$

7. Para que la aproximación vuelva a ser una igualdad, aplicamos un factor de corrección que va a ser nuestra mantisa

$$78,545 \cdot 10^{-16} = M \cdot 2^{-47}$$

De donde $M = 1,105422602$

8. Este valor es la mantisa que estamos buscando (el 1 obtenido en la parte entera ya está normalizado y no se tiene en cuenta.), pero hay que convertirla a binario. Utilizamos el método para convertir partes fraccionarias visto en el tema 2

