ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA COMPUTACION

ASIGNATURA: ICCD412 Métodos Numéricos **GRUPO:** GR2CC
TIPO DE INSTRUMENTO: Taller
FECHA DE ENTREGA LÍMITE: 04/05/2025
ALUMNO: Contreras Carrión Anthony Alexander

Tipo de Errores

OBJETIVOS

- Realizar la conversión del número decimal 263.310 a su representación en punto flotante de doble precisión (64 bits) según el estándar IEEE 754.
- Determinar el valor decimal aproximado obtenido a partir de la representación binaria generada, analizando posibles pérdidas de precisión.
- Calcular el error relativo utilizando tres cifras significativas y evaluar el impacto del redondeo en la exactitud del valor representado.

MARCO TEÓRICO

El estándar IEEE 754 [1] se utiliza para representar números en punto flotante dentro de los sistemas digitales. Fue creado por el Institute of Electrical and Electronics Engineers y busca estandarizar la forma en que los computadores almacenan y trabajan con números reales, asegurando precisión y compatibilidad entre distintos dispositivos y programas.

En esta actividad se trabaja con el formato de doble precisión (64 bits), que está compuesto por tres partes principales:

- 1 bit de signo (S): indica si el número es positivo (0) o negativo (1).
- 11 bits para el exponente (E): este valor se almacena con un sesgo o bias de 1023, lo que permite representar exponentes negativos y positivos.
- 52 bits para la mantisa (M): también llamada fracción. Representa la parte decimal del número ya normalizado. Aquí se asume un 1 implícito al inicio (excepto para ceros y subnormales).

La fórmula para obtener el valor real a partir de la representación binaria es:

Valor=(1) S $\times$ (1+Mantisa)$\times$ 2 (E1023)

Convertir un número decimal al formato IEEE 754 implica varios pasos: pasar el número a binario, normalizarlo, calcular el exponente con su sesgo, y construir la mantisa con los bits correspondientes. Durante este proceso es común que aparezcan errores por redondeo, ya que muchos números reales no se pueden representar de forma exacta en binario con una cantidad finita de bits. Por eso es importante calcular el error relativo, para analizar qué tan precisa fue la conversión.

DESARROLLO

Para el calcule del error relativo utilizamos la formula planteada en [2]

Representación Numérica a 64 bits

Pasar el número $263,3_{10}$ al formato IEEE 754 de 64 bits, una vez en binario pasarlo a decimal y calcule el error relativo a tres cifras de redondeo

	263	2
1	131,5	2
1	65,75	2
1	32,87	2
0	16,43	2
0	8,21	2
0	4,10	2
0	2,05	2
0	1,02	2
1	0,51	

100000111

$0,3 \times 2 = 0,6$ 0
$0,6 \times 2 = 1,2$ 1
$0,2 \times 2 = 0,4$ 0
$0,4 \times 2 = 0,8$ 0
$0,8 \times 2 = 1,6$ 1
$0,6 \times 2 = 1,2$ 1

$263,3 = 100000111,01\overline{0011}$

$1023 + 8 = 1031$

1031	2	
515,5	2	1
257,75	2	1
128,875	2	1
64,43	2	0
32,21	2	0
16,10	2	0
8,05	2	0
4,02	2	0
2,01	2	0
1,00	2	0
0,50		1

10000000111.

0	10000000111	00000111010011001100
S	E	11001100110011001100
		110011001100

M

Una vez en binario lo pasamos a decimal.

Signo = 0 → positivo

Exponente = 10000000111 = 1031 − 1023 = 8

Mantisa = $f = 2^{-6} + 2^{-7} + 2^{-8} + 2^{-10} + 2^{-13} + 2^{-14} + 2^{-17} + 2^{-18}$

$+ 2^{-21} + 2^{-22} + 2^{-25} + 2^{-26} + 2^{-29} + 2^{-30} + 2^{-33}$

$+ 2^{-34} + 2^{-37} + 2^{-38} + 2^{-41} + 2^{-42} + 2^{-45} + 2^{-46}$

$+ 2^{-47} + 2^{-48} + 2^{-50}$

$f = 0{,}0285$

$f + 1 = 1{,}0285$

$X = (-1)^0 \times 2^8 \times 1{,}0285$

$X = 257{,}0285$

$e_{relativo} = \left| \frac{p - p^*}{p} \right| = \left| \frac{263{,}3 - 257{,}0285}{263{,}3} \right| = 0{,}02381 = 2{,}381 \times 10$

CONCLUSIONES

El proceso de conversión del número 263.310 al formato IEEE 754 de doble precisión permitió comprender a fondo cómo se representa un número real dentro de un sistema binario estandarizado. Aunque el procedimiento puede parecer mecánico, es clave para entender las limitaciones que impone la computación al trabajar con números reales.

La representación obtenida mostró una ligera diferencia con respecto al valor original, lo cual es normal por la naturaleza del formato. Estos errores, aunque pequeños, pueden volverse críticos en contextos donde se requiere alta precisión. Por eso es importante siempre considerar el error relativo en este tipo de análisis.

RECOMENDACIONES

Recomiendo practicar este tipo de conversiones manualmente al menos unas cuantas veces para entender bien cada paso del proceso, desde la normalización hasta el armado del número en binario. Además, es útil comparar los resultados con herramientas automáticas, no para depender de ellas, sino para verificar que se está aplicando correctamente la lógica.

También es buena idea familiarizarse con el manejo del error relativo, ya que es una herramienta clave para analizar precisión en cálculos numéricos.

REFERENCIAS

[ieee754] IEEE Computer Society, "Ieee standard for floating-point arithmetic," https://standards.ieee.org/ieee/754/6210/, 2019, iEEE Std 754-2019 (Revision of IEEE 754-2008).

[burden2017] R. L. Burden, *Análisis Numérico*, 10th ed. Cengage Learning, 2017.