EJEMPLO C.1 Forma de punto flotante de cuatro números

Los siguientes números se expresan en la forma de punto flotante:

i) $\frac{1}{4} = 0.25$

ii) $2\,378 = 0.2378 \times 10^4$

iii) $-0.000816 = -0.816 \times 10^{-3}$

iv) $83.27 = 0.8327 \times 10^2$

Si el número de dígitos significativos fuera ilimitado, entonces no habría problema. Pero casi siempre que se introducen números en la computadora los errores comienzan a acumularse. Esto puede ocurrir en una de dos maneras:

i) **Truncado.** Todos los dígitos significativos después de k de ellos simplemente "se eliminan". Por ejemplo, si se trunca, se guarda $\frac{2}{3} = 0.666666\ldots$ (con $k = 8$) como $\frac{2}{3} = 0.66666666 \times 10^0$.

ii) **Redondeo.** Si $d_{k+1} \geq 5$, entonces se suma 1 a d_k y se trunca el número que resulta. De otra manera, el número simplemente se trunca. Por ejemplo, con redondeo (y $k = 8$), $\frac{2}{3} = 0.66666667 \times 10^0$.

EJEMPLO C.2 Ilustración de truncado y redondeo

Se puede ilustrar la forma en la que se almacenan algunos números truncados y redondeados con ocho dígitos significativos:

Número	Número truncado	Número redondeado
$\frac{8}{3}$	0.26666666×10^1	0.26666667×10^1
π	0.31415926×10^1	0.31415927×10^1
$-\frac{1}{57}$	-0.17543859×10^1	-0.17543860×10^1

Error de redondeo acumulado

Los errores individuales de truncado o de redondeo no parecen ser significativos. Sin embargo, cuando se realizan miles de pasos en la computadora, el **error de redondeo acumulado** puede ser devastador. Por consiguiente, al analizar cualquier esquema numérico es necesario saber no sólo si, en teoría, se obtendrá la respuesta correcta, sino también cuánto se van a acumular los errores de redondeo. Para tener un control de las cosas se definen dos tipos de error. Si x es el valor real de un número y x^* es el número que aparece en la computadora, entonces el **error absoluto** ε_a está definido por

Error absoluto

$$\boxed{\varepsilon_a = |x^* - x|} \tag{C.2}$$

Error relativo

En la mayor parte de las situaciones es más interesante el **error relativo** ε_r, definido por

$$\boxed{\varepsilon_r = \left|\frac{x^* - x}{x}\right|} \tag{C.3}$$