

Első gyakorlat

2019.02.06.

Numerikus számítások fizikusoknak

1. Mit kapunk a szokásos dupla pontosságú lebegőpontos számrendszerben a

$$(2^{53} + 1) - 2^{53}$$

műveletet elvégezve (kettes számrendszer, 52 bites mantissza, 11 bites karakterisztika, 1 előjelbit)?

2. Mi a legkisebb ábrázolható pozitív szám a dupla pontosságú lebegőpontos számrendszerben? Mi az legkisebb, 1-nél nagyobb ábrázolható szám ugyanezen számrendszerben? (a kettő közötti távolságot nevezzük gépi pontosságnak, jele: ε_1)
3. Lássuk be, hogy a $x = 0.1$ tízes számrendszerbeli szám kettes számrendszerbeli alakja a $0.000\dot{1}10\dot{0}$ szakaszos tizedes tört, ahol az utolsó négy számjegy ismétlődik.

Ezután számítsuk ki az $(x - fl(x))/x$ relatív hiba értékét ha $fl(x)$ az x szám szimpla pontosságú lebegőpontos képe (32 biten tároljuk a számokat: 1 előjelbit, 8 bit a karakterisztika és 23 bit a mantissza)!

4. **Pluszpontért beküldhető házi:** Konvergencia a $\sum_{i=1}^{\infty} \frac{1}{i}$ sor? Mi történik ha lebegőpontos számrendszerben kiszámítjuk a részletösszegeket?