Ex 19.1

Ex $19_{10} \to (\ldots)_2$

$19 = 2 \cdot 9 + 1$
$9 = 2 \cdot 4 + 1$
$4 = 2 \cdot 2 + 0$
$2 = 2 \cdot 1 + 0$
$1 = 2 \cdot 0 + 1$

$10011_2 = 19_{10}$

Check:
$10011_2 = 1 \cdot 2^4 + 1 \cdot 2^1 + 1 \cdot 2^0 = 19_{10}$

19.1_{10}

$0.1101_2 = 1 \cdot 2^{-1} + 1 \cdot 2^{-2} + 0 \cdot 2^{-3} + 1 \cdot 2^{-4}$

0.1

$0.1 = 0 \cdot 2^{-1} + 0 \cdot 2^{-2} + 0 \cdot 2^{-3} + 1 \cdot 2^{-4}$ ⊕

0.5 $\quad \frac{1}{4} = 0.25$ $\quad \frac{1}{8} = 0.125$ $\quad \frac{1}{16} = 0.0625$

$$\begin{array}{r} 0.1 \\ -\ 0.0625 \\ \hline 0.0375 \end{array}$$

⊕ $1 \cdot 2^{-5} +$

$\frac{1}{32} = 0.03125$

$19.1_{10} \approx 10011.00011_2$

$000111 \to 00100$

$000110 \to 00011$

if rounding is considered

$10^{30}+1 \to 10^{30}$ in a computer

Effect of finite length mantissa

Ex Suppose we were in decimal system with 4 digit mantissa

$$X = \pm 0.a_{-1}a_{-2}a_{-3}a_{-4} \cdot 10^{e}, \quad -M \le e \le M$$

$X_1 = 0.5006_{10} \cdot 10^0$

$X_2 = 0.1000_{10} \cdot 10^1$

$$\begin{array}{r} X_1 \\ +\ X_2 \\ \hline \end{array} \qquad \begin{array}{r} 0.5006 \\ +\ 1.0000 \\ \hline 1.5006 \end{array}$$

1.5006 rounded to $1.501_{10} = 0.1501_{10} \cdot 10^1$

Here we have a choice:

1.5006 can be <u>truncated</u> (chopped) to 1.500

1.5006 can be <u>rounded</u> to 1.501

Note: Rounding produces a smaller error.

Ex Consider the function

$$f(x) = \left(\sqrt{x+1} - \sqrt{x}\right)x$$

Compute $f(0.5)$ and $f(5000)$ using 4 digit mantissa.

$$\begin{array}{r} 0.5000 \\ +1.0000 \\ \hline 1.5000 \end{array}$$

1) $f(0,5000) = \left(\sqrt{0.5000 + 1.000} - \sqrt{0.5000}\right) \cdot (0,5000)$

$= \left(\underbrace{\sqrt{1.500}}_{1.2247448} - \underbrace{\sqrt{0.5000}}_{0.70710678 1}\right) \cdot 0.5000 \simeq$

$= (1.225 - 0.7071)(0.5000) \approx$

$= (0.5179)(0.5000) \approx \overset{0.25895}{\underline{\underline{0.2590}}}$

True answer: 0.258819 ....

Absolute error: $|0.258819\ldots - 0.2590| =$
$= 0.00018\ldots$

Relative error $= \dfrac{\text{absolute error}}{\text{exact value}} = \dfrac{0.00018\ldots}{0.258819\ldots} =$

$= 6.99 \times 10^{-4}$

2) $f(5000) = \left(\sqrt{5000 + 1.000} - \sqrt{5000}\right) \cdot 5000 =$

$= \left(\sqrt{5001} - \sqrt{5000}\right) 5000 \approx \left(70.72 - \overset{70.7177}{70,71}\right) \cdot 5000 =$

$= 0.01 \cdot 5000 = 50$

True answer: 35,35357146908... ≈ 35.35

Absolute error $= \left| 50 - 35.35\ldots \right| = 14.65\ldots$

Relative error $= \dfrac{14.65}{35.35} = .414\ldots$

41% error

Q What is the remedy:

$$f(x) = \left(\sqrt{x+1} - \sqrt{x}\right) x = \frac{\left(\sqrt{x+1} - \sqrt{x}\right)\left(\sqrt{x+1} + \sqrt{x}\right)}{\sqrt{x+1} + \sqrt{x}} \cdot x$$

$(a-b)(a+b)$ $\shortparallel$ a^2-b^2

$$= \frac{\cancel{x}+1 - \cancel{x}}{\sqrt{x+1} + \sqrt{x}} \cdot x = \frac{1}{\sqrt{x+1}+\sqrt{x}} \cdot x = \frac{x}{\sqrt{x+1} + \sqrt{x}}$$

: equivalent form of $f(x)$

$f(x) = \dfrac{x}{\sqrt{x+1} + \sqrt{x}}$ will give more correct value of $f(5000)$

$$f(5000) = \frac{5000}{\sqrt{5001} + \sqrt{5000}} \approx \frac{5000}{70.72 + 70.71} \approx \frac{5000}{141.4}$$

$$\begin{array}{r} 70.72 \\ +70.71 \\ \hline 141.43 \end{array}$$

≈ 14.14

$= 35.3606789 \approx 35.36$

True answer: $35.35\ldots$

Absolute error is 0.01

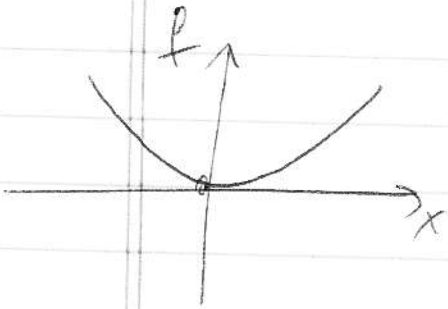
Relative error is $\dfrac{0.01}{35.35\ldots} = 2.829 \times 10^{-4}$

$2.829 \times 10^{-2}\,\% = 0.02829\%$

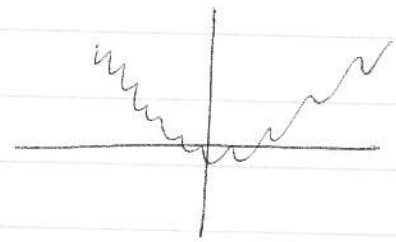
$\approx 0.03\%$

Ex Cancellation of digits

$$f(x) = e^x - \cos x - x, \quad x \text{ is near } 0$$



zoom to $-5\times10^{-8} \le x < 5\times10^{-8}$



Remedy:

We can use Taylor expansion of $f(x)$ at $x=0$.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

$$f(x) = e^x - \cos x - x = \left(\cancel{1} + \cancel{x} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots\right) -$$

$$-\left(\cancel{1} - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots\right) - \cancel{x} =$$

$$= \underbrace{2\cdot\frac{x^2}{2!} + \frac{x^3}{3!}} + \frac{x^5}{5!} + \dots$$

$$\Rightarrow f(x) \approx x^2 + \frac{x^3}{6} \quad \text{near } x=0$$