Example.

A number $x \in (-5, 5000)$.
We approx x with 3 binary sign figures (stand round). Find a number x with very low $|e_x|$ and very high $|\varepsilon_x|$

---

example $x = 35.897$

binary $35 = 32 + 2 + 1$
$= 1000\cancel{1}1_2$

$$\begin{array}{r} .897 \\ \underline{.5} \\ .397 \\ \underline{.25} \\ .147 \\ \underline{.125} \\ .022 \\ .015625 \end{array}$$

.0625
.03125
.015625

$35.897 = 100011.111001\ldots$ base 2

round

$= 100000_2$

$$\begin{array}{rl} e_x = x - \bar{x} = & 100011.111001\cdots \\ & \underline{-100000} \\ e_x = & \boxed{11.111001\cdots_2} \end{array}$$

Let $x = 1 = 1_2 = \bar{x}$

$e_x = x - \bar{x} = 0.$ $\Big]$ makes $|e_x|$ small.

$x = 5000$

$$\begin{array}{r} 5000 \\ -4096 \\ \hline 904 \\ -512 \\ \hline 392 \\ -256 \\ \hline 136 \\ -128 \\ \hline 8 \end{array}$$

$+\ 1$

$2^{12} + 2^9 + 2^8 + 2^7 + 2^3$

$x = 5000 = 1001110001000_2$

$\bar{x} = 1010000000000$

$$\begin{array}{r} 1010000000000 \\ 1001110001000_2 \\ \hline \end{array}$$

$x - \bar{x} = -0001111000_2$

$e_x = -1111000_2$

To make e_x really large

$x = 10001 \ldots 1$

$\bar{x} = 100000 \ldots 0_2$

$e_x = x - \bar{x} = 1\,0 \ldots 0_2$

$2^{10} - 1$

e_x is largest when 3rd decimal place is off by the most possible

13 12 11 | 10

$- - - \mid \uparrow \; 1^{10} 0 0 0 0 \text{———} 0_2$

most possible error.

Relative error

$$|\varepsilon_x| \le b^{-m+1} \qquad \bar{x}$$

achieved when $f = 10\ldots0$

$$\varepsilon_x = b^{-m+1}$$

m = # of sign figures.

standard rounding

$$\Rightarrow |\varepsilon_x| \le \frac{1}{2} b^{-m+1} \qquad \bar{x}$$

also achieved when $f = 10\text{—}0$

for example 3-sign binary digits

$$x = 100.01111_2 \longrightarrow 100_2$$

$$y = 100.1_2 \longrightarrow 101$$

$$\varepsilon_x = \frac{\overset{x}{100.01111} - \overset{\bar{x}}{100}}{100_2} = \frac{.01111}{100_2}$$

$$= .0001111 \approx .001_2 = \frac{1}{8}$$

$$\varepsilon_y = \frac{100.1_2 - 101_2}{101_2} = \frac{-.1_2}{101_2}$$

$$= \frac{-\frac{1}{2}}{5} = -\frac{1}{10}$$

Bound:

$$|\varepsilon_x| \le \frac{1}{2} b^{-m+1} = \frac{1}{2} 2^{-3+1} = \frac{1}{2} 2^{-2}$$

$$= 2^{-3} = \boxed{\frac{1}{8}}.$$

Bounding fractions.

$$\left|\frac{A}{B+C}\right| = \frac{|A|}{|B+C|}$$

$$\geq \frac{|A|}{|B|+|C|}$$

If $|B| > |C|$

$$\left|\frac{A}{B+C}\right| = \frac{|A|}{|B+C|} \leq \frac{|A|}{|B|-|C|}$$