CSCC37 TUT3 WEEK4
andrewyk.leung@mail.utoronto.ca

1) Representing reals in a floating point number system
2) rounding error / relative error
3) stability of formulae

① $R_b(t,s)$
- b: base
- t: size of mantissa
- s: size of exponent

ex: Represent $(72.6)_{10}$ in the FP # system
$b=3$ $t=8$ $e=[-2,+4]$
e: range of exponent

$(72)_{10} \rightarrow (2200)_3$

numerator	denominator	quotient	remainder
72	3	24	0
24	3	8	0
8	3	2	2
2	3	0	2

Stop when quotient is 0

$(.6)_{10} \longrightarrow (.\overline{1210})_3$ ← repeats infinitely

	multiplier	base	product	integral	fraction
repeats	.6	3	1.8	1	.8
	.8	3	2.4	2	.4
	.4	3	1.2	1	.2
	.2	3	0.6	0	.6
	.6	3	1.8	1	.8

Stop when fraction is 0 or cycle restarts

$\therefore (72.6)_{10} = (2200.\overline{1210})_3$

$\Longrightarrow (.22001210)_3 \times 3^{4}$ ($4 \rightarrow (11)_3$; size of exponent is 2)

mantissa; base; exponent

2) Absolute error (AE) $= X - fl(x)$

Relative error (RE) $= \dfrac{X - fl(x)}{X}$

$X = (0.2200\overline{1210})_3 \times 3^4$

$fl(x) = (0.22001210)_3 \times 3^4$

$AE = (0.00000000\overline{1210})_3 \times 3^4$

3) $1 - \cos x$

when $x \to 0$, $\cos x \to 1$ the formula suffers from

catastrophic / subtractive cancellation

Alternative: $1-\cos x \times \frac{1+\cos x}{1+\cos x}$

$$= \frac{1-\cos^2 x}{1+\cos x} \qquad [\sin^2 x + \cos^2 x = 1]$$

$$= \frac{\sin^2 x}{1+\cos x}$$

This formula is fine when $x \to 0$, but not when $x \to \pi$ (since $\cos x \to -1$)

$1-\cos x$ is fine as $x \to \pi$

Solution: Depending on the value of x, choose a formula that doesn't suffer from catastrophic / subtractive cancellation

$$\begin{cases} 1-\cos x & x \to \pi \\ \dfrac{\sin^2 x}{1+\cos x} & x \to 0 \end{cases}$$