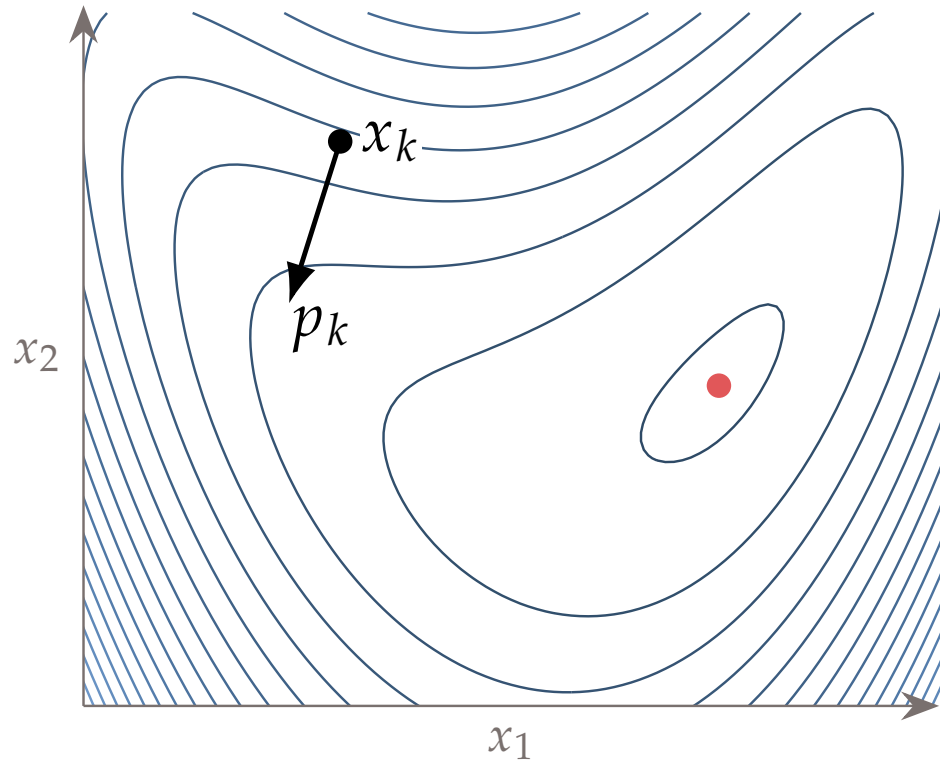


Numerical Errors

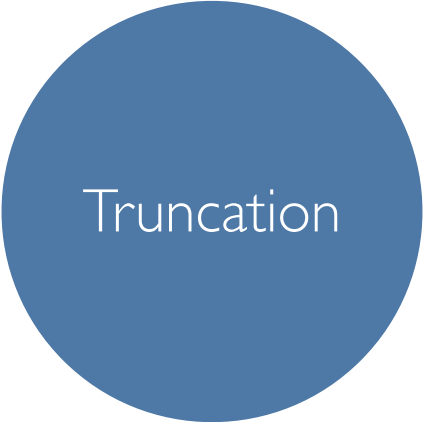


ME EN 275
Andrew Ning
aning@byu.edu

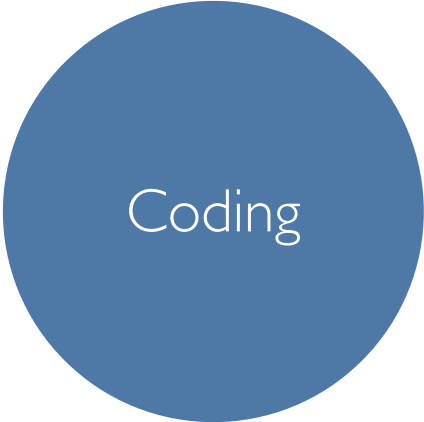
Numerical Errors



Roundoff



Truncation



Coding

Error Metrics

absolute error:

$$\mathcal{E} = |x - x^*|$$

estimate true
 ↓ ↖

relative error:

$$\mathcal{E} = \left| \frac{x - x^*}{x^*} \right| \times 100\%$$

Binary

what
number
is this?

$$\begin{array}{ccccccc} 1 & 1 & 0 & 0 & . & 1 & 0 & 1 \\ 2^3 & 2^2 & 2^1 & 2^0 & & 2^{-1} & 2^{-2} & 2^{-3} \end{array}$$

$$= 2^3 + 2^2 + 2^{-1} + 2^{-3}$$
$$= 12.625$$

$$135.2_{10}$$

$$\begin{array}{ccccccc} & 1 & & 3 & & 5 & & . & & 2 \\ \times & 10^2 & & 10^1 & & 10^0 & & & & 10^{-1} & \dots \end{array}$$

64 bit numbers

1 bit 52 bits 11 bits

↓ ↓ ↓

sign mantissa exponent

$$X = (-1)^{\text{sign}} (1 + \text{mantissa}) \times 2^{\text{exponent}}$$

Machine Precision

$$1 + \epsilon > 1$$

3 b.ts

1.

base2: 1.000

1.001



base10

1.0

1.125

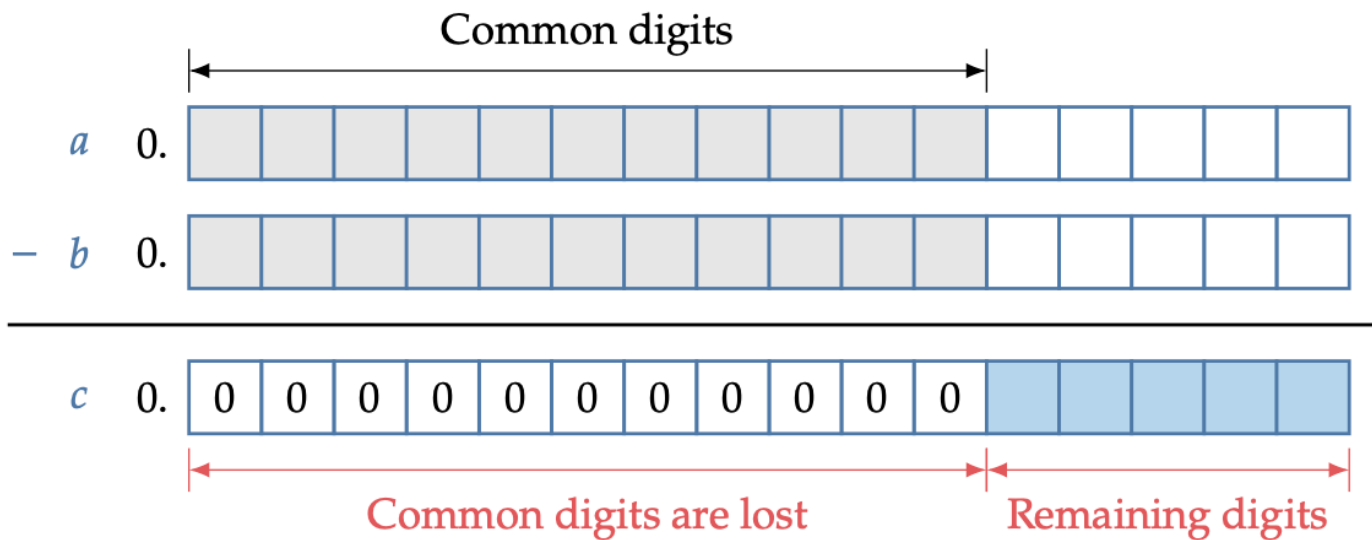
2.1.05



1.1 →

$$2^{-53} \sim \boxed{1 \times 10^{-16}}$$

Subtractive Cancellation

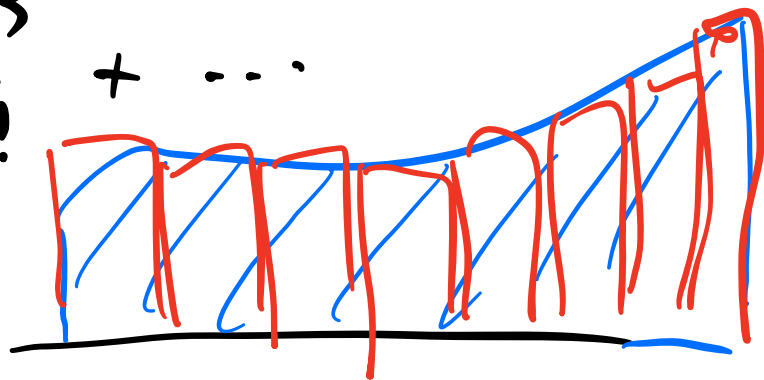


Examples

Truncation Error

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)h^2}{2!} + \dots$$

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



$$\int_a^b f(x)$$

