

## **Introduction:** Engineering problems and computational methods

- 20 cm x 30 cm rectangular plate
- Initially at temperature  $f(x,y)$ . Then, insulated at  $x=0$ , zero temperature at  $y=0$ , and convective heat transfer at  $x=20$  and at  $y=30$
- Find temperature variation with time at the center

Experimental: Advantages: No equations, Arbitrary Shape

Issues: Size, Material, Time,  $f(x,y)$

Analytical :  $\frac{\partial T}{\partial t} = \alpha \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right)$   $T(x, y, t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_{ij} e^{-\alpha(\beta_i^2 + \gamma_j^2)t} \cos(\beta_i x) \sin(\gamma_j y)$

Adv: Fast, Adaptable, Accurate; Issues: Ideal conditions

## Introduction: Engineering problems and computational methods

Numerical Method: Reduce to algebraic equations

E.g.,

		i,j+1		
	i-1,j	i,j	i+1,j	
		i,j-1		

$$\frac{\partial T}{\partial t} = \frac{T_{i,j}^{t+\Delta t} - T_{i,j}^t}{\Delta t}$$

$$\frac{\partial^2 T}{\partial x^2} = \frac{\frac{T_{i+1,j} - T_{i,j}}{\Delta x} - \frac{T_{i,j} - T_{i-1,j}}{\Delta x}}{\Delta x} = \frac{T_{i+1,j} - 2T_{i,j} + T_{i-1,j}}{\Delta x^2}$$

**Introduction:** Engineering problems and computational methods

**Results in a set of linear (or nonlinear, if  $\alpha$  is a function of temperature) equations**

**Adv:** Arbitrary shape, initial and boundary conditions; Usually fast.

**Issues:** Convergence, Accuracy, Efficiency

## **Error Analysis: Round-off and Truncation errors**

### Backward and Forward error analysis

**All computational results have errors: Computers have finite storage length and computations cannot be carried out infinitely (We do not consider Model error and Data error)**

**Round-off Error:** Due to finite storage length. E.g., with 8 significant digits:  $\sqrt{2}$  is stored as 1.4142136 (actually as  $0.14142136 \times 10^1$ ), an error of about  $3 \times 10^{-6}$  percent.

**Truncation Error:** Due to finite steps in computation. E.g.,  $\exp(2)$  is computed using the infinite series, truncated after a finite number of terms. True value is 7.389056098930650. 1 term – 1; 2 terms – 3, 3 – 5, 4 – 6.333, 5 – 7, 6 – 7.356, ..., 15 – 7.3890561.

## **Error Analysis: Round-off and Truncation errors**

### **Backward and Forward error analysis**

**Round-off Error:** Considering decimal storage, real numbers are stored as mantissa ( $m$ )  $\times 10^{\text{power } (p)}$  with  $0.1 \leq m < 1$ . If mantissa rounded-off at the  $t^{\text{th}}$  digit, maximum error =  $0.5 \times 10^{-t}$ .  
Maximum relative error =  $5 \times 10^{-t}$  called the round-off unit,  $u$ .

### **Definitions:**

*True Error ( $e$ ) = True Value – Approximate Value*

*Approximate Error ( $\varepsilon$ ) = Current Approximation – Previous Approximation*  
(e.g.,  $x = x/2 + 1/x$ ,  $x = 1, 1.5, 1.4166667, 1.4142157, 1.4142136$ )

*True Relative Error ( $e_r$ ) = (True Value – Approximate Value) / True Value*

*Approximate Relative Error ( $\varepsilon_r$ ) = (Current Approximation – Previous Approximation) / Current Approximation*

## Error Propagation

*y is a function of n independent variables,  $x_1, x_2, \dots, x_n$ , i.e. vector  $x$*

$$\mathbf{y} = \mathbf{f}(\mathbf{x})$$

Due to error in the variables ( $\Delta x_i = x_i - \tilde{x}_i$ ), there is error in y

$$\tilde{y} = f(\tilde{x})$$
$$\text{Error: } \Delta y = y - \tilde{y} = f(x) - f(\tilde{x}) = \sum_{i=1}^n \Delta x_i \left. \frac{\partial f}{\partial x_i} \right|_{\tilde{x}_i}$$

*(Neglecting higher order terms in the Taylor's series)*

## Error Propagation

$$g = \frac{2h}{t^2}$$

Height measurement accurate to a mm, time to 100<sup>th</sup> of sec.

Error in  $g$ , if  $h=100.000$  m and  $t=4.52$  s?

$$\frac{\partial g}{\partial h} = \frac{2}{t^2}; \frac{\partial g}{\partial t} = -\frac{4h}{t^3}$$

$$\Delta g \cong \frac{2}{\tilde{t}^2} \Delta h - \frac{4\tilde{h}}{\tilde{t}^3} \Delta t$$

$$= 0.0978933\Delta h - 4.33156\Delta t$$

**Max value 0.0434** (Relative error is more meaningful, roughly 0.4%.)

## Error Propagation and Condition Number of a problem

**Condition number:** Ratio of the relative change in the function,  $f(x)$ , to the relative change in the variable for a small perturbation  $\Delta x$  in the variable  $x$ .

$$C_p = \frac{\left| \frac{f(x + \Delta x) - f(x)}{f(x)} \right|}{\left| \frac{\Delta x}{x} \right|} \quad \text{or, as } \Delta x \rightarrow 0, \frac{xf'(x)}{f(x)}$$

***Well-Conditioned (<1) and Ill-conditioned (>1)***

E.g.: Hyperbola  $xy - y + 1 = 0$ . Measure  $x$ , compute  $y = f(x) = 1/(1-x)$ .

$C_p = |x/(1-x)|$  **Ill-conditioned for  $x > 0.5$**  ( $x=0.8, y=5$ ;  $x=0.808, y=5.2083$ )

Similarly,  $y = \sqrt{1+x} - 1$  for small  $x$ ;

$C_p = \left| \frac{x}{2(1+x - \sqrt{1+x})} \right|$  close to but  $< 1$ . ( $x=0.01, y=.004988$ ;  $x=0.0101, y=.005037$ )



## Error Analysis: Forward and Backward error analysis

**Forward Error:** requires the T.V.

(Recall that *True Error (e)* = *True Value* – *Approximate Value*)

E.g.:  $\sqrt{2} = 1.414213562373095$ . With 4 digit accuracy ( $u=5 \times 10^{-4}$ ), approx. value = 1.414. True error,  $e = 2.134 \times 10^{-4}$ . True relative error = 0.015%.

**Backward Error:** determines the change necessary in the input variable (or data) to explain the final error in the output (or result) obtained through errors introduced at various stages of calculation.

E.g.:  $1.414^2 = 1.999396$ . Error =  $6.04 \times 10^{-4}$ , Relative error = 0.03%.

$y = \sqrt{1+x} - 1$  with  $x=0.001616$  gives  $y = \sqrt{1.002} - 1 = 0.0009995$ , which is exact for  $x=0.002$  (error = 0.000384, 24%)

## Condition Number of an algorithm

- Condition number:** Exact input data,  $x$ ; Corresponding function value  $y$ ; Algorithm  $A$  operating on  $x$  produces  $y_A$  on a machine with precision unit  $u$ . Input data,  $x_A$ , which would produce  $y$ , if exact computations are done. *Condition number of the algorithm*

$$C_A = \frac{\left| \frac{x - x_A}{x} \right|}{u}$$

**Well-Conditioned ( $<1$ ) and Ill-conditioned ( $>1$ )**

$$y = \sqrt{1+x} - 1 \quad \text{for } x=0.001616, \text{ we get } x_A=0.002.$$

$$C_A = \frac{\left| \frac{0.001616 - 0.002}{0.001616} \right|}{5 \times 10^{-4}} = 475$$

$$y = \sqrt{1+x} - 1 = \frac{x}{\sqrt{1+x} + 1}$$

$$x=0.001616, y=0.0008076,$$

$$x_A=0.00161585, C_A=0.18$$

# Error Analysis: Vector Norms

What if the computed result is a vector?

$$[A]\{x\} = \{b\}$$

**Error vector**

$$\{\varepsilon\} = \{x\} - \{\tilde{x}\}$$

**Or, relative error vector**

$$\{\varepsilon_r\} = \left\{ \frac{x - \tilde{x}}{x} \right\}$$

**Magnitude:** In terms of the “Norm” of the vector

# Error Analysis: Vector Norms

Some properties of the vector norms are ( $\alpha$  is a scalar):

- $\|x\| = 0$  , only if  $x$  is a null vector, otherwise  $\|x\| > 0$
- $\|\alpha x\| = |\alpha| \|x\|$
- $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$

**Magnitude:** The  $L_p$  norm of an  $n$ -dimensional vector,  $\{x\}$ , *i.e.*,  $(x_1, x_2, \dots, x_n)^T$ , is given by

$$\|x\|_p = \left( |x_1|^p + |x_2|^p + |x_3|^p + \dots + |x_n|^p \right)^{1/p} \quad p \geq 1$$

# Common Vector Norms

- $p=2$ , Euclidean norm: (length)  $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$
- $p=1$ , (total distance)  $\|x\|_1 = \sum_{i=1}^n |x_i|$
- $p=\infty$ , Maximum norm  $\|x\|_\infty = \max_i |x_i|$