DSSC / MHPC / AMMA : Numerical Analysis
LM : Applied Mathematics

luca.heltai@sissa.it
gianluigi.rozza@sissa.it

---

Three main topics:
- Algorithms (approximation)
- Analysis (error)
- Implementation (python)

Numerical Approximation:
- Something "close in value", but not the same, as a desired quantity

The subjects of numerical approximation are well posed problems, i.e. problems for which a solution
- Exists
- It is unique
- It Depends continuously on problem data

A numerical approximation should not make things worse

Strategy: Replace difficult problems by easier ones.

| | | |
|---|---|---|
| infinite | $\longrightarrow$ | finite |
| differential | $\longrightarrow$ | algebraic |
| non-linear | $\longrightarrow$ | linear |

**Absolute Error:** Approximate Value − true value

**Relative Error:** $\dfrac{\text{Absolute Error}}{\text{True value}}$ (sometimes Appr. Value)

**Example:** compute the value of a function

$f : \mathbb{R} \longrightarrow \mathbb{R}$

for a given argument $x$

- $x$ : true value of input
- $f(x)$ : desired output
- $\hat{x}$ : approximate input (maybe due to Rounding)
- $\hat{f}(\hat{x})$ : function actually computed (both Rounding and Truncation)

Types of error

- Data Error (← Data)
- Computational Error (← Algorithm)
- Truncation Error (← Algorithm in exact Arithmetic)
- Rounding Error (← Finite Arithmetic)

Total Error: $\hat{f}(\hat{x}) - f(x) =$

$$\underbrace{\hat{f}(\hat{x}) - f(\hat{x})}_{\text{Computational error}} + \underbrace{f(\hat{x}) - f(x)}_{\text{propagated data error}}$$

(independent on algorithm!)

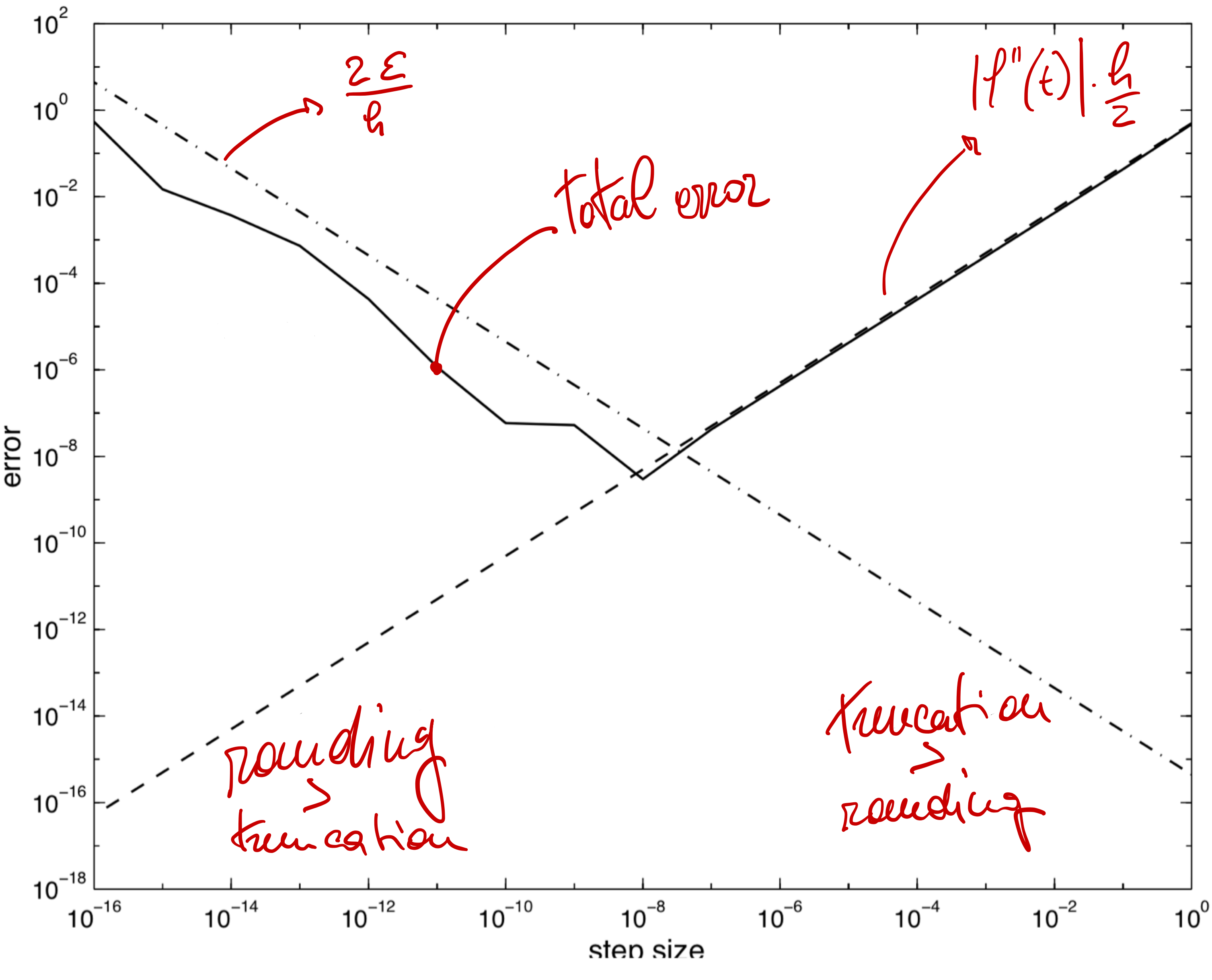
Computational error = Truncation Error + Rounding Error

Example: error in finite difference approximation

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

- Truncation error $\sim c\,|f''(t)|\cdot\frac{h}{2}$
- Rounding error $\sim c\,\frac{2\varepsilon}{h}$ ($\varepsilon$: machine precision)

Optimal choice for $h$: $\approx 2\sqrt{\frac{\varepsilon}{|f''(t)|}}$



A problem is well posed (or insensitive, or stable)
if small changes in input $\rightsquigarrow$ small changes in output

**Condition Number**

$$\kappa := \boxed{\frac{\text{Relative change in solution}}{\text{Relative change in input data}}}$$

$$\kappa = \frac{\big(f(x) - f(\hat{x})\big)/f(x)}{(x - \hat{x})/x} = \frac{\Delta y/y}{\Delta x/x}$$

$\kappa \gg 1 \rightsquigarrow$ Problem is ill-posed (or sensitive or unstable)

Absolute Condition number (if either $f(x)$ or $x$ is zero).

$$\kappa_{abs}(x) := \boxed{\frac{\Delta y}{\Delta x} = \frac{f(x) - f(\hat{x})}{(x - \hat{x})}}$$

---

---

$\hat{x} = x + \delta x \rightsquigarrow f(\hat{x}) \sim f(x) + f'(x)\,\delta x$

$\rightsquigarrow$

$$\kappa(x) \approx |f'(x)| \frac{|x|}{|f(x)|}$$

$$\kappa_{abs}(x) \approx |f'(x)|$$

A numerical approximation can be seen as a sequence of problems, that are simpler to evaluate, and converge to the original problem, i.e., $\forall n, \exists f_n$ s.t.

$$\lim_{n\to\infty} f_n(x) = f(x)$$

Each "problem" $f_n$ should be itself stable

| | | | | |
|---|---|---|---|---|
| $f$ | stable | $f_n$ | stable $\forall n$ | ✓ |
| $f$ | stable | $f_n$ | unstable $\forall n$ | ✗ |
| $f$ | un-stable | $f_n$ | stable $\forall n$ | ✗ |
| $f$ | un-stable | $f_n$ | un-stable $\forall n$ | ✗ |

the only reasonable option

$$K_n(x) \approx |f_n'(x)| \frac{|x|}{|f_n(x)|}$$

$$K_{n_{abs}}(x) \approx |f_n'(x)|$$

In general both input data or output result may belong to (infinite dimensional) vector spaces, i.e.

$$x \in \mathcal{X}, \quad y \in \mathcal{Y}$$

$$f : \mathcal{X} \longrightarrow \mathcal{Y}$$

And we assume that $f_n : \mathcal{X}_n \longrightarrow \mathcal{Y}_n$

where, in some sense

$$\lim_{n\to\infty} \mathcal{X}_n = \mathcal{X}, \quad \lim_{n\to\infty} \mathcal{Y}_n = \mathcal{Y}$$

$$\lim_{n\to\infty} f_n = f$$

We restrict ourselves to Banach spaces
( Normed and complete vector spaces )

That is $\mathbb{X}$, $\mathbb{Y}$ are normed and complete vector spaces

COMPLETENESS: Every Cauchy sequence in $\mathbb{X}$ converges to an element of $\mathbb{X}$

That is:

$$\forall \{x_n\}_{n=1}^{\infty} \text{ s.t. } \forall \varepsilon > 0, \exists \bar{n} \text{ s.t. } \forall i,j > \bar{n} \quad \|x_i - x_j\| \leq \varepsilon$$

$$\exists!\, x \in \mathbb{X} \text{ s.t. } \lim_{i \to \infty} \|x_i - x\| = 0$$

In the most general case a problem is a functional $f$ such that:

$$f : \mathbb{X} \longrightarrow \mathbb{Y}$$
$$x \longrightarrow f(x)$$

While for every $n$, there exist two sets $\mathbb{X}_n \subseteq \mathbb{X}$ and $\mathbb{Y}_n \subseteq \mathbb{Y}$ and a functional $f_n$ such that

$$f_n : \mathbb{X}_n \longrightarrow \mathbb{Y}_n$$
$$x_n \longrightarrow f_n(x_n)$$

A numerical approximation $f_n$ of a problem $f$ is convergent if

- $\lim_{n \to \infty} \|x_n - x\|_{\mathbb{X}} = 0$
- $\lim_{n \to \infty} \|\underline{\underline{f_n(x_n)}} - f(x)\|_{\mathbb{X}} = 0$

approximated data

If $x \in \mathbb{X}_n \quad \forall n$, an approximation is called consistent if

- $\lim_{n \to \infty} \|\underline{\underline{f_n(x)}} - f(x)\| = 0$

exact data

Examples : The sum of two numbers.

- $\mathcal{X} : \mathbb{R}^2 \quad , \quad \|x\|_{\mathcal{X}} := |x_1| + |x_2| = \|x\|_{\ell^1(\mathbb{R}^2)}$
- $\mathcal{Y} : \mathbb{R} \quad , \quad \|y\|_{\mathcal{Y}} := |y| = \|y\|_{\ell^1(\mathbb{R}^1)}$
- $f(x) := x_1 + x_2$

$$K_{rel} = \frac{|\Delta y|}{|\Delta x|}\frac{|x|}{|y|} \rightsquigarrow \frac{\Delta y}{\Delta x} = \frac{|\Delta x_1 + \Delta x_2|}{|\Delta x_1| + |\Delta x_2|} \leq 1 \qquad \frac{|x|}{|y|} = \frac{|x_1| + |x_2|}{|x_1 + x_2|}$$

$$K_{rel} \leqslant \frac{|x_1| + |x_2|}{|x_1 + x_2|}$$

UNSTABLE whenever $x_1 \sim -x_2 \Rightarrow K_{rel} \rightsquigarrow \infty$

BIG problem in any floating point architecture

---

FINITE DIFFERENCE

- $\mathcal{X} : C^1(I(a)) \quad , \quad a \in \mathbb{R}$ (given)
- $\mathcal{Y} : \mathbb{R}$
- $f(x) = x'(a)$ — $x$ is a continuous function with continuous first derivative

$$\|x\|_{\mathcal{X}} := \|x'\|_{L^\infty(I(a))} + \|x\|_{L^\infty(I(a))}$$

$$\Delta y = f(x + \delta x) - f(x) = \delta x'(a) \rightsquigarrow K = \frac{|\delta x'(a)|}{\|\delta x\|_{\mathcal{X}}} \frac{\|x\|_{\mathcal{X}}}{|x'(a)|}$$

- $\mathcal{X}_n \equiv \mathcal{X} \quad , \quad \mathcal{Y}_n \equiv \mathcal{Y}$
- $f_n(x) = \dfrac{x(a + 1/n) - x(a)}{1/n}$

$\mathcal{X} = \mathcal{X} \longrightarrow$ convergence and consistency are the same

But the numerical problem is unstable! (Sum of two numbers with close abs. values and opposite sign!)

# Integration

- $\mathcal{X} := L^1([0,1]) \cap C^0([0,1])$, $\|x\|_{\mathcal{X}} := \|x\|_{L^1([0,1])} = \int_0^1 |x(t)|\,dt$
- $\mathcal{Y} := \mathbb{R}$
- $f(x) := \int_0^1 x(t)\,dt$

$$f(x+\delta x) = \int_0^1 x+\delta x = \int_0^1 x + \int_0^1 \delta x$$

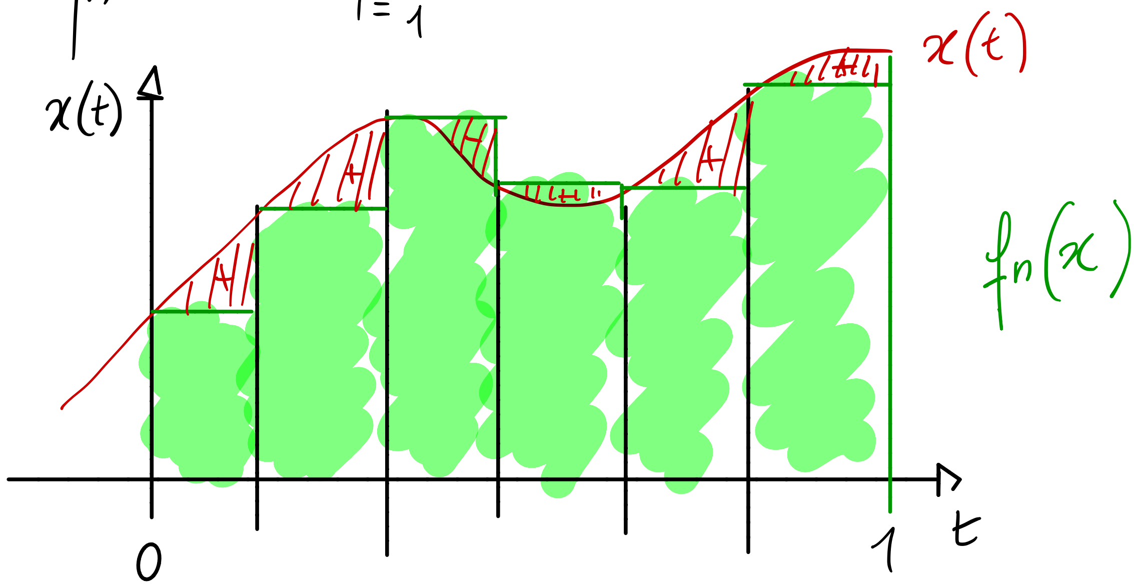
$$\frac{\|\Delta y\|}{\|\Delta x\|}\,\frac{\|x\|}{\|y\|} = \frac{\left|\int_0^1 \delta x\right|}{\int_0^1 |\delta x|}\,\frac{\int_0^1 |x|}{\left|\int_0^1 x\right|}$$

ill posed when average is close to zero.

Given a sequence $\{t_i\}_{i=1}^{n}$ $t_1=0<t_2<t_3<\dots<t_n=1$

$h = \min(t_{i+1}-t_i)$
$H = \max(t_{i+1}-t_i)$

- $f_n(x) = \sum_{i=1}^{n-1} x(t_i)(t_{i+1}-t_i)$




- $\lim_{n\to\infty} f_n(x) = f(x)$ Lebesgue integration

$$K_n(x) = \frac{\Delta y}{\Delta x}\,\frac{x}{y} = \frac{\left|\sum_{i=1}^{n-1} \delta x(t_i)(t_{i+1}-t_i)\right|}{\int_0^1 |\delta x|}\,\frac{\int_0^1 |x|}{\left|\sum_{i=1}^{n-1} x(t_i)(t_{i+1}-t_i)\right|}$$

$$\leq \frac{H\,\sum \delta x_i}{h\,\sum x_i}\,\frac{\|x\|_{L^1}}{\|\delta x\|_{L^1}}$$