# MATH 3Q03 - NUMERICAL EXPLORATIONS

## 1) Introduction & Review

<u>Numerical Analysis</u> - branch of applied mathematics studying computational algorithms designed to solve problems in analysis (e.g., representation of functions, differentiation, integration, solution of differential equations, etc. it focuses on transforming continuous (inf-dimensional) problems to a finite dimension (representable on a computer)

<u>Scientific Computing</u> - focuses on an efficient implementation of ~~numerical~~ numerical algorithms on computers; present-day computers perform the following elementary operations:

* addition / subtraction / multiplication / division
* comparison
* storage

Numerical computations thus transforms (continuous) problems in <u>analysis</u> into problems in <u>algebra</u>

## 1.1) Errors in Numerical Computations

All numerical methods are (only) approximate.

Quote: "Although this may seem a paradox, all exact science is dominated by the idea of approximation"

Bertrand Russell (1872-1970)

* in most practical situations, solutions are needed up to a certain accuracy
* even the solution of a problem is available analytically, its numerical value may only be determined up to a certain accuracy, e.g. $\pi = 3.14\ldots$
  - L how to cut a board of length $\pi$?
  - L sometimes analytical solutions may be very difficult to evaluate, so that it may be easier to approximate them numerically

Goal of Numerical Analysis - understand and characterize the structure of the errors, so that the approximate solution can be made as accurate as we wish by using more computational resources (CPU time and memory)

We are also interested in studying the <u>computational efficiency (complexity)</u> of numerical methods (number of arithmetic operations required to solve the problem with a given accuracy)

## <u>Types of systematic errors</u>

* errors in models or data

$$m\ddot{x} = f \quad \text{instead of} \quad m\ddot{x} + kx = f$$
for $k \equiv 0$

$$(m + m')\ddot{x} = f$$

→ outside the scope of ~~the~~ the course (math modelling)

* truncation errors — direct consequence of the approximation; the name comes from the analogy with truncation of a series expansion,

<u>Ex</u>

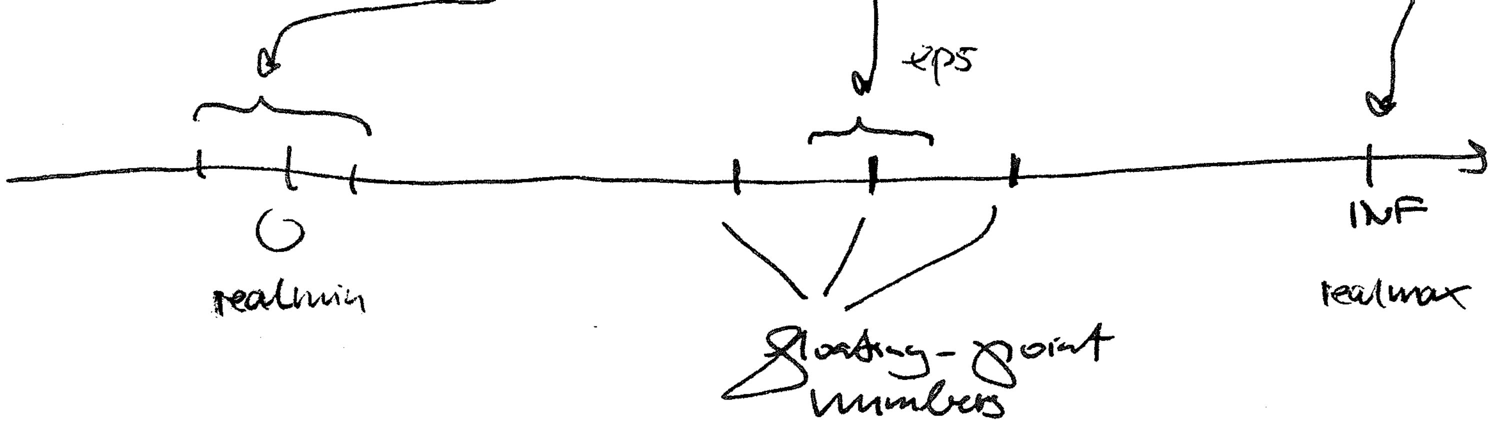
$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \ldots + \frac{x^k}{k!} + \overbrace{\sum_{n=k+1}^{\infty} \frac{x^n}{n!}}^{\text{truncation error}}$$

Note that taking $k \to \infty$ reduces ~~the~~ (eliminates) the error

* ~~~~ Round-off Errors — arise when real numbers (both rational and irrational) are represented using a finite number of "bits" in a digital computer; the so-called "floating point" arithmetic

finite arithmetic precision gives rise to:

- rounding (~~xx~~ real numbers are rounded towards the nearest floating-point number)
- underflow
- overflow




IEEE standard – single vs. ~~[illegible]~~ double precision

Both truncation and round-off errors can propagate and accumulate at various stages of the computation. The total computational error is the sum resulting from the interaction of the component errors

Absolute error = |true value − approx. value|

Relative error = $\dfrac{\text{Absolute error}}{|\text{true value}|}$, true value $\neq 0$

(usually more relevant)

~~[illegible] frequently associated with truncation errors~~

Surprise — floating-point arithmetic is not (always) associative

$$(a+b)+c \neq a+(b+c)$$

Example in MATLAB (finite_prec_arith_01.m)

---

Ex function evaluation $y = f(x)$, $f: \mathbb{R} \to \mathbb{R}$

$\hat{x}$ — approximate input (subject to floating-point errors)

$\hat{f}$ — approximate operation (subject to truncation errors)

$$\hat{f}(\hat{x}) - f(x) = [\hat{f}(\hat{x}) - f(\hat{x})] + [f(\hat{x}) - f(x)]$$

By triangle inequality $|a+b| \leq |a| + |b|$

$$|\hat{f}(\hat{x}) - f(x)| \leq \underbrace{|\hat{f}(\hat{x}) - f(\hat{x})|}_{\text{truncation error}} + \underbrace{|f(\hat{x}) - f(x)|}_{\text{round-off error}}$$

Condition number — sensitivity of a mathematical procedure to input errors

$$\varkappa = \frac{|\hat{y}-y|}{|y|} \Big/ \frac{|x-\hat{x}|}{|x|}$$

In the course we will be primarily interested in analyzing truncation errors. What about round-off errors?

Interval Arithmetics – an emerging technique for performing arithmetic operations in such a way that the result is determined together with its "confidence interval" based on the corresponding confidence intervals of the data

Well-posed Problems

Solutions:

- exist,
- are unique,
- and depend continuously on the data.

Problems which are not well-posed are ill-posed.

## 1.2 Root finding Methods for Nonlinear Equations

(Sections 8.1, 8.2 of Grasselli & Pelinovsky)

How do find "roots" (solutions, or zeros) of the equation $f(x)=0$?

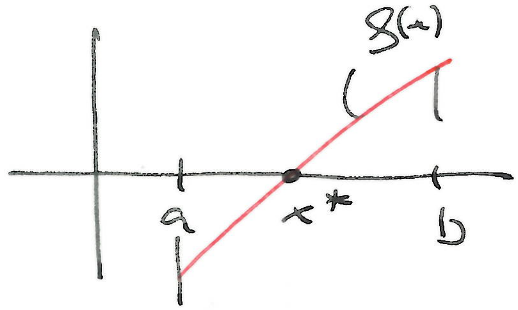
The function $f: \mathbb{R} \to \mathbb{R}$ may, or may not, be specified analytically, but will have to satisfy certain requirements (to be made precise later) e.g.,

- continuity
- existence of derivatives (differentiability)
- ...

In some cases $f(x)$ may be very difficult to evaluate, hence we are interested in method which find a good approximation of the root(s) with as few evaluations of $f(x)$ as possible.

## * BISECTION (Interval-halving) METHOD

- start with two values of $x$ that bracket the root $f(a) \cdot f(b) < 0$




- repeatedly half the interval and retain the part where $f(x)$ changes the sign, by ~~...~~ verifying the condition $f(x_i) \cdot f(x_{i+1}) < 0$

- stop when the interval containing the root shrinks to below some tolerance $\varepsilon$ prescribed