Advanced Numerical Analysis

Lecture 2
Spring 2025

Daniel Kressner

Quiz from Exercise Set 1

Definition 1.6. For $\beta \in \mathbb{N}$, $\beta \geq 2$ (**base**), $t \in \mathbb{N}$ (**length of mantissa**), and $e_{\min} < 0 < e_{\max}$ with $e_{\min}, e_{\max} \in \mathbb{Z}$ (**range of exponent**), set $\mathbb{F} = \mathbb{F}(\beta, t, e_{\min}, e_{\max}) \subset \mathbb{R}$ is defined as

$$\left\{ \pm\beta^e \left(\frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \cdots + \frac{d_t}{\beta^t} \right) : \quad \begin{array}{l} d_1, \ldots, d_t \in \{0, \ldots, \beta - 1\}, \\ d_1 \neq 0, \\ e \in \mathbb{Z},\ e_{\min} \leq e \leq e_{\max}. \end{array} \right\} \cup \{0\}.$$

a) How many distinct elements are contained in $\mathbb{F}(2, 3, -1, 1)$?

- ◯ 17
- ◯ 25
- ◯ 49
- ◯ 129

b) Let $\mathbb{F} = \mathbb{F}(\beta, t, e_{\min}, e_{\max})$ be a set of floating point numbers in the sense of Definition 1.6. Is the following statement true? If $a, b \in \mathbb{F}$, then $a + b \in \mathbb{F}$.

- ◯ True
- ◯ False

expm1

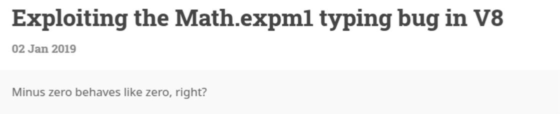
Reputable math libs have dedicated function for computing

$$\text{expm1}(x) := e^x - 1.$$

Why?

- Important in applications (compound interest rates in finance, solutions of differential equations, ...)
- Very tricky to implement[1] for $x \approx 0$ because of numerical cancellation

Examples: `numpy.expm1`, 1999 ISO C standard: `expm1`.



https://abiondo.me/2019/01/02/exploiting-math-expm1-v8/

[1]See https://www.math.utah.edu/~beebe/reports/expm1.pdf for more than you ever wanted to know about this problem.

φ-functions

φ-functions like

$$\varphi_1(x) := \frac{e^x - 1}{x}$$

play an important role when numerically solving differential equations.

For $x = 10^{-8}/9$,
$\varphi_1(x) = 1 + x/2 + x^2/6 + \cdots = 1.000000000555555\ldots$.

```
( numpy.exp(x) - 1 ) / x
```

returns 1.000000082740371.[2]

```
numpy.expm1(x) / x
```

returns 1.000000000555556.

See lecture notes for another (strange) implementation that is surprisingly accurate.

[2]The result may differ on your computer.