

# MT94 - Codage et dérivation

S. Mottelet

Université de Technologie de Compiègne

# Plan

- 1 Codage des nombres réels
- 2 Approximation d'une dérivée
- 3 Un pas de côté (dans le plan complexe)

# 1. Codage des nombres réels

Nombres réels dans une base  $b$

Pour un entier  $b \geq 2$  et  $(e_k, m_k) \in [0, b - 1]^2$ , on note

$$(e_n e_{n-1} \dots e_0 . m_1 m_2 \dots)_b = \sum_{k=0}^n e_k b^k + \sum_{k>0} m_k b^{-k}$$

Exemples :

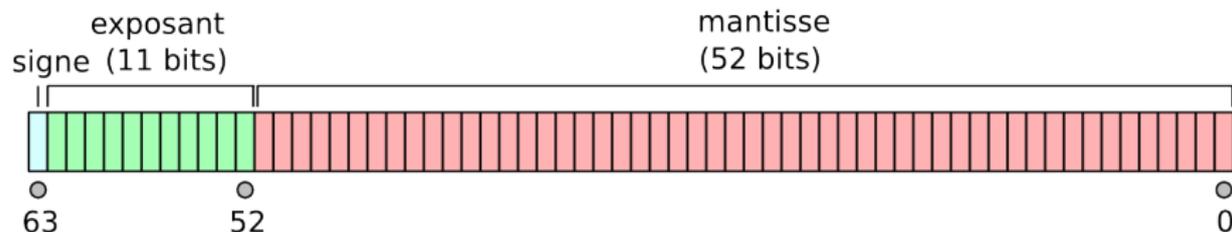
- $12.1 \stackrel{\text{def}}{=} (12.1)_{10} = 1 \times 10^1 + 2 \times 10^0 + 1 \times 10^{-1}$
- $(101.11)_2 = 2^2 + 2^0 + 2^{-1} + 2^{-2} = 5.75$
- $(1f)_{16} = 1 \times 16 + 15 = 31$

Notation :  $0.3 = (0.01001100110011\dots)_2 \stackrel{\text{def}}{=} (0.01\overline{0011})_2$

# 1. Codage des nombres réels

Norme IEEE 754 (depuis 1985)

Nombre à virgule flottante codé sur 64 bits



$$x = (-1)^s (1.m_1 m_2 \dots m_{52})_2 \times 2^{e-1023}$$

Exemple : codage de 12.1

# 1. Codage des nombres réels

Norme IEEE 754 (depuis 1985)

$$12.1 = 1.5125 \times 2^3 = (1.\overline{10000011})_2 \times 2^3$$

$$x = (0.m_1m_2\dots)_2 = \sum_{k>0} m_k 2^{-k}$$

## Algorithme

$x_1 = x$ , et pour  $k > 0$

$$m_k = \lfloor 2x_k \rfloor,$$

$$x_{k+1} = 2x_k - m_k.$$

$$0.5125 \times 2 = 1.025$$

$$0.025 \times 2 = 0.05$$

$$0.05 \times 2 = 0.1$$

$$0.1 \times 2 = 0.2$$

$$0.2 \times 2 = 0.4$$

$$0.4 \times 2 = 0.8$$

$$0.8 \times 2 = 1.6$$

$$0.6 \times 2 = 1.2$$

$$0.2 \times 2 = \dots$$

# 1. Codage des nombres réels

Norme IEEE 754 (depuis 1985)

$$12.1 = (1.10000011)2^{1026-1023}$$

Dans scilab :

```
-> bitstring(12.1)
ans =
```

```
"01000000010100000110011001100110011001100110011001100110011001100110011001100110011"
```

Attention : en pratique, l'encodage fait appel à des calculs sur des nombres entiers.

# 1. Codage des nombres réels

Calculs, précision relative

- Addition de deux nombres  $x > y$  :

$$x = (1.m_1 \dots m_{52})_2 \times 2^{e-1023},$$

$$y = (1.n_1 \dots n_{52})_2 \times 2^{f-1023},$$

On doit additionner  $e - f$  à l'exposant de  $y$  et diviser sa mantisse par  $2^{e-f}$  :

$$y \text{ est codé } (0.0 \dots 01n_1 \dots n_{52-e+f})_2 \times 2^{e-1023}$$

$$e - f = 52 :$$

$$y \text{ est codé } (0.0 \dots 01) \times 2^{e-1023}$$

$$e - f = 53 :$$

$y$  est codé 0.



# 1. Codage des nombres réels

Précision relative, dans Scilab

$$\varepsilon = 2^{-52} \approx 2.22 \times 10^{-16}$$

$\varepsilon$  = plus petit nombre tel que  $x + \varepsilon x \neq x$   
 $\varepsilon$  = précision relative

# 1. Codage des nombres réels

Perte de signification ou *catastrophic cancellation*

Calcul de  $x - y$  quand les deux nombres sont proches :

```
-> x=cos(1); y=cos(1+2^-26); bitstring([x;y;x-y])
```

ans =

```
"0011111111100001010010100010100000001111101101010000011010001100"  
"0011111111100001010010100010100000001000111110011011000101101000"  
"001111100100101011101101010101001001000000000000000000000000000000"
```

Perte de précision de 25 bits sur 52 !

## 2. Approximation d'une dérivée

### La théorie et la pratique

Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$  continûment dérivable deux fois, on a pour tout  $h \neq 0$

$$f(x+h) = f(x) + f'(x)h + \frac{h^2}{2}f''(\xi), \quad \xi \in \text{Int}(x, x+h),$$
$$f'(x) = \frac{f(x+h) - f(x)}{h} + \frac{h}{2}f''(\xi).$$

En pratique, on calcule en machine  $D(h) = \frac{f(x+h) - f(x)}{h} + \delta$  où

$$|\delta| \leq 2\varepsilon \left| \frac{f(x)}{h} \right|,$$

d'où pour  $h$ , il existe  $\xi \in \text{Int}(x, x+h)$  tel que

$$|f'(x) - D(h)| \leq 2\varepsilon \left| \frac{f(x)}{h} \right| + \left| \frac{h}{2}f''(\xi) \right|.$$

## 2. Approximation d'une dérivée

La théorie et la pratique

Le  $h$  optimal peut être approché par

$$h^* = 2 \left| \varepsilon \frac{f(x)}{M_2} \right|^{1/2},$$

où  $|f''(x)| \leq M_2$  pour tout  $x$ .

Exercice : calculez  $h^*$  pour la formule d'ordre 2 :

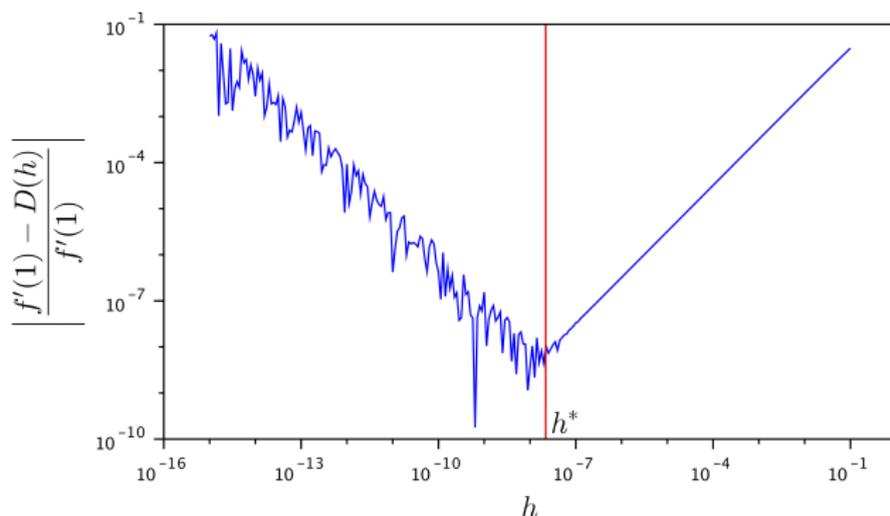
$$D_2(h) = \frac{f(x+h) - f(x-h)}{2h}$$

## 2. Approximation d'une dérivée

### Exemple

Approximation de  $f'(1)$  quand  $f(x) = \cos(x)$  :  $D(h) = \frac{\cos(1+h) - \cos(1)}{h}$

$$h^* = 2(2^{-52} \cos(1))^{1/2} \approx 2.19 \times 10^{-8}$$



### 3. Un pas de côté (dans le plan complexe)

Fonction développable en série

**Définition** : on dit que  $f : \mathbb{R} \rightarrow \mathbb{R}$  est développable en série en  $x_0$  s'il existe  $(a_n)$  et  $R > 0$  tels que si  $|x - x_0| < R$  alors

$$f(x) = \sum_{n \geq 0} a_n (x - x_0)^n.$$

Pour toute fonction développable en série on peut définir son extension/prolongement au plan complexe  $F : \mathbb{C} \rightarrow \mathbb{C}$  par

$$F(z) = \sum_{n \geq 0} a_n (z - x_0)^n.$$

L'analyse complexe est une branche des mathématiques étudiant les propriétés de telles fonctions.

### 3. Un pas de côté (dans le plan complexe)

#### Prolongement complexe

Exemples avec  $x_0 = 0$  :

$$\exp z = \sum_{n \geq 0} \frac{z^n}{n!},$$

$$\cos z = \sum_{n \geq 0} (-1)^n \frac{z^{2n}}{(2n)!},$$

qui se calculent en pratique, pour  $x$  et  $y$  réels comme

$$\exp(x + iy) = \exp x \cos y + i \exp x \sin y,$$

$$\cos(x + iy) = \cos x \cosh y - i \sin x \sinh y.$$

### 3. Un pas de côté (dans le plan complexe)

Prolongement complexe, dans Scilab

- opérateurs  $+$ ,  $*$ ,  $/$ ,  $\wedge$  sur les scalaires, vecteurs et matrices,
- fonctions usuelles (trigo., hyperboliques, réciproques, ...)

```
--> (1+%i)*(1+2*%i)
ans =
```

```
-1. + 3.i
```

```
--> cos(1+%i)
ans =
```

```
0.83373 - 0.9888977i
```

```
--> 2^(1+%i)
ans =
```

```
1.5384778 + 1.2779226i
```

### 3. Un pas de côté (dans le plan complexe)

Approximation d'une dérivée avec un pas complexe

Soit  $x_0 \in \mathbb{R}$  et  $h \in \mathbb{R}$ , on a

$$f(x_0 + ih) = f(x_0) + ihf'(x_0) - \frac{h^2}{2}f''(x_0) + iO(h^3),$$
$$\operatorname{Im}(f(x_0 + ih)) = hf'(x_0) + O(h^3),$$

$$f'(x_0) = \frac{\operatorname{Im}(f(x_0 + ih))}{h} + O(h^2),$$

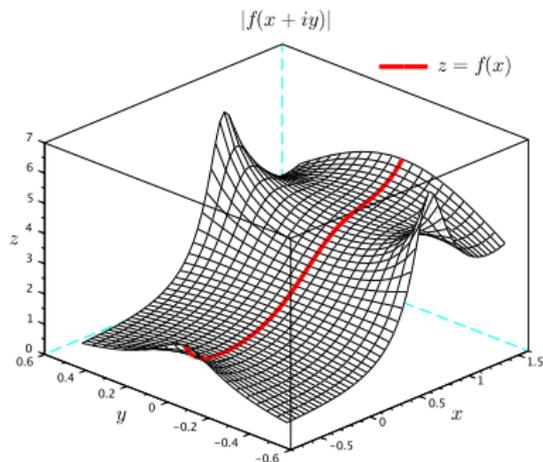
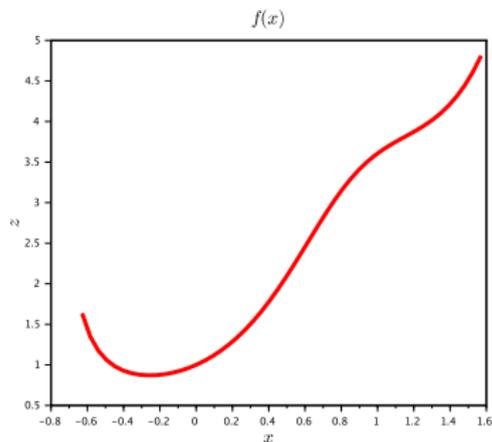
à comparer avec une autre formule d'ordre 2 :

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0 - h)}{2h} + O(h^2),$$

### 3. Un pas de côté (dans le plan complexe)

Exemple

$$f(x) = \frac{\exp(x)}{\cos^3 x + \sin^3 x}$$



### 3. Un pas de côté (dans le plan complexe)

Exemple

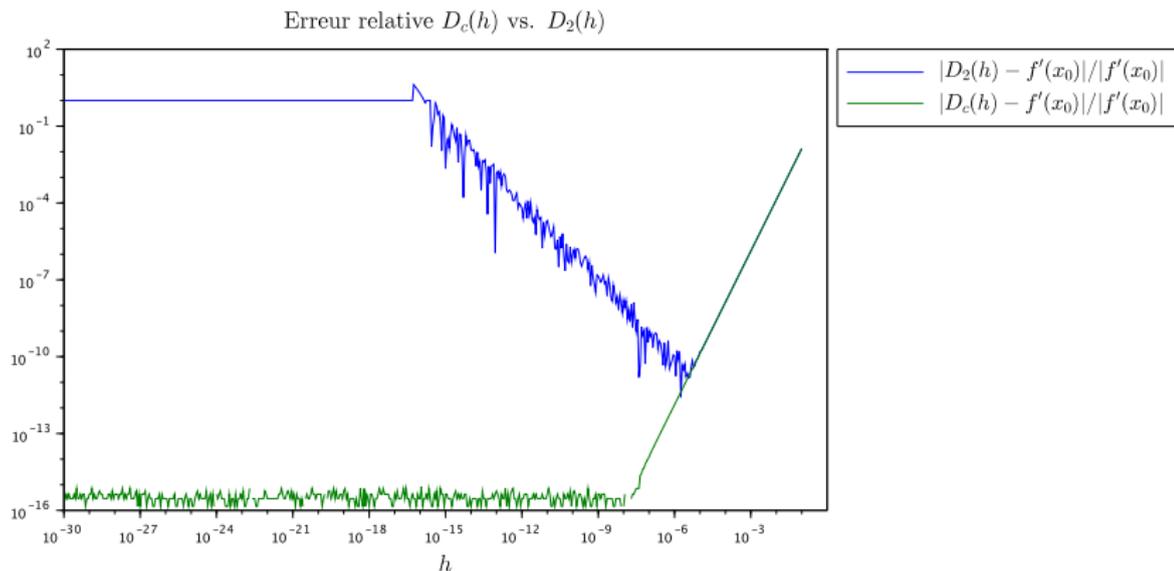
$$f'(x) = \exp(x) \frac{\cos^3 x + 3 \cos^2 x \sin x - 3 \cos x \sin^2 x + \sin^3 x}{(\cos^3 x + \sin^3 x)^2},$$

Approximation de  $f'(pi/4)$

### 3. Un pas de côté (dans le plan complexe)

Exemple

$$D_2(h) = \frac{f(x_0 + h) - f(x_0 - h)}{2h}, \quad D_c(h) = \frac{\text{Im}(f(x_0 + ih))}{h}$$



### 3. Un pas de côté (dans le plan complexe)

#### Perspectives

En prenant un minimum de précautions, il est possible de calculer, à la précision machine, les dérivées partielles d'une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  calculée par **n'importe quel programme Scilab**.

Où est cachée la dérivée dans les fonctions usuelles prolongées à  $\mathbb{C}$  ?

$$(x + iy)^2 = (x^2 - y^2 + 2ixy),$$
$$\operatorname{Im}((x + iy)^2) = 2xy$$

$$\cos(x + iy) = \cos x \cosh y - i \sin x \sinh y,$$
$$\operatorname{Im}(\cos(x + iy)) = -\sin x \sinh y$$

### 3. Un pas de côté (dans le plan complexe)

#### Perspectives

L'utilisation du pas complexe est équivalente à l'utilisation de l'algèbre des **nombre duals**.

Pour les dérivées d'ordre supérieur, on peut utiliser les **nombre hyper-duals**.

Toutes ces techniques sont liées à la **différentiation automatique**.

### 3. Un pas de côté (dans le plan complexe)

#### Références

- **Codage IEEE 754**

[https://fr.wikipedia.org/wiki/IEEE\\_754](https://fr.wikipedia.org/wiki/IEEE_754)

- **Pas complexe**

<https://hal.archives-ouvertes.fr/hal-01483287/document>

- **Nombres duals et hyper-duals**

<https://www.osti.gov/servlets/purl/1368722>

- **Différentiation automatique**

[https://en.wikipedia.org/wiki/Automatic\\_differentiation](https://en.wikipedia.org/wiki/Automatic_differentiation)