Mathématiques assistées par ordinateur

Chapitre 1: Introduction

Michael Eisermann

Mat249, DLST L2S4, Année 2008-2009 www-fourier.ujf-grenoble.fr/~eiserm/cours#mao Document mis à jour le 6 juillet 2009





Sommaire

- 1 Nombres entiers et rationnels, calcul exact
- Nombres réels, calcul approché, propagation d'erreurs
- 3 Types composés : expressions, fonctions, conteneurs, ...

Sommaire

- 1 Nombres entiers et rationnels, calcul exact
 - Numération positionnelle
 - Arithmétique des entiers
 - Nombres rationnels
- 2 Nombres réels, calcul approché, propagation d'erreurs
- 3 Types composés : expressions, fonctions, conteneurs, ...

```
DXXXVII 537
+LXXIX +79
DMXVI 616
```

| DXXXVII | 537 | MMVIII | 2008 | |
|---------|-----|---------|------|--|
| +LXXIX | +79 | -LXXIX | -79 | |
| DMXVI | 616 | MCMXXIX | 1929 | |

| DXXXVII | 537 | MMVIII | 2008 | XLIII | 43 |
|---------|-----|---------|------|------------|------|
| +LXXIX | +79 | -LXXIX | -79 | imesLXXIX | ×79 |
| DMXVI | 616 | MCMXXIX | 1929 | MMMXXXIIIX | 3397 |

Rappelons ce merveilleux outil qui est la numération positionnelle.

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 : 2008 $_{
m dec} =$ 2 \cdot 10 3 + 0 \cdot 10 2 + 0 \cdot 10 1 + 8 \cdot 10 0

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 : $\mathbf{2008}_{dec} = \mathbf{2} \cdot 10^3 + \mathbf{0} \cdot 10^2 + \mathbf{0} \cdot 10^1 + \mathbf{8} \cdot 10^0$

En base 2 : $1011_{\rm bin} = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 11_{\rm dec}$

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 :
$$\mathbf{2008}_{dec} = \mathbf{2} \cdot 10^3 + \mathbf{0} \cdot 10^2 + \mathbf{0} \cdot 10^1 + \mathbf{8} \cdot 10^0$$

En base 2 :
$$1011_{bin} = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 11_{dec}$$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 :
$$\mathbf{2008}_{dec} = \mathbf{2} \cdot 10^3 + \mathbf{0} \cdot 10^2 + \mathbf{0} \cdot 10^1 + \mathbf{8} \cdot 10^0$$

En base 2 :
$${f 1011}_{bin} = {f 1} \cdot 2^3 + {f 0} \cdot 2^2 + {f 1} \cdot 2^1 + {f 1} \cdot 2^0 = 11_{dec}$$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

$$a = 2 \cdot \text{quo} + \text{rem}$$

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 : 2008
$$_{
m dec} =$$
 2 \cdot 10 3 + 0 \cdot 10 2 + 0 \cdot 10 1 + 8 \cdot 10 0

En base 2 :
$${f 1011}_{\rm bin} = {f 1} \cdot 2^3 + {f 0} \cdot 2^2 + {f 1} \cdot 2^1 + {f 1} \cdot 2^0 = 11_{\rm dec}$$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

$$\frac{a = 2 \cdot \text{quo} + \text{rem}}{11 = 2 \cdot 5 + 1}$$

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 : 2008
$$_{
m dec} =$$
 2 \cdot 10 3 + 0 \cdot 10 2 + 0 \cdot 10 1 + 8 \cdot 10 0

En base 2 :
$${f 1011}_{bin} = {f 1} \cdot 2^3 + {f 0} \cdot 2^2 + {f 1} \cdot 2^1 + {f 1} \cdot 2^0 = 11_{dec}$$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

$$a = 2 \cdot \text{quo} + \text{rem}$$

$$11 = 2 \cdot 5 + 1$$

$$5 = 2 \cdot 2 + 1$$

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 :
$$\mathbf{2008}_{dec} = \mathbf{2} \cdot 10^3 + \mathbf{0} \cdot 10^2 + \mathbf{0} \cdot 10^1 + \mathbf{8} \cdot 10^0$$

En base 2 :
$${f 1011}_{\rm bin} = {f 1} \cdot 2^3 + {f 0} \cdot 2^2 + {f 1} \cdot 2^1 + {f 1} \cdot 2^0 = 11_{\rm dec}$$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 :
$$\mathbf{2008}_{dec} = \mathbf{2} \cdot 10^3 + \mathbf{0} \cdot 10^2 + \mathbf{0} \cdot 10^1 + \mathbf{8} \cdot 10^0$$

En base 2 :
$$1011_{bin} = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 11_{dec}$$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 : 2008 $_{
m dec}$ = $2\cdot 10^3$ + $0\cdot 10^2$ + $0\cdot 10^1$ + $8\cdot 10^0$

En base 2 : ${f 1011}_{\rm bin} = {f 1} \cdot 2^3 + {f 0} \cdot 2^2 + {f 1} \cdot 2^1 + {f 1} \cdot 2^0 = 11_{\rm dec}$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

Algorithme de conversion : On itère la division euclidienne.

$$\begin{array}{r}
 a = 2 \cdot \text{quo} + \text{rem} \\
 11 = 2 \cdot 5 + 1 \\
 5 = 2 \cdot 2 + 1 \\
 2 = 2 \cdot 1 + 0 \\
 1 = 2 \cdot 0 + 1
 \end{array}$$

Cet algorithme marche pour tout entier, en n'importe quelle base.

Rappelons ce merveilleux outil qui est la numération positionnelle.

En base 10 : 2008
$$_{
m dec} =$$
 2 \cdot 10 3 + 0 \cdot 10 2 + 0 \cdot 10 1 + 8 \cdot 10 0

En base 2 :
$${f 1011}_{\rm bin} = {f 1} \cdot 2^3 + {f 0} \cdot 2^2 + {f 1} \cdot 2^1 + {f 1} \cdot 2^0 = 11_{\rm dec}$$

Comment obtenir le développement en base 2?

$$\mathbf{1011}_{\mathrm{bin}} = \underbrace{(((\mathbf{1} \cdot 2 + \mathbf{0}) \cdot 2 + \mathbf{1}) \cdot 2}_{q \cdot 2} + \underbrace{\mathbf{1}}_{+r}$$

Algorithme de conversion : On itère la division euclidienne.

$$\begin{array}{r}
 a = 2 \cdot \text{quo} + \text{rem} \\
 11 = 2 \cdot 5 + 1 \\
 5 = 2 \cdot 2 + 1 \\
 2 = 2 \cdot 1 + 0 \\
 1 = 2 \cdot 0 + 1
 \end{array}$$

Cet algorithme marche pour tout entier, en n'importe quelle base. Le même algorithme nous servira plus tard pour les polynômes.

Proposition (division euclidienne des entiers)

Pour tout $a,b\in\mathbb{Z}$ où $b\neq 0$ il existe un unique couple $q,r\in\mathbb{Z}$ tel que

$$a = qb + r$$
 et $0 \le r < |b|$.

On appelle q =: a quo b le quotient et r =: a rem b le reste de la division euclidienne de a par b.

Proposition (division euclidienne des entiers)

Pour tout $a,b\in\mathbb{Z}$ où $b\neq 0$ il existe un unique couple $q,r\in\mathbb{Z}$ tel que

$$a = qb + r$$
 et $0 \le r < |b|$.

On appelle q =: a quo b le quotient et r =: a rem b le reste de la division euclidienne de a par b.

Démonstration. Exercice de révision!

Proposition (division euclidienne des entiers)

Pour tout $a,b\in\mathbb{Z}$ où $b\neq 0$ il existe un unique couple $q,r\in\mathbb{Z}$ tel que

$$a = qb + r$$
 et $0 \le r < |b|$.

On appelle q =: a quo b le quotient et r =: a rem b le reste de la division euclidienne de a par b.

Démonstration. Exercice de révision!

Corollaire

Soit $B \in \mathbb{Z}$ tel que $B \ge 2$. Tout entier $a \in \mathbb{Z}$ s'écrit de manière unique comme $a = \pm \sum_{k=0}^{\ell-1} a_k B^k$ tel que $a_k \in \{0, 1, \dots, B-1\}$ et $a_{\ell-1} \ne 0$.

Proposition (division euclidienne des entiers)

Pour tout $a,b\in\mathbb{Z}$ où $b\neq 0$ il existe un unique couple $q,r\in\mathbb{Z}$ tel que

$$a = qb + r$$
 et $0 \le r < |b|$.

On appelle q =: a quo b le quotient et r =: a rem b le reste de la division euclidienne de a par b.

Démonstration. Exercice de révision!

Corollaire

Soit $B \in \mathbb{Z}$ tel que $B \ge 2$. Tout entier $a \in \mathbb{Z}$ s'écrit de manière unique comme $a = \pm \sum_{k=0}^{\ell-1} a_k B^k$ tel que $a_k \in \{0, 1, \dots, B-1\}$ et $a_{\ell-1} \ne 0$.

Démonstration. Exercice de révision!

Proposition (division euclidienne des entiers)

Pour tout $a,b\in\mathbb{Z}$ où $b\neq 0$ il existe un unique couple $q,r\in\mathbb{Z}$ tel que

$$a = qb + r$$
 et $0 \le r < |b|$.

On appelle q =: a quo b le quotient et r =: a rem b le reste de la division euclidienne de a par b.

Démonstration. Exercice de révision!

Corollaire

Soit $B \in \mathbb{Z}$ tel que $B \ge 2$. Tout entier $a \in \mathbb{Z}$ s'écrit de manière unique comme $a = \pm \sum_{k=0}^{\ell-1} a_k B^k$ tel que $a_k \in \{0, 1, \dots, B-1\}$ et $a_{\ell-1} \ne 0$.

Démonstration. Exercice de révision!

On appelle $\ell \in \mathbb{N}$ la *longueur* du développement de a en base B.

Proposition (division euclidienne des entiers)

Pour tout $a,b\in\mathbb{Z}$ où $b\neq 0$ il existe un unique couple $q,r\in\mathbb{Z}$ tel que

$$a = qb + r$$
 et $0 \le r < |b|$.

On appelle q =: a quo b le quotient et r =: a rem b le reste de la division euclidienne de a par b.

Démonstration. Exercice de révision!

Corollaire

Soit $B \in \mathbb{Z}$ tel que $B \ge 2$. Tout entier $a \in \mathbb{Z}$ s'écrit de manière unique comme $a = \pm \sum_{k=0}^{\ell-1} a_k B^k$ tel que $a_k \in \{0, 1, \dots, B-1\}$ et $a_{\ell-1} \ne 0$.

Démonstration. Exercice de révision!

On appelle $\ell \in \mathbb{N}$ la *longueur* du développement de a en base B. On la note $\operatorname{len}_B(a) := \ell$; quand B = 2 on écrit $\operatorname{len}(a) := \operatorname{len}_2(a)$.

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| 01101011 | 01101011 |
|------------|------------|
| + 01001111 | - 01001111 |
| 10111010 | 00011100 |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|-------------------|---|-------------|-----|--|
| | min | | max | |
| 1 octet = 8 bits | 0 | $2^8 - 1 =$ | 255 | |
| 1 dotot = d bito | | | | |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|-------------------|---|------|-------------|-----|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^7 =$ | -128 | $2^7 - 1 =$ | 127 |
| | | | ! | |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|------|----------------|-------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^7 =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| | ı | | I | ' |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|--------|----------------|-------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^{7} =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| variante signée | $-2^{15} =$ | -32768 | $2^{15} - 1 =$ | 32767 |
| | | | • | |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|--------|----------------|------------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^{7} =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| variante signée | $-2^{15} =$ | -32768 | $2^{15} - 1 =$ | 32767 |
| 4 octets = 32 bits | | 0 | $2^{32} - 1 =$ | 4294967295 |
| | ' | | ı | |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|-------------|----------------|------------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^7 =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| variante signée | $-2^{15} =$ | -32768 | $2^{15} - 1 =$ | 32767 |
| 4 octets = 32 bits | | 0 | $2^{32} - 1 =$ | 4294967295 |
| variante signée | $-2^{31} =$ | -2147483648 | $2^{31} - 1 =$ | 2147483647 |
| | ' | | | ' |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|-------------|----------------|----------------------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^7 =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| variante signée | $-2^{15} =$ | -32768 | $2^{15} - 1 =$ | 32767 |
| 4 octets = 32 bits | | 0 | $2^{32} - 1 =$ | 4294967295 |
| variante signée | $-2^{31} =$ | -2147483648 | $2^{31} - 1 =$ | 2147483647 |
| 8 octets = 64 bits | | 0 | $2^{64} - 1 =$ | 18446744073709551615 |
| | ' | | ' | |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|----------------------|----------------|----------------------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^7 =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| variante signée | $-2^{15} =$ | -32768 | $2^{15} - 1 =$ | 32767 |
| 4 octets = 32 bits | | 0 | $2^{32} - 1 =$ | 4294967295 |
| variante signée | $-2^{31} =$ | -2147483648 | $2^{31} - 1 =$ | 2147483647 |
| 8 octets = 64 bits | | 0 | $2^{64} - 1 =$ | 18446744073709551615 |
| variante signée | $-2^{63} =$ | -9223372036854775808 | $2^{63} - 1 =$ | 9223372036854775807 |

On peut effectuer les opérations arithmétiques en base 2. Exemple :

Les microprocesseurs utilisent de tels entiers « machine » :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|----------------------|----------------|----------------------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^{7} =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| variante signée | $-2^{15} =$ | -32768 | $2^{15} - 1 =$ | 32767 |
| 4 octets = 32 bits | | 0 | $2^{32} - 1 =$ | 4294967295 |
| variante signée | $-2^{31} =$ | -2147483648 | $2^{31} - 1 =$ | 2147483647 |
| 8 octets = 64 bits | | 0 | $2^{64} - 1 =$ | 18446744073709551615 |
| variante signée | $-2^{63} =$ | -9223372036854775808 | $2^{63} - 1 =$ | 9223372036854775807 |

Avantage : les calculs sont très rapides.

On peut effectuer les opérations arithmétiques en base 2. Exemple :

Les microprocesseurs utilisent de tels entiers « machine » :

| taille en mémoire | plage des entiers qui peuvent être représentées | | | |
|--------------------|---|----------------------|----------------|----------------------|
| | | min | | max |
| 1 octet = 8 bits | | 0 | $2^8 - 1 =$ | 255 |
| variante signée | $-2^{7} =$ | -128 | $2^7 - 1 =$ | 127 |
| 2 octets = 16 bits | | 0 | $2^{16} - 1 =$ | 65535 |
| variante signée | $-2^{15} =$ | -32768 | $2^{15} - 1 =$ | 32767 |
| 4 octets = 32 bits | | 0 | $2^{32} - 1 =$ | 4294967295 |
| variante signée | $-2^{31} =$ | -2147483648 | $2^{31} - 1 =$ | 2147483647 |
| 8 octets = 64 bits | | 0 | $2^{64} - 1 =$ | 18446744073709551615 |
| variante signée | $-2^{63} =$ | -9223372036854775808 | $2^{63} - 1 =$ | 9223372036854775807 |

Avantage: les calculs sont très rapides.

Inconvénient : la taille est limitée et fixée d'avance.

Le processeur ne fournit que des entiers à 64 bits (au plus).

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué.

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué. Exemple :

 $2^{63} + 2^{63} = 0$ avec des entiers machine à 64 bits!

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué. Exemple :

 $2^{63} + 2^{63} = 0$ avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers?

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué. Exemple :

 $2^{63} + 2^{63} = 0$ avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers? Solution :

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué. Exemple :

$$2^{63} + 2^{63} = 0$$
 avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers ? Solution :

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

➤ Un grand entier est un tableau de petits entiers (chiffres).

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué. Exemple :

$$2^{63} + 2^{63} = 0$$
 avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers ? Solution :

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

- ➤ Un grand entier est un tableau de petits entiers (chiffres).
- \blacktriangleright Les opérations +, -, *, quo, rem s'implémente comme à l'école.

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué. Exemple :

$$2^{63} + 2^{63} = 0$$
 avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers? Solution :

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

- ➤ Un grand entier est un tableau de petits entiers (chiffres).
- ➤ Les opérations +, -, *, quo, rem s'implémente comme à l'école.

Tout logiciel de calcul formel implémente ainsi les grands entiers :

$$2^{63} + 2^{63} = 18446744073709551616$$

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0, \dots, 2^{64} - 1$ sera tronqué. Exemple :

$$2^{63} + 2^{63} = 0$$
 avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers? Solution:

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

- ➤ Un grand entier est un tableau de petits entiers (chiffres).
- ➤ Les opérations +, -, *, quo, rem s'implémente comme à l'école.

Tout logiciel de calcul formel implémente ainsi les grands entiers :

$$2^{63} + 2^{63} = 18446744073709551616$$

$$50! = 30414093201713378043612608166064768844377641568960512000000000000$$



Plus les entiers sont longs, plus les calculs sont coûteux.

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0,\dots,2^{64}-1$ sera tronqué. Exemple :

$$2^{63} + 2^{63} = 0$$
 avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers ? Solution :

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

- ➤ Un grand entier est un tableau de petits entiers (chiffres).
- ➤ Les opérations +, -, *, quo, rem s'implémente comme à l'école.

Tout logiciel de calcul formel implémente ainsi les grands entiers :

$$2^{63} + 2^{63} = 18446744073709551616$$

$$50! = 30414093201713378043612608166064768844377641568960512000000000000$$

- Plus les entiers sont longs, plus les calculs sont coûteux.
 - Addition et soustraction sont de complexité linéaire : $O(\ell)$.

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0, \ldots, 2^{64}-1$ sera tronqué. Exemple :

$$2^{63} + 2^{63} = 0$$
 avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers ? Solution :

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

- ➤ Un grand entier est un tableau de petits entiers (chiffres).
- \blacktriangleright Les opérations +, -, *, quo, rem s'implémente comme à l'école.

Tout logiciel de calcul formel implémente ainsi les grands entiers :

$$2^{63} + 2^{63} = 18446744073709551616$$

50! = 304140932017133780436126081660647688443776415689605120000000000000



Plus les entiers sont longs, plus les calculs sont coûteux.

- Addition et soustraction sont de complexité linéaire : $O(\ell)$.
- Multiplication et division sont de complexité quadratique : $O(\ell^2)$.

Le processeur ne fournit que des entiers à 64 bits (au plus). Tout calcul dépassant $0, \dots, 2^{64} - 1$ sera tronqué. Exemple :

$$2^{63} + 2^{63} = 0$$
 avec des entiers machine à 64 bits!

Comment calculer donc avec des grands entiers? Solution:

$$a = \sum_{k=0}^{\ell-1} a_k B^k$$
 où $B = 2^{32}$.

- ➤ Un grand entier est un tableau de petits entiers (chiffres).
- ➤ Les opérations +, -, *, quo, rem s'implémente comme à l'école.

Tout logiciel de calcul formel implémente ainsi les grands entiers :

$$2^{63} + 2^{63} = 18446744073709551616$$



- Plus les entiers sont longs, plus les calculs sont coûteux.
 - Addition et soustraction sont de complexité linéaire : $O(\ell)$.
 - Multiplication et division sont de complexité quadratique : $O(\ell^2)$.

Pour les grands entiers il existe des algorithmes sous-quadratiques (Karatsuba, FFT) donc plus efficaces que l'algorithme scolaire.

Tous les calculs numériques, aussi complexes qu'ils soient, sont basés sur les opérations élémentaires +, -, *, quo, rem des entiers.

Tous les calculs numériques, aussi complexes qu'ils soient, sont basés sur les opérations élémentaires +, -, *, quo, rem des entiers.

Voici un premier exemple célèbre :

Algorithme 1 pgcd de deux entiers selon Euclide

Entrée: deux entiers $a, b \in \mathbb{Z}$ **Sortie:** le pgcd positif de a et b

tant que $b \neq 0$ faire $r \leftarrow a \operatorname{rem} b, \quad a \leftarrow b, \quad b \leftarrow r$ fin tant que si a < 0 alors $a \leftarrow -a$

retourner a

Tous les calculs numériques, aussi complexes qu'ils soient, sont basés sur les opérations élémentaires $+,-,*,\mathrm{quo},\mathrm{rem}$ des entiers.

Voici un premier exemple célèbre :

Algorithme 1 pgcd de deux entiers selon Euclide

Entrée: deux entiers $a, b \in \mathbb{Z}$ **Sortie:** le pgcd positif de a et b

tant que $b \neq 0$ faire $r \leftarrow a \operatorname{rem} b, \ a \leftarrow b, \ b \leftarrow r$ fin tant que si a < 0 alors $a \leftarrow -a$ retourner a

Proposition

Pour tout $a,b \in \mathbb{Z}$ l'algorithme ci-dessus se termine et renvoie le pgcd positif de a et b. Il nécessite au plus $2 \operatorname{len}(a)$ itérations.

Tous les calculs numériques, aussi complexes qu'ils soient, sont basés sur les opérations élémentaires +, -, *, quo, rem des entiers.

Voici un premier exemple célèbre :

Algorithme 1 pgcd de deux entiers selon Euclide

Entrée: deux entiers $a, b \in \mathbb{Z}$ **Sortie:** le pgcd positif de a et b

```
tant que b \neq 0 faire r \leftarrow a \text{ rem } b, \ a \leftarrow b, \ b \leftarrow r fin tant que si a < 0 alors a \leftarrow -a retourner a
```

Proposition

Pour tout $a, b \in \mathbb{Z}$ l'algorithme ci-dessus se termine et renvoie le pacd positif de a et b. Il nécessite au plus $2 \operatorname{len}(a)$ itérations.

Démonstration. Exercice de révision!

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

Tout nombre rationnel $x \in \mathbb{Q}$ s'écrit comme $x = \frac{a}{b}$ avec $a, b \in \mathbb{Z}, b \neq 0$.

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

Tout nombre rationnel $x\in\mathbb{Q}$ s'écrit comme $x=\frac{a}{b}$ avec $a,b\in\mathbb{Z},\,b\neq0.$

On réduit chaque fraction de sorte que pgcd(a, b) = 1 et b > 0.

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

Tout nombre rationnel $x \in \mathbb{Q}$ s'écrit comme $x = \frac{a}{b}$ avec $a, b \in \mathbb{Z}, b \neq 0$.

On réduit chaque fraction de sorte que pgcd(a, b) = 1 et b > 0.

ightharpoonup Pour représenter x on stocke simplement a et b.

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

Tout nombre rationnel $x \in \mathbb{Q}$ s'écrit comme $x = \frac{a}{b}$ avec $a, b \in \mathbb{Z}, b \neq 0$.

On réduit chaque fraction de sorte que pgcd(a, b) = 1 et b > 0.

- ightharpoonup Pour représenter x on stocke simplement a et b.
- ➤ On implémente les opérations arithmétiques comme suit :

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad \frac{a}{b} / \frac{c}{d} = \frac{ad}{bc},$$

bien sûr suivi de la réduction des fractions (calcul du pgcd dans $\ensuremath{\mathbb{Z}}).$

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

Tout nombre rationnel $x \in \mathbb{Q}$ s'écrit comme $x = \frac{a}{b}$ avec $a, b \in \mathbb{Z}, b \neq 0$.

On réduit chaque fraction de sorte que pgcd(a, b) = 1 et b > 0.

- ightharpoonup Pour représenter x on stocke simplement a et b.
- ➤ On implémente les opérations arithmétiques comme suit :

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad \frac{a}{b} / \frac{c}{d} = \frac{ad}{bc},$$

bien sûr suivi de la réduction des fractions (calcul du pgcd dans \mathbb{Z}).

Tous les calculs numériques, aussi complexes qu'ils soient, sont basés sur les opérations élémentaires +, -, *, quo, rem des entiers.

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

Tout nombre rationnel $x \in \mathbb{Q}$ s'écrit comme $x = \frac{a}{b}$ avec $a, b \in \mathbb{Z}, b \neq 0$.

On réduit chaque fraction de sorte que pgcd(a, b) = 1 et b > 0.

- ightharpoonup Pour représenter x on stocke simplement a et b.
- ➤ On implémente les opérations arithmétiques comme suit :

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad \frac{a}{b} / \frac{c}{d} = \frac{ad}{bc},$$

bien sûr suivi de la réduction des fractions (calcul du pgcd dans \mathbb{Z}).

Tous les calculs numériques, aussi complexes qu'ils soient, sont basés sur les opérations élémentaires +, -, *, quo, rem des entiers.

Conclusion

Les calculs avec les rationnels sont exacts et a priori sans limitation.

On sait représenter les entiers $a,b\in\mathbb{Z}$ et effectuer leurs opérations.

Tout nombre rationnel $x \in \mathbb{Q}$ s'écrit comme $x = \frac{a}{b}$ avec $a, b \in \mathbb{Z}$, $b \neq 0$.

On réduit chaque fraction de sorte que pgcd(a, b) = 1 et b > 0.

- \blacktriangleright Pour représenter x on stocke simplement a et b.
- ➤ On implémente les opérations arithmétiques comme suit :

$$\frac{a}{b}\pm\frac{c}{d}=\frac{ad\pm bc}{bd},\qquad \frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd},\qquad \frac{a}{b}/\frac{c}{d}=\frac{ad}{bc},$$

bien sûr suivi de la réduction des fractions (calcul du pgcd dans \mathbb{Z}).

Tous les calculs numériques, aussi complexes qu'ils soient, sont basés sur les opérations élémentaires +, -, *, quo, rem des entiers.

Conclusion

Les calculs avec les rationnels sont exacts et a priori sans limitation. Les seuls facteurs limitant sont le temps et la mémoire disponibles.

Sommaire

- 1 Nombres entiers et rationnels, calcul exac
- 2 Nombres réels, calcul approché, propagation d'erreurs
 - La problématique illustrée par l'exemple $\sqrt{2}$
 - Nombres à virgule flottante et calcul approché
 - Propagation d'erreurs : la prudence s'impose!
- 3 Types composés : expressions, fonctions, conteneurs, ...

Un défaut des nombres rationnels

Proposition

Il n'existe pas de nombre rationnel $r \in \mathbb{Q}$ tel que $r^2 = 2$.

Un défaut des nombres rationnels

Proposition

Il n'existe pas de nombre rationnel $r \in \mathbb{Q}$ tel que $r^2 = 2$.



Les nombres rationnels ne sont pas « complets ».

Un défaut des nombres rationnels

Proposition

If n'existe pas de nombre rationnel $r \in \mathbb{Q}$ tel que $r^2 = 2$.



Les nombres rationnels ne sont pas « complets ».

Ainsi ils ne suffisent pas pour l'analyse (limites, dérivées, intégrales).

Existe-t-il des racines, en fait?

Existe-t-il des racines, en fait?

La réponse dépend d'une construction subtile : les nombres réels $\mathbb{R}\,!$

Existe-t-il des racines, en fait?

La réponse dépend d'une construction subtile : les nombres réels $\mathbb{R}\,!$

On note $[a,b]=\{x\in\mathbb{R}\mid a\leq x\leq b\}$ l'intervalle réel entre a et b.

Existe-t-il des racines, en fait?

La réponse dépend d'une construction subtile : les nombres réels \mathbb{R} ! On note $[a,b]=\{x\in\mathbb{R}\mid a\leq x\leq b\}$ l'intervalle réel entre a et b.

Théorème (des valeurs intermédiaires, TVI)

Soit $f:[a,b] \to \mathbb{R}$ une fonction continue.

Existe-t-il des racines, en fait?

La réponse dépend d'une construction subtile : les nombres réels \mathbb{R} ! On note $[a,b]=\{x\in\mathbb{R}\mid a\leq x\leq b\}$ l'intervalle réel entre a et b.

Théorème (des valeurs intermédiaires, TVI)

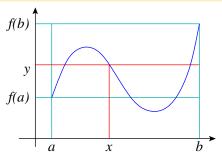
Soit $f \colon [a,b] \to \mathbb{R}$ une fonction continue. Pour tout réel y vérifiant f(a) < y < f(b) il existe (au moins) un $x \in [a,b]$ tel que f(x) = y.

Existe-t-il des racines, en fait?

La réponse dépend d'une construction subtile : les nombres réels \mathbb{R} ! On note $[a,b]=\{x\in\mathbb{R}\mid a\leq x\leq b\}$ l'intervalle réel entre a et b.

Théorème (des valeurs intermédiaires, TVI)

Soit $f: [a,b] \to \mathbb{R}$ une fonction continue. Pour tout réel y vérifiant f(a) < y < f(b) il existe (au moins) un $x \in [a,b]$ tel que f(x) = y.

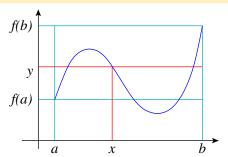


Existe-t-il des racines, en fait?

La réponse dépend d'une construction subtile : les nombres réels $\mathbb R$! On note $[a,b]=\{x\in\mathbb R\mid a\leq x\leq b\}$ l'intervalle réel entre a et b.

Théorème (des valeurs intermédiaires, TVI)

Soit $f: [a,b] \to \mathbb{R}$ une fonction continue. Pour tout réel y vérifiant f(a) < y < f(b) il existe (au moins) un $x \in [a,b]$ tel que f(x) = y.



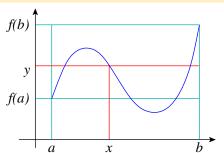
Ce résultat est caractéristique pour les nombres réels \mathbb{R} .

Existe-t-il des racines, en fait?

La réponse dépend d'une construction subtile : les nombres réels \mathbb{R} ! On note $[a,b]=\{x\in\mathbb{R}\mid a\leq x\leq b\}$ l'intervalle réel entre a et b.

Théorème (des valeurs intermédiaires, TVI)

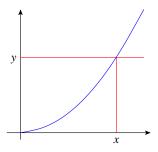
Soit $f \colon [a,b] \to \mathbb{R}$ une fonction continue. Pour tout réel y vérifiant f(a) < y < f(b) il existe (au moins) un $x \in [a,b]$ tel que f(x) = y. \square



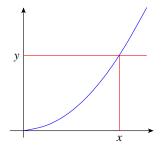
Ce résultat est caractéristique pour les nombres réels \mathbb{R} .

Il serait faux pour les rationnels, par exemple $f: \mathbb{Q} \to \mathbb{Q}$, $f(x) = x^2$.

Considérons la fonction $f \colon \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^n$.

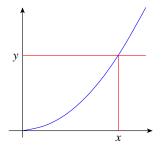


Considérons la fonction $f \colon \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^n$.



Elle est continue et strictement croissante.

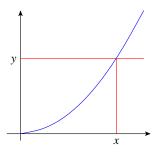
Considérons la fonction $f \colon \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^n$.



Elle est continue et strictement croissante.

Elle vérifie f(0) = 0 et $f(x) \to +\infty$ pour $x \to +\infty$.

Considérons la fonction $f \colon \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^n$.



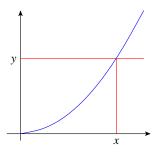
Elle est continue et strictement croissante.

Elle vérifie f(0) = 0 et $f(x) \to +\infty$ pour $x \to +\infty$.

Corollaire (existence des racines réelles)

Pour tout réel $y \in \mathbb{R}_{>0}$ il existe un unique réel $x \in \mathbb{R}_{>0}$ tel que $x^n = y$.

Considérons la fonction $f \colon \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^n$.



Elle est continue et strictement croissante.

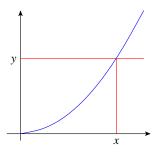
Elle vérifie f(0) = 0 et $f(x) \to +\infty$ pour $x \to +\infty$.

Corollaire (existence des racines réelles)

Pour tout réel $y \in \mathbb{R}_{\geq 0}$ il existe un unique réel $x \in \mathbb{R}_{\geq 0}$ tel que $x^n = y$.

On le note $x =: \sqrt[n]{y}$, ou bien $x =: \sqrt{y}$ dans le cas n = 2.

Considérons la fonction $f : \mathbb{R}_+ \to \mathbb{R}_+$ définie par $f(x) = x^n$.



Elle est continue et strictement croissante.

Elle vérifie f(0) = 0 et $f(x) \to +\infty$ pour $x \to +\infty$.

Corollaire (existence des racines réelles)

Pour tout réel $y \in \mathbb{R}_{>0}$ il existe un unique réel $x \in \mathbb{R}_{>0}$ tel que $x^n = y$.

On le note $x =: \sqrt[n]{y}$, ou bien $x =: \sqrt{y}$ dans le cas n = 2.

On a vu que $\sqrt{2}$ n'est pas rationnel. Donc $\sqrt{2} \in \mathbb{R}$ mais $\sqrt{2} \notin \mathbb{Q}$.

L'ordinateur peut effectuer des calculs exacts dans $\mathbb Z$ et dans $\mathbb Q.$

L'ordinateur peut effectuer des calculs exacts dans \mathbb{Z} et dans \mathbb{Q} . Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles.

L'ordinateur peut effectuer des calculs exacts dans \mathbb{Z} et dans \mathbb{Q} . Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles.

Comment effectuer de tels calculs sur ordinateur?

L'ordinateur peut effectuer des calculs exacts dans $\ensuremath{\mathbb{Z}}$ et dans $\ensuremath{\mathbb{Q}}.$

Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles.

Comment effectuer de tels calculs sur ordinateur?

Deux méthodes sont possibles :

L'ordinateur peut effectuer des calculs exacts dans $\mathbb Z$ et dans $\mathbb Q$. Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles. Comment effectuer de tels calculs sur ordinateur ?

Deux méthodes sont possibles :

■ Calcul formel : on manipule formellement des symboles et des formules, toujours exacts. Exemple :

```
entrée r:= sqrt(2) \implies sortie \sqrt{2}
entrée r^2-2 \implies sortie 0
```

L'ordinateur peut effectuer des calculs exacts dans $\mathbb Z$ et dans $\mathbb Q$. Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles. Comment effectuer de tels calculs sur ordinateur?

Deux méthodes sont possibles :

■ Calcul formel : on manipule formellement des symboles et des formules, toujours exacts. Exemple :

```
entrée r := sqrt(2) \implies sortie \sqrt{2}
entrée r^2-2 \implies sortie 0
```

■ Calcul approché : on approche tout nombre réel par un développement (binaire) fini, typiquement inexact. Exemple :

```
entrée a:= evalf(r) \Longrightarrow sortie 1.41421356237
entrée a^2-2 \Longrightarrow sortie 4.4408920985e-16
```

L'ordinateur peut effectuer des calculs exacts dans $\mathbb Z$ et dans $\mathbb Q$. Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles. Comment effectuer de tels calculs sur ordinateur?

Deux méthodes sont possibles :

■ Calcul formel : on manipule formellement des symboles et des formules, toujours exacts. Exemple :

```
entrée r:= sqrt(2) \implies sortie \sqrt{2}
entrée r^2-2 \implies sortie 0
```

■ Calcul approché : on approche tout nombre réel par un développement (binaire) fini, typiquement inexact. Exemple :

```
entrée a:= evalf(r) \Longrightarrow sortie 1.41421356237
entrée a^2-2 \Longrightarrow sortie 4.4408920985e-16
```



Chaque méthode a ses avantages et ses inconvénients.

L'ordinateur peut effectuer des calculs exacts dans $\mathbb Z$ et dans $\mathbb Q$. Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles. Comment effectuer de tels calculs sur ordinateur?

Deux méthodes sont possibles :

■ Calcul formel : on manipule formellement des symboles et des formules, toujours exacts. Exemple :

```
entrée r := sqrt(2) \implies sortie \sqrt{2}
entrée r^2-2 \implies sortie 0
```

■ Calcul approché : on approche tout nombre réel par un développement (binaire) fini, typiquement inexact. Exemple :

```
entrée a:= evalf(r) \implies sortie 1.41421356237
entrée a^2-2 \implies sortie 4.4408920985e-16
```

⚠ Chaque méthode a ses avantages et ses inconvénients.
Selon l'application que vous envisagez, c'est à vous de choisir.

L'ordinateur peut effectuer des calculs exacts dans \mathbb{Z} et dans \mathbb{Q} . Problème : certaines valeurs comme $\sqrt{2}$ ne sont pas rationnelles.

Comment effectuer de tels calculs sur ordinateur?

Deux méthodes sont possibles :

■ Calcul formel : on manipule formellement des symboles et des formules, toujours exacts. Exemple :

```
entrée r := sqrt(2) \implies sortie \sqrt{2}
entrée r^2-2 \implies sortie 0
```

■ Calcul approché : on approche tout nombre réel par un développement (binaire) fini, typiquement inexact. Exemple :

```
entrée a:= evalf(r) \Longrightarrow sortie 1.41421356237
entrée a^2-2 \Longrightarrow sortie 4.4408920985e-16
```

⚠ Chaque méthode a ses avantages et ses inconvénients. Selon l'application que vous envisagez, c'est à vous de choisir.

Pour faire un choix informé, il faut comprendre les calculs!

Comment écrire une approximation de $r := \sqrt{2}$?

Comment écrire une approximation de $r := \sqrt{2}$?

Faux — N'écrivez jamais $\sqrt{2} = 1.41421356$: c'est faux parce que $1.41421356^2 = 1.999999932878736 < 2$

Comment écrire une approximation de $r := \sqrt{2}$?

Faux — N'écrivez jamais $\sqrt{2} = 1.41421356$: c'est faux parce que

 $1.41421356^2 = 1.9999999932878736 < 2$

Très flou — $r\approx 1.41421356$ n'est pas faux, mais $r\approx 5$ n'est pas faux non plus. Problème : Ce qui compte est la marge d'erreur.

Comment écrire une approximation de $r := \sqrt{2}$?

Faux — N'écrivez jamais $\sqrt{2} = 1.41421356$: c'est faux parce que

$$1.41421356^2 = 1.9999999932878736 < 2$$

Très flou — $r\approx 1.41421356$ n'est pas faux, mais $r\approx 5$ n'est pas faux non plus. Problème : Ce qui compte est la marge d'erreur.

Moins flou — r=1.41421356... suggère que tous les chiffres indiqués sont corrects. (Vraiment?)

Comment écrire une approximation de $r := \sqrt{2}$?

Faux — N'écrivez jamais $\sqrt{2} = 1.41421356$: c'est faux parce que

$$1.41421356^2 = 1.9999999932878736 < 2$$

Très flou — $r\approx 1.41421356$ n'est pas faux, mais $r\approx 5$ n'est pas faux non plus. Problème : Ce qui compte est la marge d'erreur.

Moins flou — r=1.41421356... suggère que tous les chiffres indiqués sont corrects. (Vraiment?)

La bonne notation — Encadrement 1.41421356 < r < 1.41421357 ou approximation avec marge d'erreur : $|r - 1.41421356| \le 10^{-8}$.

Comment écrire une approximation de $r := \sqrt{2}$?

Faux — N'écrivez jamais $\sqrt{2} = 1.41421356$: c'est faux parce que $1.41421356^2 = 1.999999932878736 < 2$

Très flou — $r\approx 1.41421356$ n'est pas faux, mais $r\approx 5$ n'est pas faux non plus. Problème : Ce qui compte est la marge d'erreur.

Moins flou — $r=1.41421356\dots$ suggère que tous les chiffres indiqués sont corrects. (Vraiment?)

La bonne notation — Encadrement 1.41421356 < r < 1.41421357 ou approximation avec marge d'erreur : $|r-1.41421356| \le 10^{-8}$.

Cette écriture est honnête et lisible. En particulier c'est un énoncé précis que l'on peut vérifier. (En l'occurrence il s'avère correct.)

Comment écrire une approximation de $r := \sqrt{2}$?

Faux — N'écrivez jamais $\sqrt{2} = 1.41421356$: c'est faux parce que

$$1.41421356^2 = 1.9999999932878736 < 2$$

Très flou — $r\approx 1.41421356$ n'est pas faux, mais $r\approx 5$ n'est pas faux non plus. Problème : Ce qui compte est la marge d'erreur.

Moins flou — r = 1.41421356... suggère que tous les chiffres indiqués sont corrects. (Vraiment?)

La bonne notation — Encadrement 1.41421356 < r < 1.41421357 ou approximation avec marge d'erreur : $|r-1.41421356| \le 10^{-8}$.

Cette écriture est honnête et lisible. En particulier c'est un énoncé précis que l'on peut vérifier. (En l'occurrence il s'avère correct.)

Règle de bon sens

Avec n décimales on ne peut approcher qu'à 10^{-n} près.

Comment écrire une approximation de $r := \sqrt{2}$?

Faux — N'écrivez jamais $\sqrt{2}=1.41421356$: c'est faux parce que

$$1.41421356^2 = 1.9999999932878736 < 2$$

Très flou — $r\approx 1.41421356$ n'est pas faux, mais $r\approx 5$ n'est pas faux non plus. Problème : Ce qui compte est la marge d'erreur.

Moins flou — r = 1.41421356... suggère que tous les chiffres indiqués sont corrects. (Vraiment?)

La bonne notation — Encadrement 1.41421356 < r < 1.41421357 ou approximation avec marge d'erreur : $|r - 1.41421356| \le 10^{-8}$.

Cette écriture est honnête et lisible. En particulier c'est un énoncé précis que l'on peut vérifier. (En l'occurrence il s'avère correct.)

Règle de bon sens

Avec n décimales on ne peut approcher qu'à 10^{-n} près.

Réciproquement, si la précision est d'ordre de 10^{-n} , il est inutile, voire trompeur, de rajouter des chiffres non significatifs.

Tout réel admet un développement décimal ou binaire :

Tout réel admet un développement décimal ou binaire : $\frac{1}{3}=0.3333333333..._{\rm dec}=0.010101010101011..._{\rm bin}$ (périodique)

Tout réel admet un développement décimal ou binaire :

```
\frac{1}{3}=0.33333333333\ldots_{\rm dec}=\ 0.0101010101010101\ldots_{\rm bin} (périodique) \pi=3.1415926535\ldots_{\rm dec}=11.0010010000111111\ldots_{\rm bin} (non périodique)
```

Tout réel admet un développement décimal ou binaire :

```
\frac{1}{3}=0.3333333333\ldots_{\rm dec}=0.0101010101010101\ldots_{\rm bin} (périodique) \pi=3.1415926535\ldots_{\rm dec}=11.0010010000111111\ldots_{\rm bin} (non périodique)
```

Développement en base B

Soit $B \in \mathbb{Z}$ tel que $B \ge 2$. Tout nombre réel $x \in \mathbb{R}$ peut être développé en une série $x = \pm \sum_{k=-n}^{\infty} a_k B^{-k}$ de sorte que $a_k \in \{0, 1, \dots, B-1\}$.

Tout réel admet un développement décimal ou binaire :

```
\frac{1}{3}=0.3333333333\ldots_{\rm dec}=0.0101010101010101\ldots_{\rm bin} (périodique) \pi=3.1415926535\ldots_{\rm dec}=11.0010010000111111\ldots_{\rm bin} (non périodique)
```

Développement en base B

Soit $B \in \mathbb{Z}$ tel que $B \ge 2$. Tout nombre réel $x \in \mathbb{R}$ peut être développé en une série $x = \pm \sum_{k=-n}^{\infty} a_k B^{-k}$ de sorte que $a_k \in \{0, 1, \dots, B-1\}$.



Aucun ordinateur ne peut implémenter les nombres réels \mathbb{R} !

Tout réel admet un développement décimal ou binaire :

```
\frac{1}{3} = 0.3333333333..._{\text{dec}} = 0.0101010101010101..._{\text{bin}} (périodique)
\pi = 3.1415926535..._{\rm dec} = 11.0010010000111111..._{\rm bin} (non périodique)
```

Développement en base B

Soit $B \in \mathbb{Z}$ tel que $B \geq 2$. Tout nombre réel $x \in \mathbb{R}$ peut être développé en une série $x = \pm \sum_{k=-n}^{\infty} a_k B^{-k}$ de sorte que $a_k \in \{0, 1, \dots, B-1\}$.



 $ilde{m{ \wedge}}$ Aucun ordinateur ne peut implémenter les nombres réels $\mathbb{R}\,!$

Sur ordinateur on ne peut stocker que le début du développement :

$$\hat{x} = a_{-n}B^n + \dots + a_0B^0 + a_1B^{-1} + \dots + a_mB^{-m}$$

Tout réel admet un développement décimal ou binaire :

```
\frac{1}{3} = 0.3333333333..._{\text{dec}} = 0.0101010101010101..._{\text{bin}} (périodique)
\pi = 3.1415926535..._{\rm dec} = 11.0010010000111111..._{\rm bin} (non périodique)
```

Développement en base B

Soit $B \in \mathbb{Z}$ tel que $B \geq 2$. Tout nombre réel $x \in \mathbb{R}$ peut être développé en une série $x = \pm \sum_{k=-n}^{\infty} a_k B^{-k}$ de sorte que $a_k \in \{0, 1, \dots, B-1\}$.



igwedge Aucun ordinateur ne peut implémenter les nombres réels $\mathbb{R}\,!$

Sur ordinateur on ne peut stocker que le début du développement :

$$\hat{x} = a_{-n}B^n + \dots + a_0B^0 + a_1B^{-1} + \dots + a_mB^{-m}$$

Évidemment, une certaine erreur d'arrondi ($\leq B^{-m}$) est inévitable.

Tout réel admet un développement décimal ou binaire :

```
\frac{1}{3} = 0.33333333333..._{\text{dec}} = 0.0101010101010101..._{\text{bin}} (périodique)
\pi = 3.1415926535..._{\rm dec} = 11.0010010000111111..._{\rm bin} (non périodique)
```

Développement en base B

Soit $B \in \mathbb{Z}$ tel que $B \geq 2$. Tout nombre réel $x \in \mathbb{R}$ peut être développé en une série $x = \pm \sum_{k=-n}^{\infty} a_k B^{-k}$ de sorte que $a_k \in \{0, 1, \dots, B-1\}$.



igwedge Aucun ordinateur ne peut implémenter les nombres réels $\mathbb{R}\,!$

Sur ordinateur on ne peut stocker que le début du développement :

$$\hat{x} = a_{-n}B^n + \dots + a_0B^0 + a_1B^{-1} + \dots + a_mB^{-m}$$

Évidemment, une certaine erreur d'arrondi ($\leq B^{-m}$) est inévitable.

Réconfort : on peut très bien travailler avec des approximations. . . mais il faut les manipuler intelligemment puis interpréter les résultats.

Les nombres à virgule fixe ont un grave défaut pratique :

Les nombres à virgule fixe ont un grave défaut pratique : ils sont inefficaces pour des nombres très grands comme

Les nombres à virgule fixe ont un grave défaut pratique : ils sont inefficaces pour des nombres très grands comme

ou des nombres très petits comme

h = 0.00000000000000000000000000000000626.

Les nombres à virgule fixe ont un grave défaut pratique : ils sont inefficaces pour des nombres très grands comme

ou des nombres très petits comme

$$h = 0.00000000000000000000000000000000626.$$

Mieux adaptés sont les nombres à virgule flottante :

$$N = 6.022 \cdot 10^{23}, \qquad h = 6.626 \cdot 10^{-34}.$$

Les nombres à virgule fixe ont un grave défaut pratique : ils sont inefficaces pour des nombres très grands comme

ou des nombres très petits comme

Mieux adaptés sont les nombres à virgule flottante :

$$N = 6.022 \cdot 10^{23}, \qquad h = 6.626 \cdot 10^{-34}.$$

Le principe est de décaler la virgule en compensant par l'exposant :

$$123.456 = 1234.56 \cdot 10^{-1} = 12345.6 \cdot 10^{-2} = 123456 \cdot 10^{-3}$$

Rendons la virgule flottante!

Les nombres à virgule fixe ont un grave défaut pratique : ils sont inefficaces pour des nombres très grands comme

ou des nombres très petits comme

Mieux adaptés sont les nombres à virgule flottante :

$$N = 6.022 \cdot 10^{23}, \qquad h = 6.626 \cdot 10^{-34}.$$

Le principe est de décaler la virgule en compensant par l'exposant :

$$123.456 = 1234.56 \cdot 10^{-1} = 12345.6 \cdot 10^{-2} = 123456 \cdot 10^{-3}$$

Notation dite « scientifique » utilisée sur ordinateur

En base 10 nous écrivons 1.2345e-67 pour $1.2345 \cdot 10^{-67}$.



 \bigwedge À ne pas confondre avec $1.2345e^{-67} = 1.2345 \cdot \exp(-67)$.

Exemples de nombres flottants : $N=6.022\cdot 10^{23}$, $h=6.626\cdot 10^{-34}$.

Exemples de nombres flottants : $N=6.022\cdot 10^{23}$, $h=6.626\cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

Exemples de nombres flottants : $N = 6.022 \cdot 10^{23}$, $h = 6.626 \cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

■ $m \in \mathbb{Z}$ est appelé la *mantisse*,

Exemples de nombres flottants : $N=6.022\cdot 10^{23}$, $h=6.626\cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- lacksquare $m \in \mathbb{Z}$ est appelé la *mantisse*,
- \blacksquare B=2 ou B=10 est la *base*, et

Exemples de nombres flottants : $N=6.022\cdot 10^{23}$, $h=6.626\cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- $m \in \mathbb{Z}$ est appelé la *mantisse*,
- \blacksquare B=2 ou B=10 est la *base*, et
- \bullet $e \in \mathbb{Z}$ est *l'exposant*.

Exemples de nombres flottants : $N = 6.022 \cdot 10^{23}$, $h = 6.626 \cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- $m \in \mathbb{Z}$ est appelé la *mantisse*,
- \blacksquare B=2 ou B=10 est la *base*, et
- \bullet $e \in \mathbb{Z}$ est *l'exposant*.

Calcul arrondi : on arrondit la mantisse à une longueur fixée ℓ .

Exemples de nombres flottants : $N = 6.022 \cdot 10^{23}$, $h = 6.626 \cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- $m \in \mathbb{Z}$ est appelé la *mantisse*,
- \blacksquare B=2 ou B=10 est la *base*, et
- \bullet $e \in \mathbb{Z}$ est *l'exposant*.

Calcul arrondi : on arrondit la mantisse à une longueur fixée ℓ . On veut donc restreindre la mantisse à $m \in \{1 - B^{\ell}, \dots, B^{\ell} - 1\}$.

Exemples de nombres flottants : $N = 6.022 \cdot 10^{23}$, $h = 6.626 \cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- $m \in \mathbb{Z}$ est appelé la *mantisse*,
- \blacksquare B=2 ou B=10 est la *base*, et
- \bullet $e \in \mathbb{Z}$ est *l'exposant*.

Calcul arrondi : on arrondit la mantisse à une longueur fixée ℓ . On veut donc restreindre la mantisse à $m \in \{1-B^\ell, \dots, B^\ell-1\}$. Pour ce faire on supprime les chiffres les moins significatifs :

Exemples de nombres flottants : $N = 6.022 \cdot 10^{23}$, $h = 6.626 \cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- $m \in \mathbb{Z}$ est appelé la *mantisse*,
- \blacksquare B=2 ou B=10 est la *base*, et
- \bullet $e \in \mathbb{Z}$ est *l'exposant*.

Calcul arrondi: on arrondit la mantisse à une longueur fixée ℓ . On veut donc restreindre la mantisse à $m \in \{1-B^\ell, \dots, B^\ell-1\}$. Pour ce faire on supprime les chiffres les moins significatifs : Si $\operatorname{len}_B(m) = \ell + s$, alors $m \leftarrow m \operatorname{quo} B^s(\pm 1)$ et $e \leftarrow e + s$.

Exemples de nombres flottants : $N = 6.022 \cdot 10^{23}$, $h = 6.626 \cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- $m \in \mathbb{Z}$ est appelé la *mantisse*,
- \blacksquare B=2 ou B=10 est la *base*, et
- \bullet $e \in \mathbb{Z}$ est *l'exposant*.

Calcul arrondi : on arrondit la mantisse à une longueur fixée ℓ . On veut donc restreindre la mantisse à $m \in \{1-B^\ell, \dots, B^\ell-1\}$. Pour ce faire on supprime les chiffres les moins significatifs : Si $\operatorname{len}_B(m) = \ell + s$, alors $m \leftarrow m \operatorname{quo} B^s(\pm 1)$ et $e \leftarrow e + s$.

Exemple : le type double calcule en binaire (B=2). Il prévoit $\ell=53$ bits pour la mantisse, 11 bits pour l'exposant, 1 bit pour le signe.

Exemples de nombres flottants : $N = 6.022 \cdot 10^{23}$, $h = 6.626 \cdot 10^{-34}$.

En général, les nombres à virgule flottante sont de la forme $m \cdot B^e$ où

- $m \in \mathbb{Z}$ est appelé la *mantisse*,
- $\blacksquare B = 2$ ou B = 10 est la *base*, et
- \bullet $e \in \mathbb{Z}$ est *l'exposant*.

Calcul arrondi : on arrondit la mantisse à une longueur fixée ℓ . On veut donc restreindre la mantisse à $m \in \{1-B^\ell, \dots, B^\ell-1\}$. Pour ce faire on supprime les chiffres les moins significatifs : Si $\operatorname{len}_B(m) = \ell + s$, alors $m \leftarrow m \operatorname{quo} B^s(\pm 1)$ et $e \leftarrow e + s$.

Exemple : le type double calcule en binaire (B=2). Il prévoit $\ell=53$ bits pour la mantisse, 11 bits pour l'exposant, 1 bit pour le signe.

Erreur d'arrondi : avec mantisse de longueur ℓ on peut approcher tout nombre réel à une erreur relative de $B^{1-\ell}$ près. ($2^{-53}\approx 10^{-16}$)

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Multiplions $N=6.022\cdot 10^{23}$ et $h=6.626\cdot 10^{-34}\,!$ Résultat exact : $Nh=39.901772\cdot 10^{-11}=3.9901772\cdot 10^{-10}.$

$$\begin{split} & \text{Multiplions} \quad N = 6.022 \cdot 10^{23} \quad \text{et} \quad h = 6.626 \cdot 10^{-34} \,! \\ & \text{Résultat exact}: \quad Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}. \\ & \text{On l'arrondit à} \quad Nh \approx 3.990 \cdot 10^{-10} \quad \text{(mantisse à 4 décimales)}. \end{split}$$

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x=m_1B^{e_1}$ et $y=m_2B^{e_2}$ les opérations $x+y,\,x-y,\,x*y,\,x/y$ sont définies comme suit :

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x=m_1B^{e_1}$ et $y=m_2B^{e_2}$ les opérations $x+y,\,x-y,\,x*y,\,x/y$ sont définies comme suit : Partant du résultat exact $z=mB^e$ on arrondit à $\hat{z}=\hat{m}B^{\hat{e}}$ en réduisant la mantisse à la longueur prescrite (de ℓ chiffres).

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x=m_1B^{e_1}$ et $y=m_2B^{e_2}$ les opérations $x+y,\,x-y,\,x*y,\,x/y$ sont définies comme suit : Partant du résultat exact $z=mB^e$ on arrondit à $\hat{z}=\hat{m}B^{\hat{e}}$ en réduisant la mantisse à la longueur prescrite (de ℓ chiffres).

Avantage: la mantisse reste petite, les calculs sont rapides.

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x=m_1B^{e_1}$ et $y=m_2B^{e_2}$ les opérations $x+y,\,x-y,\,x*y,\,x/y$ sont définies comme suit : Partant du résultat exact $z=mB^e$ on arrondit à $\hat{z}=\hat{m}B^{\hat{e}}$ en réduisant la mantisse à la longueur prescrite (de ℓ chiffres).

Avantage : la mantisse reste petite, les calculs sont rapides. **Inconvénient :** chaque opération introduit une erreur d'arrondi.

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x=m_1B^{e_1}$ et $y=m_2B^{e_2}$ les opérations $x+y,\,x-y,\,x*y,\,x/y$ sont définies comme suit : Partant du résultat exact $z=mB^e$ on arrondit à $\hat{z}=\hat{m}B^{\hat{e}}$ en réduisant la mantisse à la longueur prescrite (de ℓ chiffres).

Avantage : la mantisse reste petite, les calculs sont rapides. **Inconvénient :** chaque opération introduit une erreur d'arrondi.

Perspective: calcul arrondi fiable

On a le choix d'arrondir vers le plus proche (le standard),

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x=m_1B^{e_1}$ et $y=m_2B^{e_2}$ les opérations $x+y,\,x-y,\,x*y,\,x/y$ sont définies comme suit : Partant du résultat exact $z=mB^e$ on arrondit à $\hat{z}=\hat{m}B^{\hat{e}}$ en réduisant la mantisse à la longueur prescrite (de ℓ chiffres).

Avantage : la mantisse reste petite, les calculs sont rapides. **Inconvénient :** chaque opération introduit une erreur d'arrondi.

Perspective: calcul arrondi fiable

On a le choix d'arrondir vers le plus proche (le standard), vers $-\infty$, vers $+\infty$, vers 0, ou vers $\pm\infty$ (arrondis dirigés).

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x=m_1B^{e_1}$ et $y=m_2B^{e_2}$ les opérations $x+y,\,x-y,\,x*y,\,x/y$ sont définies comme suit : Partant du résultat exact $z=mB^e$ on arrondit à $\hat{z}=\hat{m}B^{\hat{e}}$ en réduisant la mantisse à la longueur prescrite (de ℓ chiffres).

Avantage : la mantisse reste petite, les calculs sont rapides. **Inconvénient :** chaque opération introduit une erreur d'arrondi.

Perspective: calcul arrondi fiable

On a le choix d'arrondir vers le plus proche (le standard), vers $-\infty$, vers $+\infty$, vers 0, ou vers $\pm\infty$ (arrondis dirigés). Au lieu d'une approximation $\hat{z}\approx z$ sans contrôle d'erreur on garantit ainsi un encadrement $\underline{z}\leq z\leq \overline{z}$.

Multiplions $N = 6.022 \cdot 10^{23}$ et $h = 6.626 \cdot 10^{-34}$!

Résultat exact : $Nh = 39.901772 \cdot 10^{-11} = 3.9901772 \cdot 10^{-10}$.

On l'arrondit à $Nh \approx 3.990 \cdot 10^{-10}$ (mantisse à 4 décimales).

Principe du calcul arrondi

Pour les nombres à virgule flottante $x = m_1 B^{e_1}$ et $y = m_2 B^{e_2}$ les opérations x + y, x - y, x * y, x/y sont définies comme suit :

Partant du résultat exact $z=mB^e$ on arrondit à $\hat{z}=\hat{m}B^{\hat{e}}$ en réduisant la mantisse à la longueur prescrite (de ℓ chiffres).

Avantage : la mantisse reste petite, les calculs sont rapides. **Inconvénient :** chaque opération introduit une erreur d'arrondi.

Perspective: calcul arrondi fiable

On a le choix d'arrondir vers le plus proche (le standard), vers $-\infty$, vers $+\infty$, vers 0, ou vers $\pm\infty$ (arrondis dirigés). Au lieu d'une approximation $\hat{z}\approx z$ sans contrôle d'erreur on garantit ainsi un encadrement $\underline{z}\leq z\leq \overline{z}$. Cette arithmétique d'intervalles permet de faire des calculs rigoureux avec des nombres flottants.

Un exemple perturbant

```
s:= 0.0;
for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; };
if ( s = 1.0 ) print("Le compte est bon."); else
print("Vous êtes accusé de détournement de fonds !");</pre>
```

Quel en sera le résultat?

Un exemple perturbant

```
s:= 0.0;
for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; };
if ( s = 1.0 ) print("Le compte est bon."); else
print("Vous êtes accusé de détournement de fonds !");</pre>
```

Quel en sera le résultat ? Ça dépend!

Un exemple perturbant

```
s:= 0.0;
for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; };
if ( s = 1.0 ) print("Le compte est bon."); else
print("Vous êtes accusé de détournement de fonds !");</pre>
```

Quel en sera le résultat ? Ça dépend!

■ de la représentation des nombres : décimal ou binaire ?

Un exemple perturbant

```
s:= 0.0; for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; }; if ( s = 1.0 ) print("Le compte est bon."); else print("Vous êtes accusé de détournement de fonds !");
```

Quel en sera le résultat ? Ça dépend!

- de la représentation des nombres : décimal ou binaire ?
- des arrondis effectués lors des calculs approchés.

Un exemple perturbant

```
s:= 0.0; for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; }; if ( s = 1.0 ) print("Le compte est bon."); else print("Vous êtes accusé de détournement de fonds !");
```

Quel en sera le résultat ? Ça dépend!

- de la représentation des nombres : décimal ou binaire ?
- des arrondis effectués lors des calculs approchés.
- des aléas de la propagation des erreurs.

Un exemple perturbant

```
s:= 0.0; for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; }; if ( s = 1.0 ) print("Le compte est bon."); else print("Vous êtes accusé de détournement de fonds !");
```

Quel en sera le résultat ? Ça dépend!

- de la représentation des nombres : décimal ou binaire ?
- des arrondis effectués lors des calculs approchés.
- des aléas de la propagation des erreurs.

Analogie en base 10 : $\frac{1}{3} \approx 0.33333$. On a $\frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$ (exact) mais 0.33333 + 0.33333 + 0.33333 = 0.99999 < 1 (erreur d'arrondi).

Un exemple perturbant

```
s:= 0.0; for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; }; if ( s = 1.0 ) print("Le compte est bon."); else print("Vous êtes accusé de détournement de fonds !");
```

Quel en sera le résultat ? Ça dépend!

- de la représentation des nombres : décimal ou binaire ?
- des arrondis effectués lors des calculs approchés.
- des aléas de la propagation des erreurs.

Analogie en base 10 : $\frac{1}{3} \approx 0.33333$. On a $\frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$ (exact) mais 0.33333 + 0.33333 + 0.33333 = 0.99999 < 1 (erreur d'arrondi).

Exercice pratique : tester ces calculs sous Xcas puis sous C/C++ avec les types float, double, Ou long double. (Surprise!)

Un exemple perturbant

```
s:= 0.0; for( k:=1; k<=10; k:=k+1) { s:= s + 0.1; }; if ( s = 1.0 ) print("Le compte est bon."); else print("Vous êtes accusé de détournement de fonds !");
```

Quel en sera le résultat ? Ça dépend!

- de la représentation des nombres : décimal ou binaire ?
- des arrondis effectués lors des calculs approchés.
- des aléas de la propagation des erreurs.

Analogie en base 10 : $\frac{1}{3} \approx 0.33333$. On a $\frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$ (exact) mais 0.33333 + 0.33333 + 0.33333 = 0.99999 < 1 (erreur d'arrondi).

Exercice pratique: tester ces calculs sous Xcas puis sous C/C++ avec les types float, double, ou long double. (Surprise!)

Exercice mathématique : développer $0.1_{\rm dec}$ en binaire, arrondir (à 24 ou 53 bits), puis effectuer les calculs (arrondis à 24 ou 53 bits).

Deux sources d'erreurs / d'imprécisions :

Deux sources d'erreurs / d'imprécisions :

erreurs dans les données initiales (mesurées, approchées)

Deux sources d'erreurs / d'imprécisions :

- rreurs dans les données initiales (mesurées, approchées)
- 2 erreurs d'arrondi lors des calculs (nombres à virgule flottante)

Deux sources d'erreurs / d'imprécisions :

- rreurs dans les données initiales (mesurées, approchées)
- 2 erreurs d'arrondi lors des calculs (nombres à virgule flottante)

Deux sources d'erreurs / d'imprécisions :

- erreurs dans les données initiales (mesurées, approchées)
- 2 erreurs d'arrondi lors des calculs (nombres à virgule flottante)

Problème fondamental du calcul approché

En général ce diagramme ne commute pas, c'est-à-dire $f(x) \neq \hat{f}(\hat{x})$.

Le problème fondamental du calcul approché

Deux sources d'erreurs / d'imprécisions :

- rreurs dans les données initiales (mesurées, approchées)
- 2 erreurs d'arrondi lors des calculs (nombres à virgule flottante)

nombres réels
$$x \in \mathbb{R}^n \xrightarrow{f} \mathbb{R} \ni f(x) \qquad \text{résultat exact,} \\ \text{arrondi} \downarrow \qquad \qquad \uparrow \text{inclusion} \\ \text{nombres} \\ \text{machine} \qquad \hat{x} \in R^n \xrightarrow{\hat{f}} R \ni \hat{f}(\hat{x}) \qquad \text{résultat calculé,} \\ \text{typiquement erroné} \\ \text{typiquement erroné}$$

Problème fondamental du calcul approché

En général ce diagramme ne commute pas, c'est-à-dire $f(x) \neq \hat{f}(\hat{x})$. Seule la marge d'erreur donne autorité au résultat calculé $\hat{f}(\hat{x})$.

Le problème fondamental du calcul approché

Deux sources d'erreurs / d'imprécisions :

- erreurs dans les données initiales (mesurées, approchées)
- 2 erreurs d'arrondi lors des calculs (nombres à virgule flottante)

Problème fondamental du calcul approché

En général ce diagramme ne commute pas, c'est-à-dire $f(x) \neq \hat{f}(\hat{x})$. Seule la marge d'erreur donne autorité au résultat calculé $\hat{f}(\hat{x})$.

Il est indispensable de majorer l'erreur totale $|f(x) - \hat{f}(\hat{x})|$.

Le problème fondamental du calcul approché

Deux sources d'erreurs / d'imprécisions :

- rreurs dans les données initiales (mesurées, approchées)
- 2 erreurs d'arrondi lors des calculs (nombres à virgule flottante)

nombres réels
$$x\in\mathbb{R}^n \xrightarrow{f} \mathbb{R}\ni f(x) \qquad \text{résultat exact, typiquement inconnu}$$

$$\text{arrondi} \downarrow \qquad \qquad \uparrow \text{inclusion}$$

$$\text{nombres machine} \qquad \hat{x}\in R^n \xrightarrow{\hat{f}} R\ni \hat{f}(\hat{x}) \qquad \text{résultat calculé, typiquement erroné}$$

Problème fondamental du calcul approché

En général ce diagramme ne commute pas, c'est-à-dire $f(x) \neq \hat{f}(\hat{x})$. Seule la marge d'erreur donne autorité au résultat calculé $\hat{f}(\hat{x})$.

Il est indispensable de majorer l'erreur totale $|f(x) - \hat{f}(\hat{x})|$. Sinon il est inutile, voire nuisible, de se lancer dans le calcul.

Notons x, y, z les valeurs exactes et $\hat{x}, \hat{y}, \hat{z}$ des valeurs approchées.

Notons x,y,z les valeurs exactes et \hat{x},\hat{y},\hat{z} des valeurs approchées.

Les deux mesures d'erreurs

Erreur absolue $\Delta x := |\hat{x} - x|$, erreur relative $\delta x := \frac{|\hat{x} - x|}{|x|}$.

Notons x,y,z les valeurs exactes et \hat{x},\hat{y},\hat{z} des valeurs approchées.

Les deux mesures d'erreurs

Erreur absolue
$$\Delta x := |\hat{x} - x|$$
, erreur relative $\delta x := \frac{|\hat{x} - x|}{|x|}$.

Rappel : les flottants mB^e où la mantisses a ℓ chiffres approchent tout réel à une erreur relative de $B^{1-\ell}$ près.

Notons x,y,z les valeurs exactes et \hat{x},\hat{y},\hat{z} des valeurs approchées.

Les deux mesures d'erreurs

Erreur absolue
$$\Delta x := |\hat{x} - x|$$
, erreur relative $\delta x := \frac{|\hat{x} - x|}{|x|}$.

Rappel : les flottants mB^e où la mantisses a ℓ chiffres approchent tout réel à une erreur relative de $B^{1-\ell}$ près.

Comment les erreurs se propagent-elles dans les calculs?

Notons x,y,z les valeurs exactes et \hat{x},\hat{y},\hat{z} des valeurs approchées.

Les deux mesures d'erreurs

Erreur absolue
$$\Delta x := |\hat{x} - x|$$
, erreur relative $\delta x := \frac{|\hat{x} - x|}{|x|}$.

Rappel : les flottants mB^e où la mantisses a ℓ chiffres approchent tout réel à une erreur relative de $B^{1-\ell}$ près.

Comment les erreurs se propagent-elles dans les calculs?

Comparons somme exacte z=x+y et somme approchée $\hat{z}=\hat{x}+\hat{y}$:

Notons x,y,z les valeurs exactes et \hat{x},\hat{y},\hat{z} des valeurs approchées.

Les deux mesures d'erreurs

Erreur absolue
$$\Delta x := |\hat{x} - x|$$
, erreur relative $\delta x := \frac{|\hat{x} - x|}{|x|}$.

Rappel : les flottants mB^e où la mantisses a ℓ chiffres approchent tout réel à une erreur relative de $B^{1-\ell}$ près.

Comment les erreurs se propagent-elles dans les calculs?

Comparons somme exacte z=x+y et somme approchée $\hat{z}=\hat{x}+\hat{y}$:

$$\Delta z = |\hat{z} - z| = |(\hat{x} + \hat{y}) - (x + y)| = |(\hat{x} - x) + (\hat{y} - y)|$$

$$\leq |\hat{x} - x| + |\hat{y} - y| = \Delta x + \Delta y$$

Notons x, y, z les valeurs exactes et $\hat{x}, \hat{y}, \hat{z}$ des valeurs approchées.

Les deux mesures d'erreurs

Erreur absolue
$$\Delta x := |\hat{x} - x|$$
, erreur relative $\delta x := \frac{|\hat{x} - x|}{|x|}$.

Rappel : les flottants mB^e où la mantisses a ℓ chiffres approchent tout réel à une erreur relative de $B^{1-\ell}$ près.

Comment les erreurs se propagent-elles dans les calculs?

Comparons somme exacte z=x+y et somme approchée $\hat{z}=\hat{x}+\hat{y}$:

$$\Delta z = |\hat{z} - z| = |(\hat{x} + \hat{y}) - (x + y)| = |(\hat{x} - x) + (\hat{y} - y)|$$

$$\leq |\hat{x} - x| + |\hat{y} - y| = \Delta x + \Delta y$$

Observation

Lors d'une addition, les erreurs absolues s'additionnent (au pire).

Comparons le produit exact z=xy au produit approché $\hat{z}=\hat{x}\hat{y}$:

Comparons le produit exact z=xy au produit approché $\hat{z}=\hat{x}\hat{y}$:

$$\begin{split} \delta z &= \frac{|\hat{z} - z|}{|z|} = \frac{|\hat{x}\hat{y} - xy|}{|xy|} = \frac{|\hat{x}\hat{y} - x\hat{y} + x\hat{y} - xy|}{|xy|} \\ &\leq \frac{|\hat{x}\hat{y} - x\hat{y}| + |x\hat{y} - xy|}{|xy|} = \frac{|\hat{x}\hat{y} - x\hat{y}|}{|xy|} + \frac{|x\hat{y} - xy|}{|xy|} \\ &= \frac{|\hat{x} - x|}{|x|} \cdot \frac{|\hat{y}|}{|y|} + \frac{|\hat{y} - y|}{|y|} = \delta x \cdot \frac{|\hat{y}|}{|y|} + \delta y \approx \delta x + \delta y \end{split}$$

Comparons le produit exact z=xy au produit approché $\hat{z}=\hat{x}\hat{y}$:

$$\begin{split} \delta z &= \frac{|\hat{z} - z|}{|z|} = \frac{|\hat{x}\hat{y} - xy|}{|xy|} = \frac{|\hat{x}\hat{y} - x\hat{y} + x\hat{y} - xy|}{|xy|} \\ &\leq \frac{|\hat{x}\hat{y} - x\hat{y}| + |x\hat{y} - xy|}{|xy|} = \frac{|\hat{x}\hat{y} - x\hat{y}|}{|xy|} + \frac{|x\hat{y} - xy|}{|xy|} \\ &= \frac{|\hat{x} - x|}{|x|} \cdot \frac{|\hat{y}|}{|y|} + \frac{|\hat{y} - y|}{|y|} = \delta x \cdot \frac{|\hat{y}|}{|y|} + \delta y \approx \delta x + \delta y \end{split}$$

Observation

Lors d'une multiplication, les erreurs *relatives* s'additionnent.

Comparons le produit exact z=xy au produit approché $\hat{z}=\hat{x}\hat{y}$:

$$\begin{split} \delta z &= \frac{|\hat{z}-z|}{|z|} = \frac{|\hat{x}\hat{y}-xy|}{|xy|} = \frac{|\hat{x}\hat{y}-x\hat{y}+x\hat{y}-xy|}{|xy|} \\ &\leq \frac{|\hat{x}\hat{y}-x\hat{y}|+|x\hat{y}-xy|}{|xy|} = \frac{|\hat{x}\hat{y}-x\hat{y}|}{|xy|} + \frac{|x\hat{y}-xy|}{|xy|} \\ &= \frac{|\hat{x}-x|}{|x|} \cdot \frac{|\hat{y}|}{|y|} + \frac{|\hat{y}-y|}{|y|} = \delta x \cdot \frac{|\hat{y}|}{|y|} + \delta y \approx \delta x + \delta y \end{split}$$

Observation

Lors d'une multiplication, les erreurs relatives s'additionnent.

Regardons finalement une fonction $f \colon \mathbb{R} \to \mathbb{R}$, supposée dérivable :

$$|f(\hat{x}) - f(x)| = \frac{|f(\hat{x}) - f(x)|}{\Delta x} \cdot \Delta x \approx |f'(x)| \cdot \Delta x$$

Comparons le produit exact z=xy au produit approché $\hat{z}=\hat{x}\hat{y}$:

$$\begin{split} \delta z &= \frac{|\hat{z} - z|}{|z|} = \frac{|\hat{x}\hat{y} - xy|}{|xy|} = \frac{|\hat{x}\hat{y} - x\hat{y} + x\hat{y} - xy|}{|xy|} \\ &\leq \frac{|\hat{x}\hat{y} - x\hat{y}| + |x\hat{y} - xy|}{|xy|} = \frac{|\hat{x}\hat{y} - x\hat{y}|}{|xy|} + \frac{|x\hat{y} - xy|}{|xy|} \\ &= \frac{|\hat{x} - x|}{|x|} \cdot \frac{|\hat{y}|}{|y|} + \frac{|\hat{y} - y|}{|y|} = \delta x \cdot \frac{|\hat{y}|}{|y|} + \delta y \approx \delta x + \delta y \end{split}$$

Observation

Lors d'une multiplication, les erreurs relatives s'additionnent.

Regardons finalement une fonction $f \colon \mathbb{R} \to \mathbb{R}$, supposée dérivable :

$$|f(\hat{x}) - f(x)| = \frac{|f(\hat{x}) - f(x)|}{\Delta x} \cdot \Delta x \approx |f'(x)| \cdot \Delta x$$

Observation

Une application $x \mapsto f(x)$ multiplie l'erreur absolue par |f'(x)|.

Sommaire

- 1 Nombres entiers et rationnels, calcul exact
- 2 Nombres réels, calcul approché, propagation d'erreurs
- 3 Types composés : expressions, fonctions, conteneurs, ...
 - Nombres complexes
 - Polynômes et fractions rationnelles
 - Expressions et fonctions
 - Conteneurs, vecteurs et matrices

Tout nombre complexe $z \in \mathbb{C}$ s'écrit comme z = x + iy avec $x, y \in \mathbb{R}$.

Tout nombre complexe $z \in \mathbb{C}$ s'écrit comme z = x + iy avec $x, y \in \mathbb{R}$. Ici $i^2 = -1$, donc (x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx').

Tout nombre complexe $z \in \mathbb{C}$ s'écrit comme z = x + iy avec $x, y \in \mathbb{R}$. Ici $i^2 = -1$, donc (x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx'). On appelle x la partie réelle et y la partie imaginaire.

Tout nombre complexe $z \in \mathbb{C}$ s'écrit comme z = x + iy avec $x, y \in \mathbb{R}$. lei $i^2 = -1$, donc (x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx'). On appelle x la partie réelle et y la partie imaginaire.

Représentations sur ordinateur :

- $lue{}$ calcul formel : x, y sont exacts (entiers ou rationnels).
- lacktriangle calcul approché : x,y sont approchés (nombres flottants).

Tout nombre complexe $z \in \mathbb{C}$ s'écrit comme z = x + iy avec $x, y \in \mathbb{R}$. Ici $i^2 = -1$, donc (x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx'). On appelle x la partie réelle et y la partie imaginaire.

Représentations sur ordinateur :

- $lue{}$ calcul formel : x, y sont exacts (entiers ou rationnels).
- lacktriangle calcul approché : x,y sont approchés (nombres flottants).

Le comportement diffère drastiquement !

Tout nombre complexe $z\in\mathbb{C}$ s'écrit comme z=x+iy avec $x,y\in\mathbb{R}$. Ici $i^2=-1$, donc (x+iy)(x'+iy')=(xx'-yy')+i(xy'+yx').

On appelle \boldsymbol{x} la partie réelle et \boldsymbol{y} la partie imaginaire.

Représentations sur ordinateur :

- lacktriangle calcul formel : x,y sont exacts (entiers ou rationnels).
- lacktriangle calcul approché : x,y sont approchés (nombres flottants).

Le comportement diffère drastiquement!

 $solve(z^3 - 1 = 0, z)$ donne les trois racines exactes

1,
$$-\frac{1}{2} + i\frac{\sqrt{3}}{2}$$
, $-\frac{1}{2} - i\frac{\sqrt{3}}{2}$.

Tout nombre complexe $z \in \mathbb{C}$ s'écrit comme z = x + iy avec $x, y \in \mathbb{R}$. Ici $i^2 = -1$, donc (x + iy)(x' + iy') = (xx' - yy') + i(xy' + yx'). On appelle x la partie réelle et y la partie imaginaire.

Représentations sur ordinateur :

- \blacksquare calcul formel : x, y sont exacts (entiers ou rationnels).
- \blacksquare calcul approché : x, y sont approchés (nombres flottants).

Le comportement diffère drastiquement!

 $solve(z^3-1=0,z)$ donne les trois racines exactes

1,
$$-\frac{1}{2} + i\frac{\sqrt{3}}{2}$$
, $-\frac{1}{2} - i\frac{\sqrt{3}}{2}$.

solve(z^3-1.0=0.0,z) donne seulement une approximation -0.500000000004 - 0.866025403782i.

Bien entendu ce dernier résultat n'est pas exact :

$$(-0.500000000004 - 0.866025403782i)^3 - 1 \approx -3.36 \cdot 10^{-13}.$$

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

➤ On peut stocker un polynôme par la liste de ces coefficients.

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur.

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

calcul formel : les coefficients sont exacts.

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

- calcul formel : les coefficients sont exacts.
- calcul approché : les coefficients sont des flottants.

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

- calcul formel : les coefficients sont exacts.
- calcul approché : les coefficients sont des flottants.

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

- calcul formel : les coefficients sont exacts.
- calcul approché : les coefficients sont des flottants.

Opérations de base (voir le TD/TP) :

opérations arithmétiques, notamment la division euclidienne

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

- calcul formel : les coefficients sont exacts.
- calcul approché : les coefficients sont des flottants.

- opérations arithmétiques, notamment la division euclidienne
- algorithme d'Euclide pour le pgcd, comme vu pour les entiers

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

- calcul formel : les coefficients sont exacts.
- calcul approché : les coefficients sont des flottants.

- opérations arithmétiques, notamment la division euclidienne
- algorithme d'Euclide pour le pgcd, comme vu pour les entiers
- lacktriangle évaluation P(z) en un point z par la méthode de Horner

Après les nombres entiers, rationnelles, réels, complexes, les objets les plus utilisés sont les polynômes et leur fractions :

$$P(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0$$

$$\frac{P(X)}{Q(X)} = \frac{a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0}{b_m X^m + b_{m-1} X^{m-1} + \dots + b_1 X + b_0}$$

- ➤ On peut stocker un polynôme par la liste de ces coefficients.
- ➤ On peut stocker une fraction par numérateur & dénominateur. (Bien sûr, il existe aussi des polynômes à plusieurs variables.)

Comme toujours, deux approches sont possibles :

- calcul formel : les coefficients sont exacts.
- calcul approché : les coefficients sont des flottants.

- opérations arithmétiques, notamment la division euclidienne
- algorithme d'Euclide pour le pgcd, comme vu pour les entiers
- évaluation P(z) en un point z par la méthode de Horner
- \blacksquare développement en (X-z), interpolation en plusieurs points

Expressions et fonctions

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

C'est déjà le cas en mathématiques, et plus encore en calcul formel!

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

C'est déjà le cas en mathématiques, et plus encore en calcul formel!

Raison : différentes expressions peuvent définir la même fonction !

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

C'est déjà le cas en mathématiques, et plus encore en calcul formel!

Raison : différentes expressions peuvent définir la même fonction !

Exemple : Les expressions $\sin(x)^2 + \cos(x)^2$ et 1 sont différentes, tandis que les fonctions associées $g,f\colon\mathbb{R}\to\mathbb{R}$ définies par $g(x)=\sin(x)^2+\cos(x)^2$ et f(x)=1 sont identiques.

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

C'est déjà le cas en mathématiques, et plus encore en calcul formel!

Raison : différentes expressions peuvent définir la même fonction !

Exemple : Les expressions $\sin(x)^2 + \cos(x)^2$ et 1 sont différentes, tandis que les fonctions associées $g,f\colon\mathbb{R}\to\mathbb{R}$ définies par $g(x)=\sin(x)^2+\cos(x)^2$ et f(x)=1 sont identiques.

Calcul formel : l'ordinateur (comme nous d'ailleurs) travaille sur des expressions formelles comme ci-dessus. Dans des cas favorables il arrive à les simplifier, ou à montrer leur équivalence : voir le TP!

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

C'est déjà le cas en mathématiques, et plus encore en calcul formel!

Raison : différentes expressions peuvent définir la même fonction !

Exemple : Les expressions $\sin(x)^2 + \cos(x)^2$ et 1 sont différentes, tandis que les fonctions associées $g,f\colon\mathbb{R}\to\mathbb{R}$ définies par $g(x)=\sin(x)^2+\cos(x)^2$ et f(x)=1 sont identiques.

Calcul formel : l'ordinateur (comme nous d'ailleurs) travaille sur des expressions formelles comme ci-dessus. Dans des cas favorables il arrive à les simplifier, ou à montrer leur équivalence : voir le TP!

Exemple: $simplify(sin(Pi/4)^2 + cos(Pi/4)^2 - 1)$ donne 0.

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

C'est déjà le cas en mathématiques, et plus encore en calcul formel!

Raison : différentes expressions peuvent définir la même fonction !

Exemple : Les expressions $\sin(x)^2 + \cos(x)^2$ et 1 sont différentes, tandis que les fonctions associées $g,f\colon\mathbb{R}\to\mathbb{R}$ définies par $g(x)=\sin(x)^2+\cos(x)^2$ et f(x)=1 sont identiques.

Calcul formel : l'ordinateur (comme nous d'ailleurs) travaille sur des expressions formelles comme ci-dessus. Dans des cas favorables il arrive à les simplifier, ou à montrer leur équivalence : voir le TP!

Exemple: $simplify(sin(Pi/4)^2 + cos(Pi/4)^2 - 1)$ donne 0.

Calcul approché : on évalue les expressions en nombres flottants.

À noter : on distingue l'expression $\sin(x) + x^2$ et la fonction $f: \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin(x) + x^2$.

C'est déjà le cas en mathématiques, et plus encore en calcul formel!

Raison : différentes expressions peuvent définir la même fonction !

Exemple : Les expressions $\sin(x)^2 + \cos(x)^2$ et 1 sont différentes, tandis que les fonctions associées $g,f\colon\mathbb{R}\to\mathbb{R}$ définies par $g(x)=\sin(x)^2+\cos(x)^2$ et f(x)=1 sont identiques.

Calcul formel : l'ordinateur (comme nous d'ailleurs) travaille sur des expressions formelles comme ci-dessus. Dans des cas favorables il arrive à les simplifier, ou à montrer leur équivalence : voir le TP!

Exemple: $simplify(sin(Pi/4)^2 + cos(Pi/4)^2 - 1)$ donne 0.

Calcul approché: on évalue les expressions en nombres flottants. evalf ($\sin(Pi/4)^2 + \cos(Pi/4)^2 - 1$) donne 2.22044e-16.

Un conteneur regroupe des objets plus simples.

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Exemples : liste, séquence, tableau, ensemble, ... (toujours finis!)

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Exemples : liste, séquence, tableau, ensemble, ... (toujours finis!) Vecteurs et matrices sont fréquemment utilisés (algèbre linéaire).

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Exemples : liste, séquence, tableau, ensemble, ... (toujours finis!) Vecteurs et matrices sont fréquemment utilisés (algèbre linéaire).

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Exemples : liste, séquence, tableau, ensemble, ... (toujours finis!) Vecteurs et matrices sont fréquemment utilisés (algèbre linéaire).

```
P:=matrix(3,3,[8,-5,4,-5,-1,-5,-6,4,-3]); Q:= P^-1; P*Q;
```

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Exemples : liste, séquence, tableau, ensemble, ... (toujours finis!) Vecteurs et matrices sont fréquemment utilisés (algèbre linéaire).

$$P = \begin{pmatrix} 8 & -5 & -4 \\ -5 & -1 & 5 \\ -6 & 4 & 3 \end{pmatrix}, \ Q = \begin{pmatrix} \frac{23}{5} & \frac{1}{5} & \frac{29}{5} \\ 3 & 0 & 4 \\ \frac{26}{5} & \frac{2}{5} & \frac{33}{5} \end{pmatrix}, \ PQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Exemples : liste, séquence, tableau, ensemble, ... (toujours finis!) Vecteurs et matrices sont fréquemment utilisés (algèbre linéaire).

$$P = \begin{pmatrix} 8 & -5 & -4 \\ -5 & -1 & 5 \\ -6 & 4 & 3 \end{pmatrix}, \; Q = \begin{pmatrix} \frac{23}{5} & \frac{1}{5} & \frac{29}{5} \\ 3 & 0 & 4 \\ \frac{26}{5} & \frac{2}{5} & \frac{33}{5} \end{pmatrix}, \; PQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$P:=matrix(3,3,[8.0,-5,-4,-5,-1,5,-6,4,3]); Q:= P^-1; P*Q;$$

Un conteneur regroupe des objets plus simples.

On accède à ces objets par un indice (adapté au conteneur).

Exemples : liste, séquence, tableau, ensemble, ... (toujours finis!) Vecteurs et matrices sont fréquemment utilisés (algèbre linéaire).

$$P = \begin{pmatrix} 8 & -5 & -4 \\ -5 & -1 & 5 \\ -6 & 4 & 3 \end{pmatrix}, \; Q = \begin{pmatrix} \frac{23}{5} & \frac{1}{5} & \frac{29}{5} \\ 3 & 0 & 4 \\ \frac{26}{5} & \frac{2}{5} & \frac{33}{5} \end{pmatrix}, \; PQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$P:=matrix(3,3,[8.0,-5,-4,-5,-1,5,-6,4,3]); Q:= P^-1; P*Q;$$

$$P = \begin{pmatrix} 8.0 & -5 & -4 \\ -5 & -1 & 5 \\ -6 & 4 & 3 \end{pmatrix}, \; Q = \begin{pmatrix} 4.6 & 0.2 & 5.8 \\ 3.0 & 0.0 & 4.0 \\ 5.2 & 0.4 & 6.6 \end{pmatrix}, \; PQ = \begin{pmatrix} 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 3 \cdot 10^{-15} & 0.0 & 1.0 \end{pmatrix}$$

Types élémentaires :

Types élémentaires :

■ exacts : nombres entiers et rationnels, variables formelles, . . .

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

 \blacksquare polynômes $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, fractions, . . .

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

- \blacksquare polynômes $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, fractions, ...
- \blacksquare expressions x^2 , $\sin(x+x^2)$, $\frac{\exp(x)-1}{x}$, ...

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

- \blacksquare polynômes $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, fractions, ...
- \blacksquare expressions x^2 , $\sin(x+x^2)$, $\frac{\exp(x)-1}{x}$, ...
- conteneurs, listes, vecteurs, matrices, . . .

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

- \blacksquare polynômes $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, fractions, . . .
- \blacksquare expressions x^2 , $\sin(x+x^2)$, $\frac{\exp(x)-1}{x}$, ...
- conteneurs, listes, vecteurs, matrices, . . .

Un objet composé est en *mode exact* si tous ses sous-objets le sont.

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

- \blacksquare polynômes $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, fractions, . . .
- \blacksquare expressions x^2 , $\sin(x+x^2)$, $\frac{\exp(x)-1}{x}$, ...
- conteneurs, listes, vecteurs, matrices, . . .

Un objet composé est en *mode exact* si tous ses sous-objets le sont. Sinon il est en *mode approché*. Les algorithmes en dépendent!

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

- \blacksquare polynômes $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, fractions, ...
- \blacksquare expressions x^2 , $\sin(x+x^2)$, $\frac{\exp(x)-1}{x}$, ...
- conteneurs, listes, vecteurs, matrices, . . .

Un objet composé est en *mode exact* si tous ses sous-objets le sont. Sinon il est en *mode approché*. Les algorithmes en dépendent!

La commande evalf passe du mode exact au mode approché.

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

- \blacksquare polynômes $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, fractions, ...
- \blacksquare expressions x^2 , $\sin(x+x^2)$, $\frac{\exp(x)-1}{x}$, ...
- conteneurs, listes, vecteurs, matrices, . . .

Un objet composé est en *mode exact* si tous ses sous-objets le sont. Sinon il est en *mode approché*. Les algorithmes en dépendent!

La commande evalf passe du mode *exact* au mode *approché*. Elle est « sens unique » ! (Pourquoi ?)

Types élémentaires :

- exacts : nombres entiers et rationnels, variables formelles, . . .
- approchés : nombres à virgule flottante (réels, complexes)

Types composés

- \blacksquare polynômes $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, fractions, ...
- \blacksquare expressions x^2 , $\sin(x+x^2)$, $\frac{\exp(x)-1}{x}$, ...
- conteneurs, listes, vecteurs, matrices, ...

Un objet composé est en *mode exact* si tous ses sous-objets le sont. Sinon il est en *mode approché*. Les algorithmes en dépendent!

La commande evalf passe du mode *exact* au mode *approché*. Elle est « sens unique » ! (Pourquoi ?)

Passons à la pratique!

Lisez l'introduction à Xcas, puis expérimentez abondamment! Familiarisez-vous avec votre calculatrice, elle vous servira partout!

Résumé

- 1 Nombres entiers et rationnels, calcul exact
 - Numération positionnelle
 - Arithmétique des entiers
 - Nombres rationnels
- 2 Nombres réels, calcul approché, propagation d'erreurs
 - La problématique illustrée par l'exemple $\sqrt{2}$
 - Nombres à virgule flottante et calcul approché
 - Propagation d'erreurs : la prudence s'impose!
- 3 Types composés : expressions, fonctions, conteneurs, ...
 - Nombres complexes
 - Polynômes et fractions rationnelles
 - Expressions et fonctions
 - Conteneurs, vecteurs et matrices