



---

# MATHÉMATIQUES GÉNÉRALES PREMIÈRE ANNÉE

---

Version enseignant

Mourad ABOUZAIID

**ISABTP 1<sup>re</sup> année  
2020-2021**

*Dernière compilation : 19 août 2020*



# Table des matières

<b>1 Outils élémentaires</b>	<b>5</b>
Introduction . . . . .	5
1.1 Notations . . . . .	5
1.1.1 Alphabet grec . . . . .	6
1.1.2 Symboles courants . . . . .	6
1.1.3 Indices . . . . .	7
1.2 La notion de démonstration . . . . .	7
1.2.1 Proposition et tiers exclu . . . . .	7
1.2.2 Les quantificateurs . . . . .	8
1.2.3 L'implication $\Rightarrow$ . . . . .	9
1.2.4 Contraposition . . . . .	10
1.2.5 Différents modes de raisonnement . . . . .	11
1.3 Ensembles et applications . . . . .	15
1.3.1 Notion d'ensemble, d'élément . . . . .	15
1.3.2 Les différentes façons de définir un ensemble . . . . .	16
1.3.3 Inclusion . . . . .	17
1.3.4 Opération sur les ensembles . . . . .	18
1.3.5 Les parties d'un ensemble . . . . .	19
1.3.6 Cardinal d'un ensemble fini . . . . .	20
1.3.7 Produit cartésien d'ensembles . . . . .	21
1.3.8 Applications . . . . .	22
1.4 Techniques fondamentales du calcul algébrique . . . . .	31
1.4.1 Propriétés des opérations . . . . .	31
1.4.2 Le signe $\Sigma$ . . . . .	33
1.4.3 Le signe $\Pi$ . . . . .	37
1.4.4 Formule du binôme et factorisation de $a^n - b^n$ . . . . .	38
1.4.5 Résolution d'équations algébriques (une inconnue) . . . . .	41
<b>2 Fonctions d'une variable réelle</b>	<b>47</b>
Introduction . . . . .	47
2.1 Définitions . . . . .	47
2.2 Premières propriétés . . . . .	49
2.2.1 Parité, périodicité . . . . .	49
2.2.2 Fonctions bornées . . . . .	51
2.2.3 Monotonie . . . . .	52
2.2.4 Extrema . . . . .	53

2.2.5	L'ensemble $\mathcal{F}(I, \mathbb{R})$ . . . . .	54
2.3	Limites . . . . .	54
2.3.1	Limite en un point . . . . .	54
2.3.2	Limite en l'infini . . . . .	57
2.3.3	Limites et inégalités . . . . .	58
2.3.4	Limites et opérations . . . . .	59
2.3.5	Formes indéterminées . . . . .	60
2.3.6	Croissances comparées . . . . .	62
2.4	Continuité . . . . .	63
2.4.1	Définitions . . . . .	63
2.4.2	Continuité et opérations . . . . .	65
2.4.3	Prolongement par continuité . . . . .	65
2.4.4	Théorème des valeurs intermédiaires . . . . .	66
2.5	Dérivabilité . . . . .	67
2.5.1	Dérivée en un point . . . . .	67
2.5.2	Dérivée et tangente . . . . .	67
2.5.3	Fonction dérivée et dérivées successives . . . . .	69
2.5.4	Dérivées des fonctions usuelles . . . . .	71
2.5.5	Théorème de Rolle et Théorème des Accroissements finis . . . . .	74
2.6	Comparaison locale de fonctions . . . . .	75
2.6.1	Domination . . . . .	76
2.6.2	Négligeabilité . . . . .	77
2.6.3	Équivalence . . . . .	79
2.7	Développements limités . . . . .	81
2.7.1	Principes généraux . . . . .	81
2.7.2	Définitions . . . . .	82
2.7.3	Formule de Taylor-Young . . . . .	85
2.7.4	Développement limité hors de 0 . . . . .	87
2.7.5	Applications des développements limités . . . . .	89
<b>3</b>	<b>Suites numériques</b> . . . . .	<b>93</b>
	Introduction . . . . .	93
3.1	Généralités . . . . .	94
3.1.1	Définitions . . . . .	94
3.1.2	Représentation graphique . . . . .	95
3.2	Convergence d'une suite . . . . .	95
3.3	Propriétés des suites réelles . . . . .	97
3.3.1	Suites bornées . . . . .	97
3.3.2	Suites monotones . . . . .	97
3.4	Critères de convergence . . . . .	98
3.4.1	Limites et opérations . . . . .	99
3.4.2	Critères de comparaison . . . . .	101
3.5	Suites récurrentes . . . . .	105
3.5.1	Définitions . . . . .	105
3.5.2	Les suites arithmétiques et géométriques . . . . .	106
3.5.3	Suites récurrentes autonomes d'ordre 1 . . . . .	108

3.6	Vitesse de convergence . . . . .	116
3.6.1	Échelle de vitesses . . . . .	117
3.6.2	Application aux suites autonomes d'ordre 1 . . . . .	118
3.7	Résolution approchée d'équations algébriques . . . . .	118
3.7.1	Dichotomie . . . . .	118
3.7.2	Méthode de Newton-Raphson . . . . .	119
<b>4</b>	<b>Matrices et systèmes linéaires</b>	<b>123</b>
	Introduction . . . . .	123
4.1	Matrices . . . . .	123
4.1.1	Définitions . . . . .	123
4.1.2	Indexation . . . . .	124
4.1.3	Matrices particulières (carrées) . . . . .	124
4.1.4	Opérations matricielles . . . . .	126
4.1.5	Matrices inversibles . . . . .	129
4.1.6	Déterminants . . . . .	130
4.2	Systèmes linéaires . . . . .	135
4.2.1	Définitions . . . . .	135
4.2.2	Structure de l'ensemble des solutions . . . . .	138
4.2.3	Systèmes libres, systèmes liés . . . . .	145
4.2.4	Systèmes échelonnés, systèmes réduits . . . . .	148
<b>5</b>	<b>Introduction au langage Python</b>	<b>155</b>
5.1	Le langage Python . . . . .	156
5.1.1	Installation . . . . .	156
5.1.2	Objets et variables . . . . .	160
5.1.3	Fonctions numériques . . . . .	161
5.1.4	Outils de programmation . . . . .	167
5.1.5	Graphiques . . . . .	170
5.2	Résolution approchée d'équations et systèmes algébriques . . . . .	175
5.2.1	Le TVI et ses conséquences . . . . .	176
5.2.2	La méthode de Newton-Raphson . . . . .	178
5.2.3	Généralisation aux systèmes algébriques . . . . .	182
5.2.4	Compléments . . . . .	183
5.3	Intégration numérique . . . . .	186
5.3.1	Approche géométrique . . . . .	186
5.3.2	Approche analytique . . . . .	187
5.3.3	Mesure d'erreur . . . . .	188
5.3.4	Généralisation . . . . .	189
5.3.5	Traitement de données . . . . .	190



# Chapitre 1

## Outils élémentaires

### Introduction

Les mathématiques ont pour but d'étudier les différentes propriétés de nombreux objets théoriques qui représentent les éléments, les principes et les idées permettant de modéliser le monde réel. Ces objets sont dans un premier temps décrits à l'aide de définitions précises qui servent de base à un raisonnement logique. Ce raisonnement logique permet alors de valider des propositions (affirmations pouvant être vraies ou fausses) vérifiées par ces objets. Les propositions les plus importantes sont appelées *théorèmes*.

Pour étudier ces outils, les mathématiciens disposent d'un langage précis, permettant de définir des objets et des concepts très pointus et d'un ensemble de règles basées sur la logique élémentaire.

Par ailleurs, l'un des rôles des mathématiques dans les sciences dites dures est de fournir des outils d'études et de résolution pour les équations issues de la modélisation. Il existe de nombreux types d'équations, portant sur tous types d'objets mathématiques. Les équations les plus courantes sont *les équations algébriques*, liant des variables réelles ou complexes.

La résolution des équations algébriques compte parmi les plus anciens problèmes sur lesquels se sont penchés les mathématiciens. Cette recherche a produit des méthodes de résolution basées sur les propriétés élémentaires et les manipulations qu'elles autorisent des opérations spécifiques aux nombres réels et complexes : l'addition et la multiplication.

Dans ce premier chapitre, nous effectuerons un rapide survol des outils élémentaires des sciences mathématiques : quelques notions de logique élémentaires, un survol rapide de la théorie des ensembles et des applications et des notions importantes de calcul algébrique.

### 1.1 Notations

Le langage mathématique est avant tout basé sur un ensemble de symboles et notations précises, qui permettent de nommer les différents objets et de représenter leurs interactions.

### 1.1.1 Alphabet grec

Ces notations utilisent les caractères de l'alphabet latin, majuscules et minuscules, ainsi que, souvent, ceux de l'alphabet grec :

Minuscule	Majuscules	Nom	Minuscule	Majuscule	Nom
$\alpha$	$A$	alpha	$\nu$	$N$	nu
$\beta$	$B$	bêta	$\xi$	$\Xi$	xi
$\gamma$	$\Gamma$	gamma	$o$	$O$	omicron
$\delta$	$\Delta$	delta	$\pi$	$\Pi$	pi
$\epsilon$	$E$	epsilon	$\rho$	$P$	rhô
$\zeta$	$Z$	zêta	$\sigma$	$\Sigma$	sigma
$\eta$	$H$	êta	$\tau$	$T$	tau
$\theta$	$\Theta$	thêta	$\upsilon$	$\Upsilon$	upsilon
$\iota$	$I$	iota	$\varphi$	$\Phi$	phi
$\kappa$	$K$	kappa	$\psi$	$\Psi$	psi
$\lambda$	$\Lambda$	lambda	$\chi$	$X$	khi
$\mu$	$M$	mu	$\omega$	$\Omega$	omega

### 1.1.2 Symboles courants

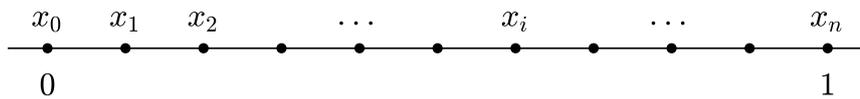
D'autre part, outre les symboles de base (+,  $\times$ , =, etc), d'autres symboles et abréviations sont utilisés pour permettre plus de concision dans les énoncés ainsi que certains symboles particuliers pour représenter les ensembles de nombres :

Symbole	Signification
$\forall$	quelque soit
$\exists$	il existe
$\in$	appartient
/	tel que
$\Rightarrow$	implique
$\Leftrightarrow$ (ou ssi)	si et seulement si
$\Sigma$	somme
$\Pi$	produit
i.e.	c'est à dire ( <i>id est</i> )
cqfd ou $\square$	ce qu'il fallait démontrer
$\mathbb{N}$	les entiers naturels ( $\geq 0$ )
$\mathbb{N}^*$	les entiers naturel non nuls
$\mathbb{Z}$	les entiers relatifs
$\mathbb{Q}$	les rationnels (fractions d'entiers)
$\mathbb{R}$	les réels
$\mathbb{R}^*$	les réels non nuls
$\mathbb{C}$	les complexes

### 1.1.3 Indices

Pour symboliser un nombre important voire indéterminé d'objets, on les distingue à l'aide d'indices.

**Exemple** : pour un entier  $n \in \mathbb{N}^*$  fixé, on découpe l'intervalle  $[0, 1]$  en  $n$  intervalles de même longueur :



Les bornes  $x_0, x_1, \dots, x_n$  de ces intervalles sont alors définies par

$$\forall i \in \{0, \dots, n\}, \quad x_i = \frac{i}{n}$$

De même, pour les éléments d'un tableau, on a recours aux doubles indices  $x_{ij}$ , le premier donnant le numéro de la ligne, le second celui de la colonne. Ainsi, un tableau à  $n$  lignes et  $p$  colonnes s'écrit :

	Col. 1	Col. 2	$\dots$	Col. $j$	$\dots$	Col. $p$
Ligne 1	$x_{11}$	$x_{12}$	$\dots$	$x_{1j}$	$\dots$	$x_{1p}$
Ligne 2	$x_{21}$	$x_{22}$	$\dots$	$x_{2j}$	$\dots$	$x_{2p}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
Ligne $i$	$x_{i1}$	$x_{i2}$	$\dots$	$x_{ij}$	$\dots$	$x_{ip}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
Ligne $n$	$x_{n1}$	$x_{n2}$	$\dots$	$x_{nj}$	$\dots$	$x_{np}$

## 1.2 La notion de démonstration

### 1.2.1 Proposition et tiers exclu

En mathématique, une *proposition* est une affirmation susceptible d'être vraie ou fausse (i.e. ayant une valeur de vérité).

La logique mathématique est basée sur le principe du *tiers exclu* : une proposition donnée est nécessairement soit vraie soit fausse et ne peut être les deux à la fois. Ainsi, si elle n'est pas vraie, elle est fausse (et réciproquement)

**Notes :**

- Ce principe est vrai pour la plupart des propositions, mais pas toujours. Ainsi, si l'on admet la proposition suivante :

$\mathcal{P}$  : "Le barbier ne rase que les hommes qui ne se rasent pas eux même."

alors la proposition

$\mathcal{Q}$  : "Le barbier se rase lui même."

n'est ni vraie ni fausse.

- Formellement, il est possible d'associer à toute proposition  $\mathcal{P}$  sa négation : la proposition (non  $\mathcal{P}$ ). Le principe du tiers exclus s'énonce alors sous la forme

Si  $\mathcal{P}$  vérifie le principe du tiers exclus, alors  $\mathcal{P}$  ou (non  $\mathcal{P}$ ) est vraie.

## 1.2.2 Les quantificateurs

On appelle *quantificateurs* les symboles  $\forall$  et  $\exists$ . Ils permettent en général d'énoncer de façon concise une proposition mathématique. Ainsi, soit  $\mathcal{P}(x)$  une proposition dépendant d'un paramètre  $x$ .

- S'il existe un élément  $x$  pour lequel cette propriété est vraie, on note

$\exists x / \mathcal{P}(x)$

- Si cette propriété est vraie pour toutes les valeurs de  $x$  d'un ensemble  $E$ , on note

$\forall x \in E, \mathcal{P}(x)$

Par ailleurs, outre la concision des énoncé mathématiques, ces quantificateurs permettent de mettre en place un précédé automatique donnant la négation d'une proposition si celle-ci est énoncée à l'aide des quantificateurs

- La négation de " $\forall x \in E, \mathcal{P}(x)$ " est

$\exists x \in E / (\text{non } \mathcal{P})$

— La négation de “ $\exists x \in E / \mathcal{P}(x)$ ” est

$$\forall x \in E, (\text{non } \mathcal{P})$$

### Notes

— ATTENTION : les quantificateurs  $\nexists$  ou  $\nforall$  n'existent pas ! Il est donc interdit de les utiliser.

— Dans une phrase mathématique (comme en français, finalement), l'ordre des quantificateurs est essentielle. En effet, les deux propositions

$$\forall x \in \mathbb{R}, \exists y \in \mathbb{R} / (x + y > 0) \quad \text{et} \quad \exists y \in \mathbb{R} / \forall x \in \mathbb{R}, (x + y > 0)$$

n'ont pas du tout le même sens. La première est vraie alors que la seconde est fausse. En effet, la première phrase affirme que “pour tout nombre réel  $x$ , il existe un nombre réel  $y$  (qui peut dépendre de  $x$ ) tel que la somme  $x + y$  soit strictement positive”. Ceci est vrai. Il suffit par exemple pour chaque  $x$  de prendre  $y = |x| + 1$ . La seconde phrase se lit “il existe un nombre réel  $y$  tel que pour tout nombre réel  $x$ , la somme  $x + y$  soit strictement positive. Ici, le même nombre  $y$  doit marcher pour tous les  $x$ ...

— Le quantificateur  $\exists$  sous entend “il existe au moins un élément ... tel que ...”. Si l'on souhaite préciser l'unicité d'un tel élément, on ajoute le symbole “!” :

$$\exists! x / P(x)$$

### 1.2.3 L'implication $\Rightarrow$

Étant données deux propositions  $\mathcal{P}$  et  $\mathcal{Q}$ , on note  $\mathcal{P} \Rightarrow \mathcal{Q}$  l'affirmation (i.e. la proposition) selon laquelle

$$\mathcal{P} \Rightarrow \mathcal{Q} : \text{"Si } \mathcal{P} \text{ est vraie, alors } \mathcal{Q} \text{ aussi"}$$

En langage courant, la proposition  $\mathcal{P} \Rightarrow \mathcal{Q}$  se dit aussi

- “ $\mathcal{Q}$  est vraie si  $\mathcal{P}$  est vraie”,
- “ $\mathcal{P}$  est vraie seulement si  $\mathcal{Q}$  est vraie”

Si la proposition  $\mathcal{P} \Rightarrow \mathcal{Q}$  est vraie, on dit également que

- La véracité de la proposition  $\mathcal{P}$  est une *condition suffisante* à celle que la proposition  $\mathcal{Q}$ ,
- La véracité de la proposition  $\mathcal{Q}$  est une *condition nécessaire* à celle que la proposition  $\mathcal{P}$ .

**Exemple** : étant donné un quadrilatère  $ABCD$ , si l'on note

$\mathcal{P}$  : " $ABCD$  est un carré" et  $\mathcal{Q}$  : " $ABCD$  est un parallélogramme"

on a bien la proposition  $\mathcal{P} \Rightarrow \mathcal{Q}$ .

En langage courant, cela peut se traduire de différentes façons :

- Si l'on sait que  $ABCD$  est un carré, cela suffit à montrer que c'est un parallélogramme.
- Pour que  $ABCD$  soit un carré, il est nécessaire que ce soit un parallélogramme.

Enfin, notons que

- Si " $ABCD$  est un parallélogramme", cela ne suffit pas à montrer que c'est un carré.
- Il n'est pas nécessaire que  $ABCD$  soit un carré pour que ce soit un parallélogramme.

Autrement dit, la proposition  $\mathcal{Q} \Rightarrow \mathcal{P}$  (appelée *reciproque* de la proposition  $\mathcal{P} \Rightarrow \mathcal{Q}$ ) est fausse.

De façon générale, il n'existe aucun lien entre une proposition et sa réciproque. Cependant, si les deux propositions  $\mathcal{P} \Rightarrow \mathcal{Q}$  et  $\mathcal{Q} \Rightarrow \mathcal{P}$  sont vraies, on dit que  $\mathcal{P}$  et  $\mathcal{Q}$  sont *équivalentes* et l'on note

$$\mathcal{P} \Leftrightarrow \mathcal{Q} \quad \text{ou} \quad \mathcal{P} \text{ ssi } \mathcal{Q}$$

### 1.2.4 Contraposition

Étant données deux propositions  $\mathcal{P}$  et  $\mathcal{Q}$ , on appelle *contraposée* de l'implication  $\mathcal{P} \Rightarrow \mathcal{Q}$  l'implication

$$(\text{non } \mathcal{Q}) \Rightarrow (\text{non } \mathcal{P})$$

**Exemple** : la contraposée de l'implication

$$(ABCD \text{ est un carré}) \Rightarrow (ABCD \text{ est un parallélogramme})$$

est

$$(ABCD \text{ n'est pas un parallélogramme}) \Rightarrow (ABCD \text{ n'est pas un carré})$$

On note que sur cet exemple, ces deux implications sont vraies. On peut montrer que, de façon générale, une implication et sa contraposée ont toujours la même valeur de vérité.

Exercice : s'en convaincre.

### 1.2.5 Différents modes de raisonnement

L'ensemble de l'édifice mathématique est basé sur la notion de *démonstration*. Précisément, une démonstration consiste à établir la véracité d'une implication  $\mathcal{P} \Rightarrow \mathcal{Q}$  à l'aide d'une suite de raisonnements logiques.

En pratique,

- La proposition  $\mathcal{P}$  sur laquelle on s'appuie est appelée *l'hypothèse*.
- La proposition  $\mathcal{Q}$  que l'on souhaite établir est appelée *la conclusion ou l'objectif*.

Notons que pour démontrer l'implication  $\mathcal{P} \Rightarrow \mathcal{Q}$ , on ne fait aucune hypothèse sur la valeur de vérité de la proposition  $\mathcal{P}$ . On se contente de démontrer que **SI**  $\mathcal{P}$  est vraie, **ALORS**  $\mathcal{Q}$  est vraie.

Selon le résultat que l'on souhaite établir, il existe alors différentes méthodes de raisonnements permettant d'aboutir au même résultat : l'implication  $\mathcal{P} \Rightarrow \mathcal{Q}$ .

#### La démonstration directe

Le *raisonnement direct* est le plus naturel (mais pas toujours le plus efficace). Il consiste à établir la propriété  $\mathcal{Q}$  directement à partir de  $\mathcal{P}$ .

Formellement, on a alors

- Hypothèse : *la proposition  $\mathcal{P}$* .
- Objectif : *la proposition  $\mathcal{Q}$* .

**Exemple** : montrer par un raisonnement direct que si  $a$  et  $b$  sont deux nombres rationnels, alors  $a + b$  est également un nombre rationnel.

On rappelle pour cela que les nombres rationnels sont les nombres de la forme  $\frac{p}{q}$  avec  $p \in \mathbb{Z}$  et  $q \in \mathbb{N}^*$ .

**Démonstration :**

Hypothèse :  $\mathcal{P}$  : "a et b sont deux nombres rationnels", i.e.

$$\exists p, p' \in \mathbb{Z}, q, q' \in \mathbb{N}^* / a = \frac{p}{q} \quad \text{et} \quad b = \frac{p'}{q'}$$

Objectif :  $\mathcal{Q}$  : "a + b est un nombre rationnel", i.e.

$$\exists p'' \in \mathbb{Z}, q'' \in \mathbb{N}^* / a + b = \frac{p''}{q''}$$

Preuve : par hypothèse,

$$a + b = \frac{p}{q} + \frac{p'}{q'} = \frac{pq' + p'q}{qq'}$$

Or,  $p, p', q$  et  $q'$  étant des entiers, on a  $pq' + p'q \in \mathbb{Z}$  et  $q$  et  $q'$  étant deux entiers naturels non nuls, on a aussi  $qq' \in \mathbb{N}^*$ .

La somme  $a + b$  est donc un nombre rationnel. □

### La démonstration par contraposition

La démonstration par contraposition s'appuie sur la remarque fait plus haut, concernant le fait qu'une implication  $\mathcal{P} \Rightarrow \mathcal{Q}$  et sa contraposée ( $\text{non } \mathcal{Q} \Rightarrow \text{non } \mathcal{P}$ ) ont la même valeur de vérité.

Dans certains cas, il est plus facile de démontrer la seconde (la contraposée).

Ainsi, pour montrer l'implication  $\mathcal{P} \Rightarrow \mathcal{Q}$  par contraposition, on a formellement,

— Hypothèse : la proposition (non  $\mathcal{Q}$ ).

— Objectif : la proposition (non  $\mathcal{P}$ ).

**Exemple** : montrer par contraposition que pour tout  $n \in \mathbb{N}$ , si  $n^2$  est pair, alors  $n$  est pair.

On rappelle un nombre entier est pair s'il peut s'écrire  $2 \times k$  pour un  $k \in \mathbb{N}$  et que tout entier non pair peut s'écrire  $2 \times k + 1$  avec  $k \in \mathbb{N}$ ).

Formellement, on note

$$\mathcal{P} : "n^2 \text{ est un entier pair}" \quad \text{et} \quad \mathcal{Q} : "n \text{ est un entier pair}"$$

On a alors

$$(\text{non } \mathcal{Q}) : "n \text{ est un entier impair}" \quad \text{et} \quad (\text{non } \mathcal{P}) : "n^2 \text{ est un entier impair}"$$

La démonstration par contraposition se formalise donc comme suit :

Hypothèse :  $(\text{non } \mathcal{Q}) : "n \text{ est un entier impair}"$ , i.e.

$$\exists k \in \mathbb{Z} / n = 2k + 1$$

Objectif :  $(\text{non } \mathcal{P}) : "n^2 \text{ est un entier impair}"$ , i.e.

$$\exists \ell \in \mathbb{Z} / n^2 = 2\ell + 1$$

Preuve : par hypothèse, il existe un entier  $k \in \mathbb{Z}$  tel que

$$n = 2k + 1$$

Mais alors

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

Or, puisque  $k$  est un entier,  $\ell = 2k^2 + 2k$  est également un entier et  $n^2 = 2\ell + 1$  est impair.  $\square$

### Démonstration par l'absurde

La méthode de raisonnement par l'absurde s'appuie directement sur la notion de tiers exclu.

Précisément, cette méthode de raisonnement consiste à démontrer une implication  $\mathcal{P} \Rightarrow \mathcal{Q}$  en supposant que les propositions  $\mathcal{P}$  et  $(\text{non } \mathcal{Q})$  sont vraies simultanément et en cherchant à aboutir à une absurdité par un raisonnement logique.

En effet, si cette démarche aboutit, c'est que les deux propriétés  $\mathcal{P}$  et  $(\text{non } \mathcal{Q})$  ne peuvent être vraies en même temps. Donc si  $\mathcal{P}$  est vrai, c'est que  $(\text{non } \mathcal{Q})$  est fausse, i.e.  $\mathcal{Q}$  est vraie.

Formellement, dans une démonstration par l'absurde de l'implication  $\mathcal{P} \Rightarrow \mathcal{Q}$ , on a donc

- Hypothèse : les propositions  $\mathcal{P}$  et (non  $\mathcal{Q}$ ).
- Objectif : aboutir à une contradiction logique.

**Exemple** : montrer que si  $a$  et  $b$  sont deux réels strictement positifs tels que  $\frac{a+1}{b} = \frac{b+1}{a}$ , alors  $a = b$ .

### Démonstration

Hypothèses :

- $\mathcal{P}$  :  $a, b > 0$  et  $\frac{a+1}{b} = \frac{b+1}{a}$
- (non  $\mathcal{Q}$ ) :  $a \neq b$

Objectif : parvenir à une absurdité.

Preuve :

$$\begin{aligned} \frac{a+1}{b} = \frac{b+1}{a} &\implies a(a+1) = b(b+1) \\ &\implies a^2 + a = b^2 + b \\ &\implies a^2 - b^2 = b - a \\ &\implies (a-b)(a+b) = -(a-b) \end{aligned}$$

Or  $a \neq b$  donc on peut simplifier l'égalité ci-dessus par  $(a-b) \neq 0$  et l'on a

$$\frac{a+1}{b} = \frac{b+1}{a} \implies a+b = -1$$

Mais  $a$  et  $b$  étant tous deux positifs, la somme ne peut être négative. Ayant abouti à une absurdité, on en conclue que nos hypothèses sont fausses.  $\square$

### Démonstration par récurrence

La démonstration par récurrence est une méthode de démonstration spécifique permettant de montrer qu'une propriété  $\mathcal{P}(n)$  est vraie pour tout  $n \in \mathbb{N}$ .

Elle se déroule en deux étapes :

1. *Initialisation* : on montre que pour un entier fixé  $n_0$  (souvent 0 ou 1), la propriété est vraie.
2. *Hérédité* : on suppose que pour un entier  $n$  donné,  $\mathcal{P}(n)$  est vraie, et l'on montre qu'alors  $\mathcal{P}(n + 1)$  est vraie également.
3. *Conclusion* : la propriété  $\mathcal{P}(n)$  étant initialisée et héréditaire, elle est vraie pour tout  $n \geq n_0$ .

**Exemple** : montrons par récurrence que pour tout  $n \in \mathbb{N}^*$ , on a

$$\mathcal{P}(n) : 1 + 2 + \dots + n = \frac{n(n+1)}{2}$$

1. *Initialisation* :  $\mathcal{P}(1)$  se vérifie assez vite, puisque  $\frac{1(1+1)}{2} = 1$ .
2. *Hérédité* : supposons qu'il existe un entier  $n$  tel que  $\mathcal{P}(n)$  est vraie et calculons  $1 + 2 + \dots + n + (n + 1)$  : comme  $\mathcal{P}(n)$  est vraie, on a

$$\begin{aligned} 1 + 2 + \dots + n + (n + 1) &= \frac{n(n+1)}{2} + (n + 1) \\ &= \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2} \end{aligned}$$

On retrouve l'expression de la propriété, au rang  $(n + 1)$  et donc  $\mathcal{P}(n + 1)$  est vraie.

3. *Conclusion* : pour tout  $n \in \mathbb{N}^*$ , la proposition  $\mathcal{P}(n)$  est vraie. □

## 1.3 Ensembles et applications

### 1.3.1 Notion d'ensemble, d'élément

Un *ensemble* est une collection d'objets ayant une propriété en commun. Pour symboliser un ensemble, on utilise les accolades ( $\{\dots\}$ ). Dans un ensemble, l'ordre n'importe pas. (Ainsi, les ensembles  $\{1, 2, 3\}$  et  $\{1, 3, 2\}$  sont égaux.)

**Exemples** :

— L'ensemble des triangles isocèles.

- L'ensemble  $\mathbb{N}$  des entiers naturels, l'ensemble  $\mathbb{Z}$  des entiers relatifs, l'ensemble  $\mathbb{R}$  des nombres réels.
- L'ensemble des nombres positifs réels  $\{x \in \mathbb{R}, x \geq 0\}$ .
- L'ensemble  $\{1, 2\}$  est l'ensemble des solutions de l'équation  $x^2 - 3x + 2 = 0$ .

Les objet d'un ensemble sont appelés ses *éléments*. Pour indiquer qu'un objet donné est ou n'est pas un élément d'un ensemble donné, on utilise le symbole  $\in$ . Ainsi, si  $x$  est un élément de l'ensemble  $E$ , on note  $x \in E$ . Sinon, on note  $x \notin E$ .

**Exemples :**

•	$5 \dots \mathbb{N}$	$5 \dots \mathbb{Z}$	$5 \dots \mathbb{Q}$	$5 \dots \mathbb{R}$	$5 \dots \mathbb{C}$
•	$-2 \dots \mathbb{N}$	$-2 \dots \mathbb{Z}$	$-2 \dots \mathbb{Q}$	$-2 \dots \mathbb{R}$	$-2 \dots \mathbb{C}$
•	$\sqrt{2} \dots \mathbb{N}$	$\sqrt{2} \dots \mathbb{Z}$	$\sqrt{2} \dots \mathbb{Q}$	$\sqrt{2} \dots \mathbb{R}$	$\sqrt{2} \dots \mathbb{C}$
•	$1 - i \dots \mathbb{N}$	$1 - i \dots \mathbb{Z}$	$1 - i \dots \mathbb{Q}$	$1 - i \dots \mathbb{R}$	$1 - i \dots \mathbb{C}$

Si  $E$  est un ensemble fini, on appelle *cardinal* de  $E$  le nombre d'éléments qu'il contient. On le note  $\text{Card}(E)$  ou  $|E|$ . On verra dans le chapitre suivant comment on peut également définir le cardinal d'un ensemble infini.

**Exemples :**

- L'ensemble vide, que l'on note  $\emptyset$  est de cardinal 0.
- Un ensemble de cardinal 1 est appelé *singleton*. Attention à ne pas confondre le singleton  $\{a\}$  et l'élément  $a$ .
- Un ensemble à deux éléments est un *doublon*, ou une *paire*. Attention à ne pas confondre la paire  $\{a, b\}$  avec le couple  $(a, b)$ .

### 1.3.2 Les différentes façons de définir un ensemble

#### La forme étendue ou explicite

La façon la plus simple de définir un ensemble est de donner explicitement tous les éléments de cet ensemble entre accolades. C'est le plus pratique, quand l'ensemble en question ne contient pas trop d'éléments (en particulier, lorsque l'ensemble contient un nombre fini d'éléments...). C'est ce que l'on appelle la représentation étendue.

**Exemples :**

- $E = \{2, 3, 5, 7, 9\}$ .

—  $E = \{a, b, c, d\}$ .

### La forme implicite

La forme explicite rend difficile la représentation des ensembles infinis. On peut par exemple représenter l'ensemble des entiers naturels sous la forme  $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ , mais l'utilisation des “...” est imprécise. D'autre part, il est des ensembles (comme  $\mathbb{R}$  ou les intervalles de  $\mathbb{R}$ ) que l'on ne peut même pas donner à l'aide des “...”. Pour de tels ensembles, on utilise alors la propriété commune à tous les éléments de cet ensemble. Ainsi, l'ensemble des éléments qui vérifient une propriété  $\mathcal{P}$  donnée est

$$E = \{x / \mathcal{P}(x)\}.$$

**Exemples :**

- L'ensemble  $E_1 = \{x \in \mathbb{N} / x \leq 11\}$ .
- L'ensemble  $E_2 = \{\text{les entiers naturels qui sont des nombres premiers}\}$ .

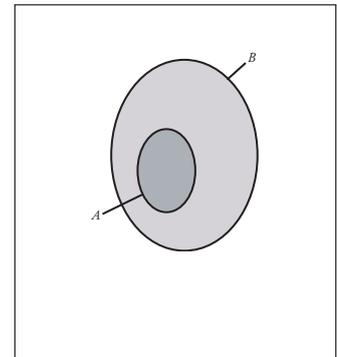
### 1.3.3 Inclusion

#### Définitions

Considérons deux ensembles  $A$  et  $B$ . On dit que  $A$  est inclus dans  $B$  si tous les éléments de  $A$  appartiennent également à  $B$ . On le note  $A \subset B$ . Ainsi,

$$(A \subset B) \iff (x \in A \Rightarrow x \in B).$$

Si  $A$  est inclus dans  $B$ , on dit que  $A$  est *un sous ensemble* ou *une partie* de  $B$ .



**Remarques :**

— N'importe quel ensemble contient l'ensemble vide et se contient lui-même :

$$\forall A, \quad \emptyset \subset A \text{ et } A \subset A.$$

— Deux ensembles  $A$  et  $B$  sont égaux s'ils chacun est inclus dans l'autre :

$$(A = B) \Leftrightarrow (A \subset B \text{ et } B \subset A).$$

Pour démontrer que deux ensembles sont égaux, on résonne donc en deux étapes : on montre que le premier ensemble est inclus dans le second, puis que le second est inclus dans le premier.

### 1.3.4 Opération sur les ensembles

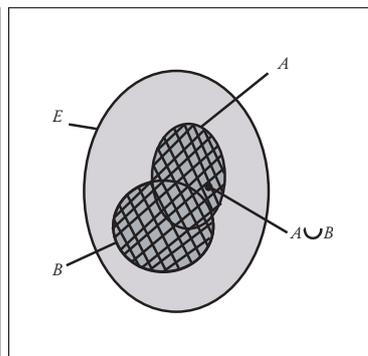
Soit  $E$  un ensemble et soient  $A$  et  $B$  deux sous-ensembles de  $E$ . On peut effectuer des opérations sur les ensembles  $A$  et  $B$  pour construire d'autres parties de  $E$  :

• L'*union* (ou *réunion*) des ensembles  $A$  et  $B$  est l'ensemble de tous les éléments appartenant à  $A$  **ou** à  $B$ . On le note  $A \cup B$  :

$$A \cup B = \{x / x \in A \text{ ou } x \in B\}.$$

**Note** : on a toujours

$$A \subset A \cup B \text{ et } B \subset A \cup B$$



• L'*intersection* des ensembles  $A$  et  $B$  est l'ensemble de tous les éléments appartenant à la fois à  $A$  **et** à  $B$ . On le note  $A \cap B$  :

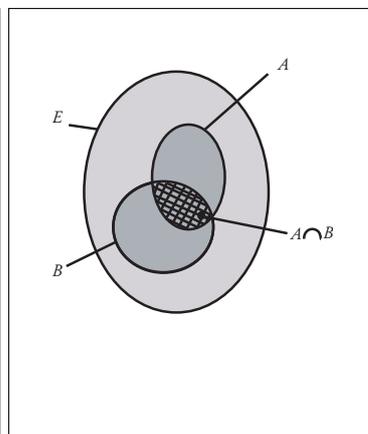
$$A \cap B = \{x / x \in A \text{ et } x \in B\}.$$

**Notes** :

— Deux ensembles  $A$  et  $B$  sont dits *disjoints* si  $A \cap B = \emptyset$ .

— On a toujours

$$A \cap B \subset A \text{ et } A \cap B \subset B$$



**Note** : à l'image des opérations bien connues dans les nombres réels, ces opérations respectent certaines règles de calcul. Ainsi, on peut montrer que les opérations  $\cap$  et  $\cup$  commutatives :

$$\forall A, B \subset E, A \cup B = B \cup A \text{ et } A \cap B = B \cap A$$

De même, on peut montrer qu'elles sont associatives :

$$\begin{aligned} \forall A, B, C \subset E, (A \cup B) \cup C &= A \cup (B \cup C) \\ (A \cap B) \cap C &= A \cap (B \cap C) \end{aligned}$$

Cela permet en particulier d'étendre ces opérations à des familles de sous-ensembles et de définir notamment la notion de partition : une famille  $A_1, A_2, \dots, A_n$  de parties de  $E$  forme une *partition* de  $E$  si et seulement si

$$\begin{cases} \forall i, j \in \{1, \dots, n\}, i \neq j \Rightarrow A_i \cap A_j = \emptyset \\ A_1 \cup A_2 \cup \dots \cup A_n = E \end{cases}$$

Enfin, on peut également montrer que l'inclusion ( $\cap$ ) est distributive sur l'union ( $\cup$ ) et réciproquement :

$$\begin{aligned} \forall A, B, C \subset E, \quad A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \\ \forall A, B, C \subset E, \quad A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \end{aligned}$$

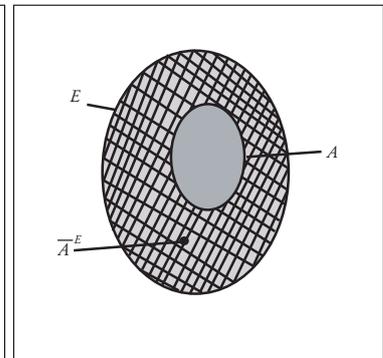
Pour finir, citons le complémentaire :

- Le *complémentaire* de  $A$  dans  $E$  est l'ensemble des éléments de  $E$  qui n'appartiennent pas à l'ensemble  $A$ . On le note  $\overline{A}^E$  :

$$\overline{A}^E = \{x \in E / x \notin A\}$$

**Note** : on a toujours

$$A \cap \overline{A}^E = \emptyset \quad \text{et} \quad A \cup \overline{A}^E = E$$



À partir de ces opérations de base, il est alors possible d'établir une arithmétique des ensembles. On peut en particulier faire une analogie forte entre ces trois opérations et les opérations sur les propositions logiques évoquées aux paragraphes précédents.

Enfin, il est également possible de définir de nouvelles opérations à partir de ces opérations élémentaires. Par exemple, on appelle *différence* l'opérations

$$B \setminus A = B \cap \overline{A}$$

Exercice : représenter graphiquement la différence  $B \setminus A$ .

### 1.3.5 Les parties d'un ensemble

Étant donné un ensemble  $E$ , on peut considérer l'ensemble de ses parties. On le note  $\mathcal{P}(E)$ . Il contient (comme son nom l'inique) tous les sous ensembles de  $E$ .

$$\mathcal{P}(E) = \{X / X \subset E\}.$$

**Note** : les éléments de  $\mathcal{P}(E)$  sont des ensembles !

**Exemple** : Si  $E = \{a, b, c\}$ , alors

$$\mathcal{P}(E) = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, E, \emptyset\}$$

Les opérations que l'on vient de voir sont donc des opérations dans  $\mathcal{P}(E)$ . Elle permettent de construire un élément de  $\mathcal{P}(E)$  à partir d'autres éléments de  $\mathcal{P}(E)$ .

De plus que l'ensemble des parties d'un ensemble  $E$  contient toujours au moins deux éléments : l'ensemble vide et l'ensemble  $E$  lui-même.

### 1.3.6 Cardinal d'un ensemble fini

#### Formule du crible

Soient  $A$  et  $B$  deux parties de  $E$ . Il est clair que si  $A$  et  $B$  sont deux ensembles finis inclus l'un dans l'autre, alors  $\text{Card}(A) \leq \text{Card}(B)$ . Ainsi on a  $\text{Card}(A \cup B) \geq \text{Card}(A)$  et  $\text{Card}(B)$ . D'autre part, si  $A$  et  $B$  sont disjoints, alors

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B).$$

La *formule du crible* donne le cas général :

#### Formule du Crible :

Soient  $A$  et  $B$  deux parties de  $E$ .

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$$

**Note** : on verra en TD qu'il est possible d'étendre cette formule à la réunion d'un nombre quelconque de parties finies.

#### Cardinal de $\mathcal{P}(E)$

Soit  $E$  un ensemble fini de cardinal  $n$ . L'ensemble  $\mathcal{P}(E)$  est encore un ensemble fini. Pour calculer le cardinal de  $\mathcal{P}(E)$ , on doit compter le nombre d'ensembles différents que l'on peut construire à partir des éléments de  $E$ . Or pour construire un sous ensemble de  $E$ , on considère un à un les éléments de  $E$ , et pour chacun, on a 2 possibilités :

- soit on le prend,
- soit on ne le prend pas.

On dénombre ainsi  $2^n$  possibilités. D'où la formule générale :

$$\text{Card}(\mathcal{P}(E)) = 2^{\text{Card}(E)}.$$

### 1.3.7 Produit cartésien d'ensembles

À partir de deux ensembles  $E$  et  $F$  donnés, on peut construire l'ensemble produit  $E \times F$  qui est l'ensemble des couples  $(e, f)$  tels que  $e$  est un élément de  $E$  et  $f$  est un élément de  $F$  :

$$E \times F = \{(e, f), e \in E, f \in F\}$$

Par exemple, si  $E = \{1, 2, 3\}$  et  $F = \{a, b\}$ , l'ensemble produit  $E \times F$  est l'ensemble

$$E \times F = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}.$$

**Attention** : un couple  $(e, f)$  n'est pas un ensemble. En particulier, l'ordre est fondamental. Ainsi,  $(e, f) \neq (f, e)$ . On dit alors que le produit cartésien n'est pas commutatif (car les ensembles  $E \times F$  et  $F \times E$  ne sont pas les mêmes).

Si les deux ensembles  $E$  et  $F$  sont donnés sous forme étendue ( $E = \{a_1, a_2, \dots, a_n\}$  et  $F = \{b_1, \dots, b_m\}$ ), le produit cartésien  $E \times F$  peut être représenté par le diagramme cartésien ci-contre. (Les éléments de l'ensemble  $E \times F$  sont les cases du tableau.)

		$\overbrace{\hspace{10em}}^F$				
		×	$b_1$	$b_2$	$\cdots$	$b_m$
{	$a_1$			$\cdots$		
	$a_2$			$\cdots$		
	$\vdots$			$\ddots$		
	$a_n$			$\cdots$		

D'après le tableau ci-dessus, on voit que si  $E$  et  $F$  sont deux ensembles finis, alors

$$\text{Card}(E \times F) = \text{Card}(E) \times \text{Card}(F).$$

On peut généraliser ce que l'on vient de voir à plus de deux ensembles. Ainsi, l'on note  $E_1, E_2, \dots, E_n$   $n$  ensembles, on peut construire le produit cartésien  $E_1 \times E_2 \times \dots \times E_n$  dont les éléments sont tous les  $n$ -uplets  $(e_1, e_2, \dots, e_n)$  tels que pour tout  $i \in \{1, \dots, n\}$ ,  $e_i \in E_i$  :

$$E_1 \times \dots \times E_n = \{(e_1, \dots, e_n) / e_i \in E_i \forall i \in \{1, \dots, n\}\}$$

Si tous les ensembles  $E_i$  sont un même ensemble  $E$ , on note  $E^n$  pour le produit  $\underbrace{E \times E \times \dots \times E}_{n \text{ fois}}$ .

**Cas particuliers : les ensembles  $\mathbb{R}^2$ ,  $\mathbb{R}^3$ ,  $\mathbb{R}^n$** 

D'après ce que l'on vient de voir, on peut définir l'ensemble  $\mathbb{R}^2$  des couples de réels :

$$\mathbb{R}^2 = \{(x, y), x \in \mathbb{R}, y \in \mathbb{R}\}.$$

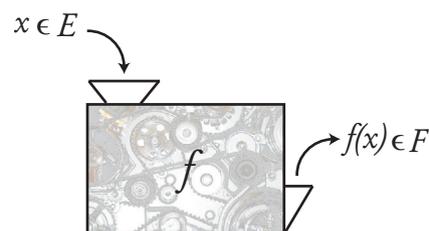
Si l'on munit le plan d'un repère, on peut alors représenter tout élément  $(x, y)$  de  $\mathbb{R}^2$  par le point de coordonnées  $x$  et  $y$  dans ce repère. Cela nous donne une représentation graphique de  $\mathbb{R}^2$  (et de tous ses sous ensembles).

De même, l'ensemble  $\mathbb{R}^3 = \{(x, y, z), x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R}\}$  peut être représenté par l'espace 3D, muni d'un repère.

Enfin, pour tout  $n \in \mathbb{N}$ , on peut définir l'ensemble  $\mathbb{R}^n = \{(x_1, \dots, x_n), x_i \in \mathbb{R}\}$ . Cet ensemble joue un rôle très particulier en mathématiques. On sera amenés à le retrouver dans la suite.

**1.3.8 Applications****Définitions**

Si  $E$  et  $F$  sont deux ensembles, une application (ou fonction)  $f$  de  $E$  dans  $F$  est une opération qui permet d'associer certains éléments de  $F$  à des éléments de  $E$ , telle que pour tout  $x$  de  $E$  il existe un unique élément  $y \in F$  qui lui soit associé. On peut également voir une fonction comme une machine transformant les éléments de  $E$  en éléments de  $F$ .



On note

$$\begin{aligned} f &: E \longrightarrow F \\ x &\longmapsto f(x) \end{aligned}$$

L'étude des fonctions est avant tout basée sur un vocabulaire précis. Ainsi,

**Définitions**

Soient  $E$  et  $F$  deux ensembles et  $f$  une application de  $E$  dans  $F$ .

- L'ensemble  $E$  est *l'ensemble source de  $f$* , ou ensemble de départ.
- L'ensemble  $F$  est *l'ensemble but de  $f$* , ou ensemble d'arrivée.

D'autre part, une fonction  $f : E \rightarrow F$  mettant en relation les ensemble  $E$  et  $F$ , elle met en relation les éléments de  $E$  et ceux de  $F$ , ainsi que les sous ensembles de  $E$  et  $F$ . Ainsi, soient  $E$  et  $F$  deux ensembles et  $f$  une application de  $E$  dans  $F$ .

#### Définitions (Image, image directe)

— Pour tout  $x \in E$ , l'élément  $f(x) \in F$  est appelé *l'image de  $x$  par  $f$* .

**Note :** tout élément de  $E$  admet exactement une image par  $f$ .

— Pour toute partie  $A \subset E$ , on appelle *image directe de  $A$  par  $f$*  le sous ensemble de  $F$  formé des images par  $f$  des éléments de  $A$ . On le note  $f(A)$  :

$$f(A) = \{f(x), x \in A\} = \{y \in F / \exists x \in A / y = f(x)\}$$

**Note :** l'ensemble  $f(E)$ , image directe de  $E$  tout entier par  $f$ , est également appelé *image de  $f$*  et parfois noté  $\text{Im}(f)$ .

#### Définitions (Antécédant, image réciproque)

— Pour tout élément  $y \in \text{Im}(f)$ , on appelle *antécédant de  $y$  par  $f$*  tout élément  $x \in E$  tel que  $f(x) = y$ .

**Note :** un élément  $y \in \text{Im}(f)$  peut avoir plusieurs antécédants dans  $E$ .

— Pour toute partie  $B \subset F$ , on appelle *image réciproque de  $B$  par  $f$*  le sous ensemble de  $E$  formé de tous les éléments de  $E$  dont l'image est dans  $B$ . On le note  $f^{-1}(B)$  :

$$f^{-1}(B) = \{x \in E / f(x) \in B\}$$

**Note :** on verra dans la suite que la notation  $f^{-1}$  peut également représenter une fonction intimement liée à  $f$ , dans le cas particulier où la fonction  $f$  est *bijective*. Notons cependant que l'image réciproque  $f^{-1}(B)$  existe même quand la fonction  $f^{-1}$  n'existe pas.

**Exercice :** Soient  $E = \{-2, -1, 0, 1, 2\}$  et  $F = \{0, 1, 2, 3, 4\}$ . On considère la fonction  $f$  définie par

$$\begin{array}{ccc} f : E & \longrightarrow & F \\ x & \longmapsto & x^2 \end{array}$$

1. Calculer l'image de chaque élément de  $E$  (on vérifiera que c'est bien à chaque fois un élément de  $F$ ) :

$$f(-2) = 4, \quad f(-1) = 1, \quad f(0) = 0, \quad f(1) = 1, \quad f(2) = 4$$

2. (a) Déterminer l'image  $\text{Im}(f)$  de  $f$  dans  $F$  :

$$\text{Im}(f) = \{0, 1, 4\}$$

(b) Déterminer l'image du sous ensemble  $A = \{-2, 0, 2\}$  :

$$f(A) = \{0, 4\}$$

3. (a) Déterminer les antécédents de chacun des éléments de  $\text{Im}(f)$  :

$$f^{-1}(\{0\}) = \{0\}, \quad f^{-1}(\{1\}) = \{-1, 1\}, \quad f^{-1}(\{2\}) = \{-2, 2\}$$

(b) Déterminer l'image réciproque de  $B = \{0, 1, 2\} \subset F$  :

$$f^{-1}(B) = \{-1, 0, 1\}$$

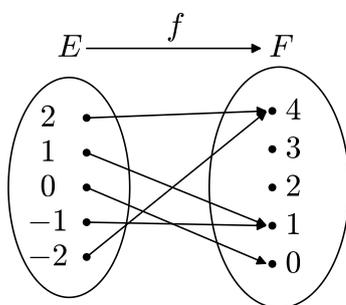
### Représentations graphiques

Selon la nature des ensembles  $E$  et  $F$ , il existe différentes représentations graphiques d'une fonction  $f : E \rightarrow F$  :

- *Le diagramme de Venn*, ou *diagramme sagittal* (ou diagramme "en patates") : les ensembles  $E$  et  $F$  sont représentés par des ovals distincts, leurs éléments sont des points et les associations  $x \mapsto f(x)$  sont représentées par des flèches.

**Exemple** : construire le diagramme de Venn de la fonction

$$f : \begin{array}{ccc} \{-2, -1, 0, 1, 2\} & \longrightarrow & \{0, 1, 2, 3, 4\} \\ x & \longmapsto & x^2 \end{array}$$

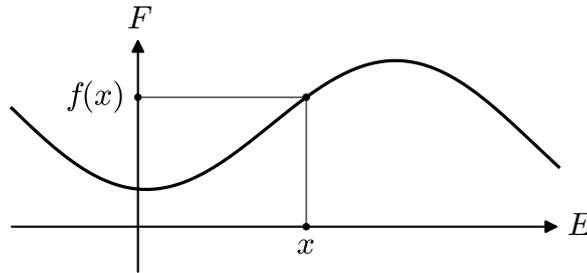


- Dans le cas particulier où  $E$  et  $F$  sont des parties de  $\mathbb{R}$ , on peut également représenter on fonction  $f$  par *son graphe* dans le plan muni d'un repère, l'ensemble  $E$  étant représenté en abscisses, l'ensemble  $F$  en ordonnées.

**Définition (graphe d'une fonction)**

On appelle *graphe de  $f$*  l'ensemble  $\mathcal{G}_f$  défini par

$$\mathcal{G}_f = \{(x, f(x)), x \in E\}$$



**Note :** formellement, on peut définir le graphe de  $f$  quelque soient les ensembles but et source. L'ensemble  $\mathcal{G}_f$  est alors un sous ensemble de  $E \times F$ .

**Quelques fonctions particulières**

**Définition (fonctions constantes)**

Soient  $E$  et  $F$  deux ensembles. On appelle *fonction constante de  $E$  dans  $F$*  toute fonctions  $f$  telle que

$$\exists c \in F / \forall x \in E, f(x) = c$$

**Définition (fonction identité)**

Soient  $E$  un ensemble. On appelle *fonction identité de  $E$*  la fonction notée  $\text{Id}_E$  définie par

$$\begin{aligned} \text{Id}_E : E &\longrightarrow E \\ x &\longmapsto x \end{aligned}$$

**Définition (égalité de deux fonctions)**

Soient  $E$  et  $F$  deux ensembles. Deux fonctions  $f$  et  $g$  de  $E$  dans  $F$  sont dite *égales* si et seulement si

$$\forall x \in E, f(x) = g(x)$$

Exercice : à quelle condition deux fonctions définies sur un même ensemble ne sont pas égales ?

**Fonctions composées****Définition (fonctions composées)**

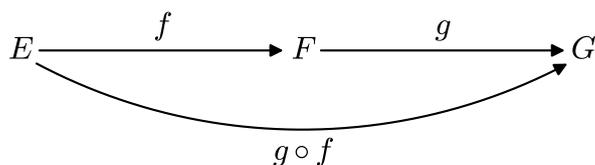
Soient  $E$ ,  $F$  et  $G$  trois ensembles et soient

$$f : E \rightarrow F \quad \text{et} \quad g : F \rightarrow G$$

deux fonctions. On appelle *fonction composée de  $f$  et  $g$*  l'application notée  $g \circ f$  de  $E$  dans  $G$  définie par

$$\begin{aligned} g \circ f &: E \longrightarrow G \\ x &\longmapsto g(f(x)) \end{aligned}$$

que l'on peut résumer par le diagramme ci-dessous.



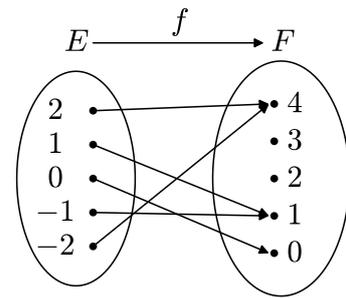
**Note** : si  $E = F$  (i.e. si  $f$  est une fonction de  $E$  dans lui même, on peut appliquer à un même  $x \in E$  la fonction  $f$  plusieurs fois. On obtient ainsi la *composée itérée* de  $f$ . On note

$$f^n = \underbrace{f \circ f \circ \dots \circ f}_{n \text{ fois}}$$

## Injections, surjections, bijections

Concernant la fonction évoquée plus haut et dont le digramme sagittal est reproduit ci-contre, on peut faire plusieurs remarques :

- Il y a des éléments de  $F$  qui ne sont pas atteints : 2 et 3.
- Il y a des éléments de  $F$  qui sont atteints plusieurs fois : 1 et 4.



Ces remarques sont à l'origine des définitions ci-dessous, qui permettent de caractériser certaines propriétés éventuelles des fonctions étudiées.

## Fonctions injectives

### Définition (fonction injective)

Une fonction  $f : E \rightarrow F$  est dite *injective* si deux éléments distincts de  $E$  ont leurs images différentes (i.e. “toutes les images sont différentes”).

Autrement dit,  $f$  est injective si

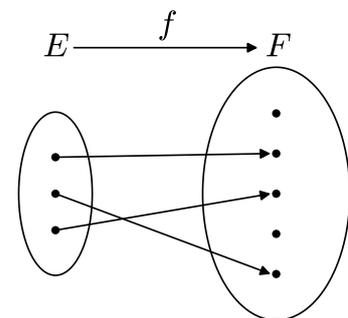
$$\forall x_1, x_2 \in E, x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

**Note** : on préfère en général travailler avec la *contraposée* de cette définition :

$$f \text{ est injective} \Leftrightarrow (\forall x_1, x_2 \in E, f(x_1) = f(x_2) \Rightarrow x_1 = x_2)$$

### Exemples :

- La fonction représentée ci-contre est injective.



- La fonction étudiée en exemple n'est pas injective car  $f(-1) = f(1)$ .

- La fonction La fonction  $\left[ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin(x) \end{array} \right]$  n'est pas injective, mais la fonction

$$\left[ \begin{array}{ccc} \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin(x) \end{array} \right] \text{ est injective.}$$

## Fonctions surjectives

### Définition (fonction surjective)

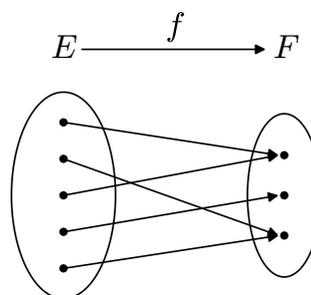
Une fonction  $f : E \rightarrow F$  est *surjective* si tout élément de  $F$  l'image par  $f$  d'un élément de  $E$ .

Ainsi,  $f : E \rightarrow F$  est surjective si et seulement si

$$\forall y \in F, \exists x \in E / y = f(x)$$

### Exemples :

— La fonction représentée ci-contre est surjective.



— La fonction étudiée en exemple n'est pas surjective car  $2 \in F$  n'est l'image d'aucun élément de  $E$ .

— La fonction  $\left[ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin(x) \end{array} \right]$  n'est pas surjective, mais la fonction  $\left[ \begin{array}{ccc} \mathbb{R} & \longrightarrow & [-1, 1] \\ x & \longmapsto & \sin(x) \end{array} \right]$  est surjective.

## Fonctions bijectives

### Définition (fonction bijective)

Une fonction  $f : E \rightarrow F$  est *bijective* si elle est à la fois injective et surjective.

On peut montrer que cela se traduit par la proposition suivante :

### Proposition

Une fonction  $f : E \rightarrow F$  est bijective si et seulement si

$$\forall y \in F, \exists! x \in E / y = f(x)$$

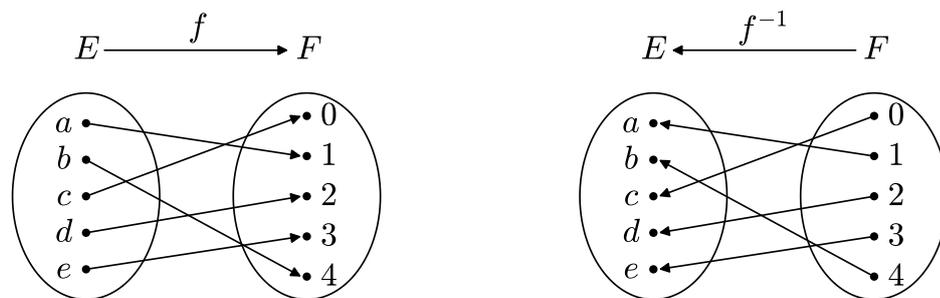
Ainsi, si  $f$  est une fonction bijective de  $E \rightarrow F$ , chaque élément de  $F$  admet un et un seul antécédent par  $f$  dans  $E$ . On peut donc considérer l'application de  $F \rightarrow E$  qui à chaque élément de  $F$  associe cet unique antécédent.

Cette application est appelée *l'application réciproque de  $f$* . Elle est notée  $f^{-1}$  et l'on peut montrer que

$$f \circ f^{-1} = \text{Id}_F \quad \text{et} \quad f^{-1} \circ f = \text{Id}_E.$$

### Exemples :

- Les deux fonctions représentées ci-dessous sont bijectives. Chacune est la bijection réciproque de l'autre.



- La fonction  $\begin{bmatrix} \mathbb{R}^+ & \longrightarrow & \mathbb{R}^+ \\ x & \longmapsto & x^2 \end{bmatrix}$  est une bijection. Sa bijection réciproque est  $\begin{bmatrix} \mathbb{R}^+ & \longrightarrow & \mathbb{R}^+ \\ x & \longmapsto & \sqrt{x} \end{bmatrix}$
- La fonction  $\exp : \mathbb{R} \rightarrow \mathbb{R}_*^+$  est une bijection. Sa bijection réciproque est la fonction  $\ln : \mathbb{R}_*^+ \rightarrow \mathbb{R}$ .

### Applications et cardinaux

Si  $E$  et  $F$  sont des ensembles finis, le caractère injectif ou surjectif d'une application de  $E$  dans  $F$  nous renseigne sur les cardinaux de chacun de ces ensembles.

Précisément,

— Si l'on peut construire une application *injective* de  $E$  dans  $F$ , c'est que

$$\text{Card}(E) \leq \text{Card}(F)$$

En effet, il doit y avoir, dans  $F$ , suffisamment d'éléments pour pouvoir associer un élément différent à chaque élément de  $E$ .

— Si l'on peut construire une application *surjective* de  $E$  dans  $F$ , c'est que

$$\text{Card}(E) \geq \text{Card}(F)$$

En effet, il doit y avoir, dans  $E$ , suffisamment d'éléments pour pouvoir en associer un à chaque élément de  $F$ .

— Si l'on peut construire une application *bijection* de  $E$  dans  $F$ , c'est que

$$\text{Card}(E) = \text{Card}(F)$$

En étendant ces idées aux ensembles infinis, il est alors possible de comparer la taille de deux ensembles infinis (on parle encore de cardinal pour ces ensembles infinis). Précisément deux ensembles infinis  $E$  et  $F$  sont dits de même cardinal s'il est possible de construire une *bijection* de l'un dans l'autre (on dit alors que  $E$  et  $F$  sont équipotents). C'est ainsi que l'on peut montrer que les ensembles  $\mathbb{N}$ ,  $\mathbb{N}^2$ ,  $\mathbb{Z}$  et  $\mathbb{Q}$  sont équipotents.

D'autre part, une étude poussée de ces notions a permis de montrer qu'il existait plusieurs tailles d'infini.

On peut en particulier montrer qu'il n'existe aucune bijection possible entre un ensemble  $E$  et l'ensemble  $\mathcal{P}(E)$  de ses parties. Autrement dit, l'ensemble  $\mathcal{P}(E)$  est toujours significativement plus grand que  $E$ .

Par ailleurs, on peut montrer (notamment à l'aide de l'écriture décimale des nombres réels) qu'il existe une bijection entre  $\mathcal{P}(\mathbb{N})$  et  $\mathbb{R}$ . Autrement dit, le cardinal de  $\mathbb{R}$  est significativement plus grand que celui de  $\mathbb{N}$  et on peut construire une échelle de tailles :

- Le cardinal de  $\mathbb{N}$  (et donc de  $\mathbb{Z}$ , de  $\mathbb{Q}$ , etc) est noté  $\aleph_0$ . C'est l'infini *dénombrable*.
- Le cardinal de  $\mathbb{R}$  est noté  $\aleph_1$ . On dit que  $\mathbb{R}$  a *la puissance du continu* et par convention, on a  $\aleph_1 = 2^{\aleph_0}$ .

**Note** : une question importante qui divise toujours la communauté des mathématiciens est de savoir s'il existe une taille d'infini comprise entre  $\aleph_0$  et  $\aleph_1$ . Autrement dit, existe-t-il un ensemble qui contient strictement plus d'éléments que  $\mathbb{N}$  mais strictement moins d'éléments que  $\mathbb{R}$ .

## 1.4 Techniques fondamentales du calcul algébrique

### 1.4.1 Propriétés des opérations

#### Propriétés élémentaires

L'addition (“+”) et la multiplication (“×” ou “.”) ont trois propriétés élémentaires :

— *La commutativité :*

$$\forall x, y \in \mathbb{C}, \quad x + y = y + x, \quad \text{et} \quad x \times y = y \times x$$

— *L'associativité :*

$$\forall x, y, z \in \mathbb{C}, \quad (x + y) + z = x + (y + z), \quad \text{et} \quad (x \times y) \times z = x \times (y \times z)$$

**Note :** c'est cette propriété d'associativité qui permet d'étendre la notion de somme à plus de deux nombres réels ou complexes.

— *La distributivité de la multiplication sur l'addition :*

$$\forall x, y, z \in \mathbb{C}, \quad (x + y) \times z = x \times z + y \times z$$

Elles possèdent également quelques propriétés plus élaborées :

— *Les éléments neutres :* pour les deux opérations sur  $\mathbb{C}$ , on appelle élément neutre les nombres “qui ne font rien”. Ainsi,

— L'élément neutre de l'addition est **le nombre 0** :

$$\forall x \in \mathbb{C}, \quad x + 0 = x$$

— L'élément neutre de la multiplication est **le nombre 1** :

$$\forall x \in \mathbb{C}, \quad 1 \cdot x = x$$

— *La notion d'inverse :* pour chaque opération, on appelle inverse d'un nombre  $x$  tout nombre qui, combiné à  $x$  ramène à l'élément neutre. Ainsi,

— *Inverse pour l'addition* : pour tout  $x \in \mathbb{C}$ , il existe un nombre  $x'$  tel que

$$x + x' = 0$$

Ce nombre est alors noté  $-x$  et est en général appelé *l'opposé*.

— *Inverse pour la multiplication* : pour tout  $x \in \mathbb{C}$  non nul, il existe un nombre  $x''$  tel que

$$x \times x'' = 1$$

Ce nombre est alors noté  $\frac{1}{x}$  ou  $x^{-1}$ .

Ce sont en particulier ces notions d'inverses qui permettent, dans une équation, d'isoler la variable inconnue et de l'exprimer ainsi en fonction des autres : dans la résolution de l'équation  $a.x + b = 0$ , les étapes sont

$$\begin{aligned} a.x + b = 0 &\Leftrightarrow a.x + \underbrace{b + (-b)}_0 = 0 + (-b) \Leftrightarrow a.x + 0 = -b \Leftrightarrow a.x = -b \\ &\Leftrightarrow \frac{1}{a} \cdot (a.x) = \frac{1}{a} \cdot (-b) \Leftrightarrow \underbrace{\left(\frac{1}{a} \cdot a\right)}_1 .x = -\frac{b}{a} \Leftrightarrow 1.x = -\frac{b}{a} \\ &\Leftrightarrow \boxed{x = -\frac{b}{a}} \end{aligned}$$

Notons enfin que le nombre 0 est dit *absorbant* pour la multiplication. Cela traduit le fait que le produit de tout nombre par 0 donne 0.

On a en outre une propriété supplémentaire concernant le nombre 0 et la multiplication : dans l'ensemble des nombres (réels ou complexes), le produit de deux nombres non nuls ne peut être nul. Sa contraposée est alors un outil puissant pour la résolution de certaines équations : un produit de nombres (réels ou complexes) est nul si et seulement si l'un de ses facteurs est nul.

## Développer, factoriser

Selon les calculs que l'on a effectuer, il peut être utile de travailler sur des sommes ou sur des produits. Le passage d'une expression à l'autre est l'une des bases du calcul algébrique, exploitant la distributivité de la multiplication sur l'addition. Ainsi, développer une expression, c'est passer d'un produit à une somme et factoriser, c'est passer d'une somme à un produit :

$$\begin{array}{ccc}
 & \xrightarrow{\text{développer}} & \\
 a.(x+y) & = & a.x + a.y \\
 & \xleftarrow{\text{factoriser}} &
 \end{array}$$

### Identités remarquables

La méthode première pour factoriser une expression algébrique donnée consiste à chercher un facteur commun ( $a$ ) à chacun des termes de la somme étudiée. Cependant, outre l'observation et la recherche de facteurs communs, la factorisation peut également se faire à l'aide d'identités remarquables (directement issues des règles de calcul ci-dessus). Ainsi, pour tout  $a, b \in \mathbb{C}$ , on a

$$\begin{aligned}
 a^2 + 2ab + b^2 &= (a + b)^2 \\
 a^2 - 2ab + b^2 &= (a - b)^2 \\
 a^2 - b^2 &= (a + b).(a - b)
 \end{aligned}$$

On verra plus loin que ces formules ne sont que quelques cas particulier d'une classe plus générale de formules.

**Exercice** : retrouver ces résultats en repartant de la définition du carré et en appliquant les règles de distribution.

### 1.4.2 Le signe $\Sigma$

#### Définitions

Le symbole  $\Sigma$  est utilisé pour symboliser des sommes constituées d'un grand nombre voire d'un nombre indéterminé de termes. Ainsi étant donnés  $n$  nombres complexes  $a_1, a_2, \dots, a_n$ , on note

$$a_1 + a_2 + \dots + a_n = \sum_{k=1}^n a_k = \sum_{1 \leq k \leq n} a_k$$

#### Définitions

Dans la somme ci-dessus,

- $a_k$  est appelé *le terme général* de la somme,
- $k$  est appelé *l'indice de sommation* de la somme.

**Exemples :**

— Pour tout  $n \in \mathbb{N}^*$ , on note

$$1 + 2 + \cdots + n = \sum_{k=1}^n k$$

— Pour tout  $x \in \mathbb{C}$  et pour tout  $n \in \mathbb{N}$ , on note

$$1 + x + x^2 + \cdots + x^n = \sum_{k=0}^n x^k$$

**Notes :**

— Bien que l'on travaille principalement avec un indice de sommation ( $k$ ) parcourant l'ensemble des entiers, l'indice de sommation, peut, de façon générale, parcourir n'importe quel type d'ensemble  $I$ . On note alors  $\sum_{k \in I} a_k$

— Par convention, on pose  $\sum_{k \in \emptyset} a_k = 0$ .

**Transformations d'écriture**

Le signe  $\Sigma$  représentant une somme, il possède certaines propriétés directement issues des propriétés de l'addition. Ainsi

— L'addition étant commutative

$$\sum_{k \in I} (a_k + b_k) = \left( \sum_{k \in I} a_k \right) + \left( \sum_{k \in I} b_k \right)$$

— Du fait de la distributivité de la multiplication sur l'addition,

$$\forall \lambda \in \mathbb{C}, \quad \sum_{k \in I} (\lambda \cdot a_k) = \lambda \cdot \left( \sum_{k \in I} a_k \right)$$

**ATTENTION** : en règle générale,

$$\sum_{k \in I} (a_k \cdot b_k) \neq \left( \sum_{k \in I} a_k \right) \left( \sum_{k \in I} b_k \right)$$

D'autre part, l'indice de sommation  $k$  est une variable dite *muette* car la somme ainsi symbolisée ne dépend pas de  $k$  :

$$\sum_{k=1}^n a_k = \sum_{\ell=1}^n a_\ell = a_1 + \cdots + a_n$$

De plus, outre le choix du symbole utilisé pour l'indice de sommation, il est également possible d'effectuer des décalages d'indices :

$$\sum_{k=1}^n a_k \stackrel{=}{=} \sum_{\ell=k-1}^{n-1} a_{\ell+1} \stackrel{=}{=} \sum_{m=2}^{n+1} a_{m-1} = a_1 + \cdots + a_n$$

Enfin, il peut être parfois utile de calculer des sommes par paquets (en séparant par exemple les indices pairs des indices impairs). Formellement, si  $I$  est la réunion disjointe de deux sous ensembles  $I_1$  et  $I_2$ , on a

$$\sum_{k \in I} a_k = \sum_{k \in I_1} a_k + \sum_{k \in I_2} a_k$$

## Sommes remarquables

### Sommes des puissances de $k$

— Somme des premiers entiers :

$$\forall n \in \mathbb{N}^*, \quad S_1 = \sum_{k=1}^n k = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

**Preuve** : en doublant la somme  $S_1$  et en la réorganisant, on a

$$\begin{aligned} 2S_1 &= 1 + 2 + \cdots + n-1 + n \\ &+ n + n-1 + \cdots + 2 + 1 \\ &= \frac{(n+1)}{2} + \frac{(n+1)}{2} + \cdots + \frac{(n+1)}{2} + \frac{(n+1)}{2} \\ &= n \times (n+1) \end{aligned}$$

D'où le résultat.  $\square$

— Somme des premiers carrés : on peut également montrer que pour tout  $n \in \mathbb{N}^*$ , on a

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

Exercice : à démontrer par récurrence.

De façon générale, il est possible d'établir des formules analogues pour toutes les

sommes

$$S_p = \sum_{k=1}^n k^p$$

pour tout  $p \in \mathbb{N}$ .

Exercice : que dire de la somme  $\sum_{k=1}^n k^0$  ?

### Somme arithmétique

Les nombres  $a_0, \dots, a_n$  forment une suite arithmétique si

$$\exists r \in \mathbb{R} / \forall k \in \{0, \dots, n\}, a_k = a_0 + k \times r$$

On peut alors montrer que

$$\sum_{k=0}^n a_k = \sum_{k=0}^n (a_0 + k.r) = (n+1) \frac{a_0 + a_n}{2}$$

Exercice : à démontrer à l'aide des propriétés du signe  $\Sigma$ .

### Somme géométrique

Les nombres  $a_0, \dots, a_n$  forment une suite géométrique si

$$\exists q \in \mathbb{R} / \forall k \in \{0, \dots, n\}, a_k = a_0 \times q^k$$

Si  $q \neq 1$ , on peut alors montrer que

$$\sum_{k=0}^n a_k = \sum_{k=0}^n (a_0 \cdot q^k) = a_0 \frac{1 - q^{n+1}}{1 - q}$$

Exercice : que dire de cette somme si  $q = 1$  ?

### Sommes télescopiques

On appelle *somme télescopique* toute somme du type

$$S = \sum_{k=1}^{n-1} (a_{k+1} - a_k)$$

Le nom “télescopique” vient des nombreuses simplifications que l'on peut faire dans l'expression ci-dessus donnée sous forme étendue :

$$\sum_{k=1}^{n-1} (a_{k+1} - a_k) = (\cancel{a_2} - a_1) + (a_3 - \cancel{a_2}) + (\cancel{a_4} - \cancel{a_3}) + \cdots + (\cancel{a_{n-1}} - \cancel{a_{n-2}}) + (a_n - \cancel{a_{n-1}})$$

$$= a_n - a_1$$

Exercice : retrouver ce résultat conservant la notation  $\Sigma$  et en exploitant les règles de calcul énoncées plus haut (séparation d'une somme en deux sommes et décalage d'indice).

### 1.4.3 Le signe $\prod$

#### Définitions

Si  $a_1, \dots, a_n$  sont des nombres complexes, on note

$$a_1 \times a_2 \times \cdots \times a_n = \prod_{k=1}^n a_k$$

et façon générale, on note  $\prod_{k \in I} a_k$  le produit des tous les nombres  $a_k$ , l'indice  $k$  parcourant l'ensemble  $I$  des indices.

**Notes** : par convention, on pose  $\prod_{k \in \emptyset} a_k = 1$

#### Transformation d'écriture

Les propriétés de la multiplication induisent des transformations d'écriture spécifiques au signe  $\prod$  :

$$\prod_{k \in I} (a_k \cdot b_k) = \left( \prod_{k \in I} a_k \right) \cdot \left( \prod_{k \in I} b_k \right) \quad \text{et} \quad \prod_{k \in I} (a_k^\alpha) = \left( \prod_{k \in I} a_k \right)^\alpha$$

**Notes** : À l'aide des formules précédentes, on peut en particulier montrer que

$$\prod_{k=1}^n (\lambda \cdot a_k) = \lambda^n \cdot \left( \prod_{k \in I} a_k \right)$$

Mais **ATTENTION**, en règle générale, on a

$$\prod_{k \in I} (a_k + b_k) \neq \left( \prod_{k \in I} a_k \right) + \left( \prod_{k \in I} b_k \right)$$

**Factoriel**

Pour tout entier  $n \in \mathbb{N}^*$ , on appelle *factoriel*  $n$  (noté  $n!$ ) le produit de tous les entiers jusqu'à  $n$  :

$$\forall n \in \mathbb{N}^*, \quad n! = 1 \times 2 \times \cdots \times n = \prod_{k=1}^n k$$

**Note** : avec la convention établie plus haut, on a  $0! = 1$ .

Il est alors possible d'exprimer à l'aide du factoriel le produit de plusieurs entiers consécutifs :

$$\forall m, n \in \mathbb{N}^*, \quad m \times (m+1) \times \cdots \times (m+n) = \prod_{k=0}^n (m+k) = \frac{(m+n)!}{(m-1)!}$$

**1.4.4 Formule du binôme et factorisation de  $a^n - b^n$** **Coefficients binomiaux****Définition (coefficients binomiaux)**

Pour tout couple  $(n, p)$  d'entiers tels que  $0 \leq p \leq n$ , on note

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}$$

le coefficient binomiale associé aux entiers  $n$  et  $p$ .

**Exemples :**

$$\binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n, \quad \binom{n}{2} = \binom{n}{n-2} = \frac{n(n-1)}{2}$$

**Notes :**

- Intuitivement, ces coefficients binomiaux comptent le nombre de familles à  $p$  éléments que l'on peut former à partir de  $n$  éléments distincts.
- On peut étendre cette notation à l'ensemble des couples d'entiers  $p \in \mathbb{Z}$  à l'aide des conventions suivantes : si  $p < 0$  ou  $p > n$ , alors  $\binom{n}{p} = 0$ .

Par ailleurs, les coefficients binomiaux vérifient deux propriétés fondamentales :

**Propriétés**

— Pour tout  $n \in \mathbb{N}$  et  $p \in \mathbb{Z}$ ,

$$\binom{n}{p} = \binom{n}{n-p}$$

qui traduit algébriquement le fait que choisir  $p$  élément parmi  $n$ , c'est aussi choisir  $n - p$  éléments que l'on ne prend pas.

— Pour tout  $n \in \mathbb{N}$  et  $p \in \mathbb{Z}$ , on a

$$\binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}$$

Cette dernière propriété permet un calcul itératif des coefficients binomiaux à l'aide du *triangle de Pascal* :

$n \backslash p$	0	1	2	3	4	...
0	1	0	0	0	0	
1	1	1	0	0	0	
2	1	·	1	0	0	
3	1	·	·	1	0	
4	1	·	·	·	1	
⋮	⋮					⋱

**Binôme de Newton**

Le nom des coefficients binomiaux vient de leur présence dans la formule du Binôme de Newton donnant, pour tout  $n \in \mathbb{N}$ , le développement du produit  $(a + b)^n$ . Ainsi, soient  $a, b \in \mathbb{R}$ . Pour tout  $n \in \mathbb{N}^*$ , on a

$$\begin{aligned} (a + b)^n &= \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \\ &= \binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \dots + \binom{n}{k} a^{n-k} b^k + \dots + \binom{n}{n-1} a b^{n-1} + \binom{n}{n} b^n \end{aligned}$$

**Exemples :**

- Pour  $n = 2$ , on retrouve, à l'aide de la troisième ligne du triangle de Pascal, les premières identités remarquables :

$$(a + b)^2 = \binom{2}{0}a^2 + \binom{2}{1}ab + \binom{2}{2}b^2 = a^2 + 2ab + b^2$$

$$(a - b)^2 = \binom{2}{0}a^2 + \binom{2}{1}a(-b) + \binom{2}{2}(-b)^2 = a^2 - 2ab + b^2$$

- À l'aide de la troisième ligne, on obtient

$$\begin{aligned} \forall x \in \mathbb{R}, \quad (x + 1)^3 &= \binom{3}{0}x^3 + \binom{3}{1}x^2 + \binom{3}{2}x + \binom{3}{3} \\ &= x^3 + 3x^2 + 3x + 1 \end{aligned}$$

- De façon générale, pour tout  $x \in \mathbb{R}$  et pour tout  $n \in \mathbb{N}$ , on a

$$(1 + x)^n = \sum_{k=0}^n \binom{n}{k} x^k$$

**Factorisation de  $a^n - b^n$** 

Il est de même possible de généraliser la dernière identité remarquable  $a^2 - b^2 = (a - b)(a + b)$  à toute différence de la forme  $a^n - b^n$  pour  $n \in \mathbb{N}^*$ . Précisément, on peut montrer que pour tout  $a, b \in \mathbb{R}$  et pour tout  $n \in \mathbb{N}^*$ , on a

$$\begin{aligned} a^n - b^n &= (a - b) \sum_{k=0}^{n-1} a^{n-k-1} b^k \\ &= (a - b)(a^{n-1} + a^{n-2}b + \dots + ab^{n-2} + b^{n-1}) \end{aligned}$$

**Note :** dans le cas  $n = 2$ , on retrouve bien l'identité remarquable évoquée.

**Preuve** : il s'agit en réalité d'une application (ou d'une autre forme) de la formule donnant la somme des premiers termes d'une suite géométrique : soient  $a$  et  $b$  deux réels distincts et  $n$  un entier naturel. On a

$$\begin{aligned} \frac{a^n - b^n}{a - b} &= \frac{a^n \left(1 - \left(\frac{b}{a}\right)^n\right)}{a \left(1 - \frac{b}{a}\right)} \\ &= a^{n-1} \frac{1 - \left(\frac{b}{a}\right)^n}{1 - \frac{b}{a}} \\ &= a^{n-1} \sum_{k=0}^{n-1} \left(\frac{b}{a}\right)^k \\ &= \sum_{k=0}^{n-1} \frac{a^{n-1}}{a^k} b^k \\ &= \sum_{k=0}^{n-1} a^{n-k-1} b^k \end{aligned}$$

On retrouve alors la formule annoncée en multipliant cette égalité par  $a - b$ .

### 1.4.5 Résolution d'équations algébriques (une inconnue)

Formellement, une équation algébrique est une égalité liant entre elles deux expressions (le membre de gauche et le membre de droite) dépendant chacune d'un certain nombre de variables dites inconnues. Chaque équation peut être vue comme une contrainte. Si un ensemble de variables est lié par une équation, on ne peut choisir librement la valeur de chacune d'entre elles. L'un des buts essentiels des théories mathématiques consiste à mettre en place des méthodes de résolutions pour ces équations, i.e. trouver des protocoles permettant d'obtenir explicitement l'ensemble des solutions d'une équation ou d'un système d'équations.

Notons que, dans ce cadre, il est illusoire d'espérer obtenir une méthode universelle, i.e. une unique méthode permettant de résoudre toutes les équations et tous les systèmes. La théorie est donc basée sur une classification des différents types d'équations avec pour objectif, d'obtenir des méthodes spécifiques à chaque classe.

On s'intéressera ici uniquement au problème constitué d'une équation à une inconnue.

Ainsi, considérons une équation portant sur une inconnue  $x \in \mathbb{R}$  (ou  $\mathbb{C}$ ). Quitte à réécrire cette équation, il est toujours possible de la présenter sous la forme

$$(E) : f(x) = 0$$

où  $f$  est une fonction connue de  $x$ .

Les méthodes de résolution de ce type d'équation dépendent alors de la nature de la fonction  $f$ , influençant notamment le nombre de solutions de l'équation  $(E)$ .

### Équations inversibles

Une équation de la forme  $f(x) = 0$  est dite *inversible* si elle admet une unique solution. Formellement, c'est le cas lorsque la fonction  $f$  réalise une bijection à valeur dans un intervalle contenant 0. L'existence de la bijection réciproque  $f^{-1}$  permet alors d'exprimer l'unique solution de  $(E)$  sous la forme

$$f(x) = 0 \iff x = f^{-1}(0)$$

**Exemple** : une équation linéaire est une équation de la forme

$$ax + b = 0$$

Elle est inversible si et seulement si  $a \neq 0$ . Formellement, la fonction  $f$  définissant cette équation est

$$f : x \longmapsto ax + b$$

Si  $a \neq 0$ ,  $f$  réalise une bijection de  $\mathbb{R}$  dans  $\mathbb{R}$  et

$$\begin{aligned} f^{-1} : \mathbb{R} &\longrightarrow \mathbb{R} \\ y &\longmapsto \frac{1}{a}(y - b) \end{aligned}$$

Via ce formalisme, on retrouve l'unique solution de  $(E)$  obtenue plus haut à l'aide des propriétés des opérations élémentaires :

$$f(x) = 0 \iff x = f^{-1}(0) = -\frac{b}{a}$$

Lorsque cela est possible, les opérations élémentaires ainsi que la théorie des fonctions permet de construire explicitement la réciproque  $f^{-1}$  de  $f$  et de donner l'unique solution cherchée. Cependant, il n'est pas rare que, bien que la théorie nous assure de l'existence d'une telle solution, il ne soit pas possible de l'exprimer explicitement. On est alors réduit à la recherche d'une solution approchée.

### Cas général

Pour les équations de la forme  $f(x) = 0$  non inversibles (i.e. admettant plusieurs solutions), la méthode générale de résolution consiste à factoriser au maximum l'expression  $f(x)$  sous la forme

$$f(x) = \prod_{k=1}^s f_k(x)$$

où chaque fonction  $f_k$  est une bijection.

Ainsi, puisqu'un produit de nombres réels (ou complexes) est nul si et seulement si l'un de ses facteurs est nul, on a

$$f(x) = 0 \Leftrightarrow \exists k \in \{1, 2, \dots, s\} / f_k(x) = 0$$

Chacune des fonctions  $f_k$  étant bijective, chaque équation  $(E_k) : f_k(x) = 0$  admet une unique solution  $x_k$  et l'ensemble des solutions de l'équation  $f(x) = 0$  est alors la réunion de toutes ces solutions :

$$f(x) = 0 \Leftrightarrow x \in \{x_1, x_2, \dots, x_s\}$$

En pratique, la factorisation de l'expression  $f(x)$  se fait à l'aide des règles de calculs évoqués plus haut, notamment via les identités remarquables et leurs différentes généralisations.

### Équations polynomiales

Parmi ces équations, on distingue les équations *polynomiales*  $P(x) = 0$  où  $P(x)$  est un polynôme :

$$P(x) = \sum_{k=1}^n a_k x^k$$

On peut en effet montrer que tout polynôme  $P(x)$  peut s'écrire sous la forme

$$P(x) = a_n \prod_{k=1}^n (x - \alpha_k)$$

Les nombres complexes  $\alpha_k$  sont appelés *les racines de P* et forment l'ensemble des solutions de l'équation  $P(x) = 0$ . Résoudre l'équation  $P(x) = 0$  consiste à déterminer les racines du polynôme  $P(x)$ , ce qui, en pratique, revient à factoriser complètement le polynôme  $P(x)$ .

### Équations polynomiales de degré 2

Dans le cas d'une équation polynomiale de degré 2

$$ax^2 + bx + c = 0$$

la *mise sous forme canonique* du polynôme  $P(x) = ax^2 + bx + c$  de l'équation permet d'exprimer l'ensemble des solutions de l'équation étudiée uniquement en fonction des coefficients  $a$ ,  $b$  et  $c$  du polynôme  $P(X)$ .

Précisément, la mise sous forme canonique consiste à voir la partie  $ax^2 + bx$  comme le début d'un carré :

$$\begin{aligned} (ax^2 + bx) + c &= a \left( \left( x + \frac{b}{2a} \right)^2 - \frac{b^2}{4a^2} \right) + c \\ &= a \left( \left( x + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right) \end{aligned}$$

En posant

$$\Delta = b^2 - 4ac$$

on peut alors voir l'expression ci-dessus comme une différence de carrés et exploiter l'identité remarquable associée. L'expression  $P(x) = ax^2 + bx + c$  s'écrit alors

$$P(x) = a(x - x_1)(x - x_2)$$

où

$$x_1 = \frac{-b + \sqrt{\Delta}}{2a} \quad \text{et} \quad x_2 = \frac{-b - \sqrt{\Delta}}{2a}$$

constituent forment l'ensemble des solutions de l'équation  $P(x) = 0$ .

#### Notes :

- Si  $\Delta = 0$ , on a  $x_1 = x_2$ . On parle alors de solution double.
- Si  $\Delta < 0$ , le terme  $\sqrt{\Delta}$  est alors un nombre complexe, ainsi que les deux solutions  $x_1$  et  $x_2$ .

#### Cas général

Il est possible d'établir des stratégies et des formules analogues pour les équations de degré 3 et 4. Cependant, la lourdeur des formules les rend peu usitées en pratique.

D'autre part, il est prouvé qu'il n'existe aucune formule donnant les solutions d'une équation de degré supérieur ou égal à cinq en fonction des coefficients du polynôme à annuler.

En pratique, la résolution exacte de telles équations passe donc, lorsque cela est possible par des méthodes de calcul consistant à “faire chuter le degré de l’équation étudiée”.

Précisément, on peut montrer que pour tout polynôme  $P(x)$  de degré  $n$ , s’il existe un nombre  $\alpha \in \mathbb{K}$  tel que  $P(\alpha) = 0$ , alors il existe un polynôme  $Q(x)$  de degré  $n - 1$  tel que

$$P(x) = (x - \alpha) \cdot Q(x)$$

Ainsi, si  $P(x)$  est un polynôme de degré  $n \geq 3$ , une technique de factorisation consiste à chercher  $n - 2$  solutions “évidentes”  $\alpha_1, \dots, \alpha_{n-2}$  de sorte que

$$P(x) = Q(x) \cdot \prod_{k=1}^{n-2} (x - \alpha_k)$$

où  $Q(x)$  est un polynôme de degré  $n - (n - 2) = 2$ .

On peut alors terminer la factorisation en appliquant la méthode spécifique aux équations de degré 2.

**Note** : en pratique, le calcul du polynôme  $Q(x)$  peut se faire soit par identification, soit en appliquant la division euclidienne des polynômes.



# Chapitre 2

## Fonctions d'une variable réelle

### Introduction

Les fonctions réelles sont un outil indispensable à la modélisation et l'étude de tout phénomène (physique, biologique, économique, etc).

Elles permettent en particulier de modéliser l'évolution de quantités numériques en fonction de certains paramètres. Dans ce chapitre, on s'intéresse précisément aux fonctions dépendant d'une seule variable réelle  $x$ .

Les mathématiques proposent alors des outils permettant l'étude détaillée de ce type de fonction, l'un des objectifs premiers étant d'étudier les variations d'une fonction donnée.

La plupart de ces outils sont basées sur deux propriétés fondamentales de l'ensemble des nombres réels :

- la représentation géométrique de l'ensemble  $\mathbb{R}$  sous la forme d'une droite, qui permet une représentation géométrique des fonctions d'une variable réelle,
- la continuité de l'ensemble  $\mathbb{R}$ , qui permet le développement d'outils particulièrement fins.

**Rappels** : on appelle *intervalle de  $\mathbb{R}$*  tout sous ensemble  $I \subset \mathbb{R}$  continu (i.e. sans trou). On dira que  $I$  est non trivial s'il est non vide et non réduit à un point.

On rappelle également que toute partie de  $\mathbb{R}$  peut être décrite par une réunion disjointe d'intervalles (éventuellement réduits à un point).

On appelle *intervalle fermé* tout intervalle de  $\mathbb{R}$  contenant ses bornes. On note

$$I = [a, b]$$

Enfin, tout intervalle non fermé est dit ouvert (à gauche, ou à droite, ou les deux).

### 2.1 Définitions

On appelle fonction d'une variable réelle toute application  $f$  définie sur une partie  $D$  de  $\mathbb{R}$  à valeurs dans  $\mathbb{R}$ , i.e. un procédé qui à chaque élément  $x \in D$  associe un unique réel  $f(x)$ .

On note

$$\begin{array}{ccc} f & : & D \subset \mathbb{R} \longrightarrow \mathbb{R} \\ & & x \longmapsto f(x) \end{array}$$

### Définitions

Soit  $f$  une fonction de  $D$  dans  $\mathbb{R}$ .

1. L'ensemble  $D$  est appelé *domaine de définition* de  $f$ .
2. Pour tout  $x \in D$ , le réel  $f(x)$  est appelé *l'image* de  $x$  par  $f$  ou la valeur de  $f$  en  $x$ .
3. On appelle *image de  $f$*  l'ensemble

$$\text{Im}(f) = \{f(x), x \in D\}$$

4. Pour tout  $y \in \text{Im}(f)$ , on appelle *antécédent de  $y$  par  $f$*  tout réel  $x \in D$  tel que  $y = f(x)$ .

**Note** : on retrouve ici les définitions vues au chapitre précédent, traitant des fonctions en général.

L'une des premières particularités des fonctions réelles est qu'il est (le plus souvent) possible de les représenter à l'aide de courbes dans le plan muni d'un repère (orthonormé). Précisément,

### Définition (graphe d'une fonction réelle)

Soit  $f : D \rightarrow \mathbb{R}$ . On appelle *graphe de  $f$*  l'ensemble de  $\mathbb{R}^2$  défini par

$$\mathcal{G}_f = \{(x, f(x)), x \in D\}$$

En assimilant  $\mathbb{R}^2$  au plan muni d'un repère orthonormé, on peut représenter le graphe de  $f$  par un ensemble de points. Les différentes propriétés d'une fonction réelle  $f$ , établies à l'aide des outils analytiques ci-dessous peuvent alors s'interpréter en termes de propriétés géométriques vérifiées par le graphe  $\mathcal{G}_f$ .

Ainsi, si  $D$  est un intervalle de  $\mathbb{R}$  et  $f$  est continue (c.f. plus loin pour une définition précise de la continuité), l'ensemble des points du graphe de  $f$  forment alors une courbe, appelée *courbe de  $f$*  et parfois notée  $\mathcal{C}_f$ .

## 2.2 Premières propriétés

Soit  $f : D \rightarrow \mathbb{R}$  une fonction réelle.

### 2.2.1 Parité, périodicité

#### Parité

##### Définition (fonctions paires - fonctions impaires)

— Une fonction  $f : D \rightarrow \mathbb{R}$  est dite *paire* si

$$\forall x \in D, \quad -x \in D \quad \text{et} \quad f(-x) = f(x)$$

— Une fonction  $f : D \rightarrow \mathbb{R}$  est dite *impaire* si

$$\forall x \in D, \quad -x \in D \quad \text{et} \quad f(-x) = -f(x)$$

Si  $f$  est paire ou impaire, son domaine de définition est centré en 0. De plus, on peut alors réduire l'ensemble d'étude de moitié en ne considérant que  $D \cap \mathbb{R}^+$  (ou  $D \cap \mathbb{R}^-$ ), l'autre moitié étant obtenue par symétrie.

D'un point de vue géométrique, le graphe d'une fonction paire ou impaire possède certaines propriétés de symétrie. Précisément :

— Si  $f$  est paire, son graphe  $\mathcal{G}_f$  dans un repère  $(xOy)$  est **symétrique par rapport à l'axe  $(Oy)$** .

— Si  $f$  est impaire, son graphe  $\mathcal{G}_f$  dans un repère  $(xOy)$  est **symétrique par rapport au centre  $O$  du repère**.

Illustrations graphiques

**Attention** : certaines fonctions ne sont ni paires ni impaires !

**Exemples** :

- Toute fonction de la forme  $x \mapsto x^{2p}$  ( $p \in \mathbb{N}$ ) est **paire**.
- Toute fonction de la forme  $x \mapsto x^{2p+1}$  ( $p \in \mathbb{N}$ ) est **impaire**.
- La fonction  $x \mapsto 2x + 1$  **n'est ni paire ni impaire**.

## Périodicité

**Définition (fonction périodique)**

Soit  $f : D \rightarrow \mathbb{R}$  une fonction réelle. Un nombre réel  $T > 0$  est *une période de  $f$*  si

$$\forall x \in D, \quad x + T \in D \quad \text{et} \quad f(x + T) = f(x)$$

**Note** : si  $T > 0$  est une période de  $f$ , alors pour tout  $n \in \mathbb{Z}$ ,  $nT$  est également une période de  $f$  (à démontrer par récurrence).

Une fonction  $f$  est dite  $T$ -périodique si  $T$  est *la plus petite période strictement positive* de  $f$ .

D'un point de vue géométrique, si  $f$  est une fonction  $T$ -périodique, son graphe est constitué d'un motif élémentaire de "longueur"  $T$  répété de proche en proche. L'étude d'une fonction  $T$ -périodique peut alors être réduite à n'importe quel intervalle de  $D$  de longueur  $T$  (par exemple  $[0, T[$  si  $0 \in D$ ).

Illustration graphique

**Exemples :**

- Les fonctions trigonométriques sinus et cosinus sont  $2\pi$ -périodiques.
- La fonction tangente est  $\pi$ -périodique.

### 2.2.2 Fonctions bornées

**Définition (fonction bornée)**

Une fonction  $f : D \rightarrow \mathbb{R}$  est dit

- majorée sur  $D$  si

$$\exists M \in \mathbb{R} / \forall x \in D, \quad f(x) \leq M$$

- minorée sur  $D$  si

$$\exists m \in \mathbb{R} / \forall x \in D, \quad f(x) \geq m$$

- bornée sur  $D$  si elle est à la fois majorée et minorée, i.e. si

$$\exists (m, M) \in \mathbb{R}^2 / \forall x \in D, m \leq f(x) \leq M$$

**Note :** on peut montrer que  $f$  est bornée si et seulement s'il existe  $A \geq 0$  telle que

$$\forall x \in D, \quad |f(x)| \leq A$$

**Exercice :** à démontrer.

### 2.2.3 Monotonie

#### Définition (fonctions monotones)

Soient  $f : D \rightarrow \mathbb{R}$  une fonction réelle et  $I \subset D$  un intervalle non vide et non réduit à un point (i.e. non trivial).  $f$  est dite

— *croissante sur  $I$*  (resp. *strictement croissante sur  $I$* ) si

$$\forall a, b \in I, \quad a < b \Rightarrow f(a) \leq f(b) \quad (\text{resp. } f(a) < f(b))$$

— *décroissante sur  $I$*  (resp. *strictement décroissante sur  $I$* ) si

$$\forall a, b \in I, \quad a < b \Rightarrow f(a) \geq f(b) \quad (\text{resp. } f(a) > f(b))$$

— *monotone sur  $I$*  (resp. *strictement monotone sur  $I$* ) si elle est *croissante sur  $I$  ou décroissante sur  $I$*  (resp. *strictement croissante sur  $I$  ou strictement décroissante sur  $I$* ).

Certaines fonctions ne sont bien sur ni croissantes ni décroissantes. Il est possible, pour certaines, de définir des intervalles de  $D$  sur lesquelles elles sont monotone, mais on peut également construire des fonctions que ne sont monotones sur aucun intervalle...

D'autre part, on a une propriété importante des fonctions strictement monotones :

#### Proposition

Soit  $I$  un intervalle non trivial de  $\mathbb{R}$ . Toute fonction  $f : I \rightarrow \mathbb{R}$  strictement monotone sur  $I$  est *injective*.

**Exercice** : à démontrer en s'appuyant sur la définition de l'injectivité.

## 2.2.4 Extrema

### Extrema globaux

#### Définitions (extrema)

Soit  $f : D \rightarrow \mathbb{R}$ .

— Un point  $x_0 \in D$  est un *maximum global* (resp. *minimum global*) de  $f$  sur  $D$  si

$$\forall x \in D, \quad f(x) \leq f(x_0) \quad (\text{resp. } f(x) \geq f(x_0))$$

On note alors

$$f(x_0) = \max_D f = \max_{x \in D} f(x) \quad (\text{resp. } f(x_0) = \min_D f = \min_{x \in D} f(x))$$

—  $x_0 \in D$  est un *extremum global* de  $f$  sur  $D$  si c'est un maximum ou un minimum global.

### Extrema locaux

#### Définition (voisinage)

Soit  $x_0 \in \mathbb{R}$ . On appelle voisinage de  $x_0$  toute partie de  $\mathbb{R}$  contenant un intervalle de la forme  $]x_0 - \delta, x_0 + \delta[$  pour un  $\delta > 0$ .

On note  $V(x_0)$  ou  $V_{x_0}$  un tel ensemble.

#### Définition (extremum local)

Soit  $f : D \rightarrow \mathbb{R}$ .

—  $x_0 \in D$  est un *maximum local* (resp. *minimum local*) de  $f$  si il existe un voisinage  $V(x_0)$  tel que

$$\forall x \in V(x_0) \cap D, \quad f(x) \leq f(x_0) \quad (\text{resp. } f(x) \geq f(x_0))$$

—  $x_0 \in D$  est un *extremum local* de  $f$  si c'est un maximum ou un minimum local.

## 2.2.5 L'ensemble $\mathcal{F}(I, \mathbb{R})$

Soit  $I$  un intervalle de  $\mathbb{R}$ . On note  $\mathcal{F}(I, \mathbb{R})$  (ou  $\mathbb{R}^I$ ) l'ensemble des fonctions réelles définies sur  $I$ . On peut alors définir un ensemble d'opérations sur cet ensemble (inspirées pour la plupart des opérations définies sur  $\mathbb{R}$ ). Ainsi, soient  $f$  et  $g$  deux fonctions de  $\mathcal{F}(I, \mathbb{R})$ .

— La fonction  $f + g \in \mathcal{F}(I, \mathbb{R})$  est la fonction définie par

$$\forall x \in I, \quad (f + g)(x) = f(x) + g(x)$$

— La fonction  $f \times g \in \mathcal{F}(I, \mathbb{R})$  est la fonction définie par

$$\forall x \in I, \quad (f \times g)(x) = f(x) \times g(x)$$

— Si l'on note  $\tilde{I} = I \setminus \{x \in I \mid g(x) = 0\}$ , la fonction  $\frac{f}{g} \in \mathcal{F}(\tilde{I}, \mathbb{R})$  est la fonction définie par

$$\forall x \in \tilde{I}, \quad \frac{f}{g}(x) = \frac{f(x)}{g(x)}$$

Si  $f$  est définie sur  $I$  et  $g$  est définie sur l'ensemble  $f(I) = \{f(x), x \in I\}$ , on appelle *composée de  $f$  et  $g$*  la fonction  $g \circ f$  définie sur  $I$  par

$$\forall x \in I, \quad g \circ f(x) = g(f(x))$$

**Note** : on peut généraliser la somme de fonctions aux *combinaisons linéaires* :

$$\forall \alpha, \beta \in \mathbb{R}, \quad \alpha.f + \beta.g : x \longmapsto \alpha.f(x) + \beta.g(x)$$

## 2.3 Limites

### 2.3.1 Limite en un point

La notion de limite, basée sur la notion de voisinage et la propriété de continuité de l'ensemble  $\mathbb{R}$  permet de mettre en place des outils très fins d'étude *locale* des fonctions réelles. Précisément, cela permet d'étudier le comportement d'une fonction  $f$  au voisinage d'un point  $a$  où  $f$  n'est pas nécessairement définie.

**Définition (adhérence)**

Soient  $D$  une partie non vide de  $\mathbb{R}$  et  $a \in \mathbb{R}$ .

1.  $a$  est dit *adhérent* à  $D$  si tout voisinage  $V(a)$  de  $a$  rencontre  $D$ .
2. On note  $\bar{D}$  l'ensemble des points adhérents à  $D$ , appelé *adhérence de  $D$*  :

$$a \in \bar{D} \Leftrightarrow \forall V(a), V(a) \cap D \neq \emptyset$$

**Exemples :**

- Tout élément de  $D$  est adhérent à  $D$  car tout voisinage  $V(a)$  de  $a$  contient  $a$ .
- L'adhérence d'un intervalle  $I = ]a, b[$  est  $\bar{I} = [a, b]$ .
- L'adhérence de  $\mathbb{R} = ]-\infty, +\infty[$ , noté  $\bar{\mathbb{R}}$  représente l'ensemble des nombres réels, infinis ou non.

**Définition (limite infinie)**

Soit  $f$  une fonction définie sur un intervalle de la forme  $]a, b]$  (resp.  $[b, a[$ ). On dit que  $f$  tend vers l'infini en  $a$  à droite (resp. à gauche) si

$$\forall A > 0, \exists \delta > 0 / x \in ]a, a + \delta[ \Rightarrow f(x) > A$$

$$\text{(resp. } \forall A > 0, \exists \delta > 0 / x \in ]a - \delta, a[ \Rightarrow f(x) > A)$$

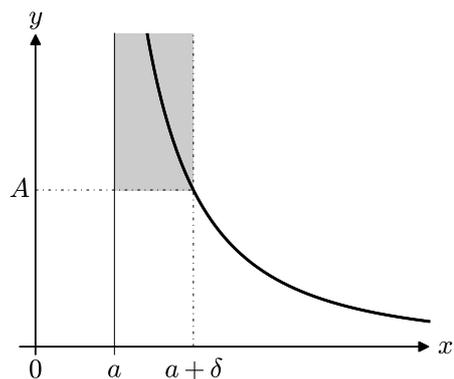
On note

$$\lim_{x \rightarrow a^+} f(x) = +\infty \text{ ou } \lim_{a^+} f = +\infty \quad \text{(resp. } \lim_{x \rightarrow a^-} f(x) = +\infty \text{ ou } \lim_{a^-} f = +\infty)$$

Géométriquement, si  $f$  admet une limite infinie en  $a \in \mathbb{R}$ , le graphe de  $f$  admet la droite d'équation

$$x = a$$

comme asymptote verticale (c.f. ci-contre).



**Exercice :**

1. Écrire les définitions analogues pour

$$\lim_{x \rightarrow a} f(x) = -\infty \quad \lim_{x \rightarrow a} f(x) = -\infty$$

2. Illustrer graphiquement les différents cas.

### Définition (limite finie)

Soient  $a \in \mathbb{R}$  et  $f$  définie au voisinage de  $a$  ( $f$  peut ici être définie ou non en  $a$ ).

1. On dit que  $f$  admet une limite finie  $\ell \in \mathbb{R}$  à droite (resp. à gauche) en  $a$  si

$$\forall \varepsilon > 0, \exists \delta > 0 / x \in ]a, a + \delta[ \Rightarrow |f(x) - \ell| < \varepsilon$$

$$\text{(resp. } \forall \varepsilon > 0, \exists \delta > 0 / x \in ]a - \delta, a[ \Rightarrow |f(x) - \ell| < \varepsilon)$$

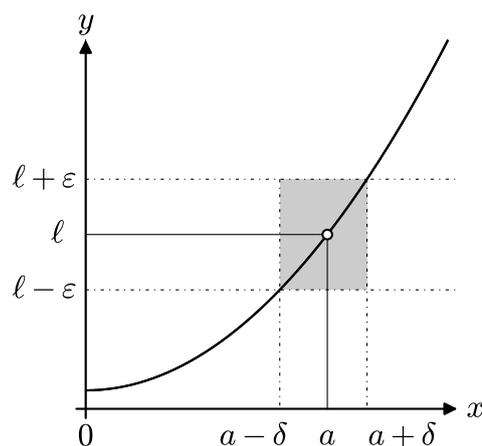
On note alors

$$\lim_{x \rightarrow a} f(x) = \ell \quad \text{(resp. } \lim_{x \rightarrow a} f(x) = \ell)$$

2. On dit que  $f$  admet une limite finie en  $a$  si elle admet une limite commune  $\ell$  à droite et à gauche. On note alors

$$\lim_{x \rightarrow a} f(x) = \ell$$

On peut illustrer cette notion à l'aide du graphe ci-dessous :



### 2.3.2 Limite en l'infini

La notion de limite permet également d'étudier le comportement d'une fonction au voisinage de l'infini.

#### Définition (limites en l'infini)

Soit  $f$  une fonction définie sur un intervalle de la forme  $[a, +\infty[$ .

— **Limite finie** :  $f$  admet une limite finie  $\ell$  en  $+\infty$  si et seulement si

$$\forall \varepsilon > 0, \exists B > 0 / x \geq B \Rightarrow |f(x) - \ell| < \varepsilon$$

On note alors  $\lim_{x \rightarrow +\infty} f(x) = \ell$ .

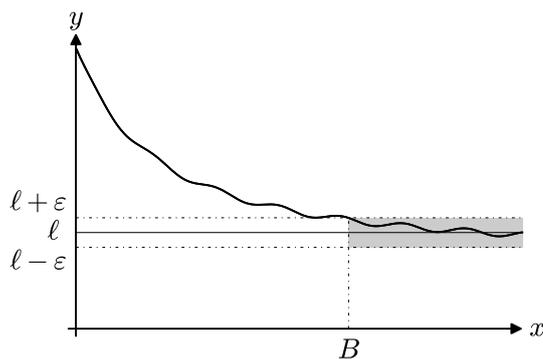
— **Limite infinie** :  $f$  tend vers  $+\infty$  en  $+\infty$  si et seulement si

$$\forall A > 0, \exists B > 0 / x \geq B \Rightarrow f(x) > A$$

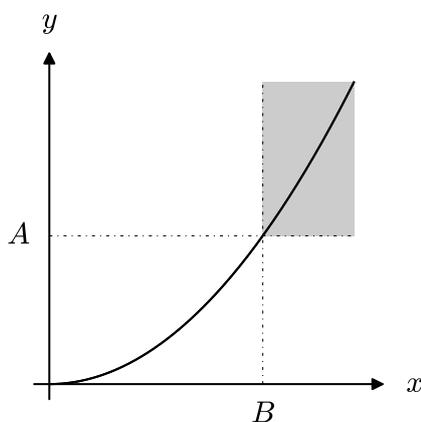
Si  $f$  admet une limite finie  $\ell$  en  $+\infty$ , son graphe admet la droite d'équation

$$y = \ell$$

comme asymptote horizontale à droite.



Si  $f$  admet une limite infinie en l'infini, on dit que le graphe de  $f$  admet une *branche infinie*.



### 2.3.3 Limites et inégalités

#### Proposition

Soient  $a \in \mathbb{R}$  et  $f$  et  $g$  deux fonctions définies sur un voisinage  $V(a)$  telles que

$$\lim_{x \rightarrow a} f(x) = \ell_1 \in \overline{\mathbb{R}} \quad \text{et} \quad \lim_{x \rightarrow a} g(x) = \ell_2 \in \overline{\mathbb{R}}$$

Si

$$\forall x \in V(a), \quad f(x) \leq g(x)$$

Alors

$$\ell_1 \leq \ell_2$$

#### Notes :

- D'après la propriété ci-dessus, il est clair que si  $f$  est une fonction bornée par  $m$  à gauche et  $M$  à droite sur un intervalle  $I \subset \mathbb{R}$  et si  $f$  admet une limite finie  $\ell$  en  $a \in \overline{I}$ , alors

$$m \leq \ell \leq M$$

- La conclusion de la proposition ci-dessus est optimale. En effet, même en remplaçant, dans l'hypothèse 3, l'inégalité large par une inégalité stricte, on ne peut avoir mieux que l'inégalité large dans la conclusion.

### Théorème (des gendarmes)

Soient  $a \in \mathbb{R}$  et  $f$ ,  $g$  et  $h$  trois fonctions définies dans un voisinage  $V(a)$  telles que

$$\forall x \in V(a), \quad f(x) \leq g(x) \leq h(x)$$

Si

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} h(x) = \ell \in \overline{\mathbb{R}}$$

alors

$$\lim_{x \rightarrow a} g(x) = \ell$$

Ce (célèbre) théorème est souvent utilisé pour montrer qu'une limite est nulle. Précisément, pour montrer que  $\lim_{x \rightarrow a} f(x) = 0$ , on s'attache à établir une inégalité de la forme

$$0 \leq f(x) \leq g(x)$$

où  $g$  est une fonction vérifiant  $\lim_{x \rightarrow a} g(x) = 0$ . Le théorème des gendarmes permet alors de conclure.

**Exemple** : montrer que  $\lim_{x \rightarrow 0} x \sin\left(\frac{1}{x}\right) = 0$ .

### 2.3.4 Limites et opérations

Via la définition, il est possible de démontrer des règles de calculs de limites compatibles avec les opérations de l'ensemble  $\mathcal{F}(I, \mathbb{R})$ . Précisément,

#### Proposition

Soient  $f$  et  $g$  deux fonctions définies au voisinage d'un point  $a$  (fini ou non) admettant respectivement des limites finies  $\ell_1$  et  $\ell_2$  en  $a$ . Alors

1.  $\lim_{x \rightarrow a} (f + g)(x) = \ell_1 + \ell_2$

2.  $\lim_{x \rightarrow a} (f \cdot g)(x) = \ell_1 \cdot \ell_2$

3. Si  $\ell_2 \neq 0$ ,  $\lim_{x \rightarrow a} \frac{f}{g}(x) = \frac{\ell_1}{\ell_2}$

**Notes :**

- L'item 2.3.4 de la propriété ci-dessus peut se généraliser aux *combinaisons linéaires* :

$$\forall \alpha, \beta \in \mathbb{R}, \quad \lim_{x \rightarrow a} (\alpha.f + \beta.g)(x) = \alpha.l_1 + \beta.l_2$$

- Dans l'item 3 ci-dessus, l'étude de la limite de  $\frac{f(x)}{g(x)}$  nécessite en théorie de s'assurer que  $g(x)$  ne s'annule pas au voisinage de  $a$ . Cependant, on peut montrer que cette propriété est vérifiée dès que  $l_2 \neq 0$ .

On peut enfin énoncer la proposition suivante concernant les limites de fonctions composées.

**Proposition**

Soient  $f$  et  $g$  deux fonctions telles que

$$\lim_{x \rightarrow a} f(x) = b \quad \text{et} \quad \lim_{y \rightarrow b} g(y) = \ell$$

Alors

$$\lim_{x \rightarrow a} g \circ f(x) = \ell$$

**2.3.5 Formes indéterminées**

Il est également possible de généraliser les résultats ci-dessus aux limites éventuellement infinies, certaines combinaisons produisant cependant des *formes indéterminées*, i.e. des limites pour lesquelles on ne peut donner de règle générale. Ainsi, soient  $f$  et  $g$  tendant respectivement vers  $l_1$  et  $l_2$  deux éléments de  $\overline{\mathbb{R}}$ . L'ensemble des résultats généraux est résumé dans les tableaux ci-dessous.

limite d'une somme  $f + g$

$l_1 \backslash l_2$	$-\infty$	$l_2 \in \mathbb{R}$	$+\infty$
$-\infty$			
$l_1 \in \mathbb{R}$			
$+\infty$			

limite d'un produit  $f.g$ 

$l_1 \backslash l_2$	$-\infty$	$l_2 < 0$	$l_2 = 0$	$l_2 > 0$	$+\infty$
$-\infty$					
$l_1 < 0$					
$l_1 = 0$					
$l_1 > 0$					
$+\infty$					

limite d'un quotient  $\frac{f}{g}$ 

$l_1 \backslash l_2$	$-\infty$	$l_2 < 0$	$l_2 = 0^-$	$l_2 = 0^+$	$l_2 > 0$	$+\infty$
$-\infty$						
$l_1 < 0$						
$l_1 = 0$						
$l_1 > 0$						
$+\infty$						

### 2.3.6 Croissances comparées

Lever une indétermination demande alors une étude plus poussée de chacune des fonctions  $f$  et  $g$  permettant de déterminer la vitesse de convergence de chacune des quantités  $f(x)$  et  $g(x)$ . Nous verrons plus loin comment définir une échelle de vitesse ainsi que des outils permettant de placer les fonctions étudiées sur cette échelle.

Rappelons pour l'instant quelques résultats classiques concernant les fonctions de références :

1. En  $+\infty$ ,

(a) n'importe quelle puissance de  $x$  l'emporte sur n'importe quelle puissance du log :

$$\forall \alpha, \beta > 0, \quad \lim_{x \rightarrow +\infty} \frac{\ln^\alpha(x)}{x^\beta} = 0$$

(b) n'importe quelle puissance de l'exponentielle l'emporte sur n'importe quelle puissance de  $x$  :

$$\forall \alpha, \beta > 0, \quad \lim_{x \rightarrow +\infty} \frac{e^{\alpha x}}{x^\beta} = +\infty$$

2. En 0, n'importe quelle puissance de  $x$  l'emporte sur n'importe quelle puissance du log :

$$\forall \alpha, \beta > 0, \quad \lim_{x \searrow 0} x^\alpha \cdot \ln^\beta(x) = 0$$

**Exemple** : on a en particulier

$$a) \lim_{x \rightarrow +\infty} \frac{\ln x}{x} = 0 \quad b) \lim_{x \rightarrow +\infty} \frac{e^x}{x} = +\infty \quad c) \lim_{x \searrow 0} x \ln(x) = 0$$

**Exercice** : déterminer les limites ci-dessous :

$$a) \lim_{x \rightarrow +\infty} x e^{-x} \quad b) \lim_{x \rightarrow +\infty} \frac{\sqrt{x}}{\ln x} \quad c) \lim_{x \searrow 0} \frac{1}{x \ln x}$$

## 2.4 Continuité

### 2.4.1 Définitions

La notion de continuité formalise la caractérisation des fonctions dont le graphe est une courbe continue. Ainsi,

#### Définition (continuité)

Soient  $I$  un intervalle de  $\mathbb{R}$  et  $f$  une fonction définie sur  $I$ .

##### 1. Continuité en un point.

Soit  $x_0 \in I$ .

(a)  $f$  est dite *continue en  $x_0$  à gauche* si et seulement si

$$\lim_{x \rightarrow x_0^-} f(x) = f(x_0)$$

(b)  $f$  est dite *continue en  $x_0$  à droite* si et seulement si

$$\lim_{x \rightarrow x_0^+} f(x) = f(x_0)$$

(c)  $f$  est dite *continue en  $x_0$*  si et seulement si

$$\lim_{x \rightarrow x_0^-} f(x) = \lim_{x \rightarrow x_0^+} f(x) = f(x_0)$$

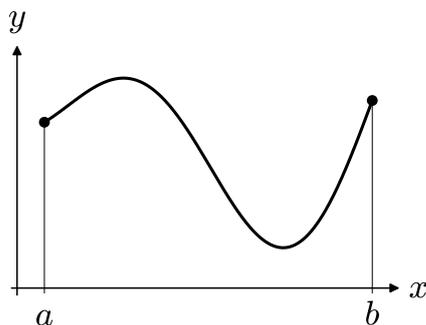
##### 2. Continuité sur un intervalle.

$f$  est dite *continue sur  $I$*  si et seulement si elle est continue en tout point  $x \in I$ .

#### Exemples

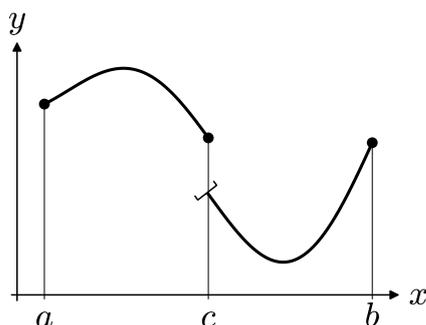
1. La plupart des fonctions usuelles sont continues sur leurs domaines de définition.
2. Cette définition permet de retrouver la notion géométrique de continuité. Ainsi,

(a) La fonction représentée ci-dessous est **continue sur l'intervalle  $[a, b]$** .



(b) La fonction représentée ci-dessous est

- continue à gauche en  $c$ .
- discontinue à droite en  $c$ .
- discontinue en  $c$ .
- discontinue sur l'intervalle  $[a, b]$ .
- continue sur l'intervalle  $[a, c]$ .
- continue sur l'intervalle  $]c, b]$ .



3. La définition ci-dessus permet également d'étudier la continuité dans les cas où le dessin n'est plus accessible. Ainsi, que peut-on dire de la fonction  $x \mapsto x \cdot \sin\left(\frac{1}{x}\right)$  au voisinage de 0 ?

**Exercice** : pour tout nombre réel  $x$ , on appelle *partie entière de  $x$*  le plus grand entier  $E(x)$  inférieur ou égal à  $x$ .

1. Tracer une esquisse du graphe de la fonction  $E$ .
2. Montrer que  $E$  est continue à gauche en chaque point  $n \in \mathbb{Z}$ .

## 2.4.2 Continuité et opérations

On verra dans la suite que de nombreuses propriétés des fonctions réelles sont spécifiques aux fonctions *continues*. Pour pouvoir s'appuyer dessus, il faudra s'assurer que la fonction étudiée est continue au(x) point(s) étudié(s).

Or, les fonctions étudiées sont définies à partir des fonctions usuelles et des opérations définies sur l'ensemble  $\mathcal{F}(\mathbb{R}, \mathbb{R})$ . La propriété ci-dessous permet d'établir la continuité des fonctions étudiées sans avoir à revenir à la définition :

### Proposition

La somme, le produit, le quotient, la composée de deux fonctions continues sont continues **partout où elles sont définies**.

Du point de vue des ensembles, pour tout intervalle  $I \subset \mathbb{R}$ , on note ainsi  $\mathcal{C}(I, \mathbb{R})$  l'ensemble des fonctions continues en tout point de  $I$ . Cet ensemble est un sous ensemble de  $\mathcal{F}(I, \mathbb{R})$  et est stable pour les opérations évoquées ci-dessus.

## 2.4.3 Prolongement par continuité

**Exemple** : la fonction  $\ln$  étant définie sur l'intervalle  $I = ]0, +\infty[$ , la fonction

$$f : x \longmapsto x \ln(x)$$

est donc également définie sur  $I$ . Cependant, on a vu plus haut que

$$\lim_{x \rightarrow 0^+} x \ln(x) = 0$$

On peut donc "étendre" le domaine de définition à l'intervalle  $\tilde{I} = [0, +\infty[$  en "ajoutant" la valeur

$$f(0) = 0$$

Formellement, on définit ainsi une seconde fonction  $\tilde{f}$ , définie par

$$\tilde{f} : [0, +\infty[ \longrightarrow \mathbb{R}$$

$$x \longmapsto \begin{cases} x \ln(x) & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}$$

On dit que la fonction  $\tilde{f}$  *prolonge* la fonction  $f$  en 0 à droite.

De façon générale, on établit la définition suivante.

**Définition (prolongements continus)**

Soient  $I = ]a, b]$  un intervalle de  $\mathbb{R}$  et  $f$  une fonction réelle, définie et continue sur  $I$ .

1. On dit que  $f$  admet un prolongement continue en  $a$  si  $f$  admet une limite finie  $\ell$  à droite en  $a$ .
2. On appelle *prolongement de  $f$  en  $a$*  la fonction  $\tilde{f}$  définie par

$$\tilde{f} : [a, b] \longrightarrow \mathbb{R}$$

$$x \longmapsto \begin{cases} f(x) & \text{si } x \in ]a, b] \\ \ell & \text{si } x = a \end{cases}$$

**2.4.4 Théorème des valeurs intermédiaires**

Le théorème fondamental relatif à la notion de continuité des fonctions réelles est *le théorème des valeurs intermédiaires* (TVI) :

**Théorème (valeurs intermédiaires)**

Soient  $I$  un intervalle de  $\mathbb{R}$  et  $f$  une fonction continue, définie sur  $I$ .

- L'image  $f(I)$  est encore un intervalle de  $\mathbb{R}$ .
- Si  $I$  est un intervalle fermé borné (i.e. de la forme  $[a, b]$ ), alors  $f(I)$  est également un intervalle fermé borné (de la forme  $[m, M]$ ).

Ce théorème formalise l'idée intuitive qu'une fonction continue prend toutes les valeurs situées entre deux valeurs données.

L'une des applications premières de cette notion de continuité se trouve dans la résolution numérique des équations algébriques. En effet, appliquée à une fonction  $f$  changeant de signe sur un intervalle  $[a, b]$  donné (i.e. telle que  $f(a)$  et  $f(b)$  sont de signes différents), ce théorème assure l'existence d'(au moins) une solution à l'équation  $f(x) = 0$ .

Par ailleurs, une autre conséquence de ce théorème est que sur un intervalle fermé  $[a, b]$  toute fonction continue est bornée et atteint ces bornes. Ce résultat est en particulier fondamental dans de nombreux problèmes d'optimisation car il permet là encore d'assurer l'existence des quantités cherchées.

Enfin, la notion de continuité est fondamentale dans la notion suivante : la notion de dérivée.

## 2.5 Dérivabilité

### 2.5.1 Dérivée en un point

La notion de dérivée constitue la pierre angulaire d'une théorie parmi les plus importantes développées par les mathématiciens : le calcul différentiel.

#### Définition (dérivée)

1. Soit  $f$  une fonction définie au voisinage d'un point  $x_0 \in \mathbb{R}$ . La fonction  $f$  est dite *dérivable en  $x_0$*  si et seulement si le taux d'accroissement

$$\frac{\Delta f}{\Delta x} = \frac{f(x) - f(x_0)}{x - x_0}$$

admet **une limite finie** quand  $x$  tend vers  $x_0$ .

Si cette limite existe, on note

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

le nombre dérivé de  $f$  en  $x_0$ .

2. Soit  $f$  une fonction définie sur un domaine  $D \subset \mathbb{R}$ . La fonction  $f$  est dite *dérivable sur  $D$*  si **elle est dérivable en chaque point  $x \in D$** .

**Note** : la notion de dérivée est basée sur une petite variation de la variable  $x$  autour du point  $x_0$  : la différence  $x - x_0$ . En notant  $h = x - x_0$ , on obtient une écriture équivalente du nombre dérivé :

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

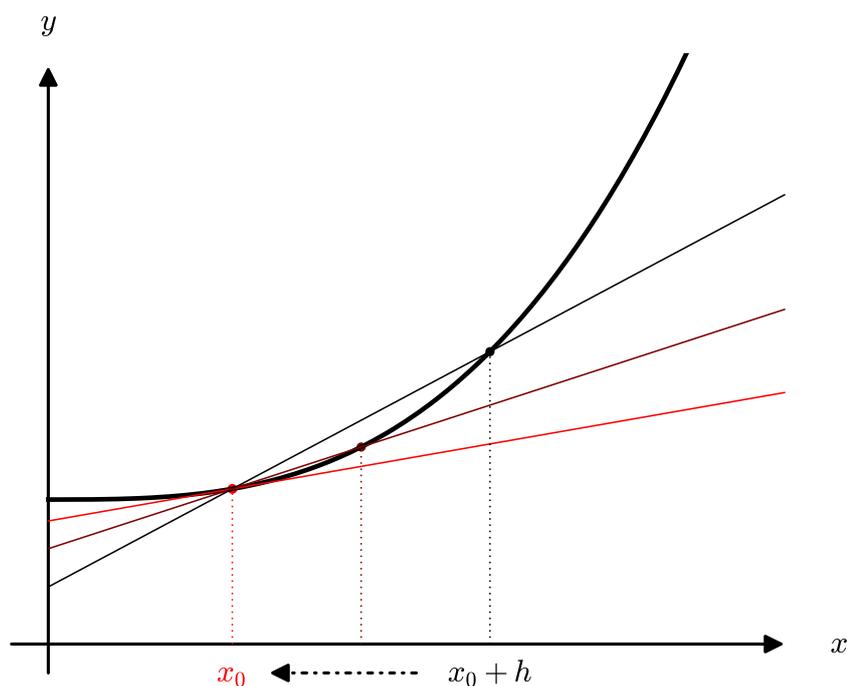
### 2.5.2 Dérivée et tangente

Il existe une interprétation fondamentale du nombre dérivé en un point d'un point de vue géométrique. Précisément, le taux d'accroissement

$$\frac{\Delta f}{\Delta x} = \frac{f(x) - f(x_0)}{x - x_0}$$

correspond à la pente de la droite du plan liant les points du graphe  $\mathcal{G}_f$  de  $f$  de coordonnées  $(x, f(x))$  et  $(x_0, f(x_0))$ .

Lorsque  $x$  tend vers  $x_0$ , cette droite tend vers la tangente à  $\mathcal{G}_f$  au point de coordonnées  $(x_0, f(x_0))$  :



Le nombre dérivé est donc, quand il existe, le *coefficient directeur* de la tangente à  $\mathcal{G}_f$  au point d'abscisse  $x_0$  lorsque cette tangente existe.

D'un point de vue théorique, il est alors possible de donner l'équation de la tangente  $T_{x_0}$  à  $\mathcal{G}_f$  au point d'abscisse  $x_0$  : cette tangente est l'unique droite du plan, de pente  $f'(x_0)$  et passant par le point  $(x_0, f(x_0))$ . Il s'agit donc de la droite d'équation

$$T_{x_0} : y = f'(x_0) \cdot (x - x_0) + f(x_0)$$

D'un point de vue pratique, l'étude de cette tangente (et donc la dérivée  $f'(x_0)$ ) est en général la première étape (et une étape fondamentale) dans l'étude locale de la fonction  $f$  au voisinage de  $x_0$ .

Ainsi, notons dans un premier temps qu'au point d'abscisse  $x_0$ , la courbe  $\mathcal{G}_f$  et la tangente  $T_{x_0}$  varient dans le même sens. Or le sens de variation de la tangente  $T_{x_0}$  est donné par le signe de son coefficient directeur  $f'(x_0)$ .

Dans un second temps, notons que la vitesse de variation de  $f$  au point d'abscisse  $x_0$  correspond là encore à l'intensité de la pente de la tangente  $T_{x_0}$ . Cette vitesse de variation est donc donnée par la quantité  $|f'(x_0)|$ .

Enfin, notons que, de façon générale, si l'on reste proche de  $x_0$ , on peut assimiler le graphe de  $f$  et sa tangente en  $x_0$ . D'un point de vue analytique, cela se traduit par une égalité du type

$$f(x) = f(x_0) + f'(x_0).(x - x_0) + e(x - x_0)$$

où  $e(x - x_0)$  est une fonction qui tend vers 0 quand  $x$  tend vers  $x_0$  et qui devient négligeable lorsque  $x$  est assez proche de  $x_0$ . Nous verrons par la suite qu'il est possible de préciser et de généraliser cette écriture.

**Note** : en note  $h = x - x_0$ , la formule ci-dessus devient

$$f(x_0 + h) = f(x_0) + h.f'(x_0) + e(h)$$

où  $e(h) \xrightarrow{h \rightarrow 0} 0$ .

**Note** : sur la même idée, il est également possible de définir la notion de *dérivée à gauche* (resp. à droite) :

Une fonction  $f$ , définie au voisinage d'un point  $x_0 \in \mathbb{R}$  est dite *dérivable à gauche* (resp. à droite) si et seulement si **la limite**

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \quad (\text{resp.} \quad \lim_{x \leftarrow x_0} \frac{f(x) - f(x_0)}{x - x_0})$$

existe.

L'existence de cette limite est alors liée à l'existence d'une *demi-tangente* à la courbe  $\mathcal{G}_f$  au point d'abscisse  $x_0$  et on peut montrer qu'une fonction est dérivable en un point  $x_0$  si et seulement si elle est dérivable à gauche et à droite et que les deux limites associées sont égales.

### 2.5.3 Fonction dérivée et dérivées successives

Comme la notion de continuité, la notion de dérivée est avant tout une notion *locale*. Cependant, si  $f$  est dérivable sur un intervalle  $I \subset \mathbb{R}$ , on peut définir, sur  $I$ , la *fonction dérivée*  $f'$  de  $f$  :

#### Définition (fonction dérivée)

Soit  $f$  une fonction réelle définie et dérivable sur un intervalle  $I \subset \mathbb{R}$ . On appelle *fonction dérivée de  $f$  sur  $I$*  la fonction  $f'$  définie par

$$f' : I \longrightarrow \mathbb{R} \\ x \longmapsto f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

**Exemples** :

1. Soient  $a, b \in \mathbb{R}$  fixés et  $f : x \mapsto ax + b$ . Montrer que  $f$  admet un nombre dérivé  $f'(x_0)$  en tout point  $x_0 \in \mathbb{R}$  et que la fonction dérivée  $f'$  est une constante à déterminer.
2. Soit  $g : x \mapsto x^2$ .
  - (a) Calculer le nombre dérivé  $g'(1)$ .
  - (b) Calculer le nombre dérivé  $g'(x_0)$  pour  $x_0 \in \mathbb{R}$ .
  - (c) En déduire la fonction  $g'$ .
3. Montrer que la fonction  $h : x \mapsto \sqrt{x}$  n'est pas dérivable en 0.

**Note** : lorsque la fonction  $f'$  existe, elle peut également admettre des nombres dérivés. Ainsi, si la fonction  $f'$  admet un nombre dérivé en  $x_0$ . C'est *la dérivée seconde de  $f$  en  $x_0$* , notée  $f''(x_0)$ . Si  $f'$  admet des dérivées en chaque point d'un intervalle  $I \subset \mathbb{R}$ , on peut construire *la fonction dérivée seconde  $f''$* .

En itérant le processus, on peut ainsi construire (quand elles existent) les dérivées successives de  $f$ , notes  $f^{(k)}$ .

#### Définition (les ensembles $\mathcal{C}^k(I, \mathbb{R})$ )

Soit  $f$  une fonction définie sur un intervalle  $I \subset \mathbb{R}$  et  $k \in \mathbb{N}$  un entier naturel.  $f$  est dite *de classe  $\mathcal{C}^k$  sur  $I$*  si

- $f$  est  $k$  fois dérivable en tout point de  $I$ .
- la fonction  $f^{(k)}$  est continue sur  $I$ .

D'autre part, on note  $\mathcal{C}^k(I, \mathbb{R})$  l'ensemble des fonctions *de classe  $\mathcal{C}^k$  sur  $I$* .

Comme pour les fonctions continues, il est alors possible de restreindre les opérations classiques sur l'ensemble des fonctions à l'ensemble des fonctions de classe  $\mathcal{C}^k$ . Précisément,

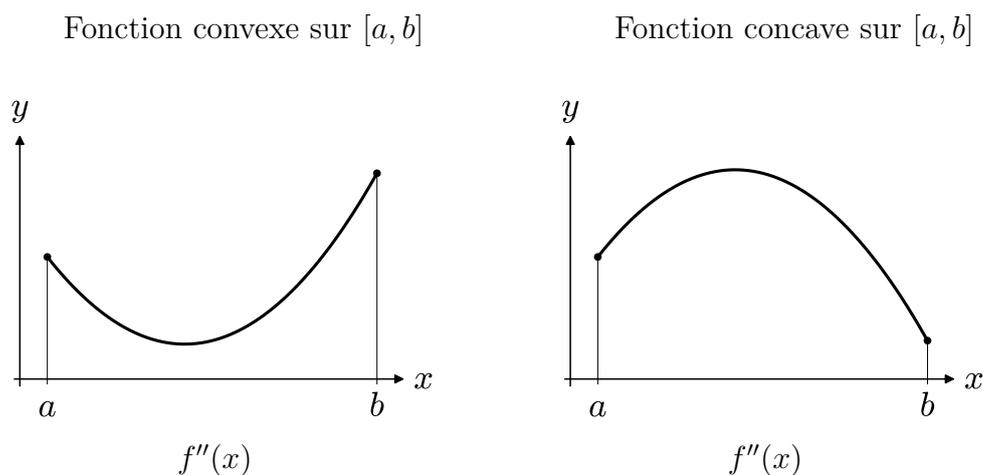
#### Proposition

Soit  $k \in \mathbb{N}$ . *La somme, le produit, le quotient, la composée de fonctions de classe  $\mathcal{C}^k$  sont de classe  $\mathcal{C}^k$  partout où elles sont définies.*

**Note** : l'étude d'une fonction réelle  $f$  dérivable passe le plus souvent par l'étude de sa dérivée  $f'$ , dont l'étude du signe permet en particulier de dresser le tableau de variations de la fonction  $f$  étudiée.

De même, la résolution de l'équation  $f'(x) = 0$  permet de déterminer les points de  $I$  auxquels la fonction  $f$  “marque un palier”. Cette résolution constitue alors la première étape dans la recherche des valeurs extrêmes de  $f$  à l'intérieur de l'intervalle étudié.

Pour une étude plus précise, on peut également se pencher sur l'étude des dérivées successives de la fonction  $f$  étudiée. Ainsi, la fonction  $f''$  permet d'obtenir de l'information sur les variations de la fonction  $f'$  et donc de préciser les variations de la fonction  $f$ . Le signe de cette dérivée seconde est en particulier lié à la notion de *concavité* de la courbe de  $f$  :



Enfin, la résolution de l'équation  $f''(x) = 0$  constitue là encore une première étape dans la recherche des *points d'inflexion* de  $f$ , i.e. les points auxquels la courbe de  $f$  change de concavité : la fonction  $f$  représentée ci-dessous

- est concave sur l'intervalle  $[a, c]$ ,
- est convexe sur l'intervalle  $[c, b]$ ,
- admet un point d'inflexion en  $c$ .

### 2.5.4 Dérivées des fonctions usuelles

Lors de l'étude d'une fonction donnée, il faut, avant de pouvoir exploiter la dérivée, établir l'existence de cette dérivée. Or là encore, les fonctions étudiées étant en général construites à partir des fonctions usuelles et des opérations élémentaires définies sur l'ensemble des fonctions, il n'est en général par nécessaire de revenir à la définition de la dérivabilité mais en s'appuyant sur les formules ci-dessous.

## Dérivées des fonctions usuelles

Fonctions	Dérivées	Domaine	Fonctions	Dérivées	Domaine
$a \in \mathbb{R}$			$x$		
$ax + b$			$x^2$		
$\sqrt{x}$			$\frac{1}{x}$		
$x^n, n \in \mathbb{N}^*$			$x^\alpha, \alpha \in \mathbb{R}$		
$\exp(x)$			$\ln(x)$		
$\sin(x)$			$\cos(x)$		
$\tan(x)$					

## Opérations et dérivées

Opérations	Dérivées
$a.u(x) + b.v(x), \quad a, b \in \mathbb{R}$	
$u(x).v(x)$	
$\frac{u(x)}{v(x)}$	
$\frac{1}{u(x)}$	
$\exp(u(x))$	
$\ln(u(x))$	
$v(ax) \quad a \in \mathbb{R}$	
$v(u(x))$	
$f^{-1}(x)$	

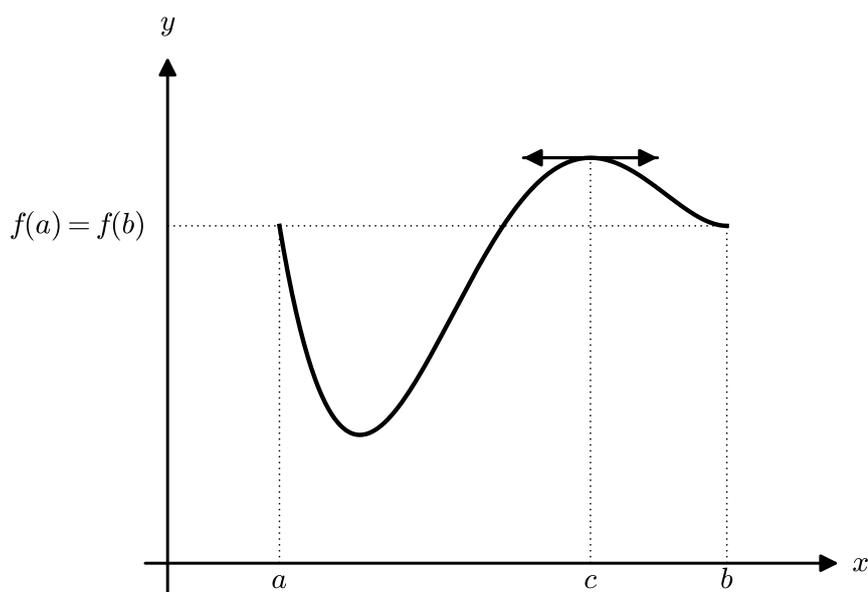
### 2.5.5 Théorème de Rolle et Théorème des Accroissements finis

Les liens forts existant entre une fonction dérivable et sa dérivée se traduisent sous la forme de deux théorèmes (célèbres). Ces résultats permettent de préciser l'idée que les variations d'une fonction dérivable sur un intervalle donné sont directement liées aux valeurs que prend la dérivée sur cet intervalle.

#### Théorème (Rolle)

Soit  $f$  une fonction continue sur un intervalle  $[a, b]$  non trivial et dérivable sur  $]a, b[$ .  
Si  $f(a) = f(b)$ , alors

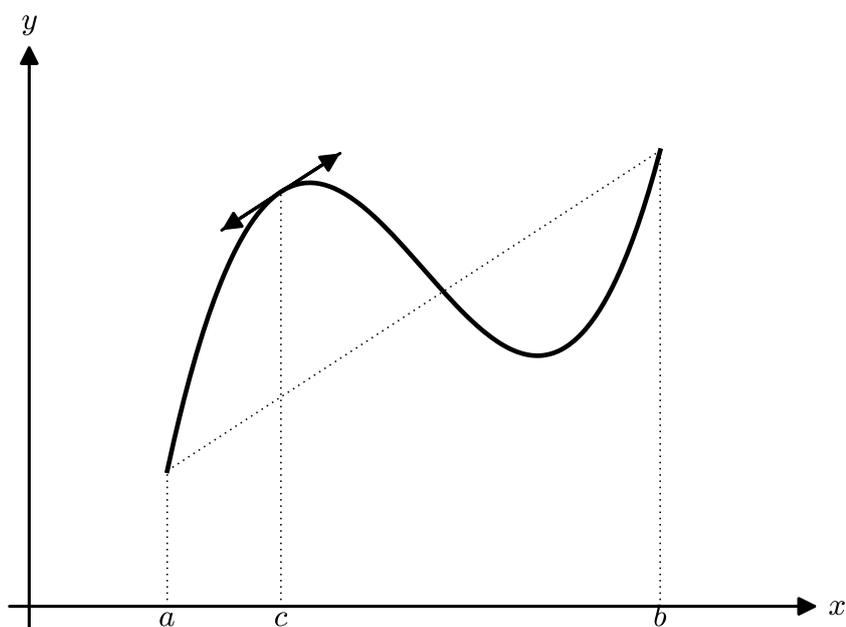
$$\exists c \in ]a, b[ \quad / \quad f'(c) = 0$$



#### Théorème (accroissements finis)

Soit  $f$  une fonction continue sur un intervalle  $[a, b]$  non trivial et dérivable sur  $]a, b[$ .  
Il existe  $c \in ]a, b[$  tel que

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$



### Corollaire (inégalité des accroissements finis)

Soit  $f$  une fonction continue sur un intervalle  $[a, b]$  non trivial et dérivable sur  $]a, b[$ .  
S'il existe  $m, M \in \mathbb{R}$  tels que

$$\forall x \in ]a, b[, \quad m \leq f'(x) \leq M$$

alors

$$m \cdot (b - a) \leq f(b) - f(a) \leq M \cdot (b - a)$$

## 2.6 Comparaison locale de fonctions

Dans cette section, nous mettons en place des outils de comparaison locale pour les fonctions réelles.

Ainsi, dans toute la suite,  $I$  désigne un intervalle non trivial de  $\mathbb{R}$ ,  $a \in \overline{\mathbb{R}}$  est un élément de  $I$  ou une borne de  $I$ , appartenant à  $I$  ou non et  $f$  et  $g$  sont deux fonctions réelles définies sur  $I$  (et donc en particulier au voisinage de  $a$ ).

## 2.6.1 Domination

**Définition (domination)**

On dit que  $f$  est dominée par  $g$  au voisinage de  $a$  s'il existe

- un voisinage  $V(a)$ ,
- une constante  $k > 0$ ,

tels que

$$\forall x \in V(a) \cap I, \quad |f(x)| \leq k \cdot |g(x)|$$

On note alors

$$f \underset{a}{=} O(g)$$

Ainsi, si  $g$  domine  $f$  au voisinage de  $a$ , alors les variations des fonctions  $f$  et  $g$  sont semblables au voisinage de  $a$ . Précisément, le théorème des gendarmes permet de relier les éventuelles limites des fonctions  $f$  et  $g$  en  $a$  :

- Si  $\lim_{x \rightarrow a} g(x) = 0$ , alors  $\lim_{x \rightarrow a} f(x) = 0$ .  
D'autre part, la fonction  $f$  tend alors vers 0 *au moins aussi vite que*  $g$ .
- Si  $\lim_{x \rightarrow a} |f(x)| = +\infty$ , alors  $\lim_{x \rightarrow +\infty} |g(x)| = +\infty$ .  
Là encore, la divergence de  $g$  vers l'infini se fait *au moins aussi vite que* celle de  $f$ .

**Note** : si la fonction  $g$  ne s'annule pas au voisinage de  $a$ , alors

$$f \underset{a}{=} O(g) \iff \frac{f}{g} \text{ borné au voisinage de } a$$

**Exemples :**

1. Déterminer quelle fonction de  $f$  ou  $g$  domine l'autre au voisinage de  $a$  dans les couples ci-dessous :

$$\begin{array}{lll} f : x \mapsto x^3 & g : x \mapsto x^2 & a = 0 \\ f : x \mapsto x^3 \sin(x) & g : x \mapsto x^3 & a = +\infty \\ f : x \mapsto x^2 & g : x \mapsto x^3 + x^2 & a = 0 \\ f : x \mapsto x^3 & g : x \mapsto x^3 + x^2 & a = 0 \end{array}$$

2. Démontrer que toute fonction est dominée par elle-même en tout point de son domaine.

## 2.6.2 Négligeabilité

Définition (négligeabilité)

On dit que  $f$  est *négligeable* devant  $g$  au voisinage de  $a$  si et seulement s'il existe

- un voisinage  $V(a)$
- une fonction  $\varepsilon$

tels que

$$\forall x \in V(a) \cap I, \quad f(x) = \varepsilon(x).g(x)$$

et

$$\lim_{x \rightarrow a} \varepsilon(x) = 0$$

On note alors

$$f \underset{a}{=} o(g)$$

**Note** : si la fonction  $g$  ne s'annule pas au voisinage de  $a$  (sauf éventuellement en  $a$ ), alors

$$f \underset{a}{=} o(g) \iff \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$$

On constate en particulier que si  $f \underset{a}{=} o(g)$ , alors  $f \underset{a}{=} O(g)$ . Là encore, il est alors possible de comparer le comportement des fonctions  $f$  et  $g$  au voisinage de  $a$  : si  $f \underset{a}{=} o(g)$ , alors

- Si  $\lim_{x \rightarrow a} g(x) = 0$ , alors  $\lim_{x \rightarrow a} f(x) = 0$ .  
D'autre part, la fonction  $f$  tend alors vers 0 *significativement plus vite que  $g$* .
- Si  $\lim_{x \rightarrow a} |f(x)| = +\infty$ , alors  $\lim_{x \rightarrow +\infty} |g(x)| = +\infty$ .  
Là encore, la divergence de  $g$  vers l'infini se fait *significativement plus vite que celle de  $f$* .

**Exemples** :

- Interpréter les résultats de croissances comparées en termes de négligeabilité.
- Montrer qu'une fonction n'est jamais négligeable devant elle même.

**Note** : par définition, une fonction  $f$  vérifie  $f \underset{a}{=} o(1)$  si et seulement si

$$\lim_{x \rightarrow a} f(x) = 0$$

Cependant, ce résultat ne donne aucune information sur la vitesse à laquelle  $f(x)$  tend vers 0.

Cette notion de négligeabilité permet de définir des classes de fonctions, dans lesquelles on peut construire une arithmétique un peu particulière :

**Proposition**

Soient  $f_1, f_2, g_1, g_2$  et  $h$  des fonctions définies sur un même intervalle  $I$  et  $a$  un élément ou une borne de  $I$  (finie ou non).

— Si  $f_1 \underset{a}{=} o(g_1)$  et  $g_1 \underset{a}{=} o(h)$ , alors  $f_1 \underset{a}{=} o(h)$ .

— Si  $f_1 \underset{a}{=} o(h)$  et  $f_2 \underset{a}{=} o(h)$ , alors  $(f_1 + f_2) \underset{a}{=} o(h)$ .

— Si  $f_1 \underset{a}{=} o(g_1)$ , alors pour tout  $\lambda \in \mathbb{R}$ ,

$$\lambda \cdot f_1 \underset{a}{=} o(g_1) \quad \text{et} \quad f_1 \underset{a}{=} o(\lambda g_1)$$

— Si  $f_1 \underset{a}{=} o(g_1)$  et  $f_2 \underset{a}{=} o(g_2)$ , alors  $f_1 \cdot f_2 \underset{a}{=} o(g_1 \cdot g_2)$ .

**Exercice** : à démontrer à l'aide de la définition.

**Note** : il arrive également que l'on note  $f \ll g$  pour  $f = o(g)$ . Le premier point de la proposition ci-dessus permet alors de construire des chaînes de négligabilité entre les fonctions usuelles, indiquant les vitesses de convergence vers 0 ou l'infini. Ainsi, classer les fonctions suivantes dans l'ordre de négligeabilité en l'infini :

$$\ln(x), e^x, 10^x, x, x^2, \sqrt{x}, e^{e^x}, x^x, \ln(\ln(x))$$

Que dire de l'inverse de ces fonctions en 0 ?

### 2.6.3 Équivalence

#### Définition (fonctions équivalentes)

Les fonctions  $f$  sont dites *équivalentes au voisinage de  $a$*  s'il existe

- un voisinage  $V(a)$ ,
- une fonction  $\varepsilon$ ,

tels que

$$\forall x \in V(a) \cap I, \quad f(x) = (1 + \varepsilon(x)).g(x) = g(x) + \varepsilon(x).g(x)$$

On note alors

$$f \underset{a}{\sim} g$$

**Note** : si  $g$  ne s'annule pas au voisinage de  $a$ , alors

$$f \underset{a}{\sim} g \iff \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1$$

C'est cette dernière définition que l'on utilise en pratique.

**Exemple** : montrer que  $x^2 + x \underset{+\infty}{\sim} x^2$ , que  $x^2 + x \underset{0}{\sim} x$  et que  $1 + \frac{1}{x} \underset{+\infty}{\sim} \frac{1}{x}$ .

On peut là encore énoncer des résultats en termes de limites. Précisément, si  $f$  et  $g$  admettent des limites (finies ou non) en  $a$  et vérifient  $f \underset{a}{\sim} g$ , alors

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x)$$

et ces convergences se font *à la même vitesse*.

Par ailleurs, on peut montrer que si  $f$  admet une limite finie non nul  $\ell$  en  $a$ , alors

$$f \underset{a}{\sim} \ell$$

**ATTENTION** : ce résultat n'est valable que si  $\ell$  est non nul. De façon générale, aucune fonction hormis la fonction nulle n'est équivalente à 0. Ainsi, si une fonction  $f$  tend vers 0 en  $a$ , elle est équivalente à une autre fonction tendant vers 0.

Enfin, il est là encore possible d'énoncer des résultats généraux concernant la notion d'équivalence et les opérations usuelles sur les fonctions :

**Proposition**

Soient  $f$ ,  $g$  et  $h$  trois fonctions définies au voisinage de  $a \in \overline{\mathbb{R}}$ .

1. Si  $f \underset{a}{\sim} g$ , alors  $g \underset{a}{\sim} f$ .
2. Si  $f \underset{a}{\sim} g$  et  $g \underset{a}{\sim} h$ , alors  $f \underset{a}{\sim} h$ .
3. Si  $f_1 \underset{a}{\sim} g_1$  et  $f_2 \underset{a}{\sim} g_2$ , alors  $f_1 \cdot f_2 \underset{a}{\sim} g_1 \cdot g_2$ .
4. Si  $f_1 \underset{a}{\sim} g_1$  et  $f_2 \underset{a}{\sim} g_2$  et  $g_1$  et  $g_2$  ne s'annulent pas au voisinage de  $a$ , alors

$$\frac{f_1}{f_2} \underset{a}{\sim} \frac{g_1}{g_2}$$

5. Si  $f = o(g)$ , alors  $f + g \underset{a}{\sim} g$ .

**ATTENTION** : il n'est pas permis d'additionner des équivalents. En effet, on peut facilement construire des fonctions  $f_1$ ,  $f_2$ ,  $g_1$  et  $g_2$  telles que

$$\begin{cases} f_1 \underset{a}{\sim} g_1 \\ f_2 \underset{a}{\sim} g_2 \end{cases} \quad \text{et} \quad f_1 + f_2 \not\underset{a}{\sim} g_1 + g_2$$

Enfin, il existe un résultat permettant de construire des équivalents dits usuels pour certaines fonctions (usuelles). Ainsi,

**Proposition (équivalents usuels)**

Si  $f$  est dérivable en  $a$  et  $f'(a) \neq 0$ , alors

$$f(x) - f(a) \underset{a}{\sim} f'(a)(x - a)$$

**Exercice** : à démontrer.

On en déduit les équivalents usuels suivants :

$$\begin{array}{lll} e^x - 1 \underset{0}{\sim} x & \frac{1}{1-x} - 1 \underset{0}{\sim} x & \ln(1+x) \underset{0}{\sim} x \\ \sin(x) \underset{0}{\sim} x & \tan(x) \underset{0}{\sim} x & (1+x)^\alpha - 1 \underset{0}{\sim} \alpha x \end{array}$$

## 2.7 Développements limités

### 2.7.1 Principes généraux

La notion de *développement limité* (DL) est dédiée à l'étude et l'approximation *locale* des fonctions réelles.

Précisément, étant donnée une fonction  $f$ , d'une variable réelle, définie et suffisamment régulière au voisinage d'un point  $x_0 \in \mathbb{R}$  et parfaitement connue en  $x_0$ , la notion de développement limité permet de construire des estimations de la quantité  $f(x)$  lorsque  $x$  est "proche" de  $x_0$ , en s'appuyant sur la notion de *polynôme*.

Dessin

**Note** : on rappelle qu'une fonction polynomiale est une fonction de la forme

$$P : x \mapsto a_0 + a_1 \cdot x + a_2 \cdot x^2 + \dots + a_n \cdot x^n = \sum_{k=0}^n a_k x^k$$

Ainsi, l'évaluation d'une telle fonction en  $x$  (i.e. le calcul de la quantité  $P(x)$  pour un  $x$  donné) nécessite uniquement des additions et multiplications de nombres réels. Ces opérations étant les seules opérations que sait effectuer un ordinateur, la notion de développement limité est alors au cœur du calcul assisté par ordinateur.

Par ailleurs, bien qu'il soit formellement possible de développer la notion de développement limité en tout point  $x_0 \in \mathbb{R}$ , le changement de variable  $x \rightsquigarrow h$  défini par

$$h = x - x_0 \iff x = x_0 + h$$

nous permet d'exposer la théorie uniquement dans le cas  $x_0 = 0$ .

Ainsi, dans la suite, on travaillera sur une fonction  $f$ , définie au voisinage de 0 et parfaitement connue en 0. Les développements limités permettent alors d'établir des approximations pour la quantité  $f(x)$  lorsque  $x$  est "petit" (en valeur absolue).

**Note** : d'un point de vue géométrique, ce changement de variable correspond à un changement de repère, consistant à déplacer le centre du repère en  $x_0$ .

## 2.7.2 Définitions

**Définition (développement limité)**

Soient  $f$  une fonction réelle, définie au voisinage de 0 et  $n$  un entier naturel fixé.

On dit que  $f$  admet un développement limité d'ordre  $n$  en 0 (noté  $DL_n(0)$ ) s'il existe un polynôme de degré  $n$

$$P_{f,n}(x) = \sum_{k=0}^n a_k x^k = a_0 + a_1 x + \dots + a_n x^n$$

tel que la différence  $f(x) - P_{f,n}(x)$  soit *négligeable devant  $x^n$*  au voisinage de 0 :

$$\forall x \in \mathbb{R}, \quad f(x) = P_{f,n}(x) + o(x^n) = \sum_{k=0}^n a_k x^k + o(x^n)$$

Dans ce cas, le polynôme  $P_{f,n}(x)$  est appelé *la partie principale* du  $DL_n(0)$  de  $f$ .

**Notes :**

- Si  $x$  est proche de 0 (on note  $x \ll 1$ ), les puissances de  $x$  sont d'autant plus négligeables que cette puissance est grande. Ainsi, le terme le plus “petit” dans l'expression

$$P_{f,n}(x) = a_0 + a_1 x + \dots + a_n x^n$$

est  $x^n$ . Dans ce cas, le terme d'erreur étant négligeable devant  $x^n$ , il l'est devant l'expression  $P_{f,n}(x)$ , ce qui justifie qu'on puisse le négliger (au sens propre du terme) lors d'éventuelles applications numériques).

- Formellement, l'égalité

$$f(x) = P_{f,n}(x) + o(x^n)$$

est valable pour tout  $x \in \mathbb{R}$ . Cependant, si  $x$  est trop “grand”, le terme d'erreur  $o(x^n)$  peut devenir prépondérant par rapport à la valeur du polynôme  $P_{f,n}(x)$  et cette égalité ne nous apporte aucune information sur la valeur de la quantité  $f(x)$ .

**Exemples :**

- Si  $f$  est dérivable en 0, alors

$$f'(0) = \lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x}$$

Autrement dit, il existe une fonction  $\varepsilon$  telle que

$$\frac{f(x) - f(0)}{x} = f'(0) + \varepsilon(x)$$

avec

$$\lim_{x \rightarrow 0} \varepsilon(x) = 0$$

Mais alors

$$f(x) = f(0) + f'(0).x + x\varepsilon(x) = f(0) + f'(0).x + o(x)$$

La fonction  $f$  admet donc un développement limité à l'ordre 1 en 0 dont la partie principale est le polynôme de degré 1

$$P_{f,1}(x) = f(0) + f'(0).x$$

et le reste est négligeable devant  $x$ .

**Note** : dans l'expression ci-dessus, on reconnaît l'équation de la tangente à la courbe  $\mathcal{C}_f$  de  $f$  en 0 :

$$T_0 : y = f(0) + f'(0).x$$

Géométriquement, le développement limité permet de formaliser l'approximation consistant à assimiler, au voisinage de 0, la courbe de  $f$  à sa tangente  $T_0$ . La notion de reste permet alors de préciser que cette approximation conduit à une erreur négligeable devant la distance  $x$  parcourue.

Dessin

— Si  $f(x) = 1 + x - x^2 + x^2 \sin(x)$ , puisque  $\sin(x) \xrightarrow{x \rightarrow 0} 0$ , on a

$$x^2 \sin(x) \underset{0}{=} o(x^2)$$

Donc

$$f(x) = 1 + x - x^2 + o(x^2)$$

et  $f$  admet un développement limité d'ordre 2 en 0 dont la partie principale est

$$P_{f,2}(x) = 1 + x - x^2$$

et dont le reste est négligeable devant  $x^2$ .

— Pour tout  $n \in \mathbb{N}$ , on a

$$1 + x + x^2 + \dots + x^n = \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x} - \frac{x^{n+1}}{1 - x}$$

Donc

$$\frac{1}{1 - x} = 1 + x + x^2 + \dots + x^n + \frac{x^{n+1}}{1 - x}$$

Or

$$\frac{x^{n+1}}{1 - x} = x^n \cdot \frac{x}{1 - x} \underset{0}{=} o(x^n)$$

La fonction  $x \mapsto \frac{1}{1 - x}$  admet donc un développement limité d'ordre  $n$  en 0 dont la partie principale est

$$P_{f,n}(x) = 1 + x + x^2 + \dots + x^n = \sum_{k=0}^n x^k$$

**Exercice** : montrer que tout polynôme admet des développements limités de tous ordres en 0.

### 2.7.3 Formule de Taylor-Young

Brook Taylor (Anglais, 1685 - 1731) a trouvé au début du XVII<sup>ème</sup> siècle la formule générale donnant la partie principale d'un développement limité à l'ordre  $n$  en 0 d'une fonction  $f$ . Comme à l'ordre 1, cette formule fait intervenir les dérivées successives de la fonction  $f$  en 0. Il faut donc en particulier que ces dérivées existent. Il faut donc que  $f$  soit de classe  $\mathcal{C}^n$ .

#### Théorème (Formule de Taylor)

Soit  $n \in \mathbb{N}$ . Toute fonction  $f$ , de classe  $\mathcal{C}^n$  au voisinage de 0 admet un développement limité d'ordre  $n$  en 0, donné par

$$\begin{aligned} f(x) &= \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + o(x^n) \\ &= f(0) + f'(0) \cdot x + \frac{f''(0)}{2} \cdot x^2 + \frac{f'''(0)}{3!} x^3 + \dots + \frac{f^{(n)}(0)}{n!} x^n + o(x^n) \end{aligned}$$

Cette formule a en particulier permis de déterminer les développements limités des fonctions usuelles en 0 (voir Table 2.1).

**Notes :**

- La formule de Taylor ci-dessus induit en particulier que toute fonction suffisamment régulière admet un *unique* développement limité d'ordre  $n$  en 0.

**ATTENTION :** la réciproque est fautive : deux fonctions différentes peuvent avoir le même développement limité à un ordre donné (chercher des exemples dans les exemples ci-dessus).

- Le développement limité des fonctions sinus et cosinus donnés dans la table 2.1 illustre une propriété des développements limités des fonctions à parité. On peut en effet montrer que, de façon générale,

- Si  $f$  est une fonction paire admettant un  $DL_n(0)$ , la partie principale de  $f$  ne contient **que des puissances paires de  $x$** .
- Si  $f$  est une fonction impaire admettant un  $DL_n(0)$ , la partie principale de  $f$  ne contient **que des puissances impaires de  $x$** .

Par ailleurs, à partir de ces développements limités usuels, on peut calculer le déve-

$e^h = 1 + \frac{h}{1!} + \frac{h^2}{2!} + \cdots + \frac{h^n}{n!} + o(h^n)$
$a^h = 1 + \frac{h \ln a}{1!} + \frac{(h \ln a)^2}{2!} + \cdots + \frac{(h \ln a)^n}{n!} + o(h^n)$
$(1+h)^\alpha = 1 + \frac{\alpha h}{1!} + \frac{\alpha(\alpha-1)h^2}{2!} + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)h^n}{n!} + o(h^n)$
$\frac{1}{1-h} = 1 + h + h^2 + h^3 + h^4 + h^5 \cdots + h^n + o(h^n)$
$\ln(1+h) = h - \frac{h^2}{2} + \frac{h^3}{3} - \frac{h^4}{4} \cdots + (-1)^{n-1} \frac{h^n}{n} + o(h^n)$
$\sin h = h - \frac{h^3}{3!} + \frac{h^5}{5!} - \frac{h^7}{7!} + \cdots + \frac{(-1)^n h^{2n+1}}{(2n+1)!} + o(h^{2n+2}) \quad \text{ATTENTION : ordre } 2n+2$
$\cos h = 1 - \frac{h^2}{2!} + \frac{h^4}{4!} - \frac{h^6}{6!} + \cdots + \frac{(-1)^n h^{2n}}{(2n)!} + o(h^{2n+1}) \quad \text{ATTENTION : ordre } 2n+1$
$\tan h = h + \frac{h^3}{3} + \frac{2h^5}{15} + \frac{17h^7}{315} + o(h^8)$
$\sinh h = h + \frac{h^3}{3!} + \frac{h^5}{5!} + \frac{h^7}{7!} + \cdots + \frac{h^{2n+1}}{(2n+1)!} + o(h^{2n+2})$
$\cosh h = 1 + \frac{h^2}{2!} + \frac{h^4}{4!} + \frac{h^6}{6!} + \cdots + \frac{h^{2n}}{(2n)!} + o(h^{2n+1})$
$\tanh h = h - \frac{h^3}{3} + \frac{2h^5}{15} - \frac{17h^7}{315} + o(h^8)$

TABLE 2.1 – Développements limités usuels

veloppement limité de très nombreuses fonctions réelles, construites à partir des fonctions usuelles et des opérations élémentaires  $+$ ,  $\times$ ,  $\circ$ , etc.

Ces méthodes de calculs reposent principalement sur le fait qu'une somme, un produit ou une composée de polynômes est encore un polynôme.

Cependant, ces opérations élémentaires sur les polynômes peuvent avoir une influence forte sur le degré des polynômes ainsi obtenus. Nous verrons en TD comment tenir compte de ces évolutions et comment déterminer le développement limité d'une fonction construite à partir des fonctions usuelles.

## 2.7.4 Développement limité hors de 0

### Développement limité en un point $x_0 \in \mathbb{R}^*$

Il est possible de généraliser la notion de développement limité en tout point  $x_0 \in \mathbb{R}$ .

#### Définition (développement limité en un point $x_0 \in \mathbb{R}$ )

Soient  $x_0 \in \mathbb{R}$  et  $f$  une fonction réelle définie au voisinage de  $x_0$ . La fonction  $f$  admet un développement limité d'ordre  $n$  en  $x_0$  s'il existe un polynôme de degré  $n$

$$P_{f,n}(X) = \sum_{k=0}^n a_k X^k$$

tel que

$$f(x) = P_{f,x_0}(x - x_0) + o((x - x_0)^n) = \sum_{k=0}^n a_k (x - x_0)^k + o((x - x_0)^n)$$

On peut alors généraliser la formule de Taylor vue en 0 à n'importe quel point  $x_0$  :

#### Théorème (formule de Taylor générale)

Soit  $f$  une fonction de classe  $\mathcal{C}^n$  au voisinage d'un point  $x_0 \in \mathbb{R}$ . Alors

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + o((x - x_0)^n)$$

$$\iff_{x=x_0+h} f(x_0 + h) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} h^k + o(h^n)$$

En pratique, pour déterminer le développement limité d'une fonction  $f$  donnée au

voisinage d'un point  $x_0 \neq 0$ , on pose le changement de variable

$$h = x - x_0 \iff x = x_0 + h$$

On peut alors étudier la fonction

$$g : h \mapsto f(x_0 + h)$$

au voisinage de 0.

**Exemple** : on souhaite établir le développement limité à l'ordre 3 en 1 de la fonction

$$f : x \mapsto \ln(x)$$

Pour cela, on pose

$$h = x - 1 \iff x = 1 + h$$

et on étudie la fonction

$$g : h \mapsto f(1 + h) = \ln(1 + h)$$

On a alors

$$\begin{aligned} g(h) &= f(1 + h) = \ln(1 + h) \\ &= h - \frac{h^2}{2} + \frac{h^3}{3} + o(h^3) \end{aligned}$$

et

$$f(x) = (x - 1) - \frac{(x - 1)^2}{2} + \frac{(x - 1)^3}{3} + o((x - 1)^3)$$

### Développements asymptotiques

Il est également possible d'étendre la notion de développement limité en l'infini.

Formellement, un développement limité en  $\pm\infty$  est appelé *développement asymptotique*.

En pratique, pour étudier une fonction  $f$  au voisinage de  $\pm\infty$ , on effectue le changement de variable

$$h = \frac{1}{x} \iff x = \frac{1}{h}$$

de sorte que

$$h \underset{x \rightarrow \pm\infty}{\longrightarrow} 0$$

On peut alors étudier, au voisinage de 0, la fonction

$$g : h \mapsto f\left(\frac{1}{h}\right)$$

**Exemple** : déterminer le développement asymptotique de la fonction  $f : x \mapsto \frac{x}{1+x}$  à l'ordre 3 en  $+\infty$ .

On pose

$$g(h) = f\left(\frac{1}{h}\right) = \frac{\frac{1}{h}}{1 + \frac{1}{h}} = \frac{1}{h+1}$$

On a alors

$$g(h) = 1 - h + h^2 - h^3 + o(h^3)$$

et

$$f(x) = 1 - \frac{1}{x} + \frac{1}{x^2} - \frac{1}{x^3} + o\left(\frac{1}{x^3}\right)$$

**Note** : à l'image de l'exemple ci-dessus, un développement asymptotique sous la forme d'un polynôme “en  $\frac{1}{x}$ ” plus un terme d'erreur négligeable devant une puissance de  $\frac{1}{x}$ .

### 2.7.5 Applications des développements limités

La notion de développement limité permet, de façon générale, d'étudier de façon précise le comportement local d'une fonction donnée.

En pratique, les applications principales portent donc sur le calcul de limites (et l'étude de formes indéterminées) et la détermination d'équivalents (donnant une information plus

précise que le simple calcul de limite).

Par ailleurs, la notion de développement limité à l'ordre 1 ou 2 peuvent également se traduire sous la forme de propriétés géométriques (toujours locales) portant sur la courbe de la fonction étudiée.

### Calcul de limites et équivalent

Le fait d'approximer les fonctions à l'aide de polynômes permet de lever les indétermination, notamment de la forme " $\frac{0}{0}$ ".

**Exemple** : on souhaite calculer la limite

$$\lim_{x \rightarrow 0} \frac{\sin(x) - x}{x^3}$$

Il s'agit d'une forme indéterminée de la forme " $\frac{0}{0}$ ". Cependant, le développement limité à l'ordre 3 de la fonction sinus permet de lever cette indétermination :

$$\sin(x) = x - \frac{x^3}{6} + o(x^3)$$

d'où

$$\begin{aligned} f(x) &= \frac{x - \frac{x^3}{6} + o(x^3) - x}{x^3} \\ &= \frac{-\cancel{x} + o(\cancel{x^3})}{\cancel{x^3}} & (*) \\ &= -\frac{1}{6} + o(1) \end{aligned}$$

La simplification effectuée à la ligne (\*) permet de lever l'indétermination et on peut conclure que

$$\lim_{x \rightarrow 0} f(x) = -\frac{1}{6}$$

Par ailleurs, les développements limités permettent également d'établir des équivalents pour les fonctions développées.

Ainsi, d'un point de vue théorique, on peut montrer qu'une fois établi le développement limité d'une fonction  $f$  en un point  $x_0 \in \overline{\mathbb{R}}$ , le terme prépondérant de ce développement limité (i.e. le terme de plus bas degré) constitue naturellement un équivalent simple de la

quantité  $f(x)$  au voisinage de  $x_0$ .

Par ailleurs, en pratique, les développements limités étant une “vraie” égalité (contrairement à la notion d'équivalent), ils permettent des opérations que ne permettent pas les équivalents (notamment l'addition).

Ainsi, dans l'exemple précédent, une fois la limite obtenue, on peut déterminer un équivalent de la quantité

$$\frac{\sin(x) - x}{x^3} + \frac{1}{6}$$

au voisinage de 0.

Pour cela, on doit “tirer” le développement limité de la fonction sinus à l'ordre 5 :

$$\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{5!} + o(x^5)$$

On a alors

$$\begin{aligned} f(x) + \frac{1}{6} &= \frac{x - \frac{x^3}{6} + \frac{x^5}{5!} + o(x^5) - x}{x^3} \\ &= \frac{-\cancel{x} + \frac{x^{\cancel{3}}}{6} + \frac{x^{\cancel{5}}}{5!} + o(x^{\cancel{5}})}{\cancel{x^3}} + \frac{1}{6} \\ &= -\frac{1}{6} + \frac{x^2}{5!} + o(x^2) + \frac{1}{6} \end{aligned}$$

D'où

$$f(x) + \frac{1}{6} = \frac{x^2}{5!} + o(x^2) \underset{0}{\sim} \frac{x^2}{120}$$

### Position relatives courbes - tangentes - asymptotes

Le développement limité à l'ordre 1 d'une fonction  $f$  étant une approximation affine de la fonction étudiée, il renseigne en particulier sur les tangentes à la courbe de  $f$ . Précisément, la partie principale du développement limité d'ordre 1 en 0

$$f(x) = f(0) + f'(0).x + o(x)$$

donne l'équation de la tangente en 0. En poussant le développement limité à l'ordre 2, on peut accéder au signe de la différence

$$f(x) - (f(0) + f'(0).x)$$

qui permet d'étudier la position relative du graphe de  $f$  et de sa tangente en 0.

**Exemple** : étudier la fonction arctan en 0.

En l'infini, la notion de tangente est remplacée par la notion de droite asymptote. Là encore le développement asymptotique permet d'accéder à l'équation de la droite asymptote éventuelle et aux positions relatives de la courbe et de la droite.

Enfin, de façon plus générale, les développements asymptotiques permettent de préciser les branches infinies d'une fonctions donnée.

**Exemple** : étudier le comportement asymptotique des fonctions

$$f : x \mapsto \sqrt{\frac{x^3}{x-1}} \quad \text{et} \quad g : x \mapsto \sqrt{\frac{x^4+1}{x-1}}$$

# Chapitre 3

## Suites numériques

### Introduction

Les suites numériques constituent l'un des outils principaux de l'approximation numérique.

Formellement, les suites numériques appartiennent à la classe des fonctions réelles, dont l'ensemble de départ est réduit à tout ou partie de l'ensemble des entiers naturels. On verra qu'il est possible d'étendre certains des outils d'analyse des fonctions d'une variable réelle à l'étude des suites numériques. Cependant, la nature discrète de l'ensemble  $\mathbb{N}$  (qui nous prive par exemple de la notion de dérivée) permet de mettre en place des outils d'études spécifiques à ces objets.

De façon moins formelle, une suite numérique est une famille infinie de nombres ordonnés et numérotés. Il existe différentes façons de construire une suite numérique. Choisir les nombres un par un au hasard permet par exemple de construire des suites *aléatoires*. Cependant, dans ce chapitre, nous étudierons exclusivement des suites construites de façon logique. Or il existe deux façons distinctes de définir une suite numérique de façon logique :

- *La forme explicite* consistant à donner explicitement  $u_n$  en fonction de  $n$ , i.e. en définissant chaque terme en fonction de sa place dans la suite.
- *La forme récurrente* : en donnant une règle logique pour passer d'un terme au suivant, i.e. en donnant  $u_n$  en fonction du ou des termes précédents  $(u_{n-1}, u_{n-2}, \dots)$ .

La forme explicite est de loin la plus pratique à étudier notamment en adaptant les outils de l'analyse des fonctions réelles. Cependant, la forme récurrente est souvent la plus adaptée pour modéliser des systèmes dynamiques (l'état d'un système à un instant donné dépend de ses états aux instants précédents). On verra donc qu'il existe des outils spécifiques à l'étude des suites récurrentes, notamment basée sur une classification des suites récurrentes en fonction de la nature de leur relation de récurrence.

Nous terminerons ce chapitre par des exemples d'applications de ces notions au *calcul numérique*, i.e. au calcul assisté par ordinateur.

## 3.1 Généralités

### 3.1.1 Définitions

Formellement, une suite numérique  $(u_n)_{n \in \mathbb{N}}$  est donc une fonction

$$u : \begin{array}{l} \llbracket n_0, +\infty \llbracket \longrightarrow \mathbb{R} \\ n \longmapsto u(n) = u_n \end{array}$$

où  $n_0$  est un entier naturel fixé et l'antécédent  $n$  étant plutôt noté en indice qu'entre parenthèses.

L'ensemble  $\mathbb{N}^{\geq n_0} = \llbracket n_0, +\infty \llbracket$  est alors appelé *l'ensemble des indices*, plutôt que l'ensemble de définition.

Le terme  $u_n$  est appelé *terme de rang  $n$*  et l'on note  $(u_n)_{n \geq n_0}$  (ou parfois  $(u_n)$  quand il n'y a pas d'ambiguïté sur l'ensemble des indices) l'ensemble des termes de la suite.

Comme on l'a évoqué en introduction, il y a deux façons distinctes de définir une suite numérique de façon logique :

— *La forme explicite* : on donne explicitement  $u_n$  en fonction de son rang  $n$ .

**Exemples :**

— Toute suite associée à une fonction réelle  $f$  définie sur  $\mathbb{N}$  (ou  $\mathbb{N}^{\geq n_0}$ ) :

$$(u_n)_{n \in \mathbb{N}} : u_n = n^2 + 1, \quad (v_n)_{n \in \mathbb{N}^*} : v_n = e^{(\frac{1}{n})}, \quad (w_n)_{n \in \mathbb{N}^*} : w_n = \sin\left(\frac{n\pi}{2}\right)$$

— Autres :

$$(x_n)_{n \in \mathbb{N}^*} : x_n = \frac{n!}{n^n}, \quad (y_n)_{n \in \mathbb{N}} : y_n = \sum_{k=0}^n 2^k, \quad (z_n)_{n \in \mathbb{N}} : z_n = \prod_{k=0}^n 2^k$$

— *La forme récurrente* : on fixe suffisamment de valeurs initiales et l'on donne une règle de calcul permettant de calculer chaque terme en fonction des précédents. La relation liant  $u_n$  aux termes précédents  $u_{n-1}$ ,  $u_{n-2}$ , etc est *la relation de récurrence*.

**Exemples :**

$$(u_n)_{n \in \mathbb{N}} : \begin{cases} u_0 = 1 \\ \forall n \geq 0, u_{n+1} = u_n + 2 \end{cases} \quad (v_n)_{n \in \mathbb{N}} : \begin{cases} v_0 = 0, v_1 = 1 \\ \forall n \geq 0, v_{n+2} = v_{n+1} + v_n \end{cases}$$

Le nombre de termes dont dépend chaque terme est *l'ordre* de la suite récurrente. Il correspond également au nombre de valeurs initiales nécessaires à l'amorçage de la suite. Ici,  $(u_n)$  est d'ordre 1,  $(v_n)$  est d'ordre 2.

### 3.1.2 Représentation graphique

Il existe différentes représentations graphiques pour une suite numérique donnée :

- en plaçant les points d'abscisse  $(u_n)$  sur une droite graduée,
- en plaçant les points de coordonnées  $(n, u_n)$  dans un repère du plan.

D'un point de vu formel, il est également possible de définir des suites à valeurs complexes  $(z_n)_{n \in \mathbb{N}}$ . Cependant, on peut montrer qu'en général, l'étude d'une telle suite se fait via l'étude des deux suites réelles associées à  $(z_n)$  :  $(u_n) = (\operatorname{Re}(z_n))$  et  $(v_n) = (\operatorname{Im}(z_n))$ . Notons tout de même qu'il est possible de représenter géométriquement une suite complexe en plaçant les différents termes  $z_n$  dans le plan complexe.

**Exemples :**

- Illustrer les deux représentations pour la suite  $(u_n)$  :  $u_n = \frac{(-1)^n}{n}$
- Représenter dans le plan complexe les premiers termes de la suite  $(z_n)$  :  $z_n = \frac{1}{n} e^{\frac{i n \pi}{3}}$

## 3.2 Convergence d'une suite

L'une des caractéristiques principales d'une suite numérique  $(u_n)$  donnée est son *comportement asymptotique*, i.e. la façon dont évolue la quantité  $u_n$  quand  $n$  tend vers  $+\infty$ .

Formellement, cela correspond à la notion de limite. Ainsi,

### Définition (Convergence d'une suite numérique)

- Une suite  $(u_n)$  est dite *convergente* s'il existe un réel  $\ell \in \mathbb{R}$  tel que

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} / n \geq n_0 \Rightarrow |u_n - \ell| < \varepsilon$$

On dit alors que  $(u_n)$  *converge vers*  $\ell$  et on note

$$\lim_{n \rightarrow +\infty} u_n = \ell \quad \text{ou} \quad \lim u_n = \ell \quad \text{ou} \quad u_n \xrightarrow[n \rightarrow +\infty]{} \ell$$

- Si  $(u_n)$  n'est pas convergente, elle est dite *divergent*.

On dira de plus qu'elle *diverge vers*  $+\infty$  (resp. vers  $-\infty$ ) si

$$\forall A > 0, \exists n_0 \in \mathbb{N} / n \geq n_0 \Rightarrow u_n > A \quad (\text{resp. } u_n < -A)$$

On note alors

$$\lim_{n \rightarrow +\infty} u_n = \pm\infty \quad \text{ou} \quad \lim u_n = \pm\infty \quad \text{ou} \quad u_n \xrightarrow[n \rightarrow +\infty]{} \pm\infty$$

**Notes :**

- Une suite  $(u_n)$  converge vers le réel  $\ell$  si l'on peut toujours trouver un rang à partir duquel *tous les termes* de la suite sont aussi proches que l'on veut de  $\ell$ .
- Une suite  $(u_n)$  diverge vers  $+\infty$  si l'on peut trouver un rang à partir duquel tous les termes de la suite sont aussi grands que l'on veut.

**Exemples** : déterminer la nature (convergente ou divergente) et la limite éventuelle des suites ci-dessous :

$$(u_n)_{n \in \mathbb{N}^*} : u_n = \frac{1}{n}, \quad (v_n)_{n \in \mathbb{N}} : v_n = n^2, \quad (w_n)_{n \in \mathbb{N}} : w_n = (-1)^n$$

La plupart des outils et propriétés que nous verrons dans ce chapitre ont pour but d'établir des critères de convergence pour les suites numériques. Ainsi, dans le cas des suites explicites définies par une relation de la forme

$$u_n = f(n)$$

où  $f$  est une fonction réelle connue, définie à partir des fonctions usuelles, les outils de l'analyse appliqués à  $f$  permettent de d'énoncer certains critères de ce type :

**Proposition**

Soit  $f$  une fonction réelle définie sur  $\mathbb{R}^+$ .

- Si la fonction  $f$  admet une limite finie  $\ell$  en  $+\infty$ , alors la suite  $(f(n))_{n \in \mathbb{N}}$  converge vers  $\ell$ .
- Si  $\lim_{x \rightarrow +\infty} f(x) = \pm\infty$ , alors  $(f(n))_{n \in \mathbb{N}}$  diverge vers  $\pm\infty$ .

**Attention** : si la fonction  $f$  n'admet pas de limite (finie ou non) en  $+\infty$ , il n'existe pas de résultat général. La suite  $(u_n)$  définie par

$$(u_n) : u_n = \sin\left(n\pi + \frac{1}{n}\right)$$

tend vers 0, alors que la fonction  $f : x \mapsto \sin\left(\pi x + \frac{1}{x}\right)$  n'admet pas de limite en  $+\infty$ .

**Exercice** : représenter graphiquement la suite  $(u_n)$  ci-dessus en s'appuyant sur le graphe de la fonction  $f$  associée.

## 3.3 Propriétés des suites réelles

### 3.3.1 Suites bornées

#### Définition (suites majorées)

Une suite numérique  $(u_n)$  est dite *majorée* si il existe un nombre  $M \in \mathbb{R}$  indépendant de  $n$  tel que

$$\forall n \in \mathbb{N}, u_n \leq M$$

#### Définition (suites minorée)

Une suite numérique  $(u_n)$  est dite *minorée* si il existe un nombre  $m \in \mathbb{R}$  indépendant de  $n$  tel que

$$\forall n \in \mathbb{N}, u_n \geq m$$

#### Définition (suites bornées)

Une suite numérique  $(u_n)$  est dite *bornée* si elle est majorée et minorée, i.e. s'il existe  $(m, M) \in \mathbb{R}^2$  tel que

$$\forall n \in \mathbb{N}, m \leq u_n \leq M$$

OU s'il existe  $A \geq 0$ , indépendant de  $n$  tel que

$$\forall n \in \mathbb{N}, |u_n| \leq A$$

**Exercice** : montrer que les deux définitions ci-dessus sont équivalentes, i.e. qu'elles déterminent le même ensemble de suites.

**Exemples** : parmi les suites ci-dessous, lesquelles sont majorée, minorée, bornée ? On donnera des exemples de bornes éventuelles.

$$\begin{array}{lll} (u_n)_{n \in \mathbb{N}} : u_n = n, & (v_n)_{n \in \mathbb{N}} : v_n = -e^n, & (w_n)_{n \in \mathbb{N}} : w_n = (-1)^n, \\ (x_n)_{n \in \mathbb{N}^*} : x_n = \frac{1}{n}, & (y_n)_{n \in \mathbb{N}} : y_n = n \cdot (-1)^n, & (z_n)_{n \in \mathbb{N}} : z_n = \sin(n^2 - n) \end{array}$$

### 3.3.2 Suites monotones

On peut également caractériser certaines suites en fonctions de leur évolution (croissantes, décroissantes) en comparant deux termes consécutifs quelconques  $u_n$  et  $u_{n+1}$ .

**Définition (suites croissantes)**

Une suite numérique  $(u_n)$  est dite *croissante* (resp. *strictement croissante*) si

$$\forall n \in \mathbb{N}, u_{n+1} \geq u_n \quad (\text{resp. } u_{n+1} > u_n)$$

**Définition (suites décroissantes)**

Une suite numérique  $(u_n)$  est dite *décroissante* (resp. *strictement décroissante*) si

$$\forall n \in \mathbb{N}, u_{n+1} \leq u_n \quad (\text{resp. } u_{n+1} < u_n)$$

**Définition (suites monotones)**

Une suite numérique  $(u_n)$  est dite *monotone* si elle est soit croissante soit décroissante.

**Définition (suites stationnaires)**

Une suite numérique  $(u_n)$  est dite *stationnaire* si

$$\exists n_0 \in \mathbb{N} / \forall n \geq n_0, u_{n+1} = u_n$$

**Note** : au vue des définitions ci-dessus, l'étude de la monotonie d'une suite passe naturellement par l'étude du signe de la différence  $u_{n+1} - u_n$ .

Dans le cas des suites à termes *strictement positifs*, il sera également possible de passer par l'étude du quotient  $\frac{u_{n+1}}{u_n}$  (que l'on comparera alors à 1).

**Attention** : toutes les suites ne sont pas monotones. Certaines sont ni croissantes, ni décroissantes.

### 3.4 Critères de convergence

Comme on l'a dit plus haut, l'étude d'une suite donnée a pour but premier de déterminer sa nature et la valeur de sa limite éventuelle.

La théorie des suites numériques consiste donc essentiellement à établir des critères

assurant (ou non) la convergence d'une suite donnée et d'obtenir un maximum d'informations sur la valeur de la limite éventuelle.

Ces critères sont de plusieurs types. Ainsi, outre les critères déjà évoqué portant sur les suites explicites de la forme  $u_n = f(n)$ , on verra qu'il existe également des critères intrinsèques, essentiellement basés sur les variations d'une suite donnée.

Enfin, on verra qu'il existe des critères de comparaison, permettant de comparer une suite donnée à un catalogue de suites dont on connaît le comportement. On verra également qu'à l'aide de ce catalogue, il est possible d'établir une échelle de vitesse de convergence et que les critères de comparaison permettent de placer une suite donnée sur cette échelle.

### 3.4.1 Limites et opérations

Comme pour les fonctions réelles, il existe des règles naturelle de convergence des suites définies à partir des opérations usuelles ( $+$ ,  $\times$ ,  $\div$ ) faisant apparaître la notion de formes indéterminées (F.I.). L'ensemble des résultats peut être résumé dans les tableaux ci-dessous : soient  $(u_n)$  et  $(v_n)$  deux suites convergentes et soient  $l_1$  et  $l_2$  leurs limites respectives. On a

limite d'une somme  $(u_n + v_n)$

$l_1 \backslash l_2$	$-\infty$	$l_2 \in \mathbb{R}$	$+\infty$
$-\infty$			
$l_1 \in \mathbb{R}$			
$+\infty$			

limite d'un produit ( $u_n \times v_n$ )

$l_1 \backslash l_2$	$-\infty$	$l_2 < 0$	$l_2 = 0$	$l_2 > 0$	$+\infty$
$-\infty$					
$l_1 < 0$					
$l_1 = 0$					
$l_1 > 0$					
$+\infty$					

limite d'un quotient ( $\frac{u_n}{v_n}$ )

$l_1 \backslash l_2$	$-\infty$	$l_2 < 0$	$l_2 = 0^-$	$l_2 = 0^+$	$l_2 > 0$	$+\infty$
$-\infty$						
$l_1 < 0$						
$l_1 = 0$						
$l_1 > 0$						
$+\infty$						

**Exemples** : déterminer les limites des fonctions suivantes :

$$(u_n)_{n \geq 2} : u_n = \frac{n^2 + 1}{n^2 - 1}, \quad (v_n)_{n \geq 1} : v_n = \sin \frac{1}{n}, \quad (w_n)_{n \geq 1} : w_n = \frac{e^{\frac{1}{n}} - 1}{n}$$

### 3.4.2 Critères de comparaison

Il est enfin possible d'étendre les notions de comparaisons locales aux suites numériques et d'en tirer des critères de convergence basés notamment sur la comparaison des suites étudiées à des familles de suites de références.

#### Limites et inégalités

Le résultat principal de cette partie est le suivant :

##### Théorème

Soient  $(u_n)$  et  $(v_n)$  deux suites convergentes telles que

$$\exists n_0 \in \mathbb{N} / \forall n \geq n_0, \quad u_n \leq v_n$$

Alors

$$\lim_{n \rightarrow +\infty} u_n \leq \lim_{n \rightarrow +\infty} v_n$$

dont on tire quelques résultats corollaires :

##### Corollaire

— **Suites bornées** : soit  $(u_n)$  une suite convergente, bornée respectivement par  $m$  et  $M$ . Alors

$$m \leq \lim_{n \rightarrow +\infty} u_n \leq M$$

— **Théorème des gendarmes** : soient  $(u_n)$ ,  $(v_n)$  et  $(w_n)$  trois suites telles que

$$\exists n_0 \in \mathbb{N} / \forall n \geq n_0, \quad u_n \leq v_n \leq w_n$$

Si  $(u_n)$  et  $(w_n)$  convergent vers une limite commune  $\ell \in \mathbb{R}$ , alors  $(v_n)$  converge également vers  $\ell$ .

On peut également étendre certains résultats aux suites admettant une limite infinie :

##### Corollaire

Soient  $(u_n)$  et  $(v_n)$  telles que pour tout  $n \geq n_0$ ,  $u_n \leq v_n$ . Alors

— Si  $(u_n)$  tend vers  $+\infty$ , alors  $(v_n)$  aussi.

— Si  $(v_n)$  tend vers  $-\infty$ , alors  $(u_n)$  aussi.

**Note** : dans tous les résultats ci-dessus, les suites  $(u_n)$ ,  $(v_n)$  et  $(w_n)$  vérifient des inégalités larges ( $\leq$ ). Il est évidemment possible de construire des suites vérifiant ces inégalités au sens strict du terme ( $<$ ). Cependant, ce changement ne peut s'étendre aux limites, qui restent, elles, ordonnées par des inégalités larges.

**Exercice** : construire une suite strictement positive dont la limite est nulle.

### Suites monotones bornées

À l'aide de la définition de la convergence, on peut montrer les propriétés suivantes :

#### Proposition

- Toute suite croissante et majorée est convergente.
- Toute suite décroissante et minorée est convergente.
- Toute suite croissante, non majorée tend vers  $+\infty$ .
- Toute suite décroissante, non minorée tend vers  $-\infty$ .

**Attention** : bien que ces résultats permettent souvent de déterminer la nature d'une suite donnée, ils ne permettent pas, en général, de déterminer la valeur de la limite. En effet, si l'on parvient à déterminer un majorant  $M$  pour tous les termes d'une suite  $(u_n)$  croissante, le seul résultat que l'on obtient concernant la valeur de la limite  $\ell$  de  $(u_n)$  est  $\ell \leq M$ .

### Suites adjacentes

#### Définition/Propriété (suites adjacentes)

Deux suites  $(u_n)$  et  $(v_n)$  sont dite *adjacentes* si

- $(u_n)$  est croissante,
- $(v_n)$  est décroissante,
- la différence  $(u_n - v_n)$  tend vers 0.

Si  $(u_n)$  et  $(v_n)$  sont deux suites adjacentes, alors elles convergent toutes les deux et ont la même limite.

### Suites extraites

Soit  $(u_n)$  une suite numérique. On appelle *suite extraite de  $(u_n)$*  toute suite construite à partir de certains termes de la suite  $(u_n)$ , pris dans l'ordre imposé par la suite  $(u_n)$ . Formellement, on peut associer toute suite  $(v_n)$  extraite de  $(u_n)$  à une fonction  $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ , strictement croissante, telle que

$$\forall n \in \mathbb{N}, \quad v_n = u_{\varphi(n)}$$

De façon générale, les suites extraites d'une suite donnée héritent du comportement de la suite globale. Précisément, si  $(u_n)$  est une suite convergente, alors toute suite extraite de  $(u_n)$  est convergente et a la même limite de  $(u_n)$ .

Ce résultat permet surtout de déterminer la nature divergente d'une suite. En effet, si l'on peut extraire d'une même suite  $(u_n)$  deux suites aux comportements différents, alors  $(u_n)$  diverge.

**Exemple :** Montrer que la suite  $(u_n) : u_n = (-1)^n$  diverge.

**Attention :** il est rarement possible de déterminer la convergence d'une suite donnée à partir de l'étude de certaines suites extraites. Ainsi, si  $(u_n) : u_n = \sin\left(\frac{n\pi}{2}\right)$ . Les suites  $(u_{4n})$  et  $(u_{4n+2})$  convergent vers 0 (elles sont même constantes), mais la suite  $(u_n)$  diverge.

On a cependant le résultat suivant :

#### Proposition (suites extraites et convergence)

Soit  $(u_n)$  une suite numérique. Si l'on peut extraire de  $(u_n)$  un ensemble de suites telles que

- toutes ces suites convergent vers la même limite  $\ell$ ,
- l'ensemble de ces suites recouvre l'ensemble des termes de la suite  $(u_n)$ .

Alors la suite  $(u_n)$  converge vers  $\ell$ .

### Comparaison asymptotiques

On peut également étendre les notions de comparaison asymptotique aux suites numériques. Ainsi

**Définition (comparaison asymptotique)**

Soient  $(u_n)$  et  $(v_n)$  deux suite numériques à termes positifs.

— La suite  $(u_n)$  est *dominée* par la suite  $(v_n)$  si

$$\exists M \geq 0, n_0 \in \mathbb{N} / \forall n \geq n_0, u_n \leq M.v_n$$

On note alors  $u_n = O(v_n)$ , qui se lit “ $u_n$  est un grand  $O$  de  $v_n$ ”.

— La suite  $(u_n)$  est *négligeable* devant  $(v_n)$  si

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} / \forall n \geq n_0, u_n \leq \varepsilon.v_n$$

On note  $u_n = o(v_n)$ , qui se lit “ $u_n$  est un petit  $o$  de  $(v_n)$ ”.

—  $(u_n)$  et  $(v_n)$  sont dites *équivalentes* si

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} / \forall n \geq n_0, |u_n - v_n| \leq \varepsilon.v_n$$

On note  $u_n \sim v_n$  qui se lit “ $u_n$  est équivalent à  $v_n$ ”.

Si la suite  $(v_n)$  ne s’annule pas, ces définitions peuvent alors s’exprimer à partir du quotient  $\frac{u_n}{v_n}$  :

**Proposition**

Soient  $(u_n)$  et  $(v_n)$  deux suites à termes positifs telles que  $(v_n)$  ne s’annule pas. Alors

—  $(u_n)$  est dominée par  $(v_n)$  si la suite  $\left(\frac{u_n}{v_n}\right)$  est bornée.

—  $(u_n)$  est négligeable devant  $(v_n)$  si la suite  $\left(\frac{u_n}{v_n}\right)$  tend vers 0.

—  $(u_n)$  et  $(v_n)$  sont équivalentes si la suite  $\left(\frac{u_n}{v_n}\right)$  tend vers 1.

**Exercice** : à démontrer à partir des définitions ci-dessus et de la définition de la convergence.

**Exemples** :

$$\sqrt{4n^2 + 1} = O(n), \quad \sqrt{4n^2 + 1} = o(n^2), \quad \sqrt{4n^2 + 1} \sim 2n$$

Enfin, ces critères de comparaison asymptotique s’étendent à la nature et aux limites éventuelles des suites numériques à termes positifs :

**Proposition**

Soient  $(u_n)$  et  $(v_n)$  deux suites à termes positifs.

- Si  $(u_n)$  est dominée par  $(v_n)$  ou négligeable devant  $(v_n)$ , alors :
  - Si  $(v_n)$  tend vers 0, alors  $(u_n)$  aussi.
  - Si  $(u_n)$  tend vers  $+\infty$ , alors  $(v_n)$  aussi.
- Si  $(u_n)$  et  $(v_n)$  sont équivalentes, alors elles sont de même nature et ont la même limite éventuelle.

Enfin, la notion de négligeabilité permet de mettre en place une échelle de comparaison permettant de déterminer le comportement asymptotique d'une suite donnée en la plaçant sur cette échelle (ici, on note  $u_n \ll v_n$  pour  $u_n = o(v_n)$ ).

**Exercice :**

- Limite infinie : classer par ordre de domination les suites définie ci-dessous :

$$1, \quad n^n, \quad \sqrt{n}, \quad n^2, \quad n, \quad 2^n, \quad \ln(n), \quad 10^n, \quad \ln(\ln(n)), \quad n!$$

- Limite nulle : classer par ordre de domination les *inverses* des suites ci-dessus.

## 3.5 Suites récurrentes

### 3.5.1 Définitions

Intuitivement, une suite récurrente est une suite que l'on construit terme à terme, chaque terme étant calculé à l'aide des précédents. Pour construire une telle suite, il faut donc

1. une série de valeurs de départ (suffisamment pour initialiser la suite) appelées *conditions initiales*,
2. une relation permettant de calculer chaque terme en fonction des précédents, appelée *relation de récurrence* de la suite.

On note

$$(u_n) : \begin{cases} u_0, u_1, \dots \text{ donné} \\ u_{n+1} = f(u_n, u_{n-1}, \dots) \end{cases}$$

L'ordre d'une suite récurrente est alors le nombre de termes nécessaires au calcul d'un terme supplémentaire. Cet ordre correspond également au nombre de conditions initiales

nécessaire à l'initialisation de la suite.

En pratique, les suites récurrentes offrent un outil idéal de modélisation de certains dispositifs dynamiques (i.e. évoluant dans le temps). En effet, l'état d'un tel dispositif à un instant donné (caractérisé par une ou plusieurs variables) dépend en partie de l'état de ce système (et donc de la valeurs des variables étudiées) aux instants précédents. L'étude d'une telle suite permet alors de connaître l'état du système étudié au bout d'un grand nombre de modifications.

Si l'on souhaite étudier une suite récurrente (i.e. trouver ses variations, sa limite éventuelle,...) l'idéal est de trouver une forme explicite. Mais hormis pour certains types de suites récurrentes, il n'est en général pas possible d'en obtenir une forme explicite. On verra alors comment les outils que l'on vient de voir (et d'autres, basés sur l'étude de la fonction  $f$  donnant la relation de récurrence) permettent de construire des protocoles d'étude efficaces.

### 3.5.2 Les suites arithmétiques et géométriques

Les suites arithmétiques et géométriques sont les suites récurrentes les plus simples que l'on puisse imaginer. Elles font partie des suites récurrentes pour lesquelles on sait trouver une forme explicite. Ainsi,

- Une suite  $(a_n)$  est dite *arithmétique* si elle vérifie une relation de récurrence de la forme

$$a_{n+1} = a_n + r$$

où  $r$  est un réel fixé, appelé *la raison*

**Exercice :**

1. Montrer par récurrence qu'une suite arithmétique  $(a_n)$  de raison  $r \in \mathbb{R}$  et de premier terme  $a_0 \in \mathbb{R}$  vérifie :

$$\forall n \in \mathbb{N}, \quad a_n = a_0 + n.r$$

2. En déduire la nature et la limite de  $(a_n)$  en fonction de  $r$ .

- Une suite  $(g_n)$  est dit *géométrique* si elle vérifie une relation de la forme

$$g_{n+1} = q.g_n$$

où  $q$  est un réel fixé, encore appelé *la raison*.

**Exercice :**

1. Montrer par récurrence si  $(g_n)$  est une suite géométrique de raison  $q$  et de premier terme  $g_0$ , elle vérifie

$$\forall n \in \mathbb{N}, \quad g_n = g_0 \cdot q^n$$

2. En déduire la nature et la limite éventuelle de  $(g_n)$  en fonction de  $q$ .
3. Montrer par récurrence que

$$\forall n \in \mathbb{N}, \quad \sum_{k=0}^{+\infty} q^k = \frac{1 - q^{n+1}}{1 - q}$$

**Note** : les suites géométriques sont un exemple particulier de suites plus générale : les *suites récurrentes linéaires à coefficients constants* vérifiant une relation de la forme

$$u_{n+p} = \alpha_n u_n + \alpha_{n+1} u_{n+1} + \dots + \alpha_{n+p-1} u_{n+p-1} = \sum_{k=n}^{n+p-1} \alpha_k u_k$$

$p \in \mathbb{N}^*$  étant l'ordre de la suite  $(u_n)$  et les réels  $\alpha_i$  étant donnés.

On verra en TD qu'il existe des méthodes permettant d'obtenir une forme explicite pour ce type de suites, basées sur les racines du *polynôme caractéristique* d'une telle suite :

$$P(X) = X^p - \sum_{k=0}^{p-1} \alpha_{n+k} \cdot X^k$$

Dans le cas  $p = 2$ , on peut en particulier établir la classification suivante : Soient  $a$  et  $b$  deux réels fixés et  $(u_n)$  une suite récurrente vérifiant la relation

$$\forall n \in \mathbb{N}, \quad u_{n+2} = au_{n+1} + bu_n$$

On note  $P(X) = X^2 - aX - b$  et  $\Delta = a^2 + 4b$ .

- Si  $\Delta > 0$ , le polynôme  $P(X)$  admet **deux racines réelles distinctes**  $r_1$  et  $r_2$ . Il existe alors deux constantes  $\lambda, \mu \in \mathbb{R}$  telles que

$$\forall n \in \mathbb{N}, \quad u_n = \lambda r_1^n + \mu r_2^n$$

- Si  $\Delta = 0$ , le polynôme  $P(X)$  admet **une unique racine double**  $r_0$ . Il existe deux constantes  $\lambda, \mu \in \mathbb{R}$  telles que

$$\forall n \in \mathbb{N}, \quad u_n = (\lambda + \mu \cdot n) r_0^n$$

- Si  $\Delta < 0$ , le polynôme  $P(X)$  admet **deux racines complexes conjuguées**

$$r_1 = \rho e^{i\theta} \quad \text{et} \quad r_2 = \bar{r}_1$$

Il existe deux constantes  $\lambda, \mu \in \mathbb{R}$  telles que

$$\forall n \in \mathbb{N}, \quad u_n = \rho^n (\lambda \cdot \cos(\theta n) + \mu \cdot \sin(\theta n))$$

**Note** : en pratique, on peut alors obtenir la forme explicite de  $(u_n)$  en déterminant les constantes  $\lambda$  et  $\mu$  à l'aide des premiers termes  $u_0$  et  $u_1$  de la suite étudiée.

### 3.5.3 Suites récurrentes autonomes d'ordre 1

On appelle suite récurrente autonome d'ordre 1 tout suite vérifiant une relation de la forme

$$(u_n) : \begin{cases} u_0 \text{ donné} \\ u_{n+1} = f(u_n) \end{cases}$$

où  $f$  est une fonction connue

On verra qu'une étude poussée de la fonction  $f$  permet de déterminer le comportement d'une telle suite, notamment en fonction de la valeur du premier terme  $u_0$ .

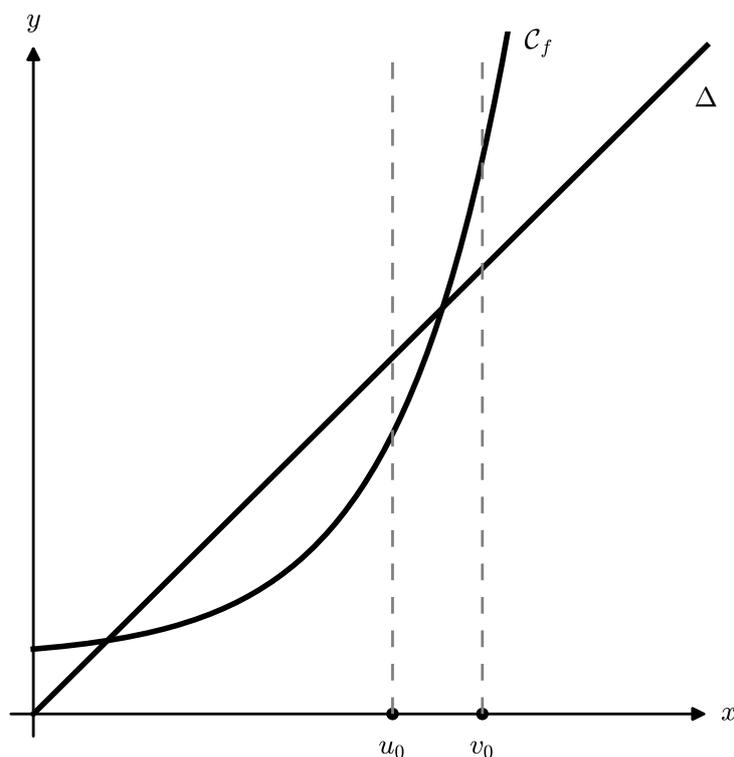
#### Représentation graphique

En traçant la courbe  $\mathcal{C}_f$  de  $f$  et la première bissectrice  $\Delta : y = x$  dans un même repère, on peut représenter l'évolution d'une suite récurrente autonome d'ordre 1. Ainsi, on considère deux suite  $(u_n)$  et  $(v_n)$  vérifiant les relations

$$u_{n+1} = f(u_n) \quad \text{et} \quad v_{n+1} = f(v_n)$$

et dont le premier terme est représenté dans le repère ci dessous.

1. À partir du point  $(u_0, 0)$  :
  - (a) Placer le point de coordonnées  $(u_0, f(u_0)) = (u_0, u_1)$ .
  - (b) Placer le point de coordonnées  $(u_1, u_1)$ .
  - (c) Placer le point de coordonnées  $(u_1, 0)$ .
  - (d) Placer le point de coordonnées  $(u_2, u_2)$ .
  - (e) Placer les points de coordonnées  $(u_n, u_n)$  pour  $n = 3$  et  $4$ .
2. Tracer le parcours de la suite  $(v_n)$  de premier point  $v_0$  et vérifiant  $v_{n+1} = f(v_n)$ .



### Points fixes

#### Définition (point fixe)

Soit  $f$  une fonction continue sur  $\mathbb{R}$ . On appelle *point fixe de  $f$*  toute solution de l'équation

$$f(x) = x$$

On verra que les points fixes d'une fonction  $f$  jouent un rôle fondamental dans l'étude des suites récurrentes vérifiant la relation  $u_{n+1} = f(u_n)$ . En particulier,

#### Proposition

Pour tout point fixe  $x^*$  de  $f$ , la suite récurrente  $(u_n)$  définie par  $\begin{cases} u_0 &= x^* \\ u_{n+1} &= f(u_n) \end{cases}$  est constante égale à  $x^*$

**Exercice** : à démontrer.

### Variations d'une suite récurrente autonome

L'idée principale de ce paragraphe est d'exposer comment les variations d'une fonction  $f$  influencent les variations des suites récurrentes vérifiant la relation  $u_{n+1} = f(u_n)$ . Ainsi,

**Proposition**

Soit  $f$  une fonction (*strictement*) *croissante*.

Toute suite récurrente  $(u_n)$  vérifiant la relation  $u_{n+1} = f(u_n)$  est (*strictement*) *monotone*.

De plus, le sens d'une variation d'une telle suite dépend alors du point de départ  $u_0$ . Précisément,

- Si  $f(u_0) \geq u_0$ , alors  $(u_n)$  est croissante.
- Si  $f(u_0) \leq u_0$ , alors  $(u_n)$  est décroissante.

**Preuve** : on montre par récurrence que la différence  $u_{n+1} - u_n = f(u_n) - u_n$  est de signe constante. Ainsi, supposons que  $u_1 \geq u_0$  et posons

$$\mathcal{P}(n) : u_{n+1} \geq u_n$$

La propriété étant vraie pour  $n = 0$  par définition, supposons qu'il existe un entier  $n \in \mathbb{N}$  tel que  $\mathcal{P}(n)$  est vraie :  $u_{n+1} \geq u_n$ .

Puisque  $f$  est croissante, elle conserve l'ordre. Donc

$$f(u_{n+1}) \geq f(u_n)$$

donc

$$u_{n+2} \geq u_{n+1}$$

La propriété est donc vraie au rang  $n + 1$ . Elle est donc héréditaire et par récurrence, elle est vraie pour tout  $n \in \mathbb{N}$ .

La démonstration du cas  $u_1 \leq u_0$  est laissée en exercice.

**Exemple** : déterminer le sens de variation des suites suivantes :

$$(u_n) : \begin{cases} u_0 = \frac{1}{2} \\ u_{n+1} = u_n^3 \end{cases} \quad ; \quad (v_n) : \begin{cases} v_0 = 2 \\ v_{n+1} = v_n^3 \end{cases}$$

**Note** : si  $f$  est une fonction décroissante, on verra que l'on peut également établir des résultats généraux, principalement basé sur la propriété suivante : si  $f$  est décroissante, alors  $f \circ f$  est croissante.

Nous supposons donc désormais (et sauf mention contraire) que la fonction  $f$  est strictement croissante. Toute suite récurrente qui lui sera associée sera donc monotone.

**Intervalles stables****Définition (intervalles stables)**

Soit  $f$  une fonction continue sur  $\mathbb{R}$  et  $x_1^*, \dots, x_r^*$  ses points fixes rangés dans l'ordre croissant :

$$x_1^* < x_2^* < \dots < x_r^*$$

On appelle *intervalles stables de  $f$*  les intervalles de la forme

$$I_0 = ]-\infty, x_1^*[, \quad I_1 = ]x_1^*, x_2^*[, \quad I_2 = ]x_2^*, x_3^*[, \quad \dots, \quad I_r = ]x_r^*, +\infty[$$

Ces intervalles ont deux propriétés principales, qui trouvent des applications directes dans l'étude et la classification des suites récurrentes vérifiant la relation  $u_{n+1} = f(u_n)$  :

**Proposition**

Soit  $f$  une fonction continue, strictement croissante et  $I_0, I_1, \dots, I_r$  ses intervalles stables.

- Chaque intervalle  $I_k$  est *stable* par  $f$ , i.e.

$$\forall x \in \mathbb{R}, \quad x \in ]x_k^*, x_{k+1}^*[ \Rightarrow f(x) \in ]x_k^*, x_{k+1}^*[$$

- Sur chaque intervalle stable  $I_k$ , la différence  $f(x) - x$  est de signe constant.

**Preuve :**

- (a) Montrer que si  $x^*$  est un point fixe de  $f$ , alors pour tout  $x \in \mathbb{R}$ ,

$$x < x^* \Rightarrow f(x) < x^*$$

- (b) Conclure.

- À faire.

Du point de vue des suites récurrentes associées à  $f$ , cela se traduit par la proposition suivante :

**Proposition**

Soient  $f$  une fonction croissante,  $x_1^* < \dots < x_r^*$  ses points fixes et  $(u_n)$  une suite numérique vérifiant la relation  $u_{n+1} = f(u_n)$ .

1. Si  $u_0$  n'est pas un point fixe de  $f$ , alors l'ensemble des termes de la suite  $(u_n)$  appartient à l'intervalle stable  $I_k$  contenant  $u_0$ .
2. Toutes les suites vérifiant la relation  $u_{n+1} = f(u_n)$  et dont les points de départ appartiennent au même intervalle stable  $I_k$  sont monotones et varient dans le même sens.

**Preuve :**

1. Immédiat par récurrence.
2. Le sens de variation de  $(u_n)$  dépend du signe de la différence  $u_1 - u_0 = f(u_0) - u_0$  et donc du signe de la différence  $f(x) - x$  sur l'intervalle stable  $I_k$  contenant  $u_0$ .

Ainsi, si  $u_0$  appartient en particulier à l'un des intervalles  $I_1, \dots, I_{r-1}$ , la suite  $(u_n)$  est bornée. Étant également monotone, elle converge.

Les résultats énoncés jusque là ne permettent pas, même en cas de convergence, de déterminer la valeur de la limite éventuelle. Cependant, une fois établie l'existence d'un réel  $\ell \in \mathbb{R}$  tel que  $\lim_{n \rightarrow +\infty} u_n = \ell$ , en passant alors à la limite  $n \rightarrow +\infty$  dans la relation

$$u_{n+1} = f(u_n),$$

on a, si  $f$  est continue

$$f(\ell) = \ell.$$

Autrement dit, si  $(u_n)$  est convergente, alors sa limite  $\ell$  est nécessairement un point fixe de  $f$  et, par encadrement, cela ne peut être que l'un des deux points fixes  $x_k^*$  ou  $x_{k+1}^*$  vérifiant  $x_k^* < u_0 < x_{k+1}^*$ . Par ailleurs, le sens de variation de la suite  $(u_n)$  (et donc le signe de  $f(x) - x$  sur l'intervalle  $I_k$ ) permet de déterminer vers lequel de ces deux points fixes la suite  $(u_n)$  converge.

**Exemple :** soit  $f : x \mapsto x^3$ .

1. Déterminer les variations et les points fixes de  $f$ .
2. Soient

$$(u_n) : \begin{cases} u_0 = \frac{1}{2} \\ u_{n+1} = u_n^3 \end{cases} \quad \text{et} \quad (v_n) : \begin{cases} v_0 = 2 \\ v_{n+1} = v_n^3 \end{cases}$$

- (a) Montrer que  $(u_n)$  est minorée. En déduire qu'elle converge et donner sa limite.
- (b) Montrer que  $(v_n)$  ne peut être convergente et en déduire sa limite.

**Résumé** : à partir des observations faites ci-dessus, on peut mettre en place à protocole d'étude pour toutes les suites récurrentes vérifiant

$$(u_n) : \begin{cases} u_0 \text{ donné} \\ u_{n+1} = f(u_n) \end{cases}$$

1. *Étude de  $f$*

- (a) Déterminer les points fixes et les intervalles stables de  $f$ .
- (b) Dresser le tableau de signes de la quantité  $f(x) - x$ .

2. *Classification*

(a) *Sens de variation*

- Toute suite  $(u_n)$  dont le point de départ  $u_0$  est un point fixe de  $f$  est constante.
- Toute suite  $(u_n)$  dont le point de départ  $u_0$  est dans un intervalle où  $f(x) - x \geq 0$  est croissante.
- Toute suite  $(u_n)$  dont le point de départ  $u_0$  est dans un intervalle où  $f(x) - x \leq 0$  est décroissante.

(b) *Convergence*

- Si la suite (monotone)  $(u_n)$  se rapproche d'un point fixe de  $f$ , alors elle converge vers ce point fixe.
- Si  $(u_n)$  s'éloigne de tous les points fixes de  $f$ , alors elle diverge vers  $\pm\infty$ .

**Exemple** : déterminer le comportement de toutes les suites récurrentes vérifiant la relation  $u_{n+1} = u_n^3$  en fonction du point de départ  $u_0$ .

### Classification des points fixes

La théorie que l'on vient d'exposer montre que si une suite récurrente autonome d'ordre 1 converge, alors elle converge nécessairement vers un point fixe de la fonction  $f$  donnant sa relation de récurrence. On a vu également que certains points fixes "attirent" les suites récurrente dont le point de départ est proche, alors que d'autres les "repoussent". On distingue ainsi, parmi les points fixes de  $f$ . Les premiers sont dits *stables*, alors que les seconds seront dits *instables*.

On peut alors relier la nature (stable ou instable) d'un point fixe  $x^*$  de  $f$  à la valeur  $f'(x^*)$ . Précisément,

**Proposition** Soit  $f$  une fonction continue et  $x^*$  un point fixe de  $f$ .

— Si  $0 < f'(x^*) < 1$ , alors  $x^*$  est stable.

Dans ce cas, toute suite récurrente vérifiant la relation  $u_{n+1} = f(u_n)$  dont le point de départ appartient à l'un des intervalles stables de  $f$  bornés par  $x^*$  converge vers  $x^*$ .

— Si  $f'(x^*) > 1$ , alors  $x^*$  est instable.

Dans ce cas, toute suite vérifiant la relation  $u_{n+1} = f(u_n)$  dont le point de départ appartient à l'un des intervalles stables bornés par  $x^*$  s'éloigne de  $x^*$ .

**Notes :** si  $f'(x^*) = 1$ , on ne peut pas conclure, i.e. tous les cas de figure sont possibles. Il est notamment possible de voir apparaître dans ce cas des points fixes *semi-stables* à gauche ou à droite. Illustrer géométriquement les différents cas si  $f'(x^*) = 1$ . On précisera comment distinguer ces cas d'un point de vue analytique.



### Cas des fonctions décroissantes

Comme cela a été évoqué ci-dessus, l'étude exhaustive des suites associées à une fonction croissante permet également de déterminer le comportement de toute suite associée à une fonction décroissante, en se basant sur le fait que si  $f$  est décroissante, alors  $f \circ f$  est, elle, croissante. Ainsi, soit  $f$  un fonction décroissante et  $(u_n)$  une suite vérifiant la relation  $u_{n+1} = f(u_n)$ . En étudiant les points fixes de la fonction  $f \circ f$  et le signe de  $f \circ f(x) - x$ , on peut déterminer le comportement des deux suites  $(v_n)$  et  $(w_n)$  définies par

$$(v_n) : \begin{cases} v_0 & = & u_0 \\ v_{n+1} & = & f \circ f(v_n) \end{cases} \quad \text{et} \quad (w_n) : \begin{cases} w_0 & = & u_1 \\ w_{n+1} & = & f \circ f(w_n) \end{cases}$$

qui sont deux les deux suites extraites de  $(u_n)$  définies par

$$\forall n \in \mathbb{N}, \quad v_n = u_{2n} \quad \text{et} \quad w_n = u_{2n+1}$$

(Exercice : à faire).

Puisque ces deux suites recouvrent à elles deux l'ensemble des termes de  $(u_n)$ , leur étude permet la nature de  $(u_n)$  et sa limite éventuelle.

En poussant encore l'étude des ces suites, on peut montrer le résultat suivant :

**Proposition**

Soit  $(u_n)$  une suite récurrente d'ordre 1 associée à une fonction  $f$  décroissante.

Les suites extraites  $(u_{2n})$  et  $(u_{2n+1})$  sont **monotones et de sens de variation différents**.

D'autre part :

- Si la distance  $|u_{2n} - u_{2n+1}|$  diminue, **alors ces deux suites sont adjacentes**.  
Dans ce cas, la suite  $(u_n)$  **converge vers l'unique point fixe de  $f$  contenu entre  $u_{2n}$  et  $u_{2n+1}$** .
- Si la distance  $|u_{2n} - u_{2n+1}|$  augmente, **alors  $(u_n)$  diverge**.

Enfin, il est possible d'étendre la définition de points fixes stables et instables au cas des fonctions décroissantes. Précisément, quelque soit le sens de variation de la fonction  $f$ , un point fixe  $x^*$  de  $f$  sera dit

- stable si  $|f'(x^*)| < 1$ ,
- instable si  $|f'(x^*)| > 1$ .

**Exemples :**

1. Déterminer le comportement de toute suite  $(u_n)$  associée à la fonction

$$f : x \mapsto 1 + \frac{1}{x}$$

2. Déterminer le comportement de toute suite  $(v_n)$  associée à la fonction

$$g : x \mapsto \frac{1}{x^2}$$

### 3.6 Vitesse de convergence

Les suites numériques et plus précisément les suites récurrentes autonomes d'ordre 1 trouvent de nombreuses applications dans l'approximation numérique. Elles permettent en effet de construire des suites qui convergent vers les quantités que l'on cherche à approcher. La théorie assure alors qu'en allant suffisamment loin dans la suite utilisée, il est possible d'obtenir une valeur approchée de la quantité cherchée à la précision souhaitée.

En pratique, on souhaite également pouvoir estimer le nombre de termes à calculer pour atteindre une précision souhaitée. D'un point de vue théorique, cette information est directement liée à la notion de *vitesse de convergence*.

### 3.6.1 Échelle de vitesses

Soit  $(u_n)$  une suite numérique qui converge et soit  $\ell$  sa limite.

1. On dit que la suite  $(u_n)$  converge à *vitesse polynomiale* vers  $\ell$  si

$$\text{il existe } A \text{ et } \alpha > 0 \text{ tels que } |u_n - \ell| \geq \frac{A}{n^\alpha}.$$

2. On dit que la suite  $(u_n)$  converge *géométriquement* vers  $\ell$  si

$$\text{il existe } A > 0 \text{ et } k \in ]0, 1[ \text{ tels que } |u_n - \ell| \leq A.k^n$$

3. On dit que la suite  $(u_n)$  converge avec une *convergence quadratique* vers  $\ell$  si

$$\text{il existe } A > 0 \text{ et } k \in ]0, 1[ \text{ tels que } |u_n - \ell| \leq A.k^{2^n}$$

4. Plus généralement, on dit que la suite  $(u_n)$  converge avec une *convergence d'ordre*  $r > 1$  vers  $\ell$  si

$$\text{il existe } A > 0 \text{ et } k \in ]0, 1[ \text{ tels que } |u_n - \ell| \leq A.k^{r^n}$$

D'un point de vue théorique, on peut alors placer une suite convergente donnée  $(u_n)$  sur cette échelle en étudiant le quotient

$$q_n = \frac{|u_{n+1} - \ell|}{|u_n - \ell|}$$

et plus précisément la limite  $k = \lim q_n$ .

On peut alors définir une nouvelle échelle de vitesse, globalement équivalente à la première :

1. Si  $k = 1$ , alors la convergence de  $(u_n)$  est *lente* (i.e. au mieux polynomiale).

2. Si  $0 < k < 1$ , alors la convergence est *géométrique*, de raison  $k$ .

3. Si  $k = 0$ , alors la convergence est *rapide* (i.e. d'ordre  $r > 1$ ).

Dans ce cas, le point fixe  $x^*$  est dit *super-attractif*.

**Note :** que dire de la suite  $(u_n)$  si  $k > 1$  ?

### 3.6.2 Application aux suites autonomes d'ordre 1

Dans le cas particulier où  $(u_n)$  est une suite récurrente vérifiant une relation de la forme  $u_{n+1} = f(u_n)$ , convergent vers un point fixe  $x^*$  de  $f$ , on peut montrer que la vitesse de convergence de  $(u_n)$  vers  $x^*$  dépend encore de la dérivée de  $f$  en son point fixe  $x^*$ . En effet,

$$\frac{|u_{n+1} - x^*|}{|u_n - x^*|} = \left| \frac{f(u_n) - f(x^*)}{u_n - x^*} \right| \longrightarrow |f'(x^*)|.$$

Autrement dit,

1. Si  $|f'(x^*)| = 1$ , la convergence vers  $x^*$  est lente.
2. Si  $|f'(x^*)| = k \in ]0, 1[$ , alors la convergence est géométrique de raison  $k$ .
3. Si  $|f'(x^*)| = 0$ , la convergence est au moins quadratique. Dans ce cas, le point fixe  $x^*$  est dit *super attractif*.

## 3.7 Résolution approchée d'équations algébriques

Une première application des suites numériques est le calcul approché de certaines valeurs réelles, via notamment la recherche de solutions approchées d'équations de la forme

$$\varphi(x) = 0$$

Il existe différentes méthodes permettant de calculer des valeurs approchées de solutions aux équations de ce type, chacune ayant ses points forts et ses points faibles, notamment concernant les conditions de convergence et la vitesse de convergence.

### 3.7.1 Dichotomie

La dichotomie est l'une des premières méthodes d'approximation pour les équations algébriques, basée principalement sur le théorème des valeurs intermédiaires. Précisément, si  $\varphi$  est une fonction continue et si l'on connaît deux nombres  $a < b$  tels que  $f(a)$  et  $f(b)$  soient de signes opposés, alors on est assuré de l'existence d'au moins un réel  $c \in [a, b]$  tel que  $f(c) = 0$ . Par ailleurs, en étudiant le signe de  $f$  au milieu  $\frac{a+b}{2}$  de l'intervalle  $[a, b]$ , on peut déterminer à quel demi-intervalle  $[a, \frac{a+b}{2}]$  ou  $[\frac{a+b}{2}, b]$  appartient cette solution.

En itérant le processus, on peut alors localiser la racine  $c$  cherchée dans un intervalle de longueur  $\varepsilon$  aussi petite que souhaitée. Tout élément de cet intervalle est une valeur approchée de  $c$  à  $\varepsilon$  près.

Formellement, la méthode de dichotomie consiste à construire deux suites  $(a_n)$  et  $(b_n)$  adjacentes, tendant toutes deux vers  $c$ . La construction de ces suites peut être confié à un ordinateur via l'algorithme ci dessous :

<p>Entrées : <math>a, b, \varepsilon</math></p> <hr style="border: 0; border-top: 1px solid black; margin: 5px 0;"/> <p><math>a_0 = a</math> <math>b_0 = b</math></p> <p>Tant que <math>b_n - a_n &gt; \varepsilon</math> :</p> <p style="padding-left: 40px;">Si <math>f(a_n) \cdot f\left(\frac{a_n+b_n}{2}\right) &gt; 0</math> :</p> <p style="padding-left: 80px;"><math>a_{n+1} = \frac{a_n+b_n}{2}</math> <math>b_{n+1} = b_n</math></p> <p style="padding-left: 40px;">Sinon :</p> <p style="padding-left: 80px;"><math>a_{n+1} = a_n</math> <math>b_{n+1} = \frac{a_n+b_n}{2}</math></p> <p>Renvoyer <math>a_{n+1}</math></p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Le principal point faible de la méthode de recherche par dichotomie est sa vitesse de convergence : on peut en effet montrer que

$$\forall n \in \mathbb{N}, \quad |a_n - c| \leq \frac{b-a}{2^n}$$

La vitesse de convergence des suites  $(a_n)$  et  $(b_n)$  vers  $c$  est donc géométrique de raison  $\frac{1}{2}$ .

### 3.7.2 Méthode de Newton-Raphson

#### Principes généraux

L'une des principales applications de la théorie ci-dessus est connue sous le nom de *Méthode de Newton-Raphson*, permettant de résoudre de façon approchée une équation algébrique de la forme

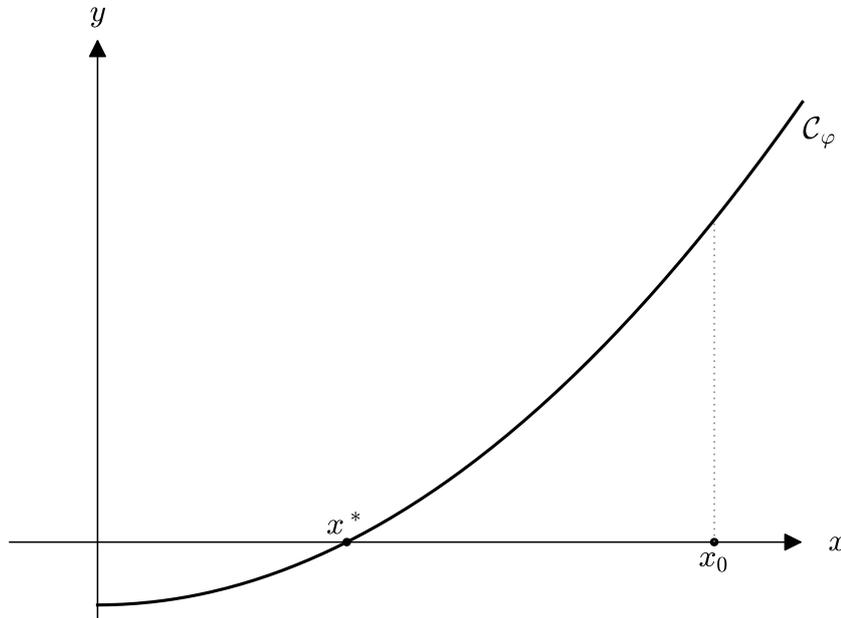
$$\varphi(x) = 0$$

Historiquement, l'approche de Newton est basée sur une observation géométrique.

Sur le dessin ci-dessous,

1. Tracer la tangente à la courbe  $\mathcal{C}_\varphi$  au point d'abscisse  $x_0$ .

2. Déterminer l'intersection de cette tangente avec l'axe des abscisses. On note  $x_1$  l'abscisse de ce point d'intersection.
3. Itérer le processus pour construire le point suivant  $(x_2, 0)$ .



On constate sur cet exemple que la suite  $(x_n)$  ainsi construite converge vers la solution  $x^*$  de l'équation  $\varphi(x) = 0$ .

En traduisant cette construction d'un point de vue analytique, on peut alors traduire l'idée de Newton sous la forme d'une suite récurrente autonome. La théorie des suites autonomes permet alors de valider l'idée de Newton, d'évaluer la qualité de l'approximation et de déterminer les conditions d'application de cette méthode d'approximation :

1. Déterminer l'équation de la droite  $T_0$ , tangente à  $\mathcal{C}_\varphi$  au point d'abscisse  $x_0$ .

$$T_{x_0} : y = \varphi'(x_0)(x - x_0) + \varphi(x_0)$$

2. En déduire l'abscisse  $x_1$  en fonction de  $\varphi(x_0)$ ,  $\varphi'(x_0)$  et  $x_0$ .

$$\varphi'(x_0)(x_1 - x_0) + \varphi(x_0) = 0 \Leftrightarrow x_1 = x_0 - \frac{\varphi(x_0)}{\varphi'(x_0)}$$

3. Donner la fonction  $f$  telle que la suite  $(x_n)$  construite en itérant le processus vérifie la relation  $x_{n+1} = f(x_n)$ .

$$f(x) = x - \frac{\varphi(x)}{\varphi'(x)}$$

4. Vérifier que la solution  $x^*$  de l'équation  $\varphi(x) = 0$  est un point fixe de  $f$  et qu'il est super attractif.

Puisque  $\varphi(x^*) = 0$ , on a

$$f(x^*) = x^* - \frac{\varphi(x^*)}{\varphi'(x^*)} = x^*$$

D'autre part,

$$f'(x) = \frac{\varphi(x)\varphi''(x)}{\varphi'(x)^2}$$

donc

$$f'(x^*) = \frac{\varphi(x^*)\varphi''(x^*)}{\varphi'(x^*)^2} = 0$$

**Note** : de façon plus précise, on peut montrer que la vitesse de convergence de la méthode de Newton dépend de la fonction  $\varphi$  étudiée, mais qu'elle est au moins quadratique :

$$\exists A > 0, k \in ]0, 1[ / \forall n \in \mathbb{N}, |x_n - x^*| \leq A.k^{2^n}$$

En pratique, cela signifie notamment qu'à chaque itération, le nombre de décimales communes entre  $x_n$  et  $x^*$  double.

### Conditions de convergence

Lorsque la suite de Newton converge, elle converge rapidement (au moins à vitesse quadratique). Cependant, certaines conditions d'application sont nécessaires pour assurer cette convergence.

Par construction, on voit en particulier que la dérivée  $\varphi'$  de la fonction  $\varphi$  ne doit pas s'annuler aux points calculés.

De façon générale, on peut établir en détails les conditions de convergence, notamment à l'aide du développement de Taylor de la fonction  $\varphi$  que l'on cherche à annuler. En pratique, on peut résumer ces conditions à deux points :

La méthode de Newton converge vers une racine  $x^*$  de la fonction  $\varphi$  à vitesse quadratique si

1.  $\varphi'(x^*) \neq 0$ ,
2. le point de départ  $x_0$  de la suite de Newton n'est pas trop loin de la solution  $x^*$  cherchée.

Afin de respecter le second point, on verra que la méthode de Newton est, en pratique, couplée à d'autres méthodes d'approximation (en général la méthode de dichotomie) moins efficaces, mais à l'aide desquelles on détermine une première approximation grossière de  $x^*$ , cette première approximation servant alors de point de départ pour la méthode de Newton.

### Méthode de la sécante

On vient de voir que la méthode de Newton nécessite le calcul de la dérivée  $\varphi'$  de  $\varphi$ . En pratique, cette dérivée peut être compliquée à évaluer. On peut alors remplacer cette dérivée par l'approximation suivante :

$$\varphi'(x_n) \approx \frac{\varphi(x_n) - \varphi(x_{n-1})}{x_n - x_{n-1}}$$

On obtient alors la méthode dite *de la sécante*. Bien que cette méthode produise une suite récurrente d'ordre 2, on peut là encore montrer que, si le point de départ  $x_0$  est assez proche de la solution  $x^*$  cherchée, la suite  $(x_n)$  vérifiant la relation

$$x_{n+1} = x_n - \frac{\varphi(x_n)(x_n - x_{n-1})}{\varphi(x_n) - \varphi(x_{n-1})}$$

converge vers  $x^*$ .

Par ailleurs, l'approximation de la dérivées  $\varphi'(x_n)$  se traduit en pratique par une perte en qualité. Précisément, on peut montrer que la méthode de la sécante est d'ordre  $r = \frac{1+\sqrt{5}}{2} \approx 1,61$ . Autrement dit, si  $x_0$  est assez proche de  $x^*$ ,

$$\exists A > 0, k \in ]0, 1[ \ / \ \forall n \in \mathbb{N}, \quad |x_n - x^*| \leq A.k \left( \frac{1+\sqrt{5}}{2} \right)^n$$

# Chapitre 4

## Matrices et systèmes linéaires

### Introduction

La notion de système linéaire est présente dans de nombreuses théories mathématiques ainsi que dans de nombreuses méthodes de modélisation.

L'objectif de ce chapitre est de présenter la théorie mise en place pour l'étude de ces systèmes, ainsi que les grands principes de résolution.

L'ensemble de ces résultats est principalement basé sur le caractère linéaire des équations étudiées et sa traduction en termes de calcul matriciel. Avant de présenter en détails les méthodes de résolutions, nous présenterons dans un premier temps les matrices et les outils de calcul matriciel. Nous verrons ensuite comment investir des nouvelles notions dans la mise en place d'outils permettant l'étude théorique et la résolution systématique des systèmes linéaires.

### 4.1 Matrices

#### 4.1.1 Définitions

Soient  $n$  et  $p$  deux entiers naturels non nuls. Une matrice à  $n$  lignes et  $p$  colonnes à coefficients réels est la donnée de  $n \times p$  nombres réels, appelés termes ou coefficients rangés dans un tableau rectangulaire à  $n$  lignes et  $p$  colonnes.

On note  $\mathcal{M}_{n,p}(\mathbb{R})$  l'ensemble des matrices à  $n$  lignes et  $p$  colonnes à coefficients réels. On peut également envisager des matrices contenant d'autres types de nombres ( $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{C}$ , etc).

— Si  $n = 1$ , on parle de *matrice ligne* ou *vecteur ligne* à  $p$  colonnes :

$$\mathcal{M}_{1,p}(\mathbb{R}) = \mathbb{R}^p$$

— Si  $p = 1$ , on parle de *matrice colonne* ou *vecteur colonne* à  $n$  lignes :

$$\mathcal{M}_{n,1}(\mathbb{R}) = \mathbb{R}^n$$

— Si  $n = p$ , on parle de *matrice carrée*.

L'ensemble des matrices carrées à  $n$  lignes et  $n$  colonnes est alors noté  $\mathcal{M}_n(\mathbb{R})$ .

### Notes :

— Les vecteurs à deux ou trois coordonnées permettent en particulier de représenter des vecteurs et des positions du plan ou de l'espace. Cette représentation permet de très nombreuses du calcul matriciel en géométrie (et permet également une généralisation des concepts géométriques aux dimensions supérieures).

Dans ce contexte, la notion de matrice permet en particulier de représenter des familles de vecteurs (en général à travers les *colonnes d'une matrice*).

— Si  $n = p = 1$ , on a simplement affaire à des nombres.

## 4.1.2 Indexation

Dans une matrice  $A \in \mathcal{M}_{n,p}(\mathbb{R})$ , le coefficient situé sur la  $i$ -ème ligne et la  $j$ -ème colonne ( $1 \leq i \leq n$ ,  $1 \leq j \leq p$ ) sera noté  $a_{ij}$  :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix} = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

**Exemple :** toute matrice de  $A \in \mathcal{M}_{24}(\mathbb{R})$  se note

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix}$$

## 4.1.3 Matrices particulières (carrées)

Soit  $n \in \mathbb{N}^*$  fixé. Dans l'ensemble  $\mathcal{M}_n(\mathbb{R})$ , on distingue certaines matrices et certains types de matrices particulières :

— La matrice *identité* (ou matrice *unité*)

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Les coefficients  $\delta_{ij}$  de la matrice identité vérifient donc

$$\forall i, j \in \{1, \dots, n\}, \quad \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

— Une matrice carrée est dite *triangulaires supérieure* (resp. *triangulaire inférieure*) si elle est de la forme

$$T_1 = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * \end{pmatrix} \quad \text{resp. } T_2 = \begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & * \end{pmatrix}$$

Ainsi,  $T = (t_{ij})$  est triangulaire supérieure (resp. triangulaire inférieure) si et seulement si

$$\forall i, j \in \{1, \dots, n\}, \quad i > j \Rightarrow t_{ij} = 0 \quad (\text{resp. } i < j \Rightarrow t_{ij} = 0)$$

— Une matrice carrée est dite *diagonale* si elle est de la forme

$$D = \begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & * \end{pmatrix}$$

Ainsi, une matrice  $D = (d_{ij})$  est diagonale si et seulement si

$$\forall i, j \in \{1, \dots, n\}, \quad i \neq j \Rightarrow d_{ij} = 0$$

**Note :** les matrices triangulaires et diagonales sont définies “en creux” :

- Une matrice carrée est triangulaire supérieurs (resp. inférieure) si tous les termes situés *sous la diagonale* (resp. *au dessus de la diagonale*) sont nuls.
- Une matrice carrée est diagonale si ses termes *non diagonaux* sont tous nuls.

### 4.1.4 Opérations matricielles

Outre la possibilité de stocker un ensemble de valeurs, il est également possible de définir des opérations sur l'ensemble des matrices. À l'aide de ces opérations, il devient alors possible d'étendre les notions de calculs et d'équations à l'ensemble des matrices.

#### Opérations linéaires

Étant données deux matrices de même taille  $A = (a_{ij})$  et  $B = (b_{ij})$  dans  $\mathcal{M}_{n,p}(\mathbb{R})$ , la matrice  $A + B$  est la matrice de  $\mathcal{M}_{n,p}(\mathbb{R})$  définie par

$$A + B = (a_{ij} + b_{ij})$$

(on dit que l'addition se fait *terme à terme*).

Étant intimement liée à l'addition des nombres réels, l'addition des matrices hérite des propriétés principale de cette addition des réels. Ainsi

— L'addition des matrices est associative :

$$\forall A, B, C \in \mathcal{M}_{n,p}(\mathbb{R}), \quad (A + B) + C = A + (B + C)$$

— L'addition des matrices est commutative :

$$\forall A, B \in \mathcal{M}_{n,p}(\mathbb{R}), \quad A + B = B + A$$

— La matrice nulle : en notant  $\mathbb{O}_{n,p}$  la matrice nulle de taille  $n.p$  (i.e. la matrice à  $n$  lignes et  $p$  colonnes dont tous les coefficients sont nuls), on a

$$\forall A \in \mathcal{M}_{n,p}(\mathbb{R}), \quad A + \mathbb{O}_{n,p} = A$$

— Matrice opposée : toute matrice  $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{R})$  admet une matrice opposée : la matrice  $-A = (-a_{ij})$ .

Il est également possible de définir une multiplication extérieure sur l'ensemble des matrices, i.e. une multiplication par un réel. Ainsi, pour toute matrice  $A \in \mathcal{M}_{n,p}(\mathbb{R})$  et pour tout  $\lambda \in \mathbb{R}$ , la matrice  $\lambda.A$  est la matrice de  $\mathcal{M}_{n,p}(\mathbb{R})$  définie par

$$\lambda.A = (\lambda.a_{ij})$$

Ce produit extérieur est alors distributif sur l'addition des matrices et compatible avec la multiplication des réels. Autrement dit,

$$\forall A, B \in \mathcal{M}_{n,p}(\mathbb{R}), \quad \forall \lambda \in \mathbb{R}, \quad \lambda.(A + B) = \lambda.A + \lambda.B$$

et

$$\forall A \in \mathcal{M}_{n,p}(\mathbb{R}), \quad \forall \lambda, \mu \in \mathbb{R}, \quad \lambda.(\mu.A) = (\lambda\mu).A = \mu.(\lambda.A)$$

### Multiplication des matrices

Il est également possible de définir une multiplication sur l'ensemble des matrices. Précisément, soient  $n, p, q \in \mathbb{N}^*$  et  $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{R})$  et  $B = (b_{ij}) \in \mathcal{M}_{p,q}(\mathbb{R})$ . Le produit  $A \times B$  est la matrice  $C = (c_{ij}) \in \mathcal{M}_{n,q}(\mathbb{R})$  définie par

$$\forall i \in \{1, \dots, n\}, \quad \forall j \in \{1, \dots, q\}, \quad c_{ij} = \sum_{k=1}^p a_{ik}b_{kj}$$

Notons en particulier que, contrairement aux deux opérations précédentes, le produit de matrices ne se fait pas terme à terme. D'autre part, il n'est pas toujours possible de multiplier entre elles deux matrices quelconques. Pour que le produit  $A \times B$  existe, il faut en particulier que le nombre de colonnes de  $A$  soit égal au nombre de lignes de  $B$ . Le produit  $A \times B$  a alors autant de lignes que  $A$  et autant de colonnes que  $B$ . On peut résumer le produit de matrice par le schéma suivant :

$$A = \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ a_{i1} & a_{i2} & \cdots & a_{ip} & \\ & & & & \end{pmatrix} \begin{pmatrix} & & & & \\ & & & & \\ & & & & \\ & & & & \\ \cdots & \cdots & \cdots & c_{ij} & \end{pmatrix} = A \times B$$

Ainsi, même si le produit  $A \times B$  existe, il n'est pas dit que le produit  $B \times A$  existe. Et même dans le cas où les deux produits  $AB$  et  $BA$  existent, les matrices obtenues ne sont pas toujours de la même taille. Seul cas particulier : le cas des matrices carrées. Précisément, si  $A$  et  $B$  sont deux matrices carrées à  $n$  lignes, les produits  $A \times B$  et  $B \times A$  existent toujours et produisent encore des matrices carrées à  $n$  lignes. Cependant, même dans ce cas, on a en général  $A \times B \neq B \times A$  (i.e. le produit de matrices n'est pas commutatif).

Hormis le défaut de commutativité, la multiplication des matrices possède tout de même quelques propriétés : soient  $A, B$  et  $C$  trois matrices telles que toutes les opérations envisagées soient possibles.

- *Associativité* :  $(A \times B) \times C = A \times (B \times C)$ .
- *Distributivité* :  $A \times (B + C) = A \times B + A \times C$ .
- Pour tout  $\lambda \in \mathbb{R}$ , on a  $(\lambda.A) \times B = A \times (\lambda.B) = \lambda.(A \times B)$ .
- *Matrice unité* :
  - Pour toute matrice  $A$  à  $n$  colonnes, on a  $A \times I_n = A$ .
  - Pour toute matrice  $B$  à  $p$  lignes, on a  $I_p \times B = B$ .

### Transposée

Étant donnée une matrice  $A \in \mathcal{M}_{n,p}(\mathbb{R})$ , on appelle *transposée* de  $A$  la matrice notée  ${}^tA$ , appartenant à  $\mathcal{M}_{p,n}(\mathbb{R})$ , dont les lignes sont les colonnes de  $A$  (et dont les colonnes sont alors les lignes de  $A$ ). Ainsi, si  $A = (a_{ij})$ , alors  ${}^tA = (a_{ji})$ .

#### Exemples :

- Si  $A = \begin{pmatrix} -1 & 1 & 2 \\ 3 & 5 & 7 \end{pmatrix}$ , alors  ${}^tA = \begin{pmatrix} -1 & 3 \\ 1 & 5 \\ 2 & 7 \end{pmatrix}$
- Si  $A = (1 \ 2 \ 3)$ , alors  ${}^tA = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$
- Si  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ , alors  ${}^tA = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$

Notons que dans le cas d'une matrice carrée, la transposée est obtenue par symétrie par rapport à sa diagonale.

Enfin, on appelle *matrice symétrique* toute matrice  $A$  vérifiant  ${}^tA = A$ . Ainsi, une matrice symétrique est nécessairement carrée et est symétrique par rapport à sa diagonale.

**Exemple** : la matrice  $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$  est une matrice symétrique  $3 \times 3$ .

On peut montrer que  ${}^t(A+B) = {}^tA + {}^tB$  (à faire) et que  ${}^t(A \times B) = {}^tB \times {}^tA \neq {}^tA \times {}^tB$  (plus dur).

**Trace**

La *trace* d'une matrice carrée  $A$  est le nombre réel obtenu en sommant ses termes diagonaux : si  $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ , alors

$$\operatorname{tr}(A) = \sum_{i=1}^n a_{ii}$$

On verra que la trace est une caractéristique importante de toute matrice carrée. On peut en particulier montrer que pour toutes matrices  $A$  et  $B$  carrées de même taille, on a

$$\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$$

On peut également montrer que, même si  $A \times B \neq B \times A$ , on a  $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ .

**4.1.5 Matrices inversibles****Définition (matrice inversible)**

Soit  $n \in \mathbb{N}^*$ . Une matrice  $A \in \mathcal{M}_n(\mathbb{R})$  est dite inversible si il existe une matrice  $B \in \mathcal{M}_n(\mathbb{R})$  telle que  $A \times B = I_n$ .

La matrice  $B$  ainsi définie est appelée *inverse de  $A$*  et est notée  $A^{-1}$ .

Par ailleurs, l'ensemble des matrices carrées inversibles à  $n$  lignes est noté  $\operatorname{GL}_n(\mathbb{R})$ .

**Propriétés**

Soit  $A \in \mathcal{M}_n(\mathbb{R})$ .

- S'il existe  $B \in \mathcal{M}_n(\mathbb{R})$  telle que  $A \times B = I_n$ , alors  $B \times A = I_n$ .
- Si  $A \in \operatorname{GL}_n(\mathbb{R})$ , alors  $A^{-1} \in \operatorname{GL}_n(\mathbb{R})$  et que  $(A^{-1})^{-1} = A$ .
- Si  $A$  et  $B$  sont deux matrices de  $\operatorname{GL}_n(\mathbb{R})$ , alors  $A \times B$  est encore dans  $\operatorname{GL}_n(\mathbb{R})$  et

$$(A \times B)^{-1} = B^{-1} \times A^{-1} \neq A^{-1} \times B^{-1}$$

Décider du caractère inversible d'une matrice donnée et calculer son inverse le cas échéant font partie des problèmes les plus complexes à traiter, même à l'aide d'un ordinateur.

En effet, même si la théorie donne des outils de décision ainsi que des méthodes théoriques pour le calcul d'inverse, aucune de ces méthodes n'est réellement efficace lorsque

la taille de la matrice étudiée augmente.

### 4.1.6 Déterminants

Le déterminant d'une matrice carrée est un nombre, calculé à partir des coefficients de la matrice en question est permettant de mettre en évidence un certain nombre de propriétés de la matrice étudiée. Pour une matrice  $A \in \mathcal{M}_n(\mathbb{R})$ , on note  $\det(A)$  ou  $|A|$  son déterminant. Sous forme étendue, si  $A = (a_{ij})$ , on note

$$\det(A) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

L'une des principales application du déterminant est lié à l'inversibilité d'une matrice donnée. Précisément, on peut montrer que pour toute matrice  $A \in \mathcal{M}_n(\mathbb{R})$ , on a

$$A \in \text{GL}_n(\mathbb{R}) \text{ (i.e. } A \text{ est inversible)} \iff \det(A) \neq 0$$

Enfin, la notion de déterminant trouve également des applications en géométrie via la notion de déterminant d'une famille de vecteurs.

Nous verrons également dans la section suivante que la notion de déterminant permet de mettre en place des propriétés théoriques importantes dans le domaine des systèmes d'équations linéaires.

#### Calcul d'un déterminant

Étant donnée une matrice  $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ , on a

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i),i}$$

où  $S_n$  représente l'ensemble des permutations possibles de l'ensemble  $\{1, \dots, n\}$ . Cette représentation permet de mettre en évidence certaines propriétés du déterminant, mais ne permet pas en pratique de calculer celui ci, notamment du fait du nombre important ( $n!$ ) de termes à calculer.

En pratique, le calcul d'un déterminant se fait par "récurrence".

Précisément, pour  $n = 2$ , si  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ , alors

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

On constate ici que

$$\det(A) = 0 \iff \frac{a}{c} = \frac{b}{d} \iff \frac{a}{b} = \frac{c}{d}$$

Autrement dit, les matrices  $2 \times 2$  inversibles sont celles dont les lignes (et donc les colonnes) ne sont pas proportionnelles.

Pour  $n = 3$ , la formule générale produit la formule de Sarrus :

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = (aei + dhc + gbf) - (gec + ahf + dbi)$$

Cependant, on verra que, même en dimension 3, cette méthode possède un certain nombre d'inconvénients majeurs qui la rendent peu exploitée en pratique, au profit d'une méthode plus générale, basée sur le *développement par rapport à une ligne ou une colonne*.

Ainsi, soit  $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$ . On commence par placer dans  $A$  un damier de signes :

$$\det(A) = \begin{vmatrix} +a & -b & +c \\ -d & +e & -f \\ +g & -h & +i \end{vmatrix}$$

On peut alors choisir de développer par rapport à la première colonne :

$$\det(A) = \begin{vmatrix} +a & b & c \\ -d & e & f \\ +g & h & i \end{vmatrix}$$

via la formule suivante :

$$\begin{aligned} \det(A) &= + \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} - \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \\ &= +a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix} \end{aligned}$$

En développant les déterminants  $2 \times 2$ , on retrouve les termes de la formule de Sarrus.

On peut également choisir de développer par rapport à la deuxième ligne :

$$\det(A) = \begin{vmatrix} a & b & c \\ -d & +e & -f \\ g & h & i \end{vmatrix}$$

$$\det(A) = - \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} - \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

$$= -d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + e \begin{vmatrix} a & c \\ g & i \end{vmatrix} - f \begin{vmatrix} a & b \\ g & h \end{vmatrix}$$

Là encore, en développant les déterminants  $2 \times 2$  obtenus, on retrouve les termes de la formule de Sarrus.

De façon générale, il est possible de développer n'importe quel déterminant  $n \times n$  par rapport à l'une de ses lignes ou l'une de ses colonnes. Théoriquement, par itérations successives, on peut alors réduire n'importe quel déterminant à une somme de déterminants  $2 \times 2$ . On verra qu'en pratique, le procédé peut vite être fastidieux et qu'il existe des manipulations possible "préparant" le développement d'un déterminant.

### Notes

- Le développement par rapport à une ligne ou une colonne permet de mettre en évidence certaines propriétés portant sur des matrices particulières. Ainsi,
  - Si  $A$  contient une ligne ou une colonne de 0, alors  $\det(A) = 0$  (et  $A$  n'est alors pas inversible).
  - Le déterminant d'une matrice triangulaire est donné par le produit de ses termes diagonaux (à démontrer). C'est en particulier le cas des matrices diagonales.
  - Une matrice triangulaire  $T$  est inversible si et seulement si aucun de ses termes diagonaux n'est nul.

— Les sous-déterminants  $(\text{co}A)_{ij} = (-1)^{i+j}$

$$\begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

intervenant dans le développement d'un déterminant sont appelés *co-facteurs*. Ils constituent les coefficients de la *co-matrice*  $\text{co}A$  de  $A$  qui est au cœur de la formule explicite de l'inverse d'une matrice  $A \in \text{GL}_n(\mathbb{R})$  :

$$A^{-1} = \frac{1}{\det(A)} {}^t(\text{co}A)$$

Si  $n = 2$ , cette formule explicite donne

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \iff A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

En dimensions supérieures, cette formule permet d'établir certains résultats théoriques sur les matrices inversibles, mais n'est que très peu exploitée dans le calcul effectif d'inverses car elle nécessite le calcul de  $n^2$  déterminants  $(n-1) \times (n-1)$  et d'un déterminant  $n \times n$ .

### Opérations élémentaires sur les lignes et les colonnes

Soit  $A$  une matrice carrée à  $n$  lignes. On appelle opérations élémentaires sur les lignes  $L_1, \dots, L_n$  (resp. les colonnes  $C_1, \dots, C_n$ ) de  $\det(A)$  les trois opérations suivantes :

— *Combinaison de lignes (resp. colonne) :*

$$L_i \leftarrow L_i + \alpha.L_j \quad (\text{resp. } C_i \leftarrow C_i + \alpha.C_j) \quad \alpha \in \mathbb{R}, \quad i \neq j$$

— *Permutation de lignes (resp. colonne) :*

$$L_i \longleftrightarrow L_j \quad (\text{resp. } C_i \longleftrightarrow C_j) \quad i \neq j$$

— *Multiplication d'une ligne (resp. colonne) :*

$$L_i \leftarrow \lambda.L_i \quad (\text{resp. } C_i \leftarrow \lambda.C_i) \quad \lambda \in \mathbb{R}$$

On peut montrer que ces opérations élémentaires ne modifient que peu, voire pas du tout le déterminant auquel elles sont appliquées. Précisément,

- Une combinaison de lignes ou de colonnes **ne modifie pas le déterminant auquel on l'applique.**
- Une permutation de lignes ou de colonnes **change le signe du déterminant auquel on l'applique.**
- Une multiplication de ligne ou de colonne par un coefficient  $\lambda \in \mathbb{R}$  **multiple par  $\lambda$  le déterminant auquel on l'applique.**

### Réduction d'un déterminant

Lors du calcul explicite d'un déterminant  $\det(A)$ , ces opérations élémentaires permettent alors de transformer  $\det(A)$  avant d'appliquer un éventuel développement.

Ainsi, les combinaisons (et permutations) de lignes et colonnes permettent de placer des zéros dans  $\det(A)$ , qui simplifient le développement.

**Exemple** : Soit  $\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ -2 & 5 & -6 \end{vmatrix}$

1. À l'aide de la ligne  $L_1$ , annuler les coefficients 2 et  $-2$  de la colonne  $C_1$ .
2. Calculer  $\det(A)$  en développant par rapport à sa première colonne le déterminant obtenu à la question précédente.

**Note** : d'un point de vue théorique, l'invariance du déterminant par combinaisons de lignes ou de colonnes permet de montrer que si une matrice carrée  $A \in \mathcal{M}_n(\mathbb{R})$  contient deux lignes égales ou proportionnelles, alors son déterminant est nul (et la matrice  $A$  n'est pas inversible). De façon générale, si l'une des lignes (ou l'une des colonnes) de  $A$  peut s'exprimer sous la forme d'une combinaison linéaire des autres, alors son déterminant est nul et la matrice  $A$  n'est pas inversible.

Enfin, on peut montrer que le déterminant est une application *multiplicative*, i.e.

$$\forall A, B \in \mathcal{M}_n(\mathbb{R}), \quad \det(A \times B) = \det(A) \times \det(B)$$

Cette propriété fondamentale permet en particulier de démontrer certains résultats importants concernant les matrices carrées :

Soient  $A$  et  $B$  deux matrices carrées de même taille.

1.  $\det(A \times B) = \det(B \times A)$ .
2. Si  $A \in \text{GL}_n(\mathbb{R})$ , alors  $\det(A^{-1}) = (\det(A))^{-1}$ .
3.  $A$  et  $B$  sont inversibles si et seulement si  $A \times B$  est inversible.

### Factorisation d'un déterminant

La multiplication d'une ligne ou d'une colonne par un coefficient peut se traduire de la façon suivante :

$$\forall \lambda \in \mathbb{R}, \quad \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \lambda \cdot a_{i1} & \cdots & \lambda \cdot a_{ij} & \cdots & \lambda \cdot a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = \lambda \cdot \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & \lambda \cdot a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & \lambda \cdot a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \lambda \cdot a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

La mise en facteur d'un peut alors permettre de rendre plus efficace l'utilisation des combinaisons dans la réduction du déterminant à calculer.

## 4.2 Systèmes linéaires

Les équations linéaires sont des équations algébriques parmi les plus simples que l'on puisse construire. On va voir comment cette forme simple permet de faire de nombreux liens avec les matrices et le calcul matriciel. Ces liens permettent alors de mettre en place des processus de résolutions théoriques et pratiques et de produire des *algorithmes de résolution*.

### 4.2.1 Définitions

#### Vocabulaire général

Soit  $p \in \mathbb{N}^*$ . Une équation linéaire à  $p$  inconnues est une équation de la forme

$$(E) : a_1x_1 + a_2x_2 + \dots + a_px_p = b$$

liant les inconnues  $x_1, \dots, x_p$ .

Les réels  $a_i$  sont les coefficients (connus) de l'équation et  $b$  est le second membre. Si  $b$  est nul, l'équation est dite *homogène*.

Un système linéaire est la donnée d'un ensemble d'équations linéaires portant sur les inconnues  $x_1, \dots, x_p$  :

$$(S) : \begin{cases} a_{11}x_1 + \dots + a_{1p}x_p = b_1 & (E_1) \\ a_{21}x_1 + \dots + a_{2p}x_p = b_2 & (E_2) \\ \vdots & \vdots \\ a_{n1}x_1 + \dots + a_{np}x_p = b_n & (E_n) \end{cases} \iff (S) = \{(E_1), \dots, (E_n)\}$$

Ici,  $p$  est le nombre d'inconnues et  $n$  est le nombre d'équations.

Les nombres réels  $a_{ij}$  sont les coefficients du système, les  $x_i$  sont les inconnues et les  $b_i$  forment le second membre. S'ils sont tous nuls, le système est dit homogène.

Une *solution* du système  $(S)$  est un  $p$ -uplet  $(x_1, \dots, x_p)$  satisfaisant l'ensemble des équations. C'est un élément de l'ensemble  $\mathbb{R}^p$ .

Résoudre  $(S)$ , c'est déterminer l'ensemble des  $p$ -uplets solutions. L'ensemble  $\Sigma$  des solutions d'un système linéaire à  $p$  inconnues est donc un sous ensemble de  $\mathbb{R}^p$ .

Enfin, on dira que deux systèmes  $(S)$  et  $(S')$  sont *équivalents* s'ils ont le même ensemble de solutions. On note  $(S) \Leftrightarrow (S')$ .

## Représentation géométrique

En petites dimensions (dans les cas  $p = 2$  et  $p = 3$ , il est possible d'aborder les systèmes linéaires d'un point de vue géométrique. Ainsi,

- En deux variables, une équation linéaire est une équation de la forme

$$ax + by = c$$

Géométriquement, l'ensemble des solutions d'une équation de ce type peut être représenté par **une droite du plan**.

De ce point de vue, tout système linéaire à deux inconnues peut être associé à **une famille de droites du plan**.

L'ensemble des solutions d'un tel système peut alors être représenté par **l'intersection de toutes ces droites**.

- En trois variables, une équation linéaire est une équation de la forme

$$ax + by + cz = d$$

Géométriquement, l'ensemble des solutions d'une équation de ce type peut être représenté par **un plan de l'espace**.

De ce point de vue, tout système linéaire à trois inconnues peut être associé à **une famille de plans de l'espace**.

L'ensemble des solutions d'un tel système peut alors être représenté par **l'intersection de tous ces plans**.

**Note** : on verra dans la suite qu'en s'appuyant sur ces représentations géométriques, on pourra mettre en évidence certaines propriétés des systèmes linéaires qui restent valables en dimensions supérieures (i.e. pour  $p \geq 4$ ). Dans cette optique, on peut généraliser la notion de droites aux espaces  $\mathbb{R}^p$ . Précisément, on appelle *hyperplan de  $\mathbb{R}^p$*  tout sous ensemble de  $\mathbb{R}^p$  correspondant à l'ensemble des solutions d'une équation linéaire.

### Représentation matricielle

L'un des premiers liens que l'on peut faire entre systèmes linéaires et matrices est une représentation des premiers à l'aide des dernières.

Soit

$$(S) : \begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1 \\ a_{21}x_1 + \cdots + a_{2p}x_p = b_2 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n \end{cases}$$

un système linéaire de  $n$  équations à  $p$  inconnues. On appelle

— La matrice de  $(S)$  est la matrice  $A \in \mathcal{M}_{n,p}(\mathbb{R})$  contenant les coefficients de  $(S)$  :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$$

— Le vecteur inconnue  $X$  de  $(S)$  est le vecteur colonne  $X \in \mathcal{M}_{p,1}(\mathbb{R})$  contenant les inconnues du système  $(S)$  :

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

— Le second membre de  $(S)$  est le vecteur colonne  $B \in \mathcal{M}_{n,1}(\mathbb{R})$  contenant les seconds membres de toutes les équations de  $(S)$  :

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

**Note** : le système  $(S)$  est alors homogène si et seulement si  $B = \vec{0}$ . On verra que les systèmes linéaires homogènes ont une place particulière dans la théorie des systèmes linéaires. La première particularité de ces systèmes est la suivante : tout système linéaire homogène admet au moins le vecteur nul de  $\mathcal{M}_{n,1}(\mathbb{R})$  comme solution.

De façon générale, le système  $(S)$  s'écrit alors sous la forme de l'égalité vectorielle

$$A \times X = B$$

L'étude des matrices  $A$  et  $B$  du système permet de mettre en place des résultats théoriques ainsi que des méthodes de résolution des systèmes linéaires associés. Dans ce cadre, on définit également la *matrice étendue du système*  $(S)$  :

$$\tilde{A} = \left( \begin{array}{ccc|c} a_{11} & \cdots & a_{1p} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{np} & b_n \end{array} \right) = \left( \begin{array}{c|c} A & B \end{array} \right)$$

Contrairement aux matrices  $A$ ,  $B$  et  $X$ , cette matrice étendue n'a pas de réel sens théorique, mais elle contient l'ensemble des informations portées par le système  $(S)$  et permet de travailler sur le système  $(S)$  en remplaçant les manipulations des équations par des manipulations sur les matrices de  $\tilde{A}$ .

### Systèmes de Cramer

On appelle *système de Cramer* ou *système inversible* tout système linéaire  $AX = B$  dont la matrice  $A$  est inversible. Ce sont en particulier des systèmes admettant autant d'inconnues que d'équations ( $n = p$ ).

On peut montrer qu'un système linéaire est inversible si et seulement si il admet une unique solution. D'autre part, le calcul matriciel permet d'exprimer cette unique solution :

$$A \times X = B \iff A^{-1} \times (AX) = A^{-1} \times B \iff X = A^{-1} \times B$$

Cependant, hormis pour des tous petits systèmes ( $n = p = 2$ ), cette méthode n'est en pratique pas utilisée pour résoudre les systèmes inversibles car en termes de temps de calcul, le calcul de la matrice  $A^{-1}$  est plus long que la résolution "directe" du système.

## 4.2.2 Structure de l'ensemble des solutions

### Contraintes vs degrés de liberté

Les méthodes de résolution pratique des systèmes linéaires (qui font l'objet des derniers chapitres de ce cours) reposent sur une étude poussée des différents ensembles de solutions  $\Sigma$  possibles d'un système linéaire  $AX = B$  donné.

Ainsi, dans un premier temps, on peut distinguer trois types de systèmes linéaires :

- Les systèmes n'admettant aucune solution :

$$\Sigma = \emptyset$$

- Les systèmes admettant une unique solution (les systèmes inversibles) :

$$\Sigma = \{A^{-1} \times B\}$$

- Les systèmes admettant **une infinité de solutions**.

Or de façon générale, tout système linéaire peut être vu comme la donnée d'un ensemble de contraintes (les équations) portant sur un certain nombre de variables (les inconnues). Le type d'ensemble de solutions d'un système linéaire donné dépend alors de l'équilibre existant entre le nombre de contraintes et le nombre de variables.

De ce point de vue, chaque variable peut être vue comme un degré de liberté et pour un système linéaire donnée  $(S)$ , on a

- $(S)$  admet une infinité de solutions si

$$\text{Nb de contraintes} < \text{Nb de degrés de liberté}$$

- $(S)$  admet une unique solution si

$$\text{Nb de contraintes} = \text{Nb de degrés de liberté}$$

- $(S)$  n'admet aucune solution si

$$\text{Nb de contraintes} > \text{Nb de degrés de liberté}$$

**Note** : le dernier cas, le système  $(S)$  est dit *sur-déterminé*. Bien qu'il n'admette aucune solution, il existe des méthodes (notamment numériques) permettant de construire une solution approchée de ce type de système, méthodes que nous n'aborderons pas dans ce cours. On supposera donc dans la suite que le nombre de contraintes sera toujours inférieur ou égal au nombre de degrés de liberté. On supposera en particulier que le nombre d'équations est au plus égale au nombre d'inconnues :

$$n \leq p \quad \text{voire} \quad n = p$$

Par ailleurs, certains systèmes peuvent n'admettre aucune solution sans être sur-déterminé. Ce sont les systèmes contenant de contraintes contradictoires, dits *incompatibles*.

**Exemple** : en deux variables, les systèmes incompatibles sont ceux dont deux de leurs équations correspondent à **des droites parallèles**

**Exercice** : dans les cas  $p = 2$  et  $p = 3$ , faire la liste des différents cas possibles en s'appuyant sur la représentation géométrique des équations à deux ou trois inconnues :

- *Cas  $p = 2$ .*
  - un système  $2 \times 2$  admet une infinité de solutions si ses deux équations correspondent à **la même droite du plan**.
  - un système  $2 \times 2$  admet une unique solution si ses deux équations correspondent à **deux droites sécantes**.
  - les systèmes à deux inconnues *sur-déterminés* correspondent à **au moins trois droites non concourantes**.

- *Cas  $p = 3$ .*
  - Système  $3 \times 3$  admettant une infinité de solutions :
    1. **trois fois le même plan  $\Rightarrow$  un plan de solutions.**
    2. **trois plans ayant une droite en commun  $\Rightarrow$  une droite de solutions.**
  - un système  $3 \times 3$  admet une unique solution si ses trois équations correspondent à **trois plans indépendants**.
  - les systèmes à trois inconnues *sur-déterminés* correspondent à **au moins quatre plans n'ayant aucun point en commun**.

### Rang d'un système linéaire

À partir des observations ci-dessus, il est clair que le nombre de contraintes contenues dans un système linéaire a une influence directe sur la "taille" de l'ensemble des solutions. Ainsi, on appelle *rang d'un système linéaire* le nombre  $r$  de contraintes effectives contenues dans celui-ci. Plus ce rang est élevé, moins il y a de solutions.

Ces contraintes étant imposées sur  $p$  degrés de libertés représentés par les variables  $(x_1, \dots, x_p)$ , on verra dans la suite que la différence

$$d = p - r$$

est un paramètre permettant de donner une mesure de l'ensemble des solutions du système étudié.

**ATTENTION** : bien que chaque équation corresponde à une contrainte, le nombre de ces contraintes n'est pas toujours égal au nombre d'équations du fait d'une possible redondance pouvant exister au sein d'un même système. Pour tout système à  $n$  équations, on a donc

$$r \leq n$$

On verra dans la suite que les méthodes de résolution (algorithmiques) des systèmes linéaires permettent, entre autre de faire le tri dans un ensemble d'équations pour en tirer un ensemble de contraintes optimal.

**Note** : en pratique, on peut montrer que le rang d'un système donné dépend uniquement *des coefficients* de celui-ci (et non de son second membre). Par extension, on appelle donc *rang d'une matrice* donnée le rang de tout système linéaire dont elle est la matrice.

Dans ce contexte, on peut en particulier montrer qu'un système ( $S$ ) est inversible si et seulement s'il est de rang  $n$ , i.e. maximal pour un système carré à  $n$  inconnues.

Du point de vue du calcul matriciel, on a alors, pour toute matrice carrée  $A$ ,

1. Si  $\det(A) \neq 0$ , alors  $\text{rg}(A) = n$ .
2. Si  $\det(A) = 0$ , alors  $\text{rg}(A) < n$ .

### Paramétrisation de l'ensemble des solutions

Soit ( $S$ ) un système linéaire de  $n$  équations à  $p$  inconnues (vérifiant  $n \leq p$ ) admettant une infinité de solutions. En notant  $r = \text{rg}(S)$ , on a  $r < n \leq p$ .

La résolution de ( $S$ ) abouti en général à la description de cet ensemble infini  $\Sigma$  de solutions, cette description se faisant sous la forme d'une *paramétrisation* de l'ensemble  $\Sigma$ .

#### Exemples :

- Si  $p = 2$ , alors  $n \leq 2$  et  $r = 1$ . Autrement dit, le système ( $S$ ) est équivalent à une unique équation linéaire à deux inconnues :

$$(S) \iff ax + by = c$$

et l'ensemble  $\Sigma$  peut être représenté par une droite du plan.

Si  $b \neq 0$ , la droite  $\Sigma$  admet une *équation réduite* de la forme

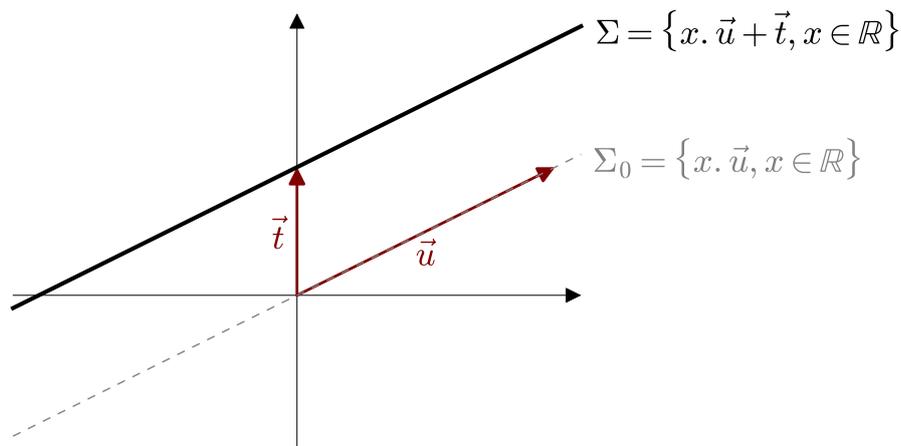
$$y = \alpha x + \beta$$

De cette équation réduite, on tire alors la description suivante :

$$\Sigma = \{(x, \alpha x + \beta), x \in \mathbb{R}\}$$

On reconnaît ici une *paramétrisation de droite* dont on peut tirer un vecteur directeur  $\vec{u}$  ainsi qu'un vecteur translation  $\vec{t}$  :

$$\Sigma = \{x \cdot (1, \alpha) + (0, \beta), x \in \mathbb{R}\} \Rightarrow \begin{cases} \vec{u} = (1, \alpha) \\ \vec{t} = (0, \beta) \end{cases}$$



**Note** : si  $a \neq 0$ , on peut tirer de l'équation ci-dessus une autre paramétrisation de la même droite  $\Sigma$  :

$$(S) \iff x = \gamma y + \delta$$

d'où

$$\Sigma = \{(\gamma y + \delta, y), y \in \mathbb{R}\} = \{y \cdot (\gamma, 1) + (\delta, 0), y \in \mathbb{R}\}$$

et

$$\begin{cases} \vec{u}' = (\gamma, 1) \\ \vec{t}' = (\delta, 0) \end{cases}$$

**Exercice** : donner et représenter géométriquement les vecteurs  $\vec{u}$ ,  $\vec{t}$ ,  $\vec{u}'$  et  $\vec{t}'$  associés à la droite d'équation

$$2x + y = 1$$

On prendra soin de vérifier que les deux descriptions obtenues correspondent bien à la même droite du plan.

— Si  $p = 3$ , alors  $n \leq 3$  et  $r \leq 2$ . On distingue alors deux cas possibles :

— Soit  $r = 1$ . Dans ce cas,  $(S)$  est équivalent à une unique équation de la forme

$$(E) : ax + by + cz = d$$

Si  $c \neq 0$ , on a

$$(E) \Leftrightarrow z = \alpha x + \beta y + \gamma$$

L'ensemble  $\Sigma$  peut alors être représenté par un plan de l'espace dont on peut obtenir une description basée sur les deux paramètres  $x$  et  $y$  :

$$\Sigma = \{(x, y, \alpha x + \beta y + \gamma), x, y \in \mathbb{R}\}$$

D'un point de vue géométrique, en notant que

$$\Sigma = \{x \cdot (1, 0, \alpha) + y \cdot (0, 1, \beta) + (0, 0, \gamma), x, y \in \mathbb{R}\}$$

on obtient une représentation paramétrique du plan  $\Sigma$  basée sur les vecteurs "générateurs"  $\vec{u}$  et  $\vec{v}$  et un vecteur translation  $\vec{t}$  :

$$\vec{u} = (1, 0, \alpha), \quad \vec{v} = (0, 1, \beta), \quad \vec{t} = (0, 0, \gamma)$$

Notons également que si les coefficients  $a$  et  $b$  sont également non nuls, il est possible d'extraire de l'équation  $(E)$  deux autres paramétrisations du plan  $\Sigma$  faisant intervenir les paramètres  $y$  et  $z$  ou  $x$  et  $z$  (à faire).

— Soit  $r = 2$ . Dans ce cas, le système  $(S)$  est équivalent à un système à deux équations, associé à un couple de plans de l'espace s'intersectant le long d'une droite. À partir des deux équations, on peut alors tirer une paramétrisation de cette droite, basée sur un seul paramètre et correspondant, d'un point de vue géométrique, à la donnée d'un vecteur directeur  $\vec{u}$  et un vecteur translation  $\vec{t}$ .

**Exemple** : soit

$$(S) : \begin{cases} x + 2y + z = 1 & (E_1) \\ y - z = 0 & (E_2) \end{cases}$$

1. On peut exprimer l'ensemble  $\Sigma$  des solutions de  $(S)$  à l'aide d'un unique paramètre, par exemple  $z$  :

$$(E_2) \Rightarrow y = z$$

$$(E_1) \Rightarrow x = 1 - 2y - z = 1 - 3z$$

$$\Sigma = \{(1 - 3z, z, z) \mid z \in \mathbb{R}\}$$

2. D'un point de vue géométrique, on a

$$\Sigma = \{(1 - 3z, z, z) \mid z \in \mathbb{R}\} = \{z \cdot (-3, 1, 1) + (1, 0, 0), \mid z \in \mathbb{R}\}$$

d'où

$$\vec{u} = (-3, 1, 1), \quad \vec{t} = (1, 0, 0)$$

**Exercice :**

- Déterminer une paramétrisation de  $\Sigma$  à l'aide de la variable  $y$ .
- Extraire de cette nouvelle représentation un vecteur directeur  $\vec{u}'$  et un nouveau point  $A'$  définissant  $\Sigma$ .
- Vérifier que la nouvelle paramétrisation obtenue correspond bien à la même droite de l'espace.

De façon générale, on peut montrer que pour tout système linéaire admettant une infinité de solution, on peut construire une *paramétrisation linéaire* de l'ensemble  $\Sigma$  de ses solutions.

Précisément, pour tout système à  $p$  inconnues de rang  $r < p$  admettant une infinité de solutions, il est possible de décrire l'ensemble  $\Sigma$  de ces solutions à l'aide d'un nombre fini  $d$  de paramètres  $t_1, \dots, t_d$  ( $d < p$ ) :

$$(x_1, \dots, x_p) \text{ vérifie } (S) \iff \begin{cases} x_1 = \alpha_{11}t_1 + \dots + \alpha_{1d}t_d + \beta_1 \\ x_2 = \alpha_{21}t_1 + \dots + \alpha_{2d}t_d + \beta_2 \\ \vdots \\ x_p = \alpha_{p1}t_1 + \dots + \alpha_{pd}t_d + \beta_p \end{cases}$$

les réels  $\alpha_{ij}$  et  $\beta_i$  étant fixés, caractéristiques du système étudié et les paramètres  $t_1, \dots, t_d$  pouvant être pris quelconques dans  $\mathbb{R}$ .

Le nombre  $d$  minimal de paramètres nécessaires à la description de l'ensemble  $\Sigma$  est appelé *la dimension* de l'ensemble  $\Sigma$  et l'on peut montrer que

$$d = p - r$$

Intuitivement, cette dimension correspond au nombre de degrés de liberté qu'il reste pour construire une solution de  $(S)$  une fois que toutes les contraintes sont satisfaites (une dans le cas d'une droite, deux dans le cas d'un plan,  $p - 1$  dans le cas d'un hyperplan).

Notons enfin que dans le cas d'un système de Cramer, on a

$$r = p = n \quad \text{et} \quad d = 0$$

On n'a donc aucune liberté pour construire une solution de  $(S)$ . La solution (unique) est imposée par les contraintes du système.

### 4.2.3 Systèmes libres, systèmes liés

#### Définitions

Au cours de la résolution d'un système linéaire, on doit donc déterminer *la dimension de l'ensemble des solutions cherché* et pour cela déterminer *le rang  $r$*  du système étudié.

En pratique, on doit donc "faire le tri" parmi les équations du système pour en tirer un système de contraintes optimal : un ensemble de  $r$  équations imposant le même ensemble de solutions que le système de départ.

Il faut pour cela caractériser rigoureusement ces systèmes optimaux, que l'on dira *libres*.

Formellement, cela passe par l'étude des éventuels liens qui peuvent exister entre ces équations. Ainsi,

Deux équations à  $p$  inconnues

$$(E_1) : a_{11}x_1 + \dots + a_{1p}x_p = b_1 \quad \text{et} \quad (E_2) : a_{21}x_1 + \dots + a_{2p}x_p = b_2$$

sont dite *liées* si elles sont proportionnelles,

i.e. s'il existe une constante  $\lambda \in \mathbb{R}$  telle que

$$(\forall j \in \{1, \dots, p\}, \quad a_{2j} = \lambda \cdot a_{1j} \quad \text{et} \quad b_2 = \lambda \cdot b_1) \Leftrightarrow (E_2) = \lambda \cdot (E_1)$$

Le système  $\{(E_1), (E_2)\}$  est alors de rang 1  
et

$$\{(E_1), (E_2)\} \Leftrightarrow \{(E_1)\} \Leftrightarrow \{(E_2)\} \quad \text{à condition que } \lambda \neq 0$$

**Exemples :**

- Deux équations à deux inconnues proportionnelles sont associées à la même droite du plan.
- Deux équations à trois inconnues proportionnelles sont associées au même plan de l'espace.

Ainsi, si un système donné contient deux équations proportionnelles, il n'est pas libre. Il est possible d'en retirer l'une des deux équations proportionnelles sans en changer l'ensemble des solutions.

Cependant, d'autres liens peuvent exister entre les équations d'un système contenant plus de deux équations.

Ainsi, de façon générale, un système linéaire  $(S) = \{(E_1), \dots, (E_n)\}$  est dit *lié* si l'une de ses équation(par exemple la première) peut s'exprimer sous la forme d'une *combinaison linéaire* des autres :

$$\exists \lambda_2, \dots, \lambda_n \in \mathbb{R}, / (E_1) = \sum_{i=2}^n \lambda_i \cdot (E_i) \quad (4.1)$$

Dans ce cas, l'équation  $(E_1)$  n'apporte pas de réelle contrainte au système. On peut la retirer sans changer l'ensemble des solutions :

$$(S) = \{(E_1), (E_2), \dots, (E_n)\} \iff \{(E_2), \dots, (E_n)\}$$

### Exemples :

- Si  $(E_1)$  et  $(E_2)$  sont deux équations à deux inconnues non proportionnelles, elles correspondent à deux droites concourantes  $\Delta_1$  et  $\Delta_2$  du plan.

Toute équation de la forme

$$(F) = \lambda \cdot (E_1) + \mu \cdot (E_2), \quad \lambda, \mu \in \mathbb{R}$$

correspond alors à une droite du plan passant par le point  $P = \Delta_1 \cap \Delta_2$ .

On a alors

$$(S) = \{(E_1), (E_2), (F)\} \iff \{(E_1), (E_2)\}$$

et l'ensemble des solutions de ces deux systèmes est réduit au point  $P$ .

Dans le repère ci-contre,

1. Tracer les droites associées aux équations suivantes :

$$\begin{aligned}(E_1) &: x + y = 2 \\ (E_2) &: 3x - 2y = 1\end{aligned}$$

2. Déterminer explicitement les équations

$$\begin{aligned}(F) &= (E_1) - (E_2) \\ \text{et} \\ (G) &= 2 \cdot (E_1) + (E_2)\end{aligned}$$

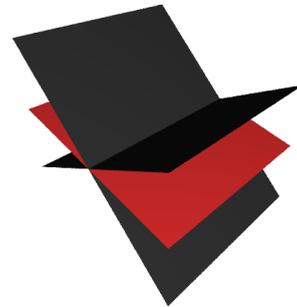
et tracer les droites associées.

Commenter.

- Si  $(E_1)$  et  $(E_2)$  sont deux équations à trois inconnues non proportionnelles, elles correspondent à deux plans de l'espace dont l'intersection est une droite  $\Delta$ . Tout équation de la forme

$$(F) = \lambda \cdot (E_1) + \mu \cdot (E_2), \quad \lambda, \mu \in \mathbb{R}$$

correspond alors à un plan de l'espace contenant la droite  $\Delta$ .



On a alors

$$\{(E_1), (E_2), (F)\} \iff \{(E_1), (E_2)\}$$

et l'ensemble des solutions de ces deux systèmes est constitué de la droite  $\Delta$ .

#### Notes :

- Dans les exemples géométriques ci-dessus, qu'il est à chaque fois possible de conserver deux des trois ou quatre équations présentes sans changer l'ensemble des solutions.

D'un point de vue analytique, cela traduit le fait que dans chaque cas, l'égalité liant les équations du système représenté peuvent se mettre sous la forme

$$(E_1) = \dots \quad \text{ou} \quad (E_2) = \dots$$

De façon générale, si l'un des coefficients  $\lambda_2, \dots, \lambda_n$  de l'égalité (4.1) est non nul (par exemple  $\lambda_2$ ), il est alors possible d'exprimer l'équation  $(E_2)$  associée en fonction

des autres. On a alors

$$\begin{aligned} (S) = \{(E_1), (E_2), (E_3), \dots, (E_n)\} &\iff \{(E_1), (E_3), \dots, (E_n)\} \\ &\iff \{(E_2), (E_3), \dots, (E_n)\} \end{aligned}$$

En pratique, cela signifie qu'il existe plusieurs tris possibles pour un système donné. Il existe donc plusieurs système optimaux équivalents à un même système de départ. Tous ces systèmes optimaux ont cependant un point en commun : le nombre d'équations qu'il contiennent, correspondant au rang du système de départ.

- Dans le cas d'un système incompatible, on peut montrer qu'il existe également une relation linéaire liant les équations du système. Cependant, cette relation n'est vérifiée que par les coefficients  $a_{ij}$  du système et non les seconds membres  $b_i$ . C'est ce défaut qui rend impossible le fait que toutes les équations du système soient vraies en même temps. C'est pour cela qu'un tel système n'admet pas de solution.

#### 4.2.4 Systèmes échelonnés, systèmes réduits

##### Définitions

Les systèmes linéaires les plus rapides à résoudre sont les systèmes *triangulaire* et notamment les systèmes triangulaires inversibles.

Précisément, un système linéaire carré  $(T) : AX = B$  est dit *triangulaire* si sa matrice (carrée) est triangulaire :

$$(T) : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 & (E_1) \\ \phantom{a_{11}x_1} + a_{22}x_2 + \dots + a_{2n}x_n = b_2 & (E_2) \\ \phantom{a_{11}x_1} \phantom{+ a_{22}x_2} + \dots + a_{nn}x_n = b_n & (E_n) \end{cases}$$

avec

$$\forall i \in \{1, \dots, n\}, \quad a_{ii} \neq 0$$

La condition ci-dessus assure en particulier que la matrice  $A$  (et donc le système  $(T)$ ) est inversible. On obtient par ailleurs cette unique solution en "remontant" le système, la dernière équation donnant la valeur de  $x_n$  et chaque équation permettant d'exprimer une nouvelle inconnue en fonction des inconnues déjà calculées :

$$\begin{aligned}
 (E_n) &\Rightarrow x_n = \frac{b_n}{a_{nn}} \\
 (E_{n-1}) &\Rightarrow x_{n-1} = \frac{1}{a_{n-1,n-1}}(b_{n-1} - a_{n-1,n}x_n) \\
 &\vdots \\
 (E_1) &\Rightarrow x_1 = \frac{1}{a_{11}} \left( b_1 - \sum_{k=2}^n a_{1k}x_k \right)
 \end{aligned}$$

En étudiant de près les propriétés de ces systèmes triangulaires, on constate que cette notion n'est qu'un cas particulier (le cas inversible) d'un cas plus général : la notion de système échelonné.

Ainsi, soient  $(S)$  un système linéaire à  $n$  équations et  $p$  inconnues et  $A$  sa matrice.

- On appelle  $i$ -ème pivot de  $(S)$  ou pivot de la  $i$ -ème ligne le premier coefficient non nul de la  $i$ -ème équation de  $(S)$  et donc le premier coefficient non nul de la  $i$ -ème ligne de  $A$
- Le système  $(S)$  (ou la matrice  $A$ ) est dit(e) *échelonné(e)* si le pivot de chacune de ses lignes se trouve strictement à droite du pivot de la ligne précédente.
- Le système  $(S)$  (ou la matrice  $A$ ) est dit(e) *réduit(e)* si
  - il (elle) est échelonné(e),
  - chacun de ses pivots vaut 1,
  - chaque pivot est l'unique coefficient non nul de sa colonne.

**Note :** étant donné le lien fort existant entre un système linéaire et sa matrice, on parlera indifféremment de système ou de matrice échelonné(e) ou réduit(e).

### Exemples :

- Déterminer le pivot de chacune des lignes des matrices ci-dessous et préciser celles qui sont échelonnées, celles qui sont réduites.

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad A_2 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix} \quad A_3 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A_4 = \begin{pmatrix} -1 & 3 & 4 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad A_5 = \begin{pmatrix} -1 & 3 & 4 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix} \quad A_6 = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- À partir de cette définition, il est clair que les systèmes de Cramer échelonnés sont

précisément les systèmes triangulaires inversibles évoqués plus haut. D'autre part, si un tel système est réduit, on a

$$\forall i, j \in \{1, \dots, n\}, \quad \begin{cases} i \neq j \Rightarrow a_{ij} = 0 \\ i = j \Rightarrow a_{ij} = 1 \end{cases}$$

Autrement dit, les seuls systèmes de Cramer réduits sont ceux dont la matrice est la matrice identité. Chaque équation est alors

$$(E_i) : x_i = b_i$$

De façon générale, les systèmes échelonnés et/ou réduits sont les plus simples à lire et à résoudre. On peut en particulier montrer qu'un système échelonné est nécessairement libre. Son rang correspond donc au nombre d'équations qu'il contient.

Par ailleurs, d'un point de vue pratique, la notion de système échelonné généralise bien la notion de système triangulaire aux systèmes non inversibles puisqu'en les "remonter", on peut là encore déterminer une paramétrisation de l'ensemble des solutions (la dimension de ce dernier étant directement accessible puisque l'on connaît son rang ainsi que le nombre de ses inconnues).

**Exemple** : déterminer l'ensemble des solutions des systèmes homogènes associés à chacune des matrices échelonnées et/ou réduites étudiées à l'exemple précédent.

### Opérations élémentaires et élimination

De façon générale, les méthodes systématiques de résolutions d'un système linéaire consistent à transformer le système de départ ( $S$ ) en un système ( $S'$ ) *équivalent* (i.e. admettant le même ensemble de solutions) mais plus simple à résoudre. Ces méthodes sont en particulier rendues possibles grâce opérations élémentaires ci-dessous, qui ne modifient pas l'ensemble des solutions du système auquel elles sont appliquées :

— Combinaison d'équations :

$$(E_i) \leftarrow (E_i) + \lambda.(E_j), \quad i \neq j, \lambda \in \mathbb{R}$$

— Permutation d'équations :  $(E_i) \leftrightarrow (E_j)$

— Multiplication d'un équation par une constante non nulle :

$$(E_i) \leftarrow \lambda.(E_i), \quad \lambda \neq 0$$

La combinaison (appliquée indifféremment aux équations du système ( $S$ ) ou aux lignes de sa matrice étendue) permettent en particulier d'opérer *l'élimination de variables* dans un maximum d'équations de ( $S$ ). Une organisation systématique de cette élimination, couplée aux deux autres opérations élémentaires, permet alors un traitement systématique de n'importe quel système linéaire. Précisément

### Théorème (Gauss-Jordan)

Soient  $n$  et  $p$  deux entiers non nuls tels que  $n \leq p$ .

1. À l'aide de combinaisons et de permutations d'équations, on peut transformer tout système linéaire de  $n$  équations à  $p$  inconnues en un système échelonné équivalent.
2. À l'aide de la multiplication d'équations, on peut transformer tout système linéaire de  $n$  équations à  $p$  inconnues en un système réduit.

**Exemple** : échelonner puis réduire la matrice  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ .

### Applications

**Résolution de systèmes linéaires.** La systématisation du processus d'élimination évoquée ci-dessus permet de mettre en place des protocoles *algorithmiques* produisant une forme échelonnée de tout système linéaire. Couplée à des méthodes algorithmiques de résolution des systèmes échelonnés, cela permet de construire des algorithmes de résolution complets pour les systèmes linéaires.

**Note** : appliqué à un système ( $S$ ) non inversible, l'algorithme d'échelonnage permet également de "faire le tri" dans l'ensemble des équations. Le système échelonné étant en effet libre, les équations "inutiles" du système ( $S$ ) sont remplacées par des lignes de 0 au bas de la matrice  $A$ . Si le système étudié admet des solutions, ces lignes de 0 correspondent à des équations de la forme  $0 = 0$  qui peuvent être retirées du système. Si le système est incompatible, ces lignes correspondent à des équations de la forme  $0 = \alpha$  avec  $\alpha \neq 0$ . Ces équations n'étant jamais satisfaites, le système n'admet aucune solution.

**Calcul de déterminant.** On peut noter que les opérations élémentaires évoquées ci-dessus correspondent aux opérations élémentaires de la manipulation des déterminants. Précisément, à l'aide des combinaisons de lignes (qui ne changent pas un déterminant) et les permutations de lignes (qui ne changent que le signe du déterminant), on peut associer à toute matrice carrée  $A$ , une matrice échelonnée ayant le même déterminant que  $A$  (au signe près). Cette matrice échelonnée étant triangulaire, son déterminant est égale au produit de ses coefficients diagonaux. On peut alors construire des processus algorithmiques efficaces de calcul de déterminant.

**Calcul d'inverse de matrice.** Enfin, le protocole de réduction, basé sur les trois opérations élémentaires de lignes, appliqué aux matrices inversibles permet également de

mettre en place protocole de calcul d'inverse de matrice, se traduisant lui aussi sous la forme d'un algorithme, plus efficace que l'algorithme "naïf" consistant à utiliser la formule générale. En effet, appliquer le protocole de réduction à une matrice inversible transforme cette dernière en la matrice identité. Ainsi, en appliquant le protocole de réduction à la double matrice

$$\left( \begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & \cdots & 0 & 1 \end{array} \right) = \left( \begin{array}{c|c} A & I \end{array} \right)$$

on obtient une matrice de la forme

$$\left( \begin{array}{c|c} I & A' \end{array} \right)$$

et on peut montrer que la matrice  $A'$  ainsi construite est exactement la matrice  $A^{-1}$ .

**Exemple :** Montrer que la matrice

$$A = \begin{pmatrix} 0 & 2 & 1 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

est inversible et calculer  $A^{-1}$ .

$$\left( \begin{array}{ccc|ccc} 0 & 2 & 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \xleftrightarrow[L_1 \leftrightarrow L_2]{} \left( \begin{array}{ccc|ccc} -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right)$$

$$\xleftrightarrow[L_3 \leftarrow L_3 - \frac{1}{2}L_2]{} \left( \begin{array}{ccc|ccc} -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 1 \end{array} \right)$$

$$\xleftrightarrow[L_1 \leftarrow -L_1, L_2 \leftarrow \frac{1}{2}L_2, L_3 \leftarrow 2L_3]{} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 2 \end{array} \right)$$

$$\xleftrightarrow[L_2 \leftarrow L_2 - \frac{1}{2}L_3]{} \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 2 \end{array} \right)$$

D'où

$$A^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ -1 & 0 & 2 \end{pmatrix}$$



# Chapitre 5

## Introduction au langage Python

### Introduction

Les méthodes numériques regroupent l'ensemble des méthodes algorithmiques permettant de construire des solutions pour des problèmes mathématiques, physiques, biologiques, etc, et de tout problème pouvant se traduire en termes d'équations algébriques et/ou différentielles.

Nous aborderons ici les problèmes suivants :

- Résolution d'équations et de systèmes algébriques
- Calcul intégral
- Résolution d'équations et de systèmes différentiels

Quelque soient les problèmes auxquelles elles sont destinées à répondre, toutes les méthodes numériques seront basées sur le même modèle : déterminer de façon approchée les grandeurs caractéristiques qui définissent la ou les solutions du problème posé.

En pratique la recherche de ces grandeurs caractéristiques se fait "de proche en proche", en partant d'un ensemble de données de base et en affinant l'approximation de façon dynamique.

Formellement, la notion de *suite numérique* et notamment de suite *récurrente* est au cœur des méthodes numériques.

**Rappel** : une suite récurrente est une suite  $(u_n)$  définie par

- un ensemble de données initiales  $u_0, u_1, \dots$  permettant d'amorcer le calcul,
- une relation de récurrence  $u_n = \varphi(u_{n-1}, u_{n-2}, \dots)$  permettant de calculer chaque terme en fonction des précédents.

Ainsi, toutes les méthodes que l'on va voir sont basées sur la construction d'une ou plusieurs suites de ce type qui *convergent* vers les grandeurs caractéristiques que l'on cherche à approcher.

Nous reviendrons en détails sur cette notion de convergence, mais notons dès à présent

qu'elle contient notamment l'idée générale suivante : pour une suite donnée répondant aux caractéristiques voulues, plus on va "loin" dans la suite en question, plus l'approximation que l'on obtient est précise.

Autrement dit, pour obtenir une "bonne" précision, on doit calculer de nombreux termes pour les suites étudiées, puisque dans le cas d'une suite récurrente, le calcul d'un terme nécessite le calcul de tous les termes précédents.

C'est essentiellement pour cette raison que les méthodes numériques (dont le développement a commencé entre le XVIIIème et le XVIIIème siècle) sont devenues très populaires avec l'avènement de l'informatique.

Dans un premier temps, nous commencerons par présenter le logiciel `Python` et son langage de programmation.

Nous aborderons ensuite différentes méthodes numériques avec, pour chacune, une approche pratique (basée sur l'écriture d'algorithmes `Python` et de traitement d'exemples) et une approche théorique, permettant notamment de mesurer *la qualité* des différentes méthodes présentées.

## 5.1 Le langage Python

Le langage `Python` est un langage de programmation permettant de manipuler des quantités *numériques* et différents types d'objets mathématiques définis à partir de ces quantités numériques : fonctions, vecteurs, matrices, etc.

En pratique, le logiciel `Python` s'utilise principalement au travers de *scripts*, i.e. d'algorithmes (écrits dans n'importe quel éditeur de texte) qui sont ensuite lus et interprétés au travers du *shell*.

`Python` peut ainsi être installé sur toutes les plateformes et peut être utilisé au travers de nombreuses interfaces spécifiquement adaptées.

### 5.1.1 Installation

`Python` un logiciel libre, largement utilisé et donc la communauté est très active, notamment sur Internet.

Il évolue donc continuellement et notons que l'on trouve actuellement *deux versions* distinctes : la version 2 et la version 3. Chacune de ces versions évoluent en parallèles et ne sont compatibles qu'au prix d'un certain nombre de modifications. Dans ce cours, nous utiliserons exclusivement la version 2.

Ainsi, dans ce cours, nous travaillerons au travers de l'interface IDLE (GUI).

## Shell vs Script

À l'ouverture du logiciel, on obtient le shell :

```
*****CAPTURE SHELL*****
```

Bien qu'il soit possible d'entrer directement des instructions dans cette fenêtre, nous privilégierons l'utilisation d'un script annexe :

1. Dans l'onglet `File`, ouvrir une nouvelle fenêtre (`New window`).
2. Enregistrer cette nouvelle fenêtre avec l'extension `.py` dans un dossier de travail (`File` > `Save as`).

```
*****CAPTURE SHELL et SCRIPT*****
```

On peut alors écrire des instructions dans cette seconde fenêtre et les faire lire par le logiciel (i.e. les *compiler*) : `Module` > `Run module`.

**Note** : avant chaque compilation, il est nécessaire d'enregistrer les dernières modifications. En effet, Python compile uniquement la version enregistrée (qui peut parfois être différente de la version affichée à l'écran).

## Bibliothèques et imports

Comme de nombreux logiciels de programmation, Python est organisé en bibliothèques indépendantes, contenant chacune des fonctions pré-écrites spécifiques à un domaine d'utilisation donné.

Selon les outils dont on pourra avoir besoin, il est donc nécessaire d'importer ces outils et/ou bibliothèques.

Bien que ces imports puissent se faire au cours du calcul, il est d'usage d'importer en tête l'ensemble des fonctions que l'on utilise dans la suite.

En Python, il existe deux méthodes permettant d'importer des commandes et/ou des bibliothèques complètes :

— La commande

```
from bibliothèque import commande(s)
```

est en général utilisée pour importer quelques commandes depuis une bibliothèque donnée.

— La commande

```
import bibliothèque
```

permet, elle, d'importer *l'ensemble* des commande depuis une bibliothèque donnée.

Notons qu'avec cette méthode, à chaque fois que l'on souhaitera utiliser une commande de la bibliothèque ainsi importée, il faudra indiquer au logiciel où aller chercher cette commande en indiquant *en préfixe* le nom de cette bibliothèque :

```
bibliothèque.commande(arguments)
```

Pour plus de souplesse, il est alors possible, au moment de l'import, d'associer *un alias* à la bibliothèque choisie :

```
import bibliothèque as alias
```

Les commandes ainsi importées sont alors accessibles via l'alias choisi :

```
alias.commande(arguments)
```

**Notes :** il est également possible d'importer l'ensemble des commandes d'une bibliothèque donnée à l'aide de la première syntaxe via la ligne de commandes

```
from bibliothèque import *
```

Les commandes ainsi importées peuvent alors être utilisées sans alias.

Cependant, en cas d'imports multiples, cette méthode peut générer certains conflits (si par exemples deux commandes différentes possèdent le même noms dans leurs bibliothèques respectives).

Nous préférons donc la seconde méthode, plus contraignante, mais plus stable.

Dans l'ensemble des programmes que nous écrirons dans le cadre de ce cours, nous chargerons par défaut les deux bibliothèques suivantes :

— La bibliothèque `numpy` importée sous l'alias `np` :

```
import numpy as np
```

qui contient de nombreux outils indispensables (fonctions usuelles, constantes universelles, vecteurs, matrices, etc).

— La bibliothèque `matplotlib.pyplot` importée sous l'alias `plt` :

```
import matplotlib.pyplot as plt
```

qui contient tous les outils nécessaires aux graphiques (2D).

## Mise en page et commentaires

De façon général, il est important, en programmation, d'écrire des scripts claires et bien commentés.

Dans ce cadre, nous verrons que les spécificités du langage `Python` oblige, dans une certaine mesure, à adopter les conventions de mise en page usuelles en programmation. Mais notons dès à présent qu'il est possible d'insérer *des commentaires* à tous les niveaux d'un script.

Précisément, un commentaire est un ensemble de caractères figurant dans le script, mais qui ne seront pas lus par la machine. Cela permet d'indiquer le plan général d'un algorithme, de préciser les conditions d'application et plus généralement de fournir à un.e éventuel.lecteur ou lectrice tous les détails nécessaires à la bonne compréhension du script en question.

En `Python`, il existe deux modes 'commentaires' distincts :

- Il est possible d'agrémenter un script de commentaires *ponctuels* (titre, sens d'une ligne, etc) à l'aide du caractère `#`. Précisément, il peut être placé n'importe où dans une ligne de code. Tout ce qui est placé après ce `#`, sur la même ligne, ne sera pas lu lors de la compilation.

```
print 2+2 # une somme
```

- Il est possible d'insérer de "grands" blocs de commentaires à l'aide de `'''` ou `"""`. Précisément, trois symboles de ce type permettent de marquer le début d'un bloc de commentaires et la répétition de ces trois symboles permet de fermer ce bloc. Lors de la compilation, tout ce qui sera contenu entre le premier et le second triplet ne sera pas lu.

```
"""
Les lignes ci-dessous affichent
differentes sommes d'entiers pairs
"""
print 2+2 # une somme
print 2+4 # une autre
print 4+6 # une autre
```

**Note** : comme de nombreux logiciel développés essentiellement par les anglo-saxons, les accents et autres caractères spéciaux ne sont pas toujours bien gérés par `Python` . Ils sont donc à éviter en général, y compris dans les commentaires.

Notons également qu'en `Python`, le séparateur décimal est le point. La virgule est, elle, dédiée à séparer différents objets.

## 5.1.2 Objets et variables

### Opérations mathématiques

La première utilisation de Python peut être celui d'une calculatrice, Python étant capable d'effectuer les opérations de base sur les nombres réels :

```
# Les operations usuelles :
print 2+3      # une somme
print 2*3      # un produit
print 2**3     # une puissance (eviter le "chapeau" ^ )
print 2/3      # un quotient d'entiers  $\frac{2}{3}$ 
print 2.0/3    # un quotient de nombres reels
```

**Note** : Python permet également de définir et manipuler les nombres complexes à l'aide de la commande `complex(a,b)` :

```
print complex(-1,2)
print (complex(0,1))**2
```

### Notion de variables

Comme on l'a évoqué plus haut, Python permet de manipuler de nombreux objets mathématiques basés sur une représentation numérique. En pratique, ces objets on peut donner des noms à ces différents objets afin de simplifier leur manipulation.

Donner un nom à un objet donné, ou *affecter une valeur à une variable donnée* se fait à l'aide du symbole `=`. Précisément, la ligne

```
nom = valeur
```

permet d'affecter la valeur `valeur` à la variable `nom`. Exemple :

```
a = 2
b = 3
print a+b
```

Notons ici que le signe `=`, en Python n'est pas une égalité au sens mathématique du terme, mais plutôt une flèche ( $\leftarrow$ ) permettant de demander à la machine de mettre ce qu'il y a à droite de la flèche dans ce qu'il y a à gauche. Cette action étant en particulier asymétrique, les lignes suivantes sont acceptables (et permettent souvent d'*incrémenter* une variable donnée) :

```
a = 2
print a
a = a+1
print a
```

Notons également que lors d'une affectation, la valeur associée au nom choisi est dé-

terminée à partir des valeurs déjà en mémoire *au moment de l'affectation*. Ainsi :

```
a = -2
b = 3
c = b-a
print c
a = 1
print c
print b-a
```

Notons enfin qu'il est possible d'effectuer plusieurs affectations simultanées :

```
a,b = -2,3
print a
print b
print a,b
```

En résumé, une variable est une chaîne de caractères sous laquelle on désigne un objet donné. Il est possible d'utiliser tout type de chaîne de caractères, à condition de respecter les quelques règles ci-dessous :

- Ne pas utiliser de caractère spécial (symboles, parenthèses, apostrophes, accents, cédilles, etc) ni d'espace.
- Ne jamais commencer par un chiffre.

### 5.1.3 Fonctions numériques

La bibliothèque `numpy` contient l'ensemble des fonctions réelles usuelles. Ainsi si cette bibliothèque est importée sous l'alias `np`, Python connaît alors les fonctions suivantes :

Fonction	Commande
sinus	<code>np.sin</code>
cosinus	<code>np.cos</code>
exponentielle	<code>np.exp</code>
racine carrée	<code>np.sqrt</code>
logarithme néperien	<code>np.log</code>
logarithme base 2	<code>np.log2</code>
logarithme base 10	<code>np.log10</code>
factoriel	<code>np.factorial</code>
...	

La bibliothèque `numpy` contient également quelques constantes universelles et notamment la constante

$$\pi \approx 3.1416$$

Ainsi,

```
print np.sin(np.pi/2)
```

A partir de ces fonctions usuelles et des opérations de base, il est alors possible de construire n'importe quelle fonction réelle d'une ou plusieurs variables, à l'aide de la commande `lambda` :

```
f = lambda x : x**2+np.exp(x)
print f(0)
print f(1)
```

ou encore

```
g = lambda x,y : np.sqrt(x**2+y**2)
print f(-1,1)
a,b = np.cos(1),np.sin(1)
print f(a,b)
```

**Note** : la fonction  $f$  ainsi définie ne peut être appliquée qu'à des valeurs *numériques* ou des variables définies au préalable :

```
f = lambda x : x**2+np.exp(x)
print f(x)
```



## Matrices et vecteurs

La bibliothèque `numpy` (là encore importée sous l'alias `np`) contient également l'ensemble des commandes spécifiques à l'algèbre linéaire et permettant de définir et de manipuler les matrices et les vecteurs.

Ainsi, la commande `np.array` permet, selon les arguments donnés, de construire un vecteur ou une matrice :

— La commande

```
U = np.array([0.5,3,-1])
```

permet de construire un vecteur à trois composantes.

— La commande

```
M = np.array([[0.2,1,-4],[1.5,0,3]])
```

permet de créer une matrice à deux lignes et trois colonnes.

Avant de voir les différentes manipulation spécifiques aux vecteurs et matrices, notons que la commande `len` (pour "length" = "longueur") permet d'accéder à la longueur d'un vecteur. Appliquée à une matrice, la commande `len` renvoie le nombre de lignes.

Une fois définis des matrices et/ou des vecteurs, il est possible des manipuler via les

outils de l'algèbre linéaire. Ainsi,

- L'addition (+) et la multiplication scalaire (\*) effectue la combinaison linéaire "terme à terme" (à condition que les tailles l'autorise).

```
U = np.array([0.5,3,-1])
V = np.array([101,300.0,-200])
print 2*U + 3*V
```

- Le produit matriciel  $M \times N$  se fait, lui, via la commande `np.dot(M,N)` :

```
M = np.array([[0.2,1,-4],[1.5,0,3]])
N = np.array([[1.2,1],[1.5,0],[3,3]])
print np.dot(M,N)
print np.dot(N,M)
```

**ATTENTION** : le produit  $M*N$  peut également exister, mais il effectue les multiplications terme à terme ; ce qui ne correspond en rien à la multiplication matricielle au sens algébrique du terme.

De façon générale, l'ensemble des opérations classiques sur les matrices sont définies dans la bibliothèque `numpy` (et dans sa sous-bibliothèque `np.linalg`).

Cependant, dans les programmes que nous allons aborder dans ce cours, nous utiliserons principalement les matrices et vecteurs pour stocker les valeurs numériques spécifique à nos modèles. Plutôt que les opérations de l'algèbre linéaire, nous utiliserons en pratique certaines propriétés spécifiques aux vecteurs de Python.

Il est ainsi possible d'accéder à un coefficient ou une partie des coefficients d'une matrice donnée, à l'aide des symboles `[]` et `..`. Précisément,

- Les coefficients d'un vecteur :

```
U = np.array([0.5,3,-1])
print U[1] # acces au coefficient 1
print U[0] # acces au coefficient 0
```

- Les coefficients d'une matrice :

```
M = np.array([[0.2,1,-4],[1.5,0,3]])
print M[1,2] # acces au coefficient (1,2)
print M[0,0] # acces au coefficient (0,0)
```

- Un sous-vecteur :

```
U = np.array([0.0,0.5,3,-1,-2,1,4])
print U[2:5] # du coeff 2 au coeff 4
print U[:4] # du coeff 0 au coeff 3
print U[3:] # du coeff 3 au dernier
```

— Une sous-matrice :

```
M = np.array([[0.2,1,-4],[1.5,0,3]])
print M[0:2,0:2] # une sous matrice 2x2
print U[0,:]     # une ligne
print U[:,2]     # une colonne
```

**Note** : on aura remarqué dans les exemples ci-dessus quelques règles caractéristiques du langage Python :

1. La première place, dans un vecteur ou une matrice est la numéro 0. De façon générale, l'énumération d'entiers en Python commence par défaut à 0. Notons qu'alors, la dernière coordonnée d'un vecteur à  $n$  coordonnées est la numéro  $n - 1$ .
2. Lorsque l'on indique une place de fin, le compte s'arrête toujours un coup avant.

Nous verrons à l'usage que ces deux spécificité sont en réalité intimement liées.

Pour finir, donnons quelques commandes permettant de construire des matrices et/ou vecteurs particuliers :

```
print np.eye(4)           # une matrice identite
print np.zeros(3)        # un vecteur nul
print np.zeros((3,2))    # une matrice nulle
print np.ones(4)         # un vecteur de 1
print np.ones((2,3))     # une matrice de 1
```

Dans la même famille de commande, notons également la commande

```
np.linspace(a,b,n)
```

qui permet construit un vecteur à  $n$  coordonnées régulièrement réparties le long de l'intervalle  $[a, b]$ .

Exemple :

```
I = np.linspace(0,1,5)
print I
```

Enfin, notons qu'il est possible d'appliquer une fonction numérique (par exemple construite avec la commande `lambda`) à l'ensemble des coefficient d'un vecteur donné en appliquant directement la fonction souhaitée au vecteur voulu :

```
V = np.array([-1,2,-3])
print V**2
```

ou

```
V = np.array([-1,2,-3])
f = lambda x : np.sqrt(x**2)
print f(V)
```

## Listes et chaînes de caractères

Les listes et les chaînes de caractères sont deux outils supplémentaires que l'on utilise fréquemment en programmation. Bien qu'ils soient de nature et d'utilité radicalement différents, on verra qu'ils se manipulent globalement de la même façon. Ainsi,

- Une *liste* est l'un des outils élémentaires permettant de stocker des objets dans un même ensemble. Précisément, il s'agit d'une collection d'objets, donnés entre crochets [...].

```
L = [0.2,1,4,8.5,7,-2]
print L
```

Comme pour les vecteurs et matrices, il est possible d'isoler un élément ou une sous-liste d'une liste donnée, à l'aide des symboles [ : ] :

```
L = [0.2,1,4,8.5,7,-2]
print L[0]
print L[-1]
print L[1:4]
```

Enfin, outre le fait que, contrairement aux matrices et vecteurs, il est possible de stocker tous les types d'objets que connaît Python (et pas seulement des valeurs numériques), l'une des différences principales entre vecteurs et listes se trouve au niveau des opérations que l'on peut effectuer à partir de ces objets. Précisément, la somme de deux liste  $L$  et  $L_2$  renvoie *la liste concaténées* faite de  $L$  et  $L_2$  :

```
L = [0.2,1,4,8.5,7,-2]
L2 = [1,-7,4,0]
print L+L2
print L2+L
```

De ce fait, il est également possible de multiplier une liste donnée par un entier  $n$ . On obtient alors une liste contenant  $n$  fois la liste de départ.

Notons que cette différence de traitement empêche en particulier d'appliquer une fonction numérique à une liste.

Deux commandes spécifiques au listes pour finir :

- La commande `range` permet de construire des suites d'entiers sous forme de listes :

```
print range(1,10)
print range(10,20,2)
print range(20)
```

- La commande `append` permet d'ajouter un élément au bout d'une liste donnée :

```
L = [1,2,3]
print L
L.append(-9)
print L
```

- Une *chaîne de caractères* est un ensemble de caractères alphanumériques permettant d'afficher et/ou de manipuler du texte. En `Python`, une chaîne de caractères est définie à l'aides des guillemets (" ") ou d'apostrophes (' ') :

```
a = 'Bonjour'
print a
```

Comme pour les listes, il est alors possible d'isoler un caractère ou une sous-chaîne d'une chaîne donnée :

```
a = 'Bonjour'
print a[0]
print a[2]
print a[:3]
print a[3:]
```

Et toujours comme pour les liste, il est possible de concaténer deux chaînes de caractères en les additionnant :

```
a = 'Bonjour'
b = 'Python'
print a+b
```

Cette notion de chaîne de caractère permet entre autres de manipuler des données écrites (au format `txt`, par exemple). De notre côté, nous la pratiquerons surtout pour agrémenter les retours de `Python` de petits messages aidant à la lecture. Pour cela, il pourra être utile de transformer une valeur (numérique) donnée à l'aide de la commande `str` (pour "string" = "chaîne de caractère"). Ainsi, plutôt que

```
a = 2
print "La valeur de a est " + a 
```

on notera

```
a = 2
print "La valeur de a est " + str(a)
```

car on ne peut additionner deux objets qui ne sont pas de même nature (un entier (`int`) et une chaîne de caractères (`str`)).

Notons tout de même que, puisqu'il est possible d'afficher un ensemble de résultats, séparés par des virgules, on peut également utiliser la syntaxe ci-dessous :

```
a = 2
print "La valeur de a est ",a
```

### 5.1.4 Outils de programmation

Les principaux outils de Python (comme pour tout langage de programmation) sont ses outils de programmation : ses boucles et ses tests ainsi que la possibilité de construire ses propres commandes.

#### Boucle for

Parmi les outils élémentaires de programmation, la notion de *boucle* désigne un ensemble d'instructions permettant de demander à la machine de répéter plusieurs fois un ensemble d'opérations données.

En pratique, il existe deux façons différentes de créer une boucle. La méthode la plus souple est la boucle for dont la syntaxe générale est la suivante :

```
for indice in Ensemble d'indices :
    Ensemble des actions
    à effectuer à chaque
    tour de boucle
```

Précisément,

1. Les mots `for` et `in` ainsi que les `:` sont spécifiques et indispensables.
2. En guise d'indice, on peut choisir n'importe quelle chaîne de caractères (avec les mêmes contraintes que pour les variables).
3. L'ensemble d'indice doit être une liste, un vecteur ou une chaîne de caractères.

Une fois ces choix effectués, on doit indiquer l'ensemble des opérations que l'on souhaite effectuer à *chaque tour* en les décalant par rapport à la limite gauche de la fenêtre.

Ce décalage, appelé *indentation* est l'une des particularités du langage Python. Il fait partie intégrante de la syntaxe d'une boucle Python. Outre le fait que cette règle oblige à une mise en page rigoureuse, elle permet également de se passer d'un *marqueur de fin de boucle* : une fois l'ensemble des opérations de la boucle indiquées, il suffit de revenir en bord de page pour continuer le programme.

Une fois qu'une telle boucle est écrite et validé, l'ordinateur va répéter l'ensemble des actions demandées pour chacune des valeur de l'ensemble d'indice, l'indice prenant, à chaque tour, la valeur suivante dans l'ensemble.

```
for c in "Bonjour Python" :
    print c
```

**Note** : dans une boucle `for`, le nombre de tours correspond au *cardinal* de l'ensemble des indices. Nous verrons en pratique qu'il est toujours possible de construire une boucle *for* sur un intervalle d'entiers (à l'aide de la commande `range`, donc). Notons également que par définition, on ne peut utiliser une boucle `for` que si l'on sait au préalable le nombre de tours que l'on souhaite faire.

### Exemples :

1. Afficher les 10 premiers carrés.
2. Construire et afficher la liste des 10 premiers cubes.
3. Construire la somme des 100 premiers entiers  $S = \sum_{n=0}^{100} k$ .

### Boucle while

La seconde méthode permettant de construire une boucle est la boucle `while` qui permet de demander à l'ordinateur de répéter un ensemble d'actions *tant que* telle condition est vérifiée.

En Python, la syntaxe est la suivante :

```
while condition d'arrêt non atteinte :
    Ensemble des actions
    à effectuer à chaque
    tour de boucle
```

Dans cette syntaxe, outre le mot-clé `while` qui remplace les mots-clés `for` et `in`, on retrouve ici les " : " et l'indentation.

Notons que, dans une boucle `while`, le nombre de tour n'est pas nécessairement connu. Par ailleurs, parmi les actions à effectuer à chaque tour, l'une d'entre elles au moins doit permettre de progresser vers la condition d'arrêt (sous peine d'avoir une boucle infinie...).

### Exemple :

1. À l'aide d'une boucle `while`, construire et afficher la liste des carrés plus petits que 200.
2. Déterminer le plus petit entier dont le logarithme népérien est supérieur à 10.

Enfin, notons que les conditions d'arrêt peuvent notamment s'exprimer à l'aide des opérateurs booléens

Symbole	Sens
<	strictement inférieur
>	strictement supérieur
<=	inférieur ou égal
>=	supérieur ou égal
==	égal
!= ou <>	différent
in	est dans
not in	n'est pas dans
or	ou
and	et
...	

### Test if

Le troisième pilier de tout langage de programmation est *le test if*. Précisément, à l'aide des opérations booléennes ci-dessus, il est possible de demander à un ordinateur d'effectuer différentes actions, selon que telle ou telle condition est vérifiée.

En Python, la syntaxe est la suivante :

```

if condition 1 vérifiée :
    Ensemble des actions
    associées à la condition 1
elif condition 2 vérifiée :
    Ensemble des actions
    associées à la condition 2
    :
elif condition n vérifiée :
    Ensemble des actions
    associées à la condition n
else :
    Ensemble des actions
    par défaut

```

On retrouve ici les spécificités de la syntaxe Python : l'utilisation de l'indentation et les ":".

### Notes :

- Les entrées `elif` et `else` sont facultatives. Si elles apparaissent, elles doivent être alignées (verticalement) avec l'entrée `if` dont elles dépendent.
- Si plusieurs conditions sont vérifiées, seules les actions associées à la première condition vraie rencontrée sont effectuées.
- Les actions associées à `else` sont effectuées si aucune des conditions précédentes n'est vérifiée.

## Construire ses propres fonctions Python

Enfin, Python offre également la possibilité de construire ses propres commandes. On peut ainsi construire notre propre fonction `racine_carree` ou construire un programme qui détermine les racines d'un polynôme de degré 2, ou encore construire une fonction qui renvoie le plus grand élément dans une liste de valeurs numériques, etc

La syntaxe permettant de construire ses propres fonctions Python est la suivante :

```
def Nom_de_la_fonction(arguments) :
    Ensemble des actions à effectuer,
    à partir des arguments,
    pour obtenir le résultat souhaité
    return le résultat souhaité
```

Les mots clés de cette commande sont donc `def` et `return` et l'on retrouve les ":" et le système d'indentation permettant de marquer l'intérieur du programme.

Dans le code ci-dessus, les *arguments* sont des variables représentant l'ensemble des objets sur lesquels on souhaite faire travailler la fonction que l'on construit (le nombre positif  $a$  dont on veut la racine carrée, les coefficients  $a$ ,  $b$  et  $c$  du trinôme  $ax^2 + bx + c$  dont on veut les racines, la liste  $L$  dont veut le plus grand élément, etc).

Notons qu'il s'agit de variables *locales*. Il n'est en particulier pas nécessaire de les initialiser pour les utiliser dans un tel environnement.

Notons enfin que lorsqu'un ensemble d'instructions de ce type est validé, aucune opération n'est effectuée par la machine. Elle se contente en effet d'*associer* l'ensemble des actions à effectuer au nom choisi. Ces actions ne seront effectuées que lorsque cette nouvelle fonction sera *utilisée*.

Par ailleurs, lors de l'utilisation de cette nouvelle fonction, le résultat obtenu sera celui donné après le mot clé `return`.

**Exemple** : construire une fonction `racines(a,b,c)` qui renvoie la liste des racines (réelles ou complexes) d'un trinôme du second degré.

### 5.1.5 Graphiques

Python permet également de construire une large gamme d'objets graphiques (2D et 3D).

Dans ce paragraphe, nous présentons rapidement les éléments de base de la manipulation des graphiques avec Python.

Les principaux outils de dessin sous Python sont contenus dans les bibliothèques `numpy` et `matplotlib.pyplot`.

```
import numpy as np
import matplotlib.pyplot as plt
```

Les fonctions du module `numpy` sont désormais accessibles à l'aide du préfixe `np`, celles du module `matplotlib.pyplot`, à l'aide du préfixe `plt`.

De façon générale, quand on souhaite construire un dessin sous Python, on doit, à partir d'une figure initialement vide, ajouter progressivement les différents éléments graphiques (graphes, axes, légendes, etc) que l'on souhaite voir apparaître.

Pour créer une figure vide, on s'appuie sur la commande

```
plt.figure()
```

Cette fonction possède un certain nombre de paramètres qui permettent de régler les paramètres de la fenêtre. Dans un premier temps, les valeurs par défaut sont suffisantes, mais l'on retiendra le paramètre `figsize` qui permet de fixer la taille de la fenêtre graphique (en pouces) la première ligne d'un code créant un graphe peut ainsi être :

```
plt.figure(figsize = (10,10)) # un fenetre carree
```

Après avoir lu cette ligne, la machine se prépare à enregistrer un ensemble d'objets graphiques.

On indique alors les différents éléments graphiques que l'on souhaite voir apparaître, puis on appelle l'affichage du graphe à l'aide de la fonction `plt.show()`.

À la lecture de cette ligne, la machine ouvre une fenêtre annexe, contenant le graphe souhaité, sous la forme souhaitée et supprime de sa mémoire les différents éléments.

```
plt.figure(figsize = (10,10)) # un fenetre carree
Calculs et construction
des différents objets graphiques
plt.show()
```

## Tracé simple

Si  $X = [x_1, \dots, x_n]$  et  $Y = [y_1, \dots, y_n]$  sont deux listes de même taille, la commande `plt.plot(X,Y)` définit la ligne brisée reliant les points de coordonnées  $[x_i, y_i]$ .

### Exemple :

1. Entrer les lignes ci-dessous pour tracer votre premier graphe en Python :

```
n = 5
w = 2*np.pi/n
X = [np.cos(w*t) for t in range(n)]
Y = [np.sin(w*t) for t in range(n)]

plt.figure(figsize = (10,10))
plt.plot(X, Y)
plt.show()
```

2. Quel figure construit-on à l'aide de ce code ?
3. Modifier le code pour obtenir la figure "complète".
4. Modifier la valeur de  $n$ . Que reconnaît-on pour de grandes valeurs de  $n$  ?

### Graphe d'une fonction

Notons que Python (comme n'importe quel langage de programmation) est incapable de tracer de façon exacte la courbe d'une fonction donnée. Python ne peut que construire une approximation sous la forme d'un ensemble fini de points reliés par des segments de droites.

De ce fait, la commande `plt.plot` permet également de tracer le graphe d'une fonction  $f$  sur un intervalle  $[a, b]$ . Il suffit pour cela de construire les listes  $X$  et  $Y$  correspondantes.

**Exemple** : entrer et valider les lignes de commandes suivantes.

```
X = np.linspace(-2*np.pi, 2*np.pi, 250)
C = np.cos(X)
S = np.sin(X)

plt.figure(figsize = (10,10))
plt.plot(X, C)
plt.plot(X, S)
plt.show()
```

### Options de tracé

On constate que lorsque l'on superpose plusieurs tracés, Python attribue une couleur différente à chaque courbe. On peut imposer soi-même ses couleurs et imposer une multitude de paramètres précisant chacune des courbes.

Modifier le code ci-dessus en remplaçant les lignes correspondantes par les lignes ci-dessous :

```
plt.plot(X, C, color="blue", linewidth=2.5)
plt.plot(X, S, color="red", marker = "o", linestyle=" ")
```

On peut également associer une légende à chacune des courbes à l'aide de l'option `label`. La fonction `plt.legend` permet alors de préciser les détails de l'affichage de la légende. L'argument `loc` permet en particulier de préciser où on souhaite la placer :

```
plt.plot(X, C, color="blue", linewidth=2.5, label = "cos")
plt.plot(X, S, color="red", marker = "o", linestyle=" ", label = "sin")
plt.legend(loc='upper left')
```

## Compléments

La commande `plt.plot` adapte par défaut pour les axes aux valeurs extrémales des listes `X` et `Y`. Il est cependant possible de préciser ces valeurs extrémales à l'aide des commandes

```
plt.xlim(xmin,xmax) # les limites de l'axe des x
plt.ylim(ymin,ymax) # les limites de l'axe des y
```

On peut également imposer le découpage des deux axes donnant, dans les commandes `plt.xticks` et `plt.yticks`, les listes des jalons que l'on veut voir apparaître.

```
plt.xticks(L1) # la liste des jalons en x
plt.yticks(L2) # la liste des jalons en y
```

On peut également ajouter une légende pour chacun des axes à l'aide des commandes

```
plt.xlabel("L'axe des x")
plt.ylabel("L'axe des y")
```

Enfin,

— La commande

```
plt.grid()
```

permet d'ajouter une grille (légère) au graphe, associée aux jalons des axes.

— La commande

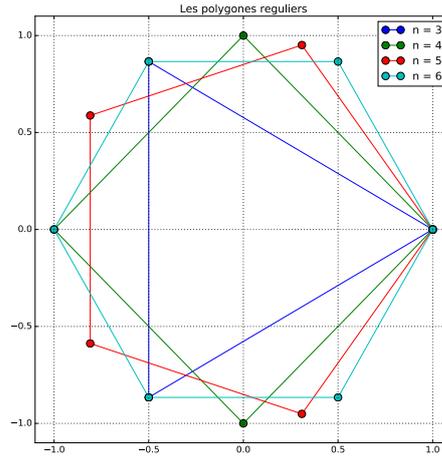
```
plt.title("Titre du graphe")
```

permet d'ajouter un titre à l'ensemble de la figure.

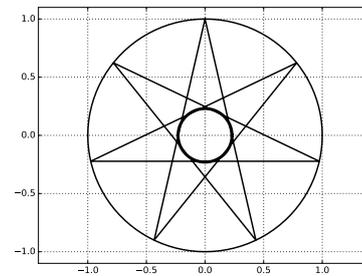
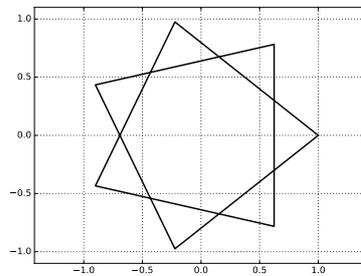
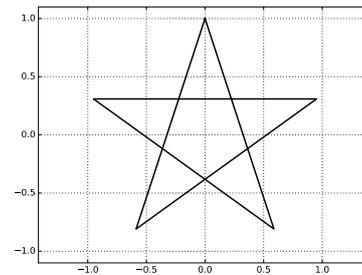
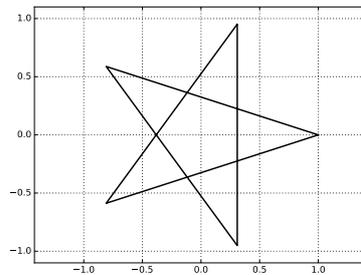
## Exercices

### Exercice 1 (*Les polygônes*)

1. Construire le dessin ci-dessous



2. En s'inspirant du programme écrit à la question précédente, construire un programme qui trace les étoiles ci-dessous.



## Exercice 2

1. (a) Tracer sur l'intervalle  $[-\pi, \pi]$  les courbes les fonctions suivantes.

$$\left[ x \mapsto \sin\left(\frac{x}{10}\right) \right], \quad \left[ x \mapsto \sin\left(\frac{x}{3}\right) \right], \quad \left[ x \mapsto \sin\left(\frac{x}{2}\right) \right]$$

$$\left[ x \mapsto \sin(x) \right], \quad \left[ x \mapsto \sin(2x) \right], \quad \left[ x \mapsto \sin(3x) \right]$$

**Note :** pour plus de clarté, on ajoutera une légende au dessin.

- (b) Que dire de l'influence du paramètre  $\omega$  dans la fonction  $x \mapsto \sin(\omega x)$  ?
2. (a) Tracer sur un même graphe les fonctions suivantes, pour  $x \in [0, 4\pi]$ .

$$[x \mapsto \sin(x)], \quad \left[ x \mapsto \sin\left(x + \frac{\pi}{6}\right) \right], \quad \left[ x \mapsto \sin\left(x - \frac{\pi}{2}\right) \right],$$

$$[x \mapsto \sin(x + \pi)], \quad [x \mapsto \sin(x + 2\pi)]$$

- (b) Que dire du paramètre  $\varphi$  dans la fonction  $x \mapsto \sin(x + \varphi)$  ?

### Exercice 3

1. Sur un même graphe, tracer les courbes des fonctions ci-dessous :

$$\left[ x \mapsto \frac{\sin(x)}{x} \right], \quad \left[ x \mapsto \frac{1}{x} \right], \quad \left[ x \mapsto -\frac{1}{x} \right]$$

2. (a) Tracer le graphe de la fonction  $\left[ x \mapsto \sin\left(\frac{1}{x}\right) \right]$  pour  $x \in [-0.1, 0.1]$

Que dire de la limite  $\lim_{x \rightarrow 0} \sin\left(\frac{1}{x}\right)$  ?

- (b) Sur un même graphe, superposer les courbes des fonctions

$$\left[ x \mapsto x \sin\left(\frac{1}{x}\right) \right], \quad [x \mapsto x], \quad [x \mapsto -x]$$

Que dire de la limite  $\lim_{x \rightarrow 0} x \sin\left(\frac{1}{x}\right)$  ?

- (c) Sur un même graphe, superposer les courbes des fonctions

$$\left[ x \mapsto x^2 \sin\left(\frac{1}{x}\right) \right], \quad [x \mapsto x^2], \quad [x \mapsto -x^2]$$

3. Étudier le comportement asymptotique des fonctions  $[x \mapsto x \sin\left(\frac{1}{x}\right)]$  et  $[x \mapsto x^2 \sin\left(\frac{1}{x}\right)]$ .

## 5.2 Résolution approchée d'équations et systèmes algébriques

L'un des problèmes fréquemment rencontrés lors d'une modélisation est la résolution d'une ou plusieurs équations algébriques.

Nous verrons dans un premier temps comment, à l'aide de méthodes "naïves", on

peut, sans trop de formalisme, mettre en place un premier type de méthode relativement efficace. Nous verrons ensuite comment la notion de suite récurrente permet d'améliorer grandement ces premières méthodes naïves et de les généraliser aux *systèmes* d'équations algébriques.

### 5.2.1 Le TVI et ses conséquences

Soit

$$(E) : f(x) = 0$$

une équation algébrique où  $f$  est une fonction connue, que l'on supposera continue et dérivable autant de fois que nécessaire.

Les premières méthodes de résolution approchée de ce type d'équation sont basées principalement (voire uniquement) sur le *Théorème des valeurs intermédiaires* (TVI) qui assure que, si  $f$  est continue, alors entre deux points  $a$  et  $b$  de l'axe réel, la quantité  $f(x)$  prend toutes les valeurs entre  $f(a)$  et  $f(b)$ .

Cette idée permet dans un premier temps de localiser une solution  $x^*$  de l'équation  $f(x) = 0$ , à condition de trouver deux valeurs  $a$  et  $b$  telles que  $f(a)$  et  $f(b)$  sont de signes distincts.

En s'appuyant toujours sur le TVI, il est alors possible de réduire petit à petit l'intervalle dans lequel est contenue la solution  $x^*$  cherchée. Une fois que celle-ci est "coincée" dans un intervalle de longueur  $\varepsilon > 0$  fixé, chaque valeur de cet intervalle est une valeur approchée de  $x^*$  à  $\varepsilon$  près.

**Note :** dans tous les exercices portant sur cette partie, on appliquera les méthodes proposées à la résolution de l'équation

$$x^2 - 2 = 0$$

sur  $\mathbb{R}^+$ . Il est clair que cette équation admet  $x^* = \sqrt{2}$  comme solution exacte.

Le fait d'appliquer nos méthodes à ce problème dont on connaît la solution exacte va en particulier nous permettre de *comparer* les résultats obtenus avec les différentes méthodes à cette solution exacte.

#### Exercice 4 (La méthode du scan)

On souhaite résoudre l'équation

$$(E) : x^2 - 2 = 0$$

sur  $\mathbb{R}^+$  par la méthode "du scan".

1. Montrer que l'équation  $x^2 - 2 = 0$  admet une unique solution dans l'intervalle  $[0, 2]$ .

2. Construire une suite  $(x_n)$  qui scanne l'intervalle  $[0, 2]$  par pas de 0.1 et déterminer une valeur approchée de  $\sqrt{2}$  à  $10^{-1}$  près. On fera apparaître la valeur obtenue ainsi que l'erreur commise et le nombre d'itérations nécessaires au calcul.
3. Déterminer de même une valeur approchée de  $\sqrt{2}$  à  $10^{-3}$  près,  $10^{-5}$  près,  $10^{-7}$  près.
4. Tracer dans un repère les points de coordonnées  $(n, \varepsilon_n)$ ,  $\varepsilon_n$  donnant la distance

$$\varepsilon_n = |x_n - \sqrt{2}|$$

correspondant à chaque itération (on reviendra pour cela à une erreur de  $10^{-3}$ ).

À l'aide d'un minimum de réflexion et sans outils théoriques supplémentaires, il est possible d'améliorer de façon significative l'algorithme précédent, i.e. de construire un algorithme basé sur les mêmes outils, mais nécessitant beaucoup moins de calculs pour arriver à la même précision : la dichotomie.

Le principe de la dichotomie, toujours basé sur le TVI, consiste à diviser par deux à chaque itération l'intervalle contenant la solution  $x^*$  cherchée, en étudiant pour cela le signe de  $f$  au milieu de l'intervalle d'étude.

### Exercice 5 (*La dichotomie*)

On reprend l'équation  $x^2 - 2 = 0$  dont l'une des solutions est contenue dans l'intervalle  $[0, 2]$ .

1. À l'aide d'un test `if`, déterminer un intervalle  $[a_1, b_1]$  de longueur 1 contenant la solution  $x^*$  cherchée.
2. Déterminer de même un intervalle  $[a_2, b_2]$  de longueur 0.5 contenant la solution  $x^*$  cherchée.
3. À l'aide d'une boucle `while`, déterminer un intervalle  $[a_n, b_n]$  de longueur  $10^{-3}$  contenant  $x^*$ . En déduire une valeur approchée de  $\sqrt{2}$  à  $10^{-3}$  près.
4. Modifier la boucle précédente pour qu'elle renvoie la valeur obtenue ainsi que l'erreur commise et le nombre d'itérations nécessaires au calcul.
5. Déterminer de même une valeur approchée de  $\sqrt{2}$  à  $10^{-5}$  près, à  $10^{-10}$  près, à  $10^{-15}$  près.
6. Tracer dans un repère les erreurs  $e_n = |a_n - x^*|$ .
7. Comparer ce graphe à celui obtenu par la méthode du scan.
8. Montrer sur le papier que

$$\forall n \in \mathbb{N}, \quad |a_n - x^*| \leq \frac{b - a}{2^n}$$

9. Dans un repère semi-log en  $y$ , tracer l'erreur de la méthode de dichotomie et les points de coordonnées  $(n, \frac{1}{2^n})$ .

### 5.2.2 La méthode de Newton-Raphson

La méthode de Newton est une méthode d'approximation des solutions d'une équation de la forme

$$f(x) = 0$$

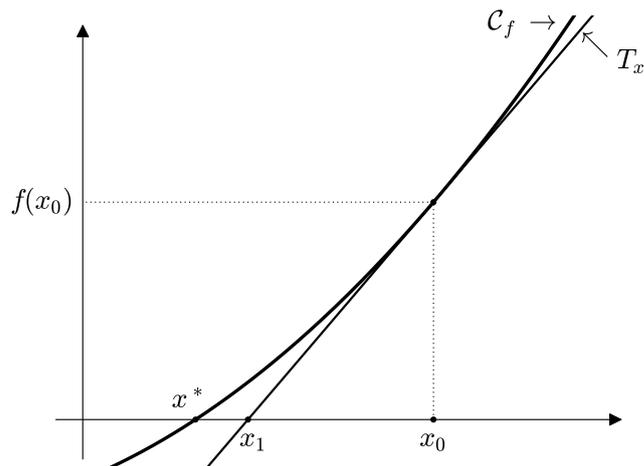
basée sur la relation de récurrence

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

On peut montrer que cette méthode est bien plus efficace que les méthodes "naïves" vues jusque là, à condition que le point de départ  $x_0$  ne soit pas trop éloigné de la solution  $x^*$  cherchée.

#### Mise en place

Historiquement, cette méthode, développée par I. Newton est basée sur l'observation graphique suivante :



#### Exercice 6 (*La méthode de Newton*)

1. *La relation de récurrence*

- (a) Rappeler l'équation de la tangente  $T_{x_0}$  à  $C_f$  au point d'abscisse  $x_0$ .
- (b) En déduire une expression de  $x_1$  en fonction de  $x_0$ .

2. *L'algorithme*

- (a) Calculer une valeur approchée de  $\sqrt{2}$  à  $10^{-4}$  près à l'aide de la méthode de Newton initialisée à  $x_0 = 2$ . On comptera là encore le nombre d'itérations nécessaires au calcul et l'on fera apparaître les erreurs

$$e_n = |x_n - \sqrt{2}|$$

à chaque itération.

- (b) Tracer dans un repère semi-log en  $y$  les erreurs des méthodes de dichotomie et de Newton.

### Vitesse de convergence

L'étude de la qualité d'une méthode telle que la méthode de Newton repose sur l'évaluation de la distance  $|x_n - x^*|$  en fonction de  $n$ .

Bien qu'il soit exclu de pouvoir déterminer de façon exacte cette distance, il est néanmoins possible de l'estimer, notamment en établissant une relation de la forme

$$|x_n - x^*| \leq E(n)$$

où  $E$  est une fonction connue, dépendant des données du problème (i.e.  $f$  et  $x_0$ ) et qui tend explicitement vers 0 quand  $n$  tend vers  $+\infty$ .

Si l'on parvient à établir une telle inégalité, on pourra alors déterminer une valeur de  $n$  pour laquelle la borne  $E(n)$  est inférieure à une précision  $\varepsilon$  souhaitée. D'après l'inégalité en question, on sera alors sûr que la valeur  $x_n$  est une valeur approchée de  $x^*$  à  $\varepsilon$  près.

Or pour établir une inégalité de ce type concernant la méthode de Newton, il faut prendre un peu de hauteur et étudier en détails la convergence des suites numériques et plus précisément la convergence des suites récurrentes vérifiant une relation de la forme  $x_{n+1} = \varphi(x_n)$ .

Notons dans un premier temps que, si  $(x_n)$  est une suite numérique qui converge vers  $x^*$ , la vitesse de convergence de cette suite peut être caractérisée par le quotient

$$\tau_n = \frac{|x_{n+1} - x^*|}{|x_n - x^*|}$$

et plus précisément

$$\tau = \lim_{n \rightarrow +\infty} \tau_n$$

Intuitivement, à  $n \in \mathbb{N}$  fixé, la quantité  $\tau_n$  peut être vu comme un "taux de progression" vers  $x^*$  entre le terme  $x_n$  et le terme  $x_{n+1}$ .

La limite  $\tau$  permettant quant à elle d'étudier ce taux de progression au voisinage de la limite  $x^*$ .

Ainsi, si l'on parvient par exemple à établir une inégalité de la forme

$$\forall n \in \mathbb{N}, \quad \tau_n \leq k$$

pour une constante  $k > 0$  fixée, on peut alors montrer que la distance  $|x_n - x^*|$  vérifie une relation de la forme

$$\forall n \in \mathbb{N}, \quad |x_n - x^*| \leq C.k^n$$

**Exercice** : à démontrer. On précisera la valeur de la constante  $C > 0$ .

Si en outre la constante  $k > 0$  déterminée lors de l'étude de  $\tau_n$  vérifie  $k < 1$ , cela signifie que la suite  $(x_n)$  converge vers  $x^*$  à vitesse *géométrique*, i.e. au moins aussi vite que la vitesse à laquelle la suite géométrique  $(k^n)$  tend vers 0.

**Note** : formellement, pour établir une inégalité du type  $\tau_n \leq k$ , il suffit d'étudier la limite  $\tau$ . On peut en particulier montrer qu'une telle relation est vérifiée si et seulement si  $\tau < 1$ .

De façon générale, on peut également montrer que plus cette limite est petite, plus la convergence de la suite  $(x_n)$  sera rapide. On peut en particulier montrer que si  $\tau = 0$ , cette vitesse de convergence est significativement plus rapide que dans le cas géométrique. Précisément, on peut montrer qu'il existe un réel  $r > 1$  tel que

$$\forall n \in \mathbb{N}, \quad |x_n - x^*| \leq C.k^{r^n}$$

On parle alors de *vitesse de convergence d'ordre  $r$* .

Enfin, pour étudier la vitesse de convergence de la suite de Newton à travers ces indicateurs, rappelons un résultat important concernant les suites récurrentes vérifiant une relation du type

$$x_{n+1} = \varphi(x_n)$$

où  $\varphi$  est une fonction connue (une telle suite est dite *autonome*).

Précisément, on peut montrer que si une telle suite converge, alors sa limite  $x^*$  vérifie  $\varphi(x^*) = x^*$  (un tel point est appelé *point fixe* de la fonction  $\varphi$ ).

**Exercice** :

1. Déterminer la fonction  $\varphi$  associée à la suite de Newton.
2. Montrer que la solution  $x^*$  de l'équation  $f(x) = 0$  est un point fixe de  $\varphi$ .

Or dans ce cas, le taux de progression  $\tau_n$  vérifie

$$\tau_n = \frac{|x_{n+1} - x^*|}{|x_n - x^*|} = \left| \frac{\varphi(x_n) - \varphi(x^*)}{x_n - x^*} \right|$$

et

$$\tau = \lim_{n \rightarrow +\infty} \tau_n = |\varphi'(x^*)|$$

Autrement dit, la vitesse de convergence d'une suite autonome associée à  $\varphi$  dépend directement de la valeur de la dérivée  $\varphi$  au niveau de la limite.

**Exercice** : déterminer la limite  $\tau$  dans le cas d'une suite de Newton.

Enfin, pour déterminer précisément la vitesse de convergence d'une suite de Newton, on peut faire appel aux *formules de Taylor*. On peut alors montrer que la convergence des

suites de Newton est *quadratique*, i.e. d'ordre 2. En pratique, cela signifie que le nombre de décimales exactes double à chaque itération :

$$\begin{aligned} |x_{n+1} - x^*| &\leq |\varphi'(x^*)| \cdot |x_n - x^*| + \left| \frac{\varphi''(\xi_n)}{2} (x_n - x^*)^2 \right| = \left| \frac{\varphi''(\xi_n)}{2} (x_n - x^*)^2 \right| \\ &\leq M \cdot |x_n - x^*|^2 \end{aligned}$$

où  $M = \frac{1}{2} \max_{[x_n, x^*]} |\varphi''|$ .

### Critères de convergence

Le principal inconvénient de la méthode de Newton est *sa sensibilité aux conditions initiales*. En effet, on peut montrer qu'une suite de Newton converge vers la solution  $x^*$  cherchée si la condition initiale  $x_0$  est "assez proche" de  $x^*$ .

Exemples :  $f(x) = 4 \frac{\sin(x)}{x}$ ,  $x_0 = 0.845$  ou  $x_0 = 0.83$ .

Formellement, cette limite de la méthode se retrouve dans la relation de récurrence

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

En effet, il est clair que le terme  $x_{n+1}$  n'existe que si  $f'(x_n) \neq 0$ .

En pratique, cela signifie que la méthode de Newton n'aboutit pas toujours si  $f'(x^*) = 0$ .

Dans le cas contraire (i.e. si  $f'(x^*) \neq 0$ ) et si  $f$  est suffisamment régulière, on peut montrer qu'il existe un intervalle de la forme  $]x^* - \delta, x^* + \delta[$  sur lequel la fonction  $f'$  ne s'annule pas. La méthode de Newton permet alors d'approcher  $x^*$  à condition de prendre la condition initiale  $x_0$  dans cet intervalle.

**Note** : on retrouve également cette contrainte dans l'estimation de l'erreur faite plus haut. En effet, de la relation

$$|x_{n+1} - x^*| \leq M \cdot |x_n - x^*|^2$$

on peut tirer une majoration de l'erreur  $\varepsilon_n = |x_n - x^*|$  de la forme

$$|x_n - x^*| \leq \frac{(M \cdot \varepsilon_0)^{2^n}}{M}$$

On constate ici que cette erreur tend vers 0 à vitesse quadratique à *condition* que la constante  $M\varepsilon_0 < 1$ . On voit là encore que cela est vérifié à condition que l'erreur initiale  $\varepsilon_0 = |x_0 - x^*|$  soit suffisamment petite, i.e. à condition que le point de départ  $x_0$  soit suffisamment proche de la solution  $x^*$  cherchée.

Pour palier à ce défaut, il est possible de combiner la méthode de Newton avec une dichotomie. Précisément, à partir d'un intervalle donné, on détermine une première approximation grossière de  $x^*$  à l'aide de la dichotomie, puis on utilise cette première approximation comme point de départ pour appliquer la méthode de Newton.

### Exercice 7 (*La fonction Racine\_carree*)

Construire un algorithme qui calcule toute racine carrée à  $10^{-13}$  près, à l'aide d'une combinaison de la dichotomie et de la méthode de Newton. On pourra utiliser la commande

```
a = input('Entrer un nombre positif a')
```

### 5.2.3 Généralisation aux systèmes algébriques

La méthode de Newton-Raphson se généralise à la résolution approchée de *systèmes* d'équations algébriques

$$(S) : \begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, \dots, x_n) = 0 \end{cases}$$

Formellement, un tel système est associé à un *champ de vecteurs*, i.e. une fonction

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^n \\ X = (x_1, \dots, x_n) \longmapsto F(X) = (f_1(X), \dots, f_n(X))$$

Résoudre (S) consiste alors à déterminer un vecteur  $X^* \in \mathbb{R}^n$  tel que  $F(X^*) = \vec{0}$ .

Or on peut adapter la méthode de Newton en remplaçant la dérivée  $f'(x)$  par la *matrice jacobienne*  $\nabla F(X)$  du champ de vecteurs  $F$  :

$$\nabla F(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X) & \cdots & \frac{\partial f_1}{\partial x_n}(X) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(X) & \cdots & \frac{\partial f_n}{\partial x_n}(X) \end{pmatrix}$$

On obtient alors la formule de récurrence suivante :

$$\forall n \in \mathbb{N}, \quad X_{n+1} = X_n - [\nabla F(X_n)]^{-1} \times X_n$$

Et on peut montrer que si la matrice  $\nabla F(X^*)$  est inversible et si  $X_0$  est assez proche de  $X^*$ , la suite de vecteurs  $(X_n)$  ainsi construite vérifie

$$\lim_{n \rightarrow +\infty} \|X_n - X^*\| = 0$$

avec, là encore, une vitesse de convergence quadratique.

## 5.2.4 Compléments

**Exercice 8 (La suite logistique)**

On appelle *suite logistique* toute suite récurrente  $(x_n)$  dont le point de départ  $x_0$  est dans l'intervalle  $]0, 1[$  et vérifiant une relation de récurrence de la forme

$$x_{n+1} = \mu \cdot x_n \cdot (1 - x_n)$$

où  $\mu \in [0, 4]$  est un paramètre fixé.

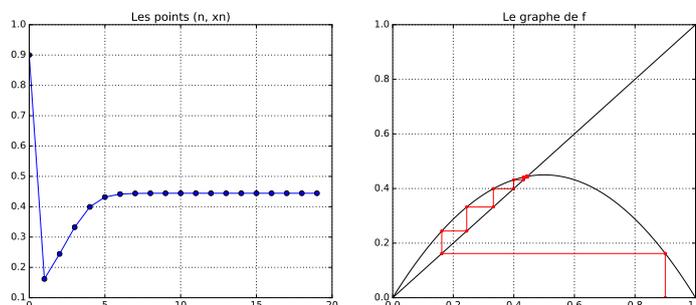
L'objectif de cet exercice est d'illustrer comment ce type de suite, issu des modèles de dynamique des populations, permet une première approche de la notion de chaos en faisant varier la valeur du paramètre  $\mu$ .

Dans toute la suite, on fixera  $x_0 = 0.9$  (bien que, hormis pour quelques valeurs particulières, cette valeur de départ n'a pas de réelle influence sur le comportement de la suite  $(x_n)$ ).

## 1. Représentations graphiques

Il existe deux méthodes permettant de représenter les suites récurrentes :

- en traçant et reliant les points  $(n, x_n)$ ,
- en s'appuyant sur le graphe de la fonction  $f_\mu : x \mapsto \mu x(1 - x)$  et la première bissectrice.



- (a) Construire une fonction `graphe1(mu, x0, n)` qui trace et relie entre eux les points  $(0, x_0)$  à  $(n, x_n)$  de la suite logistique de paramètre  $\mu$ .
- (b) Construire une fonction `graphe2(mu, x0, n)` qui trace le parcours des  $n$  premiers points de la suite logistique de paramètre  $\mu$  en s'appuyant sur la courbe de  $f_\mu$  et la première bissectrice.

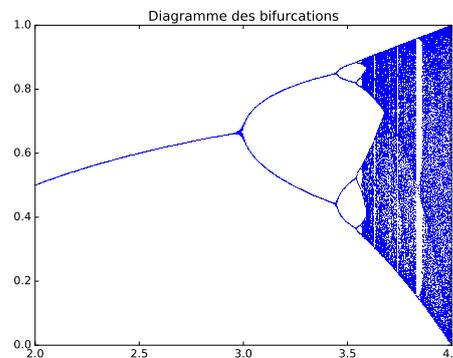
2. Étude graphique de l'influence du paramètre  $\mu$ 

- (a) Étudier graphiquement le comportement de  $(x_n)$  lorsque  $\mu \in ]0, 1]$ . On pourra s'appuyer sur les deux fonctions construites à la question précédente.

- (b) Cas  $\mu \in ]1, 3]$ .
- Étudier graphiquement le comportement de  $(x_n)$  lorsque  $\mu \in ]1, 3]$ .
  - Déterminer par le calcul la valeur de  $\lim x_n$  dans ces cas.
  - Quelle différence observe-t-on entre les cas  $\mu \in ]1, 2]$  et  $\mu \in ]2, 3]$  ?
- (c) Qu'observe-t-on dans le cas  $\mu = 3.05$  ? Dans le cas  $\mu = 3.5$  ?
- (d) Observer le cas  $\mu = 3.86$ .
3. Pour  $\mu = 4$ , le système dynamique obtenu est chaotique : une infime variation de la valeur de départ  $x_0$  induit de grandes variations de la valeur de  $x_n$  après quelques itérations seulement. Illustrer cela pour quelques valeurs de  $x_0$  très proches les unes des autres.

#### 4. Diagramme des bifurcations

Il est possible de démontrer que lorsque  $\mu$  est compris entre 3 et  $1 + \sqrt{6} \approx 3.45$ , la suite  $(x_n)$  finit par osciller entre deux valeurs dépendantes de  $\mu$  mais pas de  $x_0$ . Entre 3.45 et 3.54 environ, la suite finit par osciller entre quatre valeurs. Au delà de cette valeur la suite oscille entre huit valeurs, puis seize, etc. À partir de  $\mu = 3.57$  environ, le chaos s'installe et, mis à part pour certaines valeurs, il n'est plus possible de décrire le comportement asymptotique de la suite. Tout ceci peut être résumé par un *diagramme de bifurcations* représentant, en fonction de  $\mu$ , les valeurs distinctes que prend la suite  $(x_n)$  pour de grandes valeurs de  $n$  :



Le diagramme ci-dessus a été obtenu en faisant varier  $\mu$  entre 2 et 4 par pas de 0.002. Pour chacune des valeurs de  $\mu$  sont représentées en ordonnée les valeurs distinctes à  $10^{-3}$  près de  $x_{101}, x_{102}, \dots, x_{200}$ .

Rédiger un script Python qui réalise le tracé de ce diagramme.

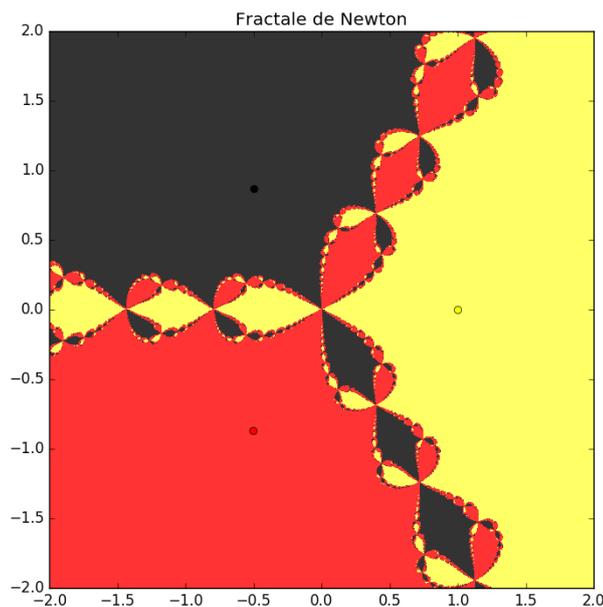
#### Notes :

- On peut tronquer un nombre flottant à l'aide de la commande `round`.
- Le tracé a été obtenu avec les options `marker=','`, `linestyle=""`.

## Fractales de Newton

En appliquant la méthode de Newton à des *polynômes complexes*, il est possible de générer des *fractales*.

### Exercice 9 (*Fractale de Newton*)



La fractale de Newton est une autre illustration de la notion de chaos construite à partir des suites récurrentes autonomes, étendue aux nombres complexes. Ainsi, le polynôme  $P(x) = x^3 - 1$  admet trois racines complexes  $z_1, z_2, z_3$  distinctes (lesquelles?) et l'on peut montrer que toute suite complexe  $(z_n)$  vérifiant la relation de Newton converge vers l'une de ces trois racines. Cependant, si le point de départ  $z_0$  d'une telle suite n'est pas situé directement à proximité de l'une des trois racines, il peut être difficile (voire impossible) de déterminer vers laquelle de ces trois racines elle va converger.

1. Construire un programme Python qui parcourt le carré  $[-2, 2]^2$  par pas de  $h = 0.005$  à l'aide de deux boucles imbriquées.

Pour chacun des nombres complexes  $z_{ij}$  associés aux points du carré  $[-2, 2]$ , on appliquera le protocole suivant :

- (a) Calculer les 100 premiers points de la suite de Newton associée à  $P(x)$  et partant du point d'affixe  $z_{ij}$ .
- (b) Déterminer, parmi les points d'affixes  $z_1, z_2, z_3$  le point le plus proche de  $z_{100}$ .

- Si c'est  $z_1$ , colorier le point d'affixe  $z_{ij}$  en jaune.
- Si c'est  $z_2$ , colorier le point d'affixe  $z_{ij}$  en noir.
- Si c'est  $z_3$ , colorier le point d'affixe  $z_{ij}$  en rouge.

2. Reprendre le protocole

- (a) sur le carré  $[-1, 1]^2$  parcouru par pas de  $h = 0.002$ ,
- (b) sur le carré  $[-0.5, 0.5]^2$  parcouru par pas de  $h = 0.001$ ,
- (c) sur le carré  $[-0.2, 0.2]^2$  parcouru par pas de  $h = 0.0005$ ,
- (d) sur le carré  $[-0.1, 0.1]^2$  parcouru par pas de  $h = 0.0001$ .

3. Commenter

### 5.3 Intégration numérique

L'intégration numérique apporte une réponse (approximative) au problème consistant à calculer une intégrale, portant sur une fonction dont on ne connaît pas de primitives.

Précisément, étant donnée une fonction  $f$ , d'une variable réelle, définie sur un intervalle  $[a, b]$  fixé, on souhaite estimer la quantité numérique

$$I = \int_a^b f(x)dx$$

Comme on va le voir dans l'exercice ci-dessous, on peut rapidement mettre en place des méthodes d'intégration numérique, en s'appuyant notamment sur une interprétation géométrique de la quantité  $I$  à calculer :  $I$  donne une mesure de l'aire du plan située "sous la courbe" représentant  $f$ .

Une approche plus théorique des méthodes abordées ci-dessous nous permettra ensuite d'en étudier la qualité et de les améliorer.

#### 5.3.1 Approche géométrique

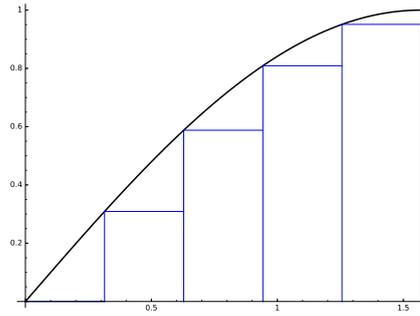
##### Exercice 10 (Calculs approchés d'une intégrale)

Pour évaluer la qualité des approximations obtenues, on va appliquer ces méthodes à l'intégrale suivante :

$$I = \int_0^{\frac{\pi}{2}} \sin(x)dx$$

1. *Méthode des rectangles*

Pour approcher l'intégrale  $\int_a^b f(x)dx$ , la méthode des rectangles consiste à découper l'intervalle  $[a, b]$  en  $n$  intervalles de même longueur et à approcher, sur chaque intervalle, l'aire sous la courbe par l'aire d'un rectangle :

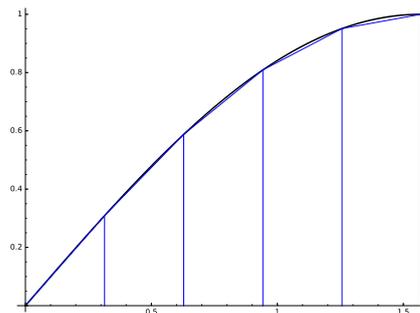


L'aire sous la courbe (et donc l'intégrale de  $f$  entre  $a$  et  $b$ ) peut alors être approchée par la somme des aires des rectangles considérés.

Calculer par la méthode des rectangles l'intégrale  $I$  pour différentes valeurs de  $n$  et comparer avec le résultat théorique.

### 2. Méthode des trapèzes

Il est possible d'affiner la méthode des rectangles en remplaçant les rectangles par des trapèzes (toujours basés sur un découpage régulier de l'intervalle d'intégration) :



Calculer par cette méthode l'intégrale  $I$  et comparer là encore le résultat obtenu à la valeur exacte.

## 5.3.2 Approche analytique

Afin d'étudier les caractéristiques des méthodes numériques ci-dessus, commençons par les modéliser à l'aide d'outils mathématiques.

Dans les deux cas, le calcul s'appuie avant tout sur un découpage régulier de l'intervalle  $[a, b]$  en  $n$  intervalles de même longueur,  $n$  étant un entier naturel fixé par l'utilisateur.

Formellement, ce découpage, appelé *discrétisation* de l'intervalle  $[a, b]$  est représenté par un ensemble de  $n + 1$  points

$$x_0 = a < x_1 < x_2 < \dots < x_n = b$$

tels que

$$\forall i \in \{0, \dots, n-1\}, \quad x_{i+1} - x_i = \frac{b-a}{n}$$

On note  $h = \frac{b-a}{n}$  le pas de discrétisation.

Dans un second temps, ces méthodes consistent ensuite à approcher la fonction  $f$  par des fonctions plus simples sur chaque intervalle  $[x_i, x_{i+1}]$  et à remplacer l'intégrale de  $f$  par l'intégrale de la fonction associée sur chaque sous-intervalle  $[x_i, x_{i+1}]$ . En ajoutant ces approximations "locales", on obtient dans chaque cas une valeur approchée de  $I$ .

Précisément, à  $n$  fixé,

— La valeur approchée obtenue par la méthodes des rectangles est la somme

$$I_n = \sum_{i=0}^{n-1} h \cdot f(x_i)$$

— La valeur approchée obtenue par la méthode des trapèzes est la somme

$$I_n = \sum_{i=0}^{n-1} \frac{h}{2} (f(x_i) + f(x_{i+1}))$$

### 5.3.3 Mesure d'erreur

Comme lors de la résolution des équations algébriques, les méthodes d'intégration numérique consistent formellement à construire des suites  $(I_n)$  qui convergent vers la valeur  $I$  cherchée (chaque terme  $I_n$  étant associé à un découpage de l'intervalle d'étude).

Pour valider la méthode, on doit en particulier vérifier que chacune de ces suites tend vers la valeur  $I$  cherchée. De plus, on peut là encore évaluer la qualité de chacune des méthodes en estimant la distance

$$\varepsilon_n = |I_n - I|$$

D'un point de vue théorique, la convergence de la méthode des rectangles est "évidente" car elle correspond en réalité à la *définition* de la notion d'intégrale.

Concernant la mesure de l'erreur, il est là encore possible d'estimer la distance  $\varepsilon_n = |I - I_n|$  en s'appuyant sur le développement de Taylor (ou le théorème des accroissements finis).

En effet, sur chaque sous-intervalle  $J_i = [x_i, x_{i+1}]$ , on approche la quantité  $I^{(i)} = \int_{x_i}^{x_{i+1}} f(x) dx$  par la quantité  $I_n^{(i)} = h \cdot f(x_i) = \int_{x_i}^{x_{i+1}} f(x_i) dx$ .

On peut alors exprimer l'erreur locale  $\varepsilon_n^{(i)}$  associée à l'intervalle  $J_i$  :

$$\varepsilon_n^{(i)} = \left| \int_{x_i}^{x_{i+1}} f(x) dx - \int_{x_i}^{x_{i+1}} f(x_i) dx \right| = \left| \int_{x_i}^{x_{i+1}} f(x) - f(x_i) dx \right|$$

et l'inégalité triangulaire appliquée aux intégrales permet alors une première majoration :

$$\varepsilon_n^{(i)} \leq \int_{x_i}^{x_{i+1}} |f(x) - f(x_i)| dx$$

À cette étape, le théorème des accroissements finis nous permet là encore d'obtenir une estimation de la différence  $f(x) - f(x_i)$  pour tout  $x \in J_i$ . En effet, si  $f$  est assez régulière, pour tout  $x \in [x_i, x_{i+1}]$ , il existe un réel  $c_{x,i} \in ]x_i, x[$  tel que

$$\frac{f(x) - f(x_i)}{x - x_i} = f'(c_{x,i}) \iff f(x) - f(x_i) = f'(c_{x,i}) \cdot (x - x_i)$$

D'où

$$|f(x) - f(x_i)| = |f'(c_{x,i})| \cdot (x - x_i) \leq \underbrace{\max_{[a,b]} |f'|}_M \cdot (x - x_i)$$

et

$$\varepsilon_n^{(i)} \leq \int_{x_i}^{x_{i+1}} M \cdot (x - x_i) dx = M \cdot \left[ \frac{(x - x_i)^2}{2} \right]_{x_i}^{x_{i+1}} = M \cdot \frac{(x_{i+1} - x_i)^2}{2}$$

En notant que  $x_{i+1} - x_i = h = \frac{b-a}{n}$ , on obtient la majoration suivante pour l'erreur locale :

$$\varepsilon_n^{(i)} \leq M \cdot \frac{h^2}{2} = M \cdot \frac{(b-a)^2}{2n^2}$$

Dans un second temps, on obtient l'erreur globale  $\varepsilon_n$  en additionnant les  $n$  erreurs locales :

$$\varepsilon_n = \sum_{i=0}^n \varepsilon_n^{(i)} \leq n \cdot M \cdot \frac{(b-a)^2}{2n^2} = M \cdot \frac{(b-a)^2}{2n}$$

Ainsi, la méthode des rectangles introduit une erreur  $\varepsilon_n = O\left(\frac{1}{n}\right)$ . La méthode est dite *d'ordre 1*.

En s'appuyant sur le même type d'outils et de calculs, il est possible de montrer que la méthode des trapèzes est, elle, d'ordre 2, i.e. l'erreur induite par cette méthode vérifie

$$\varepsilon_n = O\left(\frac{1}{n^2}\right)$$

### 5.3.4 Généralisation

D'un point de vue théorique, on peut voir les deux méthodes d'intégration ci-dessus comme deux cas particuliers d'une méthode plus générale, basée sur l'approximation des fonctions étudiées à l'aide de *polynômes*.

En effet,

- La méthode des rectangles consiste à approcher, sur chaque intervalle  $J_i = [x_i, x_{i+1}]$  la fonction  $f$  par la constante  $f(x_i)$ , i.e. par un polynôme de degré 0.

- La méthode des trapèzes consiste à approcher, sur chaque intervalle  $J_i$  la fonction  $f$  par une fonction affine, i.e. par un polynôme de degré 1.

En s'appuyant sur ces observations, il est alors possible de généraliser le processus en s'appuyant sur des approximations locales de la fonction  $f$  par des polynômes de différents degrés.

Et à l'image de ce que l'on a observé pour les deux premières méthodes, on peut montrer que plus le degré des polynômes utilisés est élevé, plus la méthode est précise.

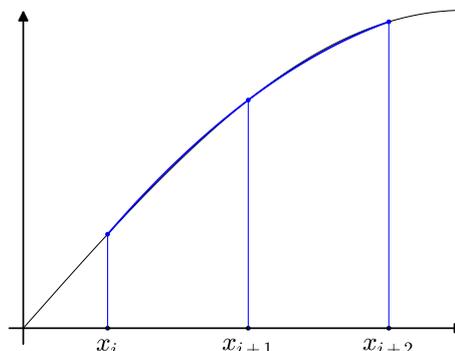
Par ailleurs, en étudiant de près le calcul intégral appliqué aux polynômes, on peut montrer que, dans chaque cas, il est possible d'exprimer les aires locales en s'appuyant uniquement sur les valeurs de la fonction  $f$  aux points  $x_i$  de discrétisation.

Ainsi, citons la *méthode de Simpson*, basée sur une approximation de  $f$  par des polynômes de degrés 2.

On peut montrer que l'approximation associée est obtenue par la somme suivante :

$$I_n = \sum_{\substack{i=0 \\ i \text{ pair}}}^{n-2} 2h \cdot \left( \frac{1}{6}f(x_i) + \frac{2}{3}f(x_{i+1}) + \frac{1}{6}f(x_{i+2}) \right)$$

On note sur le dessin ci-contre qu'il faut adapter un peu le calcul et la somme à effectuer pour tenir compte du fait qu'il nous faut trois points (et donc deux intervalles) pour définir une parabole.



**Exercice** : calculer cette somme pour différentes valeurs de  $n$  pour approcher l'intégrale  $I = \int_0^{\frac{\pi}{2}} \sin(x)dx$  et vérifier que l'erreur commise est de l'ordre de grandeur de  $\frac{1}{n^4}$ .

### 5.3.5 Traitement de données

Les méthodes d'intégration numérique point par point permettent également d'étendre le calcul intégral aux fonctions issues de l'expérimentation. Précisément, les méthodes expérimentales produisent en général des données sous la forme de *nuages de points* dont les coordonnées correspondent à des valeurs relevées par les capteurs lors de l'expérience.

Dans ce cas, il est rare que l'on puisse tirer de l'expérience une expression explicite de la quantité  $f(x)$  que l'on étudie.

Cependant, à partir du nuage de point, il est malgré tout possible d'estimer l'intégrale de cette fonction.

#### Exercice 11

On considère un ensemble de données expérimentales rassemblées dans le fichier `donnees_expe.txt`

1. Récupérer le fichier dans le dossier contenant le script sur lequel on travaille.
2. À l'aide de la commande

```
D = np.loadtxt("donnees_expe.txt")
```

importer les données dans une matrice  $D$ .

3. À l'aide de la commande `plt.plot`, tracer la courbe associée à ces données.
4. Calculer une valeur approchée de l'aire sous cette courbe à l'aide de la méthode des rectangles puis de la méthode des trapèzes.