

Cours de mathématiques

PCSI 2

Lycée Bellevue

Année 2025-2026

*Ce document a pour but de donner un aperçu des notions essentielles vues en cours.
Il ne se substitue pas aux notes prises en cours, et ne remplace pas une attention soutenue en classe.*

Les sciences elles-mêmes, et les mathématiques en particulier, consistent d'abord en des textes. On a souvent l'idée selon laquelle les mathématiques consistent en des formules. Effectivement, le langage symbolique joue un rôle ; il est devenu indispensable dans le développement de la discipline. Cependant, aujourd'hui comme hier, ces formules ne prennent un sens qu'à l'intérieur d'un texte. Un article de mathématiques est d'abord, et avant tout, un texte ; c'est un récit. Les mathématiciens, et de façon plus générale les scientifiques, sont des écrivains. Les textes ont toujours une structure de récit, c'est-à-dire une structure linéaire. On raconte une sorte d'histoire, dans laquelle sont introduits progressivement des protagonistes qui vont interagir entre eux. Les mathématiques et les sciences sont une forme particulière de littérature.

Laurent Lafforgue

Table des matières

1 Logique et raisonnement	1
1 Assertions	1
1.1 Définition	1
1.2 Opérations	2
1.3 Quantificateurs	4
2 Types de raisonnement	5
2.1 Montrer une implication	5
2.2 Montrer une équivalence	5
2.3 Quelques méthodes de raisonnement	5
2 Calculs algébriques	7
1 L'ensemble ordonné \mathbb{R}	7
1.1 La relation \leq	7
1.2 La valeur absolue	9
2 Méthodes de résolution d'équations et d'inéquations dans \mathbb{R}	9
2.1 Remarques générales	9
2.2 Équations du second degré	10
2.3 Avec des valeurs absolues	10
2.4 Avec des racines	10
3 Sommes et produits	10
3.1 Notations	11
3.2 Cas particuliers	12
4 Coefficients binomiaux	12
3 Fonctions d'une variable réelle	15
1 Généralités	15
1.1 Définir une fonction	15
1.2 Opérations sur les fonctions	15
1.3 Décrire une fonction	16
4 Nombres complexes	18
1 Généralités	18
1.1 L'ensemble \mathbb{C}	18
1.2 Conjugaison complexe	19
1.3 Interprétation géométrique	20
1.4 Module	20
1.5 Équations du second degré	21
2 Nombres complexes et trigonométrie	21
2.1 Le cercle unité	21
2.2 L'exponentielle complexe	24

2.3	Racines de l'unité	25
2.4	Racines n ^{èmes}	26
3	Géométrie du plan	26
3.1	Alignment et orthogonalité	26
3.2	Transformations du plan	26
5	Fonctions usuelles	28
1	Logarithmes, exponentielles, puissances	28
1.1	Logarithmes	28
1.2	Exponentielles	29
1.3	Puissances	30
1.4	Croissances comparées	31
2	Fonctions trigonométriques	32
3	Fonctions hyperboliques	33
6	Ensembles et applications	35
1	Ensembles	35
1.1	Définition	35
1.2	Opérations	36
1.3	Produit cartésien	37
2	Applications	37
2.1	Définition	37
2.2	Injections, surjections, bijections	38
7	Arithmétique	40
1	Nombres premiers et divisibilité	40
2	PGCD et PPCM	43
3	Division euclidienne	43
8	Primitives	45
1	Primitives et intégrales	45
1.1	Primitives	45
1.2	Intégrales	46
1.3	Théorème fondamental de l'analyse	47
2	Calcul de primitives et d'intégrales	47
2.1	Calcul direct	48
2.2	Cas des fonctions trigonométriques	48
2.3	Cas des fractions rationnelles	48
2.4	Intégration par parties	48
2.5	Changement de variable	49
2.6	Cas des fonctions $x \mapsto \frac{1}{ax^2 + bx + c}$	50
9	Équations différentielles	51
1	Généralités	51
2	Équations différentielles linéaires d'ordre 1	52
2.1	À coefficients constants	52
2.2	À coefficients non constants	52
3	Équations différentielles linéaires d'ordre 2	54
3.1	Équation homogène	54
3.2	Solution particulière	55

10 Calcul matriciel et systèmes linéaires	58
1 Systèmes linéaires	58
1.1 Définitions	58
1.2 Opérations élémentaires	59
1.3 Algorithme du pivot de Gauss-Jordan	59
1.4 Résolution	60
1.5 Vers les matrices	61
2 Ensembles de matrices	62
2.1 Définition	62
2.2 Opérations matricielles	62
3 Matrices carrées	64
3.1 Puissances	64
3.2 Inversion	66
4 Opérations élémentaires	67
4.1 Matrices d'opérations élémentaires	67
4.2 Méthode du pivot	68
5 Transposition et symétrie	70
11 Suites numériques	72
1 Généralités	72
2 Suites récurrentes	73
2.1 Suites récurrentes linéaires d'ordre 1	73
2.2 Suites récurrentes linéaires d'ordre 2	73
2.3 Suites récurrentes générales d'ordre 1	74
3 Suites convergentes	74
3.1 Définition	74
3.2 Opérations sur les limites	76
3.3 Résultats d'existence d'une limite	76
4 Suites extraites	78
5 Cas des suites complexes	79
6 Relations de comparaison	80
6.1 Équivalents	80
6.2 Petits- <i>o</i>	81
6.3 Grands- <i>O</i>	81
12 Polynômes	83
1 L'ensemble $\mathbb{K}[X]$	83
1.1 Définitions	83
1.2 Opérations	84
2 Arithmétique dans $\mathbb{K}[X]$	85
2.1 Divisibilité	85
2.2 Division euclidienne	85
3 Dérivation	86
4 Racines et décomposition	87
4.1 Factorisation	87
4.2 Décomposition en facteurs irréductibles	89
4.3 Relations coefficients-racines	90

13 Limites et continuité	91
1 Limites et fonctions	91
1.1 Définitions	91
1.2 Ce qui ne change pas par rapport aux suites	92
1.3 Ce qui change (un peu) par rapport aux suites	94
2 Continuité en un point	94
3 Continuité sur un intervalle	95
4 Fonctions complexes	97
14 Dérivabilité	98
1 Nombre dérivé, fonction dérivée	98
1.1 Dérivabilité	98
1.2 Opérations sur les dérivées	99
2 Propriétés des fonctions dérivables	99
2.1 Dérivées et extrema	99
2.2 Conséquences du théorème des accroissements finis	100
3 Fonctions de classe C^n	102
4 Fonctions convexes	103
5 Fonctions complexes dérivables	104
15 Espaces vectoriels	105
1 Espaces vectoriels	105
1.1 Définition	105
1.2 Exemples fondamentaux	106
1.3 Combinaison linéaire	106
2 Sous-espaces vectoriels	106
2.1 Définition	106
2.2 Sous-espaces engendrés	107
2.3 Opérations sur les sous-espaces vectoriels	107
3 Familles de vecteurs	108
3.1 Familles génératrices	108
3.2 Familles libres	108
3.3 Bases	109
16 Applications linéaires	111
1 Généralités	111
1.1 Définitions	111
1.2 Opérations sur les applications linéaires	111
1.3 Image et noyau	112
2 Isomorphismes	113
3 Endomorphismes remarquables	113
3.1 Homothéties et rotations	113
3.2 Projecteurs	114
3.3 Symétries	114
17 Développements limités	115
1 Formules de Taylor	115
2 Développements limités	116
2.1 Définition et premières propriétés	116
2.2 Développements limités usuels	118
2.3 Opérations sur les développements limités	118

3	Utilisations des DL	120
3.1	Calcul de limites et d'équivalents	120
3.2	Position d'une courbe par rapport à sa tangente	120
3.3	Approximation locale de fonctions réciproques	121
4	Développements asymptotiques	122
18	Probabilités	126
1	Expériences aléatoires et univers	126
2	Espaces probabilisés	127
3	Probabilités conditionnelles	127
4	Événements indépendants	128
19	Dimension	130
1	Dimension d'un espace vectoriel	130
1.1	Définition	130
1.2	Propriétés fondamentales	131
2	Dimension et sous-espaces vectoriels	131
3	Rang	132
3.1	Rang d'une famille de vecteurs	132
3.2	Rang d'une application linéaire	132
20	Représentation matricielle et déterminants	134
1	Représentation matricielle	134
1.1	Définition	134
1.2	Propriétés	135
2	Changements de bases	135
2.1	Matrices de passage	135
2.2	Formules de changement de base	136
3	Rang d'une matrice	136
4	Déterminant d'une matrice	137
4.1	Définition	137
4.2	Propriétés	137
5	Calcul de déterminants	139
5.1	Par l'algorithme du pivot	139
5.2	Par développement par rapport à une ligne ou une colonne	139
6	Utilisations du déterminant	140
6.1	Déterminant d'un endomorphisme	140
6.2	Résolution de systèmes linéaires	140
21	Intégration	141
1	Définition	141
1.1	Fonctions en escalier	141
1.2	Fonctions continues	142
2	Sommes de Riemann	143
3	Théorème fondamental de l'analyse	143
4	Intégration des fonctions complexes	144

22 Variables aléatoires	145
1 Variables aléatoires	145
1.1 Définition	145
1.2 Espérance, variance, écart-type	146
1.3 Opérations sur les variables aléatoires	147
2 Lois usuelles	148
2.1 Loi uniforme	148
2.2 Loi de Bernoulli	148
2.3 Loi binomiale	148
3 Couples de variables aléatoires	149
3.1 Propriétés	149
3.2 Indépendance	149
23 Séries numériques	151
1 Généralités	151
1.1 Définition	151
1.2 Propriétés	152
1.3 Cas particuliers	152
2 Séries à termes positifs	152
2.1 Comparaison entre séries	153
2.2 Comparaison série-intégrale	154
2.3 Séries de Riemann	154
2.4 Développement décimal d'un réel	154
3 Convergence absolue	155
24 Espaces préhilbertiens	156
1 Produit scalaire	156
1.1 Définition	156
1.2 Norme	157
2 Orthogonalité	158
2.1 Vecteurs et sous-espace orthogonaux	158
2.2 Familles orthogonales	159
3 Espaces euclidiens	159
3.1 Bases orthonormées	159
3.2 Orthonormalisation	160
3.3 Projecteurs orthogonaux	160
4 Représentation matricielle du produit scalaire	161
25 Fonctions de deux variables	163
1 Généralités	163
1.1 Exemples	163
1.2 Ouverts de \mathbb{R}^2	164
2 Continuité	164
3 Dérivabilité	165

Notations usuelles

ALPHABET GREC

Lettre	minuscule	Majuscule	Correspondance
alpha	α	A	a
bêta	β	B	b
gamma	γ	Γ	g
delta	δ	Δ	d
epsilon	ϵ ou ε	E	e
zêta	ζ	Z	z
êta	η	H	ê
thêta	θ	Θ	th
iota	ι	I	i
kappa	κ	K	c ou k
lambda	λ	Λ	l
mu	μ	M	m
nu	ν	N	n
ksi ou xi	ξ	Ξ	x
omicron	\circ	O	o
pi	π ou ϖ	Π	p
rho	ρ	P	r
sigma	σ	Σ	s
tau	τ	T	t
upsilon	υ	Υ	y
phi	ϕ ou φ	Φ	ph
khi	χ	X	ch ou kh
psi	ψ	Ψ	ps
omega	ω	Ω	\hat{o}

SYMBOLES MATHÉMATIQUES

Symbole	Signification
\forall	« pour tout »
$\exists, \exists!$	« il existe », « il existe un unique »
\in	« appartient à »
\subset	« est inclus dans »
\cap, \cup	Intersection, union
\neg, \wedge, \vee	« non », « et », « ou »
$\{\dots\}, \emptyset$	Ensemble, ensemble vide
\Rightarrow	« implique », « si... alors... »
\Leftrightarrow	« équivaut à », « si et seulement si »
\int, \sum, \prod	Intégrale, somme, produit
∞	« infini »

ENSEMBLES DE NOMBRES

Symbole	Signification
\mathbb{N}	Entiers naturels
\mathbb{Z}	Entiers relatifs
\mathbb{D}	Nombres décimaux (à développement décimal fini)
\mathbb{Q}	Nombres rationnels (de la forme $\frac{p}{q}$, $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$)
\mathbb{R}	Nombres réels
\mathbb{C}	Nombres complexes
$\mathbb{N}^*, \mathbb{Z}^*, \mathbb{D}^*, \mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*$	Mêmes ensembles privés de 0
\mathbb{U}	Nombres complexes de module 1
\mathbb{U}_n	Racines n -ièmes de l'unité ($n \in \mathbb{N}^*$)

ENSEMBLES DE FONCTIONS

Symbole	Signification
$\mathcal{F}(A, B)$ ou B^A	Fonctions $A \rightarrow B$
$\mathbb{R}^{\mathbb{N}}$ ou $\mathbb{C}^{\mathbb{N}}$	Suites réelles ou complexes à indices dans \mathbb{N}
$\mathcal{C}(I, \mathbb{R})$ ou $\mathcal{C}(I, \mathbb{C})$	Fonctions continues sur l'intervalle I
$\mathcal{D}(I, \mathbb{R})$ ou $\mathcal{D}(I, \mathbb{C})$	Fonctions dérivables sur I
$\mathcal{C}^1(I, \mathbb{R})$ ou $\mathcal{C}^1(I, \mathbb{C})$	Fonctions dérivables à dérivée continue sur I
$\mathcal{C}^n(I, \mathbb{R})$ ou $\mathcal{C}^n(I, \mathbb{C})$	Fonctions n fois dérivables avec dérivée n ^{ème} continue
$\mathcal{C}^{\infty}(I, \mathbb{R})$ ou $\mathcal{C}^{\infty}(I, \mathbb{C})$	Fonctions n fois dérivables pour tout $n \in \mathbb{N}$

Formulaire de trigonométrie

Valeurs remarquables

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos \theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
$\sin \theta$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\tan \theta$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	non défini

Relations fondamentales

- $\cos^2(\theta) + \sin^2(\theta) = 1$,
- $\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$,
- $1 + \tan^2(\theta) = \frac{1}{\cos^2(\theta)}$.

Périodes et symétries

- $\cos(\theta + 2\pi) = \cos(\theta)$ et $\sin(\theta + 2\pi) = \sin(\theta)$,
- $\tan(\theta + \pi) = \tan(\theta)$,
- $\cos(-\theta) = \cos(\theta)$ et $\sin(-\theta) = -\sin(\theta)$,
- $\cos(\theta + \pi) = -\cos(\theta)$ et $\sin(\theta + \pi) = -\sin(\theta)$,
- $\cos(\pi - \theta) = -\cos(\theta)$ et $\sin(\pi - \theta) = \sin(\theta)$,
- $\cos\left(\theta + \frac{\pi}{2}\right) = -\sin(\theta)$ et $\sin\left(\theta + \frac{\pi}{2}\right) = \cos(\theta)$,
- $\cos\left(\frac{\pi}{2} - \theta\right) = \sin(\theta)$ et $\sin\left(\frac{\pi}{2} - \theta\right) = \cos(\theta)$,
- $\tan(-\theta) = -\tan(\theta)$ et $\tan\left(\frac{\pi}{2} - \theta\right) = \frac{1}{\tan(\theta)}$.

Résolution d'équations

- $(\cos a = \cos b) \iff (a = \pm b \bmod 2\pi)$,
- $(\sin a = \sin b) \iff (a = b \text{ ou } \pi - b \bmod 2\pi)$,
- $(\tan a = \tan b) \iff (a = b \bmod \pi)$,

Addition et duplication

- $\cos(a + b) = \cos(a)\cos(b) - \sin(a)\sin(b)$,
 - $\cos(a - b) = \cos(a)\cos(b) + \sin(a)\sin(b)$,
 - $\sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)$,
 - $\sin(a - b) = \sin(a)\cos(b) - \sin(b)\cos(a)$,
 - $\tan(a + b) = \frac{\tan(a) + \tan(b)}{1 - \tan(a)\tan(b)}$,
 - $\tan(a - b) = \frac{\tan(a) - \tan(b)}{1 + \tan(a)\tan(b)}$.
-
- $\cos 2a = \cos^2 a - \sin^2 a = 2\cos^2 a - 1 = 1 - 2\sin^2 a$,
 - $\sin 2a = 2\sin a \cos a$,
 - $\tan 2a = \frac{2\tan a}{1 - \tan^2 a}$.

Linéarisation et factorisation

- $\cos a \cos b = \frac{1}{2} (\cos(a - b) + \cos(a + b))$,
 - $\sin a \sin b = \frac{1}{2} (\cos(a - b) - \cos(a + b))$,
 - $\sin a \cos b = \frac{1}{2} (\sin(a - b) + \sin(a + b))$.
-
- $\cos^2 a = \frac{1}{2} (1 + \cos(2a))$,
 - $\sin^2 a = \frac{1}{2} (1 - \cos(2a))$,
 - $\cos a \sin a = \frac{1}{2} \sin(2a)$.

- $\cos a + \cos b = 2 \cos\left(\frac{a-b}{2}\right) \cos\left(\frac{a+b}{2}\right)$,
- $\cos a - \cos b = -2 \sin\left(\frac{a-b}{2}\right) \sin\left(\frac{a+b}{2}\right)$,
- $\sin a + \sin b = 2 \cos\left(\frac{a-b}{2}\right) \sin\left(\frac{a+b}{2}\right)$,
- $\sin a - \sin b = 2 \sin\left(\frac{a-b}{2}\right) \cos\left(\frac{a+b}{2}\right)$,

Analyse

- $\frac{1 - \cos x}{x} \xrightarrow{x \rightarrow 0} 0$,
- $\frac{\sin x}{x} \xrightarrow{x \rightarrow 0} 1$,
- $\sin' = \cos$,
- $\cos' = -\sin$,
- $\tan' = 1 + \tan^2 = \frac{1}{\cos^2}$

Dérivées et primitives usuelles

$f(x)$	D_f	$f'(x)$	D'_f
x^α ($\alpha \in \mathbb{R}$)	\mathbb{R}_+^*	$\alpha x^{\alpha-1}$	\mathbb{R}_+^*
x^n ($n \in \mathbb{N}^*$)	\mathbb{R}	nx^{n-1}	\mathbb{R}
$\frac{1}{x^n}$ ($n \in \mathbb{N}^*$)	\mathbb{R}^*	$-\frac{n}{x^{n+1}}$	\mathbb{R}^*
\sqrt{x}	\mathbb{R}_+	$\frac{1}{2\sqrt{x}}$	\mathbb{R}_+^*
e^x	\mathbb{R}	e^x	\mathbb{R}
$\ln x$	\mathbb{R}_+^*	$\frac{1}{x}$	\mathbb{R}_+^*
$\sin x$	\mathbb{R}	$\cos x$	\mathbb{R}
$\cos x$	\mathbb{R}	$-\sin x$	\mathbb{R}
$\tan x$	$\mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$	$1 + \tan^2 x = \frac{1}{\cos^2 x}$	$\mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$
$\arcsin x$	$[-1, 1]$	$\frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$
$\arccos x$	$[-1, 1]$	$-\frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$
$\arctan x$	\mathbb{R}	$\frac{1}{1+x^2}$	\mathbb{R}
$\operatorname{ch} x$	\mathbb{R}	$\operatorname{sh} x$	\mathbb{R}
$\operatorname{sh} x$	\mathbb{R}	$\operatorname{ch} x$	\mathbb{R}
$g \circ u(x)$	$\{x \mid u(x) \in D_g\}$	$u'(x)g' \circ u(x)$	$\{x \mid u'(x) \in D_{g'}\}$
$u^{-1}(x)$	$u(D_u)$	$\frac{1}{u' \circ u^{-1}(x)}$	$\left\{ x \mid \begin{array}{l} u^{-1}(x) \in D_{u'} \\ u' \circ u^{-1}(x) \neq 0 \end{array} \right\}$

On donne à chaque fois **une** primitive F de f . Les autres sont les $F + c$, avec $c \in \mathbb{R}$.

$f(x)$	D_f	$\int f$
x^α ($\alpha \in \mathbb{R}$)	\mathbb{R}_+^*	$\frac{x^{\alpha+1}}{\alpha+1}$
x^n ($n \in \mathbb{N}^*$)	\mathbb{R}	$\frac{x^{n+1}}{n+1}$
$\frac{1}{x^n}$ ($n \in \mathbb{N}^*, n \neq 1$)	\mathbb{R}^*	$-\frac{1}{n-1} \frac{1}{x^{n-1}}$
$\frac{1}{x}$	\mathbb{R}^*	$\ln x $
e^x	\mathbb{R}	e^x
$\cos x$	\mathbb{R}	$\sin x$
$\sin x$	\mathbb{R}	$-\cos x$
$\tan x$	$\mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$	$-\ln \cos x $
$\frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$	$\arcsin x$
$-\frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$	$\arccos x$
$\frac{1}{1+x^2}$	\mathbb{R}	$\arctan x$
$\operatorname{sh} x$	\mathbb{R}	$\operatorname{ch} x$
$\operatorname{ch} x$	\mathbb{R}	$\operatorname{sh} x$
$u'(x)g' \circ u(x)$	$\{x \mid u'(x) \in D_{g'}\}$	$g \circ u(x)$

Développements limités usuels

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^n}{n!} + \underset{x \rightarrow 0}{o}(x^n) \\
\operatorname{sh}(x) &= x + \frac{x^3}{6} + \frac{x^5}{120} + \cdots + \frac{x^{2k+1}}{(2k+1)!} + \underset{x \rightarrow 0}{o}(x^{2k+2}) \\
\operatorname{ch}(x) &= 1 + \frac{x^2}{2} + \frac{x^4}{24} + \cdots + \frac{x^{2k}}{(2k)!} + \underset{x \rightarrow 0}{o}(x^{2k+1})
\end{aligned}$$

$$\begin{aligned}
\frac{1}{1+x} &= 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + \underset{x \rightarrow 0}{o}(x^n) \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n-1} \frac{x^n}{n} + \underset{x \rightarrow 0}{o}(x^n) \quad (x > -1)
\end{aligned}$$

$$\begin{aligned}
\sin(x) &= x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots + (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \underset{x \rightarrow 0}{o}(x^{2k+2}) \\
\cos(x) &= 1 - \frac{x^2}{2} + \frac{x^4}{24} + \cdots + (-1)^k \frac{x^{2k}}{(2k)!} + \underset{x \rightarrow 0}{o}(x^{2k+1}) \\
\tan(x) &= x + \frac{x^3}{3} + \frac{2x^5}{15} + \underset{x \rightarrow 0}{o}(x^6)
\end{aligned}$$

$$\begin{aligned}
(1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{6} x^3 + \cdots \\
&\quad + \frac{\alpha(\alpha-1)(\alpha-2) \cdots (\alpha-n+1)}{n!} x^n + \underset{x \rightarrow 0}{o}(x^n) \quad (\alpha \in \mathbb{R}) \\
\sqrt{1+x} &= 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + \underset{x \rightarrow 0}{o}(x^3) \\
\sqrt[3]{1+x} &= 1 + \frac{1}{3}x - \frac{1}{9}x^2 + \frac{5}{81}x^3 + \underset{x \rightarrow 0}{o}(x^3)
\end{aligned}$$

$$\begin{aligned}
\arctan(x) &= x - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + (-1)^k \frac{x^{2k+1}}{2k+1} + \underset{x \rightarrow 0}{o}(x^{2k+2}) \\
\arcsin(x) &= x + \frac{x^3}{6} + \frac{3x^5}{40} + \cdots + \frac{(2k)!}{2^{2k}(k!)^2(2k+1)} x^{2k+1} + \underset{x \rightarrow 0}{o}(x^{2k+2})
\end{aligned}$$

Chapitre 1

Logique et raisonnement

Ce premier chapitre est consacré au fonctionnement des mathématiques en tant que science de la démonstration. Sans prétendre à l'exhaustivité, nous présentons ici les principaux outils de construction logique que nous utiliserons tout au long de l'année.

1 Assertions

1.1 Définition

Définition 1.1.1. Une *assertion* ou *proposition* est une phrase, qui peut être soit vraie, soit fausse ; c'est-à-dire qui possède une *véracité* ou encore une *valeur de vérité*.

Exemple 1.1.2. Parmi les assertions suivantes, certaines sont toujours vraies, certaines sont toujours fausses, et la véracité des autres dépend du contexte dans lequel elles sont énoncées :

- « 2 est un entier. »
- « Le triangle ABC est rectangle en A . »
- « Il a fait beau tous les jours de la semaine. »
- « $\exists n \in \mathbb{N}, 2n < n^2 + 1$. »
- « $\forall n \geq 3, \forall (x, y, z) \in (\mathbb{N}^*)^3, x^n + y^n \neq z^n$. » (*théorème de Fermat*)

Remarque 1.1.3.

- L'activité mathématique consiste à vérifier si des assertions sont vraies ou fausses. On met pour cela en œuvre une *démonstration*, c'est-à-dire une suite logique d'assertions établies comme vraies.
- Toute assertion est vraie ou fausse dans une *théorie* donnée (arithmétique, géométrie...), et sa véracité découle des assertions constitutives de cette théorie, admises sans démonstration, appelées *axiomes* ou *postulats*. Par exemple, un axiome de l'arithmétique est que « tout entier a un successeur. » Un postulat de la géométrie du plan est que « par un point donné, il passe une et une seule droite parallèle à une droite donnée. » Ce postulat ne convient pas contre pas à la géométrie sphérique.
- Dans toute théorie, il y a des assertions dont la véracité n'est pas encore connue, les *conjectures* ; par exemple, la *conjecture de Goldbach* en arithmétique affirmant que « tout entier pair est somme de deux nombres premiers » ou encore l'*hypothèse de Riemann*¹ en analyse complexe. Il existe enfin, inévitablement (d'après le *théorème de Gödel*), des assertions *indécidables*, c'est-à-dire dont la véracité *ne peut pas* être établie à partir des axiomes de la théorie considérée. Ainsi de l'*hypothèse du continu* en théorie des ensembles selon laquelle « il existe un infini intermédiaire entre celui des entiers et celui des réels ».

1. présentée par exemple dans cette vidéo de David Louapre : <https://www.youtube.com/watch?v=Kvcu1Wl-jhE>

1.2 Opérations

Définition 1.2.1. La *négation* d'une assertion P est l'assertion, notée $\neg P$ (« non P »), dont la valeur de vérité est contraire à celle de P .

On peut représenter les valeurs de vérité prises par $\neg P$ dans une *table de vérité*, où V correspond à « vrai » et F à « faux ».

P	$\neg P$
V	F
F	V

Exemple 1.2.2. Les négations des assertions de l'exemple 1.1.2 sont les assertions suivantes :

- « 2 n'est pas un entier. »
- « Le triangle ABC n'est pas rectangle en A . »
- « Il y a au moins un jour de la semaine où il n'a pas fait beau. »
- « $\forall n \in \mathbb{N}, 2n \geq n^2 + 1$. »
- « $\exists n \geq 3, \exists (x, y, z) \in (\mathbb{N}^*)^3, x^n + y^n = z^n$. »

Proposition 1.2.3. Soit P une assertion. On a : $\neg(\neg P) = P$.

Définition 1.2.4. La *conjonction* de deux assertions P et Q est l'assertion, notée $P \wedge Q$ (« P et Q »), vraie si et seulement si P est vraie **et** Q est vraie.

P	Q	$P \wedge Q$
V	V	V
V	F	F
F	V	F
F	F	F

Exemple 1.2.5. L'assertion $(x \geq 2) \wedge (x \leq 5)$ est l'assertion $(2 \leq x \leq 5)$.

Définition 1.2.6. La *disjonction* de deux assertions P et Q est l'assertion, notée $P \vee Q$ (« P ou Q »), vraie si et seulement si P est vraie **ou** Q est vraie.

P	Q	$P \vee Q$
V	V	V
V	F	V
F	V	V
F	F	F

Exemple 1.2.7. L'assertion $(x \geq 2) \vee (x \leq -2)$ est l'assertion $(|x| \geq 2)$.

Ces opérations se combinent de la façon suivante :

Proposition 1.2.8 (Lois de De Morgan). Soient P et Q des assertions.

- $\neg(P \wedge Q) = (\neg P) \vee (\neg Q)$,
- $\neg(P \vee Q) = (\neg P) \wedge (\neg Q)$.

Démonstration.

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$	$\neg P$	$\neg Q$	$(\neg P) \vee (\neg Q)$
V	V	V	F	F	F	F
V	F	F	V	F	V	V
F	V	F	V	V	F	V
F	F	F	V	V	V	V

P	Q	$P \vee Q$	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$(\neg P) \wedge (\neg Q)$
V	V	V	F	F	F	F
V	F	V	F	F	V	F
F	V	V	F	V	F	F
F	F	F	V	V	V	V

□

De plus, les conjonctions et disjonctions ont les propriétés suivantes :

Proposition 1.2.9. Soient P , Q et R des assertions. Les conjonctions et disjonctions sont :

- chacune associative :

$$(P \wedge Q) \wedge R = P \wedge (Q \wedge R),$$

$$(P \vee Q) \vee R = P \vee (Q \vee R),$$

- distributives l'une par rapport à l'autre :

$$P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R),$$

$$P \vee (Q \wedge R) = (P \vee Q) \wedge (P \vee R).$$

Démonstration. Avec des tables de vérité ! (Exercice)

□

On définit également les opérations suivantes :

Définition 1.2.10. L'*implication* de l'assertion P vers l'assertion Q est l'assertion, notée $P \Rightarrow Q$ (« P implique Q » ou « si P alors Q »), égale à $(\neg P) \vee Q$.

P	Q	$P \Rightarrow Q$
V	V	V
V	F	F
F	V	V
F	F	V

Remarque 1.2.11. Si $P \Rightarrow Q$ est vraie, on dit que P est une *condition suffisante* pour avoir Q , et que Q est une *condition nécessaire* pour avoir P .

Exemple 1.2.12. Soit x un réel ; notons P l'assertion « x est un entier » et Q l'assertion « $2x$ est un entier ». On a bien $P \Rightarrow Q$, et par exemple :

- si $x = 2$, alors P et Q sont vraies,
- si $x = \sqrt{2}$, alors P et Q sont fausses,
- si $x = \frac{3}{2}$, alors P est fausse et Q est vraie.

Proposition 1.2.13. Soient P et Q des assertions.

- $(P \Rightarrow Q) = ((\neg Q) \Rightarrow (\neg P))$,
- $\neg(P \Rightarrow Q) = P \wedge (\neg Q)$.

On dit que l'*implication* $(\neg Q) \Rightarrow (\neg P)$ est la *contraposée* de l'*implication* $P \Rightarrow Q$. D'après l'égalité ci-dessus, ces deux implications ont même véracité.

Démonstration.

- $(P \Rightarrow Q) = (\neg P) \vee Q = \neg(\neg Q) \vee (\neg P) = ((\neg Q) \Rightarrow (\neg P))$,
- D'après les lois de De Morgan, $\neg(P \Rightarrow Q) = \neg((\neg Q) \Rightarrow (\neg P)) = P \wedge (\neg Q)$.

□

Exemple 1.2.14. Soit P l'assertion « J'ai des pommes » et Q l'assertion « J'ai des fruits ». L'assertion $P \Rightarrow Q$ est vraie : « Si j'ai des pommes, alors j'ai des fruits ». La contraposée est l'assertion « Si je n'ai pas de fruits, alors je n'ai pas de pommes ».

Définition 1.2.15. L'équivalence de deux assertions P et Q est l'assertion, notée $P \Leftrightarrow Q$ (« P équivaut à Q »), égale à $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$.

P	Q	$P \Rightarrow Q$	$Q \Rightarrow P$	$P \Leftrightarrow Q$
V	V	V	V	V
V	F	F	V	F
F	V	V	F	F
F	F	V	V	V

Remarque 1.2.16. L'assertion $P \Leftrightarrow Q$ est donc vraie si et seulement si P et Q ont même valeur de vérité. On dit que P est une *condition nécessaire et suffisante* pour avoir Q , et inversement.

1.3 Quantificateurs

La plupart des assertions font intervenir un ou plusieurs *paramètres*. Par exemple, l'assertion P : « x^2 est positif » dépend du paramètre x . On peut alors la noter $P(x)$. Sa véracité dépend de la valeur du paramètre : par exemple, l'assertion $P(-3)$: « $(-3)^2$ est positif » est vraie, tandis que l'assertion $P(i)$ est fausse. Pour décrire la véracité d'une assertion à paramètres, on utilise deux *quantificateurs* :

Définition 1.3.1. Le *quantificateur universel*, noté \forall , est la locution « quel que soit » ou « pour tout ».

Exemple 1.3.2. L'assertion $P(x)$: « x^2 est positif » est vraie pour tout nombre réel x . En langage mathématique, on écrit : $\forall x \in \mathbb{R}, P(x)$.

Définition 1.3.3. Le *quantificateur existentiel*, noté \exists , est la locution « il existe ».

Exemple 1.3.4. Il existe un nombre complexe tel que l'assertion $P(x)$: « x^2 est positif » est fausse. En langage mathématique, on écrit : $\exists x \in \mathbb{C}, \neg P(x)$.

Remarque 1.3.5. On rencontre également parfois le quantificateur « il existe un unique », noté $\exists!$.

Remarque 1.3.6. Dans les exemples précédents, les assertions « $\forall x \in \mathbb{R}, P(x)$ » et « $\exists x \in \mathbb{C}, \neg P(x)$ » sont de *nouvelles assertions* formées à partir de l'assertion $P(x)$. **Elles ne dépendent plus du paramètre x** : celui-ci est devenu une *variable muette* interne à la nouvelle assertion.

Proposition 1.3.7. Soit $P(x)$ une assertion dépendant d'un paramètre x appartenant à un ensemble I . On a :

- $\neg(\forall x \in I, P(x)) = (\exists x \in I, \neg P(x))$,
- $\neg(\exists x \in I, P(x)) = (\forall x \in I, \neg P(x))$.

Exemple 1.3.8.

- $\neg(\forall x \in \mathbb{R}, x^2 \geq 0) = (\exists x \in \mathbb{R}, x^2 < 0)$.
- $\neg(\text{« Tous les canards sont verts }}) = \text{« Il existe un canard qui n'est pas vert. »}$

Remarque 1.3.9. Quand plusieurs quantificateurs interviennent, **leur ordre est fondamental**. Par exemple, étant donné une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$:

- dans l'assertion « $\exists A \in \mathbb{R}, \forall x \in \mathbb{R}, f(x) \leq A$ », la variable A est indépendante de la variable x , donc l'assertion est vraie si et seulement si f est majorée,
- dans l'assertion « $\forall x \in \mathbb{R}, \exists A \in \mathbb{R}, f(x) \leq A$ », la variable A dépend de la variable x , donc l'assertion est toujours vraie (en prenant $A = f(x)$).

Enfin, revenons sur les méthodes de démonstration correspondant aux assertions à paramètres. Pour démontrer :

- « $\forall x \in I, P(x)$ » : on *introduit* un x quelconque, et on montre $P(x)$: « Soit x dans I , montrons $P(x)$ ».
- « $\exists x \in I, P(x)$ » : on cherche un *exemple* x pour lequel $P(x)$ est vraie.
- « $\neg(\forall x \in I, P(x))$ » : on cherche un *contre-exemple* x pour lequel $P(x)$ est fausse.
- « $\neg(\exists x \in I, P(x))$ » : on *introduit* un x quelconque, et on montre $\neg P(x)$: « Soit x dans I , montrons $\neg P(x)$ ».

2 Types de raisonnement

La plupart des assertions mathématiques (théorèmes, propositions, lemmes, formules...) prennent la forme d'implications ou d'équivalences. On dispose de plusieurs méthodes pour tenter de les démontrer, dont l'efficacité dépendra du contexte :

2.1 Montrer une implication

Définition 2.1.1. Si une assertion se présente sous la forme d'une implication $P \Rightarrow Q$, on dit que P est l'*hypothèse* (ou la *prémissse*) de l'assertion, et que Q est sa *conclusion*.

Remarque 2.1.2. Lorsqu'on utilise un théorème T : $(P \Rightarrow Q)$, il est fondamental de bien faire apparaître ces différents éléments : « L'hypothèse P est vérifiée, donc, d'après le théorème T , la conclusion Q est vérifiée. »

Pour montrer une implication $P \Rightarrow Q$, on peut :

- la montrer directement : Dans ce cas, on suppose P , et on montre Q .

Exemple 2.1.3. $\forall n \in \mathbb{N}, (6 \text{ divise } n) \Rightarrow (3 \text{ divise } n)$.

- montrer sa contraposée (cf. la proposition 1.2.13) : on suppose $\neg Q$, et on montre $\neg P$.

Exemple 2.1.4. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ affine non constante : $\forall (x, y) \in \mathbb{R}^2, (x \neq y) \Rightarrow (f(x) \neq f(y))$.

- raisonner par l'absurde : Dans ce cas, on suppose P vraie et Q fausse, pour aboutir à une contradiction. (On montre donc $\neg(P \wedge \neg Q)$ (cf. la proposition 1.2.13)).

Exemple 2.1.5. $\forall x \in \mathbb{R}, (x^2 = 2) \Rightarrow (x \notin \mathbb{Q})$.

2.2 Montrer une équivalence

Pour montrer une équivalence $P \Leftrightarrow Q$, on peut :

- raisonner par équivalences : on construit une chaîne logique d'équivalences partant de P et aboutissant à Q .

Exemple 2.2.1. $\forall z \in \mathbb{C}, (z \in \mathbb{U}) \Leftrightarrow (\bar{z} = z^{-1})$.

- montrer une double implication : on montre $P \Rightarrow Q$ puis $Q \Rightarrow P$.

Exemple 2.2.2. $\forall (x, y) \in \mathbb{R}^2, (x^2 + y^2 = 0) \Leftrightarrow (x = y = 0)$.

- raisonner par disjonction de cas : dans ce cas, on montre $P \Rightarrow Q$ puis $\neg P \Rightarrow \neg Q$ (c'est-à-dire la contraposée de l'implication réciproque).

Exemple 2.2.3. $\forall n \in \mathbb{N}, (n \text{ est impair}) \Leftrightarrow (n^2 \text{ est impair})$.

2.3 Quelques méthodes de raisonnement

Certaines méthodes de raisonnement sont particulièrement efficaces dans certains cas :

1. *L'analyse-synthèse* pour montrer une équivalence $P \Leftrightarrow Q$ lorsque Q n'est pas connue :

- l'*analyse* consiste à supposer P et à en déduire une condition nécessaire Q (on montre donc $P \Rightarrow Q$),
- la *synthèse* consiste à vérifier que la condition Q est suffisante à P (on montre donc $Q \Rightarrow P$).

Exemple 2.3.1. Déterminer les fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ polynomiales de degré 2 telles que $f(0) = 1$, $f(1) = 0$ et $f(2) = 2$.

2. La *disjonction de cas* déjà vue plus haut,

Exemple 2.3.2. $\exists (a, b) \in (\mathbb{R} \setminus \mathbb{Q})^2$, $a^b \in \mathbb{Q}$.

3. La *récurrence* pour montrer une assertion à variable entière : « $\forall n \in \mathbb{N}$, $P(n)$ ». On distingue :

- la *récurrence simple* : On montre alors $P(0)$ (*initialisation*) et $\forall n \in \mathbb{N}$, $P(n) \Rightarrow P(n + 1)$ (*hérité*).

Exemple 2.3.3. Soit $(u_n)_{n \in \mathbb{N}}$ la suite définie par $u_0 = 2$ et $\forall n \in \mathbb{N}$, $u_{n+1} = 1 + \frac{1}{1 + u_n}$. Alors : $\forall n \in \mathbb{N}$, $1 \leq u_n \leq 2$.

- la *récurrence double* : On montre alors $P(0)$, $P(1)$ (*initialisation*) et $\forall n \in \mathbb{N}$, $(P(n) \text{ et } P(n + 1)) \Rightarrow P(n + 2)$ (*hérité*).

Exemple 2.3.4. Soit $(u_n)_{n \in \mathbb{N}}$ la *suite de Fibonacci* définie par $u_0 = 0$, $u_1 = 1$ et $\forall n \in \mathbb{N}$, $u_{n+2} = u_n + u_{n+1}$.

Alors : $\forall n \in \mathbb{N}$, $u_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right)$.

- la *récurrence forte* : On montre alors $P(0)$ (*initialisation*) et $\forall n \in \mathbb{N}$, $(\forall k \leq n, P(k)) \Rightarrow P(n + 1)$ (*hérité*).

Exemple 2.3.5. Soit $(u_n)_{n \in \mathbb{N}}$ la suite définie par $u_0 = 1$ et $\forall n \in \mathbb{N}$, $u_{n+1} = \sum_{k=0}^n u_k$. Alors : $\forall n \in \mathbb{N}^*$, $u_n = 2^{n-1}$.

Chapitre 2

Calculs algébriques

1 L'ensemble ordonné \mathbb{R}

Les principaux ensembles de nombres utilisés en mathématiques sont les suivants, représentés en inclusion (\subset) :

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{D} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

On rappelle les définitions des premiers d'entre eux :

- $\mathbb{N} = \{0, 1, 2, \dots\}$ est l'ensemble des *entiers naturels*,
- $\mathbb{Z} = \mathbb{N} \cup -\mathbb{N}$ est l'ensemble des *entiers relatifs*,
- $\mathbb{D} = \left\{ \frac{p}{10^q} \mid p \in \mathbb{Z}, q \in \mathbb{N} \right\}$ est l'ensemble des *nombres décimaux*,
- $\mathbb{Q} = \left\{ \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}^* \right\}$ est l'ensemble des *nombres rationnels*.

Historiquement, des signes d'utilisation des entiers naturels existent dès le Paléolithique. L'apparition des nombres rationnels est à peu près contemporaine de celle de l'écriture, au IV^{ème} millénaire av. J.-C. en Mésopotamie. Celle des nombres décimaux correspond à l'utilisation de la numération décimale, à partir du III^{ème} millénaire av. J.-C. en Égypte. Les nombres négatifs ne seront utilisés qu'à partir du II^{ème} siècle av. J.-C. en Chine, et très progressivement en Europe à partir du XV^{ème} siècle de notre ère.

La découverte des nombres *irrationnels* se situe autour du V^{ème} av. J.-C. en Grèce, notamment avec le cas de $\sqrt{2}$. Une construction formelle de \mathbb{R} , ensemble des *nombres réels*, regroupant nombres rationnels et irrationnels, est obtenue par Richard Dedekind en 1872. La construction de \mathbb{R} est délicate et hors programme. On mentionne néanmoins la propriété suivante :

Proposition 1.0.1. *L'ensemble \mathbb{Q} est dense dans \mathbb{R} , c'est-à-dire qu'entre deux nombres réels distincts il y a toujours un nombre rationnel.*

En langage mathématique, avec les quantificateurs \forall ("pour tout") et \exists ("il existe"), le symbole \in ("qui appartient à") et l'abréviation t.q. ("tel que"), cette propriété peut s'écrire ainsi :

$$\forall(a, b) \in \mathbb{R}^2 \text{ t.q. } a < b, \exists r \in \mathbb{Q}, a < r < b.$$

Les ensembles \mathbb{D} et $\mathbb{R} \setminus \mathbb{Q}$ (l'ensemble des nombres irrationnels) sont également denses dans \mathbb{R} .

Enfin, l'ensemble des *nombres complexes* \mathbb{C} sera étudié dans un chapitre ultérieur.

1.1 La relation \leq

Une des caractéristiques majeures de l'ensemble \mathbb{R} est d'être un ensemble *ordonné*, au sens suivant :

Définition 1.1.1. Une relation binaire (=reliant deux éléments) \triangleleft sur un ensemble E est une *relation d'ordre* si elle est :

- *réflexive* : $\forall x \in E, x \triangleleft x$,
- *antisymétrique* : $\forall (x, y) \in E^2$, si $(x \triangleleft y \text{ et } y \triangleleft x)$, alors $x = y$,
- *transitive* : $\forall (x, y, z) \in E^3$, si $(x \triangleleft y \text{ et } y \triangleleft z)$, alors $x \triangleleft z$.

On peut facilement vérifier que la relation \leq est une relation d'ordre sur \mathbb{R} (contrairement à la relation $<$, qui n'est pas réflexive !). Cet ordre est *total*, c'est-à-dire que tout couple de réels est lié par \leq : $\forall (x, y) \in \mathbb{R}^2, x \leq y \text{ ou } y \leq x$. Autrement dit :

Proposition 1.1.2. *L'ensemble (\mathbb{R}, \leq) est totalement ordonné.*

Un autre exemple de relation d'ordre est donné sur l'ensemble $\mathcal{P}(\mathbb{N})$, ensemble des parties (=sous-ensembles) de \mathbb{N} , par la relation d'inclusion \subset . Cette relation n'est pas totale : par exemple, les éléments $2\mathbb{N}$ (le sous-ensemble des entiers pairs) et $2\mathbb{N} + 1$ (le sous-ensemble des entiers impairs) ne sont pas liés par \subset . On dit que $(\mathcal{P}(\mathbb{N}), \subset)$ est *partiellement ordonné*.

Attention : La manipulation de la relation \leq est délicate. Comme indiqué dans la proposition ci-dessous, la multiplication d'inégalités ne fonctionne bien que lorsque les termes sont **positifs**. Il ne faut **en aucun cas** soustraire ou diviser deux inégalités entre elles.

Proposition 1.1.3 (Opérations sur \leq).

1. $\forall (a, b, c, d) \in \mathbb{R}^4$, si $(a \leq b \text{ et } c \leq d)$, alors $a + b \leq c + d$.
2. $\forall (a, b) \in \mathbb{R}^2$, si $a \leq b$, alors $-b \leq -a$.
3. $\forall (a, b, c, d) \in \mathbb{R}_+^4$, si $(a \leq b \text{ et } c \leq d)$, alors $ab \leq cd$.
4. $\forall (a, b) \in \mathbb{R}^2, \forall \lambda \in \mathbb{R}_+$, si $a \leq b$, alors $\lambda a \leq \lambda b$.
5. $\forall (a, b) \in (\mathbb{R}_+^*)^2$, si $a \leq b$, alors $\frac{1}{b} \leq \frac{1}{a}$.

La structure donnée à \mathbb{R} par la relation \leq met en avant une classe particulière de parties de \mathbb{R} , les *intervalles* :

Définition 1.1.4. Pour (a, b) dans \mathbb{R}^2 tel que $a < b$, on note $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$.

On dit qu'une partie I de \mathbb{R} est un *intervalle* si : $\forall (a, b) \in I^2$ t.q. $a < b$, $[a, b] \subset I$.

Autrement dit, les intervalles sont les parties "d'un seul tenant", ou *convexes*, de \mathbb{R} . On peut les décrire explicitement :

Proposition 1.1.5. *Tout intervalle est d'une des formes suivantes, où a et b désignent des réels tels que $a < b$:*

- \emptyset (l'ensemble vide, qui ne contient aucun élément),
- $\{a\}$,
- $[a, b]$,
- $[a, b[= \{x \in \mathbb{R} \mid a \leq x < b\}$,
- $]a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$,
- $]a, b[= \{x \in \mathbb{R} \mid a < x < b\}$,
- \mathbb{R} .
- $[a, +\infty[= \{x \in \mathbb{R} \mid a \leq x\}$,
- $]a, +\infty[= \{x \in \mathbb{R} \mid a < x\}$,
- $] - \infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$,
- $] - \infty, b[= \{x \in \mathbb{R} \mid x < b\}$,

Définition 1.1.6. Soient I et J des intervalles. On note :

- $I \cap J = \{x \in \mathbb{R} \mid x \in I \text{ et } x \in J\}$ l'*intersection* de I et J ,
- $I \cup J = \{x \in \mathbb{R} \mid x \in I \text{ ou } x \in J\}$ l'*union* de I et J .

À noter que l'intersection de deux intervalles est toujours un intervalle, ce qui n'est pas le cas pour l'union (par exemple, l'union de $[0, 1]$ et $[2, 3]$ n'est pas un intervalle).

1.2 La valeur absolue

La structure donnée à \mathbb{R} par la relation \leq nous amène à nous intéresser à la *distance* entre deux réels a et b : celle-ci vaut $b - a$ si $a \leq b$ et $a - b$ si $a \geq b$. Plus synthétiquement, elle vaut toujours $|b - a|$ en utilisant la notation suivante :

Définition 1.2.1. Pour tout réel x , on appelle *valeur absolue* de x , et on note $|x|$, la distance à 0 de x :

$$\forall x \in \mathbb{R}, |x| = \begin{cases} x & \text{si } x \geq 0, \\ -x & \text{si } x \leq 0. \end{cases}$$

De façon équivalente, on peut écrire $|x| = \max(x, -x)$, où $\forall (a, b) \in \mathbb{R}^2$, $\max(a, b) = \begin{cases} a & \text{si } a \geq b, \\ b & \text{sinon.} \end{cases}$.

On peut également décrire la valeur absolue en utilisant la racine carrée, comme ci-dessous ; on rappelle que pour tout réel *positif* a , la *racine carrée de a* , notée \sqrt{a} , est l'unique solution *positive* de l'équation $x^2 = a$.

Proposition 1.2.2. $\forall x \in \mathbb{R}$, $|x| = \sqrt{x^2}$.

Proposition 1.2.3. $\forall (x, M) \in \mathbb{R} \times \mathbb{R}_+$, $(|x| \leq M) \Leftrightarrow (-M \leq x \leq M) \Leftrightarrow (x \in [-M, M])$.

Remarque 1.2.4. On a donc : $\forall (x, a, \varepsilon) \in \mathbb{R}^2 \times \mathbb{R}_+$, $(|x - a| \leq \varepsilon) \Leftrightarrow (x \in [a - \varepsilon, a + \varepsilon])$.

On dit alors que x est ε -proche de a .

Proposition 1.2.5 (Propriétés de la valeur absolue).

1. $\forall x \in \mathbb{R}$, $|x| \geq 0$. *De plus, si $|x| = 0$, alors $x = 0$ (on dit que la valeur absolue est définie positive).*
2. $\forall x \in \mathbb{R}$, $-|x| \leq x \leq |x|$.
3. $\forall (x, y) \in \mathbb{R}^2$, $|xy| = |x| \cdot |y|$ (*la valeur absolue est multiplicative*).
4. Inégalité triangulaire : $\forall (x, y) \in \mathbb{R}^2$, $|x + y| \leq |x| + |y|$ (*la valeur absolue est sous-additive*).
5. Inégalité triangulaire renversée : $\forall (x, y) \in \mathbb{R}^2$, $||x| - |y|| \leq |x - y|$.

Démonstration.

1. à 3. : exercice !

4. Soit (x, y) dans \mathbb{R}^2 . D'après 3., $-|x| \leq x \leq |x|$ et $-|y| \leq y \leq |y|$, donc $-|x| - |y| \leq x + y \leq |x| + |y|$. Donc, d'après la proposition précédente, $|x + y| \leq |x| + |y|$.

5. : Soit (x, y) dans \mathbb{R}^2 . D'après l'inégalité triangulaire appliquée à y et $z = x - y$, on a $|y + z| \leq |y| + |z|$, c'est-à-dire $|x| \leq |y| + |x - y|$, donc $|x| - |y| \leq |x - y|$. Par symétrie, on a également $|y| - |x| \leq |y - x| = |x - y|$, donc $||x| - |y|| = \max(|x| - |y|, |y| - |x|) \leq |x - y|$.

□

Les équivalences suivantes nous seront très utiles en pratique :

- Proposition 1.2.6.** 1. $\forall (a, b) \in \mathbb{R}^2$, $|a| = |b| \Leftrightarrow a^2 = b^2$.
2. $\forall (a, b) \in \mathbb{R}^2$, $|a| \leq |b| \Leftrightarrow a^2 \leq b^2$.

2 Méthodes de résolution d'équations et d'inéquations dans \mathbb{R}

2.1 Remarques générales

Définition 2.1.1. Soit $k \in \mathbb{N}^*$. Une *équation* à k inconnues est une égalité mathématique faisant intervenir k variables x_1, \dots, x_k .

Une *solution* d'une équation est un k -uplet (x_1, \dots, x_k) pour lequel l'équation est vraie.

Résoudre une équation consiste à déterminer l'ensemble de ses solutions.

Une inégalité du même type est une *inéquation*.

Remarque 2.1.2. Une équation est donc une *assertion* dont la valeur de vérité dépend de la valeur d'un ou de plusieurs paramètres.

Définition 2.1.3. Le *domaine de définition* d'une équation est l'ensemble des valeurs des variables pour lequel l'équation a un sens.

2.2 Équations du second degré

Deux méthodes voisines permettent de résoudre les équations du second degré :

- la mise sous forme canonique,
- la recherche du discriminant, puis l'application de la formule idoine.

Remarque 2.2.1. On prendra garde à distinguer les notions de *solution*, de *racine* et de *point d'annulation* d'une part; d'*équation polynomiale*, de *polynôme* et de *fonction polynomiale* d'autre part.

2.3 Avec des valeurs absolues

Exemple 2.3.1. $|x - 1| \leq |2x - 3|$

- Première méthode : Distinguer les cas

On représente les différents cas possibles pour les valeurs absolues dans un tableau :

x	$] -\infty, 1]$	$\left[1, \frac{3}{2} \right]$	$\left[\frac{3}{2}, +\infty \right[$
$ x - 1 $	$1 - x$	$x - 1$	$x - 1$
$ 2x - 3 $	$3 - 2x$	$3 - 2x$	$2x - 3$
$ x - 1 \leq 2x - 3 $	$1 - x \leq 3 - 2x$ $\Leftrightarrow x \leq 2$	$x - 1 \leq 3 - 2x$ $\Leftrightarrow x \leq \frac{4}{3}$	$x - 1 \leq 2x - 3$ $\Leftrightarrow x \geq 2$
S	$S_1 =] -\infty, 1]$	$S_2 = \left[1, \frac{4}{3} \right]$	$S_3 = [2, +\infty[$

$$\text{donc } S = S_1 \cup S_2 \cup S_3 = \left] -\infty, \frac{4}{3} \right] \cup [2, +\infty[.$$

- Deuxième méthode : Élever au carré

$$\text{On a : } |x - 1| \leq |2x - 3| \iff (x - 1)^2 \leq (2x - 3)^2 \iff 3x^2 - 10x + 8 \geq 0 \iff x \in \left] -\infty, \frac{4}{3} \right] \cup [2, +\infty[.$$

2.4 Avec des racines

Exemple 2.4.1. $\sqrt{x - 1} \leq \sqrt{2x + 5}$

- Déterminer le domaine de définition de l'équation

La fonction racine étant définie sur \mathbb{R}_+ , l'équation est définie pour x tel que $x - 1 \geq 0$ et $2x + 5 \geq 0$, donc $D = [1, +\infty[$.

- Élever au carré

Pour $x \in D$, on a $\sqrt{x - 1} \leq \sqrt{2x + 5} \iff x - 1 \leq 2x + 5 \iff x \geq -6$. Donc $S = [-6, +\infty[\cap D = [1, +\infty[$.

3 Sommes et produits

On introduit ici les notations \sum et \prod .

3.1 Notations

Soit (n_0, n) dans \mathbb{N}^2 tel que $n_0 \leq n$. On note $\llbracket n_0, n \rrbracket = \{k \in \mathbb{N} \mid n_0 \leq k \leq n\}$.

Définition 3.1.1. Soit $(a_k)_{k \in \llbracket n_0, n \rrbracket} = (a_{n_0}, a_{n_0+1}, \dots, a_n)$ une suite finie de réels (ou de complexes). On note $\sum_{k=n_0}^n a_k$ ou $\sum_{k \in \llbracket n_0, n \rrbracket} a_k$ la somme des éléments de cette suite :

$$\sum_{k=n_0}^n a_k = \sum_{k \in \llbracket n_0, n \rrbracket} a_k = a_{n_0} + a_{n_0+1} + \dots + a_n.$$

Exemple 3.1.2. $\sum_{k=2}^5 k^3 = 2^3 + 3^3 + 4^3 + 5^3 = 224$.

Définition 3.1.3. Avec les mêmes notations, on note $\prod_{k=n_0}^n a_k$ ou $\prod_{k \in \llbracket n_0, n \rrbracket} a_k$ le produit des éléments de A :

$$\prod_{k=n_0}^n a_k = \prod_{k \in \llbracket n_0, n \rrbracket} a_k = a_{n_0} \times a_{n_0+1} \times \dots \times a_n.$$

Remarque 3.1.4. Remplacer la variable k par une autre lettre (ou n'importe quel symbole) ne change rien au résultat de la somme (ou du produit) : c'est une *variable muette*.

$$\prod_{k=2}^5 k^3 = \prod_{j=2}^5 j^3 = \prod_{\alpha=2}^5 \alpha^3 = 2^3 \times 3^3 \times 4^3 \times 5^3 = 1\,728\,000.$$

Proposition 3.1.5. Soit c dans \mathbb{R} . On a $\sum_{k=n_0}^n c = (n - n_0 + 1)c$. En particulier, $\sum_{k=1}^n c = nc$.

De même, $\prod_{k=n_0}^n c = c^{n-n_0+1}$.

Proposition 3.1.6 (Changement d'indice). On a $\sum_{k=n_0}^n a_k = \sum_{p=0}^{n-n_0} a_{n_0+p}$.

Proposition 3.1.7 (Linéarité). Soient $(a_k)_{k \in \llbracket n_0, n \rrbracket}$ et $(b_k)_{k \in \llbracket n_0, n \rrbracket}$ deux suites de réels (ou de complexes).

$$1. \quad \sum_{k=n_0}^n a_k + b_k = \sum_{k=n_0}^n a_k + \sum_{k=n_0}^n b_k.$$

$$2. \quad \forall \lambda \in \mathbb{C}, \quad \sum_{k=n_0}^n \lambda a_k = \lambda \sum_{k=n_0}^n a_k.$$

$$\text{De même, } \prod_{k=n_0}^n a_k \cdot b_k = \left(\prod_{k=n_0}^n a_k \right) \cdot \left(\prod_{k=n_0}^n b_k \right).$$

Proposition 3.1.8 (Sommation par paquets). Soient A et B deux sous-ensembles de $\llbracket n_0, n \rrbracket$ tels que $A \cup B = \llbracket n_0, n \rrbracket$ et $A \cap B = \emptyset$ (on dit que A et B partitionnent $\llbracket n_0, n \rrbracket$). Alors $\sum_{k=n_0}^n a_k = \sum_{k \in A} a_k + \sum_{k \in B} a_k$.

Exemple 3.1.9. $\sum_{k=0}^{2n+1} (-1)^k k$.

3.2 Cas particuliers

- Sommes télescopiques

Proposition 3.2.1. Soit $(u_k)_{k \in \llbracket n_0, n+1 \rrbracket}$ une suite de réels (ou de complexes). On a :

$$\sum_{k=n_0}^n (u_{k+1} - u_k) = u_{n+1} - u_0.$$

De même, si (u_k) ne s'annule pas : $\prod_{k=n_0}^n \frac{u_{k+1}}{u_k} = \frac{u_{n+1}}{u_0}$.

Exemple 3.2.2. $\sum_{k=1}^n \frac{1}{k(k+1)}, \prod_{k=1}^n \left(1 + \frac{1}{k}\right)$.

- Sommes arithmétiques

Proposition 3.2.3. Lorsque $(a_k)_{k \in \llbracket n_0, n \rrbracket}$ est une suite arithmétique de raison r (c'est-à-dire : $\forall k \in \llbracket n_0, n-1 \rrbracket, a_{k+1} = a_k + r$), on a : $\sum_{k=n_0}^n a_k = (n - n_0 + 1) \frac{a_{n_0} + a_n}{2}$.

- Sommes géométriques

Proposition 3.2.4. Lorsque $(a_k)_{k \in \llbracket n_0, n \rrbracket}$ est une suite géométrique de raison q (c'est-à-dire : $\forall k \in \llbracket n_0, n-1 \rrbracket, a_{k+1} = q \times a_k$), on a : $\sum_{k=n_0}^n a_k = a_{n_0} \frac{1 - q^{n-n_0+1}}{1 - q}$.

Remarque 3.2.5. Il n'est pas forcément conseillé d'apprendre ces formules telles quelles : les notations changent à chaque exercice ! Il est par contre fondamental de bien les comprendre et de savoir les mettre en pratique. On peut par exemple retenir les propositions 3.2.3 et 3.2.4 sous la forme suivante :

Somme des termes d'une suite arithmétique = (Nombre de termes) \times (Moyenne des termes extrémaux).

Somme des termes d'une suite géométrique = (Premier terme) \times $\frac{1 - \text{raison}^{\text{Nombre de termes}}}{1 - \text{raison}}$.

4 Coefficients binomiaux

On rappelle que pour n dans \mathbb{N}^* , $n!$ (« factorielle n ») est défini par $n! = \prod_{k=1}^n k = 1 \times 2 \times \dots \times n$. Par convention, on pose également $0! = 1$. On peut également utiliser la définition récursive :

$$\begin{cases} \forall n \geq 1, n! &= n \times (n-1)! \\ 0! &= 1. \end{cases}$$

Définition 4.0.1. Pour (n, p) dans \mathbb{N}^2 , le coefficient binomial $\binom{n}{p}$ (« p parmi n ») est défini par :

$$\binom{n}{p} = \begin{cases} \frac{n!}{p!(n-p)!} & \text{si } p \leq n, \\ 0 & \text{si } p > n. \end{cases}$$

Remarque 4.0.2. Il s'agit du nombre de parties à p éléments d'un ensemble à n éléments.

En particulier : $\forall n \in \mathbb{N}$, $\binom{n}{0} = \binom{n}{n} = 1$, et $\binom{n}{1} = \binom{n}{n-1} = n$.

Les coefficients binomiaux ont la propriété remarquable suivante :

Proposition 4.0.3 (Formule du triangle de Pascal). $\forall (n, p) \in \mathbb{N}^2$ tel que $1 \leq p \leq n$, $\binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}$.

Démonstration. Le calcul est direct :

$$\binom{n-1}{p-1} + \binom{n-1}{p} = \frac{(n-1)!}{(p-1)!(n-p)!} + \frac{(n-1)!}{p!(n-p-1)!} = \frac{p(n-1)! + (n-p)(n-1)!}{p!(n-p)!} = \frac{n!}{p!(n-p)!} = \binom{n}{p}.$$

□

Cette propriété a deux conséquences :

- Pour n dans \mathbb{N}^* , les coefficients binomiaux de rang n (les $\binom{n}{k}$ où $1 \leq k \leq n$) peuvent être obtenus à partir des coefficients binomiaux de rang $n-1$. Il est donc judicieux de les représenter sous la forme suivante, dite *triangle de Pascal*, où chaque coefficient à partir du rang 2 est obtenu comme la somme des deux coefficients qui le surplombent :

$$\begin{array}{ccccccccc}
 & & \binom{0}{0} & = & 1 & & & & \\
 & & \binom{1}{0} & = & 1 & & \binom{1}{1} & = & 1 \\
 & & \binom{2}{0} & = & 1 & & \binom{2}{1} & = & \binom{1}{0} + \binom{1}{1} = 2 & & \binom{2}{2} & = & 1 \\
 \binom{3}{0} & = & 1 & & \binom{3}{1} & = & \binom{2}{0} + \binom{2}{1} = 3 & & \binom{3}{2} & = & \binom{2}{1} + \binom{2}{2} = 3 & & \binom{3}{3} & = & 1 \\
 & & \binom{4}{0} & = & 1 & & \binom{4}{1} & = & 4 & & \binom{4}{2} & = & 6 & & \binom{4}{3} & = & 4 & & \binom{4}{4} & = & 1 \\
 \binom{5}{0} & = & 1 & & \binom{5}{1} & = & 5 & & \binom{5}{2} & = & 10 & & \binom{5}{3} & = & 10 & & \binom{5}{4} & = & 5 & & \binom{5}{5} & = & 1 \\
 & \dots
 \end{array}$$

- Les coefficients binomiaux interviennent dans la formule fondamentale suivante :

Théorème 4.0.4 (Formule du binôme de Newton).

$$\forall n \in \mathbb{N}, \forall (a, b) \in \mathbb{C}^2, \quad (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Démonstration. La preuve se fait par récurrence : Soit $(a, b) \in \mathbb{C}^2$. Pour n dans \mathbb{N} , notons P_n la proposition :

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

La proposition P_0 : $(a+b)^0 = \binom{0}{0} a^0 b^0$ est trivialement vraie.

Soit n dans \mathbb{N} , supposons P_n . Alors :

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= (a+b) \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \quad \text{par hypothèse de récurrence} \\
&= \sum_{k=0}^n \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^n \binom{n}{k} a^k b^{n-k+1} \quad \text{par distribution} \\
&= \sum_{l=1}^{n+1} \binom{n}{l-1} a^l b^{n-l+1} + \sum_{k=0}^n \binom{n}{k} a^k b^{n-k+1} \quad \text{par changement d'indice } l = k + 1 \\
&= \sum_{l=1}^n \binom{n}{l-1} a^l b^{n-l+1} + \binom{n}{n} a^{n+1} + \sum_{k=1}^n \binom{n}{k} a^k b^{n-k+1} + \binom{n}{0} b^{n+1} \quad \text{par séparation} \\
&= \sum_{k=1}^n \left[\binom{n}{k-1} + \binom{n}{k} \right] a^k b^{n-k+1} + a^{n+1} + b^{n+1} \quad \text{par factorisation} \\
&= \sum_{k=1}^n \binom{n+1}{k} a^k b^{n-k+1} + \binom{n+1}{n+1} a^{n+1} + \binom{n+1}{0} b^{n+1} \quad \text{d'après la formule du triangle de Pascal} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n-k+1} \quad \text{par regroupement.}
\end{aligned}$$

On obtient P_{n+1} .

On a donc P_0 et, $\forall n \in \mathbb{N}$, $P_n \implies P_{n+1}$. Donc par récurrence, P_n est vraie pour tout n dans \mathbb{N} . \square

Exemples 4.0.5.

- $\sum_{k=0}^n \binom{n}{k} = \sum_{k=0}^n \binom{n}{k} 1^k 1^{n-k} = (1+1)^n = 2^n$,
- $\sum_{k=0}^n (-1)^k \binom{n}{k} = \sum_{k=0}^n \binom{n}{k} (-1)^k 1^{n-k} = (-1+1)^n = 0$.

Chapitre 3

Fonctions d'une variable réelle

1 Généralités

1.1 Définir une fonction

Soient E et F deux parties de \mathbb{R} non vides. On rappelle que l'*ensemble produit* de E par F est $E \times F = \{(x, y) \mid x \in E, y \in F\}$. On note également « $\exists!$ » le quantificateur « il existe un unique ».

Définition 1.1.1. Une *fonction* f de E dans F est une relation entre E et F qui, à tout élément x de E , associe un élément y de F .

De manière équivalente, c'est la donnée d'une partie Γ de $E \times F$ telle que :

$$\forall x \in E, \exists! y \in F, (x, y) \in \Gamma.$$

On note alors $y = f(x)$.

- On note $f : \begin{cases} E & \rightarrow F \\ x & \mapsto f(x) \end{cases}$
- E est l'*ensemble de départ* de la fonction f et F est l'*ensemble d'arrivée*
- Γ est le *graphe* de la fonction f
- y est l'*image* de x par f , ou encore la *valeur* prise par f en x ,
- x est un *antécédent* de y par f
- L'ensemble $f(E) = \{f(x) \mid x \in E\}$ est l'*ensemble image* de E par f . C'est un sous-ensemble de F .

Remarques 1.1.2.

- Tout élément de E a une et une seule image par f ; mais un élément de F peut avoir un, plusieurs ou aucun antécédent par f .
- Si la correspondance $x \mapsto f(x)$ est donnée sans préciser E , on appelle *domaine de définition* de f l'ensemble de départ maximal (dans \mathbb{R}) de f .
Si E est précisé et que $E \subset D_f$, on dit que f est *bien définie*.

Définition 1.1.3. La *représentation graphique* ou la *courbe* d'une fonction f est son tracé dans le plan (Oxy) . L'*équation cartésienne* de cette courbe est l'équation $y = f(x)$.

1.2 Opérations sur les fonctions

Définition 1.2.1. On définit les fonctions standard suivantes :

- la *fonction nulle* $0_E : \begin{cases} E & \rightarrow \mathbb{R} \\ x & \mapsto 0 \end{cases}$,

- la *fonction identité* $\text{Id}_E : \begin{cases} E & \rightarrow & E \\ x & \mapsto & x \end{cases}$,
- la *fonction indicatrice de E* : $\mathbb{1}_E : \begin{cases} \mathbb{R} & \rightarrow & \{0, 1\} \\ x & \mapsto & 1 \text{ si } x \in E, 0 \text{ sinon} \end{cases}$.

Soit E une partie de \mathbb{R} , et soient f et g deux fonctions de E dans \mathbb{R} . On note $f + g$ la *somme* de f et g : $f + g : \begin{cases} E & \rightarrow & \mathbb{R} \\ x & \mapsto & f(x) + g(x) \end{cases}$. De même, on note fg le *produit* de f et g : $fg : \begin{cases} E & \rightarrow & \mathbb{R} \\ x & \mapsto & f(x) \cdot g(x) \end{cases}$. On définit de la même manière le produit λf pour λ dans \mathbb{R} et le quotient $\frac{f}{g}$ si g ne prend pas de valeur nulle : $\forall x \in E, g(x) \neq 0$.

Remarque 1.2.2. Le produit de deux fonctions peut être nul sans que l'une des deux soit nulle : par exemple, $\mathbb{1}_{\mathbb{R}_-} \mathbb{1}_{\mathbb{R}_+^*} = 0_{\mathbb{R}}$.

Définition 1.2.3. Soient $f : E \rightarrow \mathbb{R}$ et $g : F \rightarrow \mathbb{R}$ deux fonctions telles que $f(E) \subset F$. On appelle *composée de f par g* et on note $g \circ f$ la fonction $g \circ f : \begin{cases} E & \rightarrow & \mathbb{R} \\ x & \mapsto & g(f(x)) \end{cases}$.

Remarque 1.2.4. Lorsqu'elles existent, en général $f \circ g \neq g \circ f$. Par exemple, pour $f : x \mapsto x + 1$ et $g : x \mapsto x^2$, $f \circ g : x \mapsto x^2 + 1$ et $g \circ f : x \mapsto (x + 1)^2 = x^2 + 2x + 1$. On dit que la composition *n'est pas commutative*. En revanche, si trois fonctions $f : A \rightarrow \mathbb{R}$, $g : B \rightarrow \mathbb{R}$ et $h : C \rightarrow \mathbb{R}$ vérifient $h(C) \subset B$ et $g(B) \subset A$, alors $(f \circ g) \circ h = f \circ (g \circ h)$. On dit que la composition est *associative*.

Proposition 1.2.5. Soit $f : E \rightarrow F$. Alors $\text{Id}_F \circ f = f \circ \text{Id}_E = f$.

Démonstration. Immédiat. □

En particulier, la composition par une fonction affine se manifeste graphiquement comme suit :

Proposition 1.2.6. Soient $f : E \rightarrow F$ et $a \in \mathbb{R}$.

- Le graphe de $x \mapsto f(x) + a$ est le translaté de Γ_f par le vecteur $a \vec{j}$.
- Le graphe de $x \mapsto f(x + a)$ est le translaté de Γ_f par le vecteur $-a \vec{i}$.
- Le graphe de $x \mapsto af(x)$ est le dilaté de Γ_f par la transformation $(x, y) \mapsto (x, ay)$.
- Si $a \neq 0$, le graphe de $x \mapsto f(ax)$ est le dilaté de Γ_f par la transformation $(x, y) \mapsto \left(\frac{x}{a}, y\right)$.

1.3 Décrire une fonction

On donne ici des éléments de vocabulaire permettant de décrire une fonction.

Parité, périodicité

Définition 1.3.1. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$. On dit que :

- f est *paire* si : $\forall x \in \mathbb{R}, f(-x) = f(x)$,
- f est *impaire* si : $\forall x \in \mathbb{R}, f(-x) = -f(x)$,
- f est *T-périodique*, où $T > 0$, si : $\forall x \in \mathbb{R}, f(x + T) = f(x)$.

Proposition 1.3.2 (Interprétation graphique).

- Si f est paire, Γ_f est symétrique par rapport à l'axe des ordonnées
- Si f est impaire, Γ_f est symétrique par rapport à l'origine du plan. En particulier, $f(0) = 0$.
- Si f est T-périodique, Γ_f est invariant par translation de vecteur $T\vec{i}$.

Proposition 1.3.3. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$. Il existe une unique fonction paire g et une unique fonction impaire h telles que $f = g + h$.

Démonstration. On raisonne par analyse-synthèse :

- Analyse : Supposons qu'il existe une fonction paire g et une fonction impaire h telles que $f = g + h$. Alors : $\forall x \in \mathbb{R}$,

$$f(x) = g(x) + h(x) \text{ et } f(-x) = g(-x) + h(-x) = g(x) - h(x).$$

En sommant puis en soustrayant ces deux égalités, on obtient :

$$\forall x \in \mathbb{R}, g(x) = \frac{f(x) + f(-x)}{2} \text{ et } h(x) = \frac{f(x) - f(-x)}{2}.$$

En particulier, g et h sont uniques.

- Synthèse : Soient $g : x \mapsto \frac{f(x) + f(-x)}{2}$ et $h : x \mapsto \frac{f(x) - f(-x)}{2}$. Alors : $\forall x \in \mathbb{R}$, $g(-x) = g(x)$, donc g est paire, et : $\forall x \in \mathbb{R}$, $h(-x) = -h(x)$, donc h est impaire. De plus : $\forall x \in \mathbb{R}$, $g(x) + h(x) = f(x)$, donc $f = g + h$. Donc g et h existent.

Par analyse-synthèse, il existe bien g paire et h impaire telles que $f = g + h$; et g et h sont uniques. \square

Remarque 1.3.4. On dit que g est la *partie paire* de f et que h est sa *partie impaire*. Par exemple, la partie paire de la fonction exponentielle est la fonction *cosinus hyperbolique*, notée ch , définie sur \mathbb{R} par : $\text{ch}(x) = \frac{e^x + e^{-x}}{2}$. Sa partie impaire est la fonction *sinus hyperbolique*, notée sh , définie sur \mathbb{R} par : $\text{sh}(x) = \frac{e^x - e^{-x}}{2}$.

Majoration, monotonie

Définition 1.3.5. On dit qu'une fonction $f : E \rightarrow F$ est :

- *majorée* si : $\exists M \in \mathbb{R}$, $\forall x \in E$, $f(x) \leq M$,
- *minorée* si : $\exists m \in \mathbb{R}$, $\forall x \in E$, $f(x) \geq M$,
- *bornée* si f est majorée et minorée,
- admet un *maximum* si : $\exists a \in E$, $\forall x \in E$, $f(x) \leq f(a)$,
- admet un *minimum* si : $\exists a \in E$, $\forall x \in E$, $f(x) \geq f(a)$.

Proposition 1.3.6. Une fonction $f : E \rightarrow \mathbb{R}$ est bornée si et seulement si la fonction $|f| : \begin{cases} E & \rightarrow \mathbb{R} \\ x & \mapsto |f(x)| \end{cases}$ est majorée.

Définition 1.3.7. On dit qu'une fonction $f : E \rightarrow F$ est :

- *croissante* sur E si f conserve l'ordre des éléments de E : $\forall (x, y) \in E^2$, $x \leq y \Rightarrow f(x) \leq f(y)$,
- *strictement croissante* sur E si : $\forall (x, y) \in E^2$, $x < y \Rightarrow f(x) < f(y)$,
- *décroissante* sur E si f inverse l'ordre des éléments de E : $\forall (x, y) \in E^2$, $x \leq y \Rightarrow f(x) \geq f(y)$,
- *strictement décroissante* sur E si : $\forall (x, y) \in E^2$, $x < y \Rightarrow f(x) > f(y)$,
- *monotone* sur E si f est croissante sur E ou si f est décroissante sur E .

Exemples 1.3.8.

- La fonction $x \mapsto \sqrt{x}$ est croissante sur \mathbb{R}_+ ,
- La fonction \ln est croissante sur \mathbb{R}_+^* ,
- La fonction $x \mapsto e^{-x}$ est décroissante sur \mathbb{R} ,

- La fonction $f : x \mapsto \frac{1}{x}$ **n'est pas** décroissante sur \mathbb{R}^* (par exemple, $-2 \leq \frac{1}{3}$ mais $f(-2) = -\frac{1}{2} < 3 = f\left(\frac{1}{3}\right)$). Elle est par contre décroissante sur \mathbb{R}_- , et décroissante sur \mathbb{R}_+ .

Remarque 1.3.9. une fonction à la fois croissante sur E **et** décroissante sur E est constante sur E .

Proposition 1.3.10 (Interprétation graphique).

- Si f est croissante, Γ_f « monte » lorsqu'on se déplace vers la droite du repère (i.e. dans le sens des x croissants)
- Si f est décroissante, Γ_f « descend » lorsqu'on se déplace vers la droite du repère.

2 Continuité

Définition 2.0.1. Soient I un intervalle de \mathbb{R} , $f : I \rightarrow \mathbb{R}$ et $a \in I$. On dit que f est *continue en a* si $f(x) \xrightarrow{x \rightarrow a} f(a)$. On dit que f est *continue sur I* si pour tout a dans I , f est continue en a .

Remarque 2.0.2. Une fonction est donc continue en a si sa *limite* en a est égale à sa *valeur* en a . Nous ne détaillerons pas ici cette définition, qui sera formalisée au second semestre. On peut pour l'instant retenir la caractérisation suivante :

Proposition 2.0.3 (Caractérisation séquentielle de la limite). Soient $f : \mathbb{R} \rightarrow \mathbb{R}$, $a \in \mathbb{R} \cup \{\pm\infty\}$ et $l \in \mathbb{R} \cup \{\pm\infty\}$. On dit que f tend vers l en a , et on note $f(x) \xrightarrow{x \rightarrow a} l$, si pour toute suite (x_n) convergente vers a , la suite $(f(x_n))$ converge vers l .

Exemples 2.0.4.

- La fonction $f : x \mapsto x^2$ est continue en -2 , puisque $f(x) \xrightarrow{x \rightarrow -2} 4 = f(-2)$. La fonction f est plus généralement continue sur \mathbb{R} .
- La plupart des fonctions usuelles sont continues sur leur domaine de définition, à l'exception notable de la fonction *partie entière* $E : \mathbb{R} \rightarrow \mathbb{R}$, définie ainsi : pour tout x réel, $E(x)$ désigne le plus grand entier inférieur ou égal à x : $E(x) = \{p \in \mathbb{Z} \mid p \leq x < p + 1\}$. La fonction partie entière est alors discontinue en tout $p \in \mathbb{Z}$ puisque : $\forall n \in \mathbb{N}^*$, $E\left(p - \frac{1}{n}\right) = p - 1 \xrightarrow{n \rightarrow +\infty} 0 \neq E(p)$.

Les fonctions usuelles continues étant connues, la continuité de la plupart des fonctions rencontrées sera établie grâce à la proposition suivante :

Proposition 2.0.5. Soient $f, g : I \rightarrow \mathbb{R}$ continues. Alors :

1. $f + g$ et fg sont continues,
2. $\forall \lambda \in \mathbb{R}$, λf est continue,
3. si g ne s'annule pas, $\frac{f}{g}$ est continue,
4. si $g(I) \subset I$, alors $f \circ g$ est continue.

Mentionnons pour finir le célèbre :

Théorème 2.0.6 (Théorème des valeurs intermédiaires). Soit $f : I \rightarrow \mathbb{R}$ continue, soient $a, b \in I$ avec $a < b$. Toute valeur intermédiaire à $f(a)$ et $f(b)$ est atteinte par f entre a et b :

$$\forall y \in [f(a), f(b)], \exists x \in [a, b], y = f(x).$$

Démonstration. Admis (pour l'instant !) □

Remarque 2.0.7. Ce théorème ne s'applique pas si I n'est pas un intervalle, comme le montre l'exemple de la fonction inverse entre -1 et 2 .

Corollaire 2.0.8. Soit $f : [a, b] \rightarrow \mathbb{R}$ continue.

- Si f est croissante, alors $f([a, b]) = [f(a), f(b)]$.
- Si f est décroissante, alors $f([a, b]) = [f(b), f(a)]$.

3 Dérivation

Un moyen efficace d'étudier les variations d'une fonction (c'est-à-dire sa monotonie, dont dépend par exemple sa bijectivité) consiste, lorsque c'est possible, à calculer sa *dérivée*.

3.1 Définition

Soient I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{R}$.

Définition 3.1.1. Pour a dans I , on dit que f est *dérivable en a* si $\frac{f(x) - f(a)}{x - a}$ a une limite finie lorsque $x \rightarrow a$. On note alors $f'(a)$ cette limite.

Si f est dérivable en tout point a de I , on peut alors définir une fonction $f' : \begin{cases} I & \rightarrow \mathbb{R} \\ a & \mapsto f'(a) \end{cases}$. C'est la fonction *dérivée* de f .

Remarque 3.1.2. On prendra garde au fait que la dérivation s'opère sur une **fonction** et non un **nombre**. On écrit donc $f'(x)$ et non pas $(f(x))'$.

Remarque 3.1.3 (Interprétation graphique). La quantité $\frac{f(x) - f(a)}{x - a}$ est le *taux d'accroissement* de f entre a et x ; graphiquement, c'est la pente de la corde au graphe de f entre a et x , c'est-à-dire du segment reliant les points $(a, f(a))$ et $(x, f(x))$. En passant à la limite $x \rightarrow a$, $f'(a)$ est donc la pente de la *tangente* au graphe de f en a .

Exemples 3.1.4.

- La fonction $f : x \mapsto x^2$ est dérivable sur \mathbb{R} . En effet, pour tout a dans \mathbb{R} : $\frac{f(x) - f(a)}{x - a} = \frac{x^2 - a^2}{x - a} = x + a \xrightarrow{x \rightarrow a} 2a$. On a donc $f' : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ a & \mapsto 2a \end{cases}$.
- La fonction $f : x \mapsto \sqrt{x}$ est dérivable sur \mathbb{R}_+^* . En effet, pour tout a dans \mathbb{R}_+^* : $\frac{f(x) - f(a)}{x - a} = \frac{\sqrt{x} - \sqrt{a}}{x - a} = \frac{1}{\sqrt{x} + \sqrt{a}} \xrightarrow{x \rightarrow a} \frac{1}{2\sqrt{a}}$. On a donc $f' : \begin{cases} \mathbb{R}_+^* & \rightarrow \mathbb{R} \\ a & \mapsto \frac{1}{2\sqrt{a}}$. On observe de plus que $\frac{f(x) - f(0)}{x - 0} = \frac{1}{\sqrt{x}} \xrightarrow{x \rightarrow 0} +\infty$: la fonction racine n'est pas dérivable en 0. Graphiquement, le fait que le taux d'accroissement tende vers l'infini se manifeste par le fait que le graphe de f admet une tangente verticale en 0.
- La fonction $f : x \mapsto |x|$ est dérivable sur \mathbb{R}^* . En effet, pour tout a dans \mathbb{R}^* : $\frac{f(x) - f(a)}{x - a} = \frac{|x| - |a|}{x - a}$. Si a est strictement positif, alors tout x suffisamment proche de a l'est également, et alors $\frac{|x| - |a|}{x - a} = \frac{x - a}{x - a} = 1 \xrightarrow{x \rightarrow a} 1$. De même, si a est strictement négatif, alors tout x suffisamment proche de a l'est également, et alors $\frac{|x| - |a|}{x - a} = \frac{-x + a}{x - a} = -1 \xrightarrow{x \rightarrow a} -1$. On a donc $f' : \begin{cases} \mathbb{R}^* & \rightarrow \mathbb{R} \\ a & \mapsto \begin{cases} 1 & \text{si } a > 0 \\ -1 & \text{si } a < 0 \end{cases} \end{cases}$. On observe de plus que $\frac{f(x) - f(0)}{x - 0} = \frac{|x|}{x}$ n'a pas de limite en 0 (la limite à gauche valant -1 et la limite à droite 1) : f n'est pas dérivable en 0. Graphiquement, le graphe de f n'admet pas de tangente en 0, et y présente une « pointe ».

Remarque 3.1.5. Attention, le *domaine de dérivabilité* de f est a priori différent du domaine de définition de f' : $D'_f \neq D_{f'}$, comme le montre l'exemple de la fonction \ln .

Comme on l'a dit, l'intérêt premier de la dérivée est de donner les variations de la fonction :

Proposition 3.1.6. Si f est dérivable sur I , alors : (f est croissante (resp. décroissante) sur I) \iff (f' est positive (resp. négative) sur I).

Par conséquent : (f est constante sur I) \iff ($f' = 0$ sur I).

Démonstration. \implies : Supposons f croissante sur I . Alors, pour a dans I : $\forall x \geq a$, $f(x) \geq f(a)$ et $\forall x \leq a$, $f(x) \leq f(a)$. Donc : $\forall x \neq a$, $\frac{f(x) - f(a)}{x - a} \geq 0$, donc par passage à la limite $x \rightarrow a$: $f'(a) \geq 0$. Donc $f' \geq 0$ sur I . Le cas f décroissante se traite de la même façon.

\impliedby : Admis pour l'instant. \square

Proposition 3.1.7. Si f est dérivable en $a \in I$, alors son graphe admet une tangente en a , d'équation cartésienne :

$$y = f'(a)(x - a) + f(a).$$

La dérivabilité d'une fonction entraîne par ailleurs sa continuité :

Proposition 3.1.8. Si f est dérivable sur I , alors f est continue sur I .

Démonstration. Supposons f dérivable sur I . Soient x et a distincts dans I , on a : $f(x) = \frac{f(x) - f(a)}{x - a}(x - a) + f(a)$, où $\frac{f(x) - f(a)}{x - a} \xrightarrow{x \rightarrow a} f'(a)$ et $(x - a) \xrightarrow{x \rightarrow a} 0$, donc $f(x) \xrightarrow{x \rightarrow a} f(a)$. Donc f est continue en a . \square

Remarque 3.1.9. La réciproque est fausse : ainsi, la fonction $x \mapsto |x|$ est continue en 0 mais pas dérivable en 0.

Comme pour la continuité, la dérivabilité des fonctions rencontrées sera le plus souvent justifiée par la proposition suivante et par la dérivabilité des fonctions usuelles :

Proposition 3.1.10. Soient $f, g : I \rightarrow \mathbb{R}$ dérивables. Alors :

1. $f + g$ et fg sont dérивables,
2. $\forall \lambda \in \mathbb{R}$, λf est dérivable,
3. si g ne s'annule pas, $\frac{f}{g}$ est dérivable,
4. si $g(I) \subset I$, alors $f \circ g$ est dérivable.

Dérivées d'ordre supérieur

Définition 3.1.11. Soit $f : I \rightarrow \mathbb{R}$. Lorsque cela est possible, on note $f^{(n)} = f$ si $n = 0$, $(f^{(n-1)})'$ si $n \geq 1$. C'est la *dérivée $n^{\text{ème}}$* de f .

On note $D^n(I, \mathbb{R})$ l'ensemble des fonctions n fois dérивables sur I , et $C^n(I, \mathbb{R})$ l'ensemble de celles telles que $f^{(n)}$ est continue.

Définition 3.1.12. On dit que f est *convexe* si sa courbe est en-dessous de ses cordes et au-dessus de ses tangentes, *concave* si c'est le contraire.

Proposition 3.1.13. Si f est deux fois dérivable, on a : (f est convexe) \iff ($f'' \geq 0$).

Exemple 3.1.14. $\forall x \in \mathbb{R}$, $e^x \geq 1 + x$; $\forall x \in \mathbb{R}_+$, $\sin(x) \leq x$.

3.2 Opérations sur les dérivées

Pour calculer une dérivée, on ne revient pas systématiquement au taux d'accroissement : il suffit de connaître les dérivées de quelques fonctions de base (cf. la section suivante) et d'utiliser les règles de calcul ci-après :

Proposition 3.2.1. Soient $f, g : I \rightarrow \mathbb{R}$ dérivables sur I et $\lambda \in \mathbb{R}$. Alors :

1. $(f + g)' = f' + g'$
2. $(\lambda f)' = \lambda f'$
3. $(fg)' = f'g + fg'$ (Formule de Leibniz)
4. Si g ne prend pas la valeur 0, $\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}$
5. Si g ne prend pas la valeur 0, $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$
6. $(f \circ g)' = g' \times (f' \circ g)$

Démonstration. On considère à chaque fois le taux d'accroissement au voisinage de $a \in I$:

1. $\frac{(f + g)(x) - (f + g)(a)}{x - a} = \frac{f(x) - f(a)}{x - a} + \frac{g(x) - g(a)}{x - a} \xrightarrow{x \rightarrow a} f'(a) + g'(a)$
2. $\frac{(\lambda f)(x) - (\lambda f)(a)}{x - a} = \frac{\lambda f(x) - \lambda f(a)}{x - a} \xrightarrow{x \rightarrow a} \lambda f'(a)$
3. $\frac{(fg)(x) - (fg)(a)}{x - a} = \frac{f(x)g(x) - f(a)g(a)}{x - a} = \frac{f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)}{x - a} \xrightarrow{x \rightarrow a} f'(a)g(a) + f(a)g'(a)$
4. $\frac{\left(\frac{1}{g}\right)(x) - \left(\frac{1}{g}\right)(a)}{x - a} = \frac{1}{x - a} \frac{g(a) - g(x)}{g(x)^2} \xrightarrow{x \rightarrow a} f \frac{g'(a)}{g(a)^2}$
5. D'après les propriétés précédentes, $\left(\frac{f}{g}\right)' = \left(f \times \frac{1}{g}\right)' = f' \times \frac{1}{g} - f \times \frac{g'}{g^2} = \frac{f'g - fg'}{g^2}$
6. $\frac{(f \circ g)(x) - (f \circ g)(a)}{x - a} = \frac{f(g(x)) - f(g(a))}{x - a} = \frac{g(x) - g(a)}{x - a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \xrightarrow{x \rightarrow a} g'(a) \times f'(g(a))$

□

3.3 Dérivées usuelles

Le tableau ci-après présente les dérivées usuelles à connaître, avec les domaines de définition et de dérivabilité correspondants :

$f(x)$	D_f	$f'(x)$	D'_f
x^α ($\alpha \in \mathbb{R}$)	\mathbb{R}_+^*	$\alpha x^{\alpha-1}$	\mathbb{R}_+^*
x^n ($n \in \mathbb{N}^*$)	\mathbb{R}	nx^{n-1}	\mathbb{R}
$\frac{1}{x^n}$ ($n \in \mathbb{N}^*$)	\mathbb{R}^*	$-\frac{n}{x^{n+1}}$	\mathbb{R}^*
\sqrt{x}	\mathbb{R}_+	$\frac{1}{2\sqrt{x}}$	\mathbb{R}_+^*
e^x	\mathbb{R}	e^x	\mathbb{R}
$\ln x$	\mathbb{R}_+^*	$\frac{1}{x}$	\mathbb{R}_+^*
$\sin x$	\mathbb{R}	$\cos x$	\mathbb{R}
$\cos x$	\mathbb{R}	$-\sin x$	\mathbb{R}
$\tan x$	$\mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$	$1 + \tan^2 x = \frac{1}{\cos^2 x}$	$\mathbb{R} \setminus \left(\frac{\pi}{2} + \pi\mathbb{Z}\right)$
$\arcsin x$	$[-1, 1]$	$\frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$
$\arccos x$	$[-1, 1]$	$-\frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$
$\arctan x$	\mathbb{R}	$\frac{1}{1+x^2}$	\mathbb{R}
$\operatorname{ch} x$	\mathbb{R}	$\operatorname{sh} x$	\mathbb{R}
$\operatorname{sh} x$	\mathbb{R}	$\operatorname{ch} x$	\mathbb{R}

4 Bijections

Soit une fonction $f : E \rightarrow F$. On rappelle que tout élément x de E a exactement une image $y \in F$ par f ; mais qu'à l'inverse, un élément y de F peut avoir un, plusieurs ou aucun antécédent par f .

Définition 4.0.1.

- Si tout élément y de F a **au moins** un antécédent par f , on dit que f est *surjective*,
- Si tout élément y de F a **au plus** un antécédent par f , on dit que f est *injective*.

Exemple 4.0.2.

- La fonction $\begin{cases} \mathbb{R} & \rightarrow \mathbb{R}_+ \\ x & \mapsto x^2 \end{cases}$ est surjective et non injective,
- La fonction $\begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto e^x \end{cases}$ est injective et non surjective.

Définition 4.0.3. Si f est à la fois injective et surjective, on dit que f est *bijective*. Tout élément y de F a alors exactement un antécédent par f :

$$\forall y \in F, \exists! x \in E \text{ t.q. } y = f(x).$$

Cela définit une fonction $F \rightarrow E$, appelée *fonction réciproque* de f , également bijective, notée f^{-1} , définie par :

$$f^{-1} : \begin{cases} F & \rightarrow E \\ y & \mapsto (\text{l'unique}) x \text{ t.q. } f(x) = y. \end{cases}$$

Exemple 4.0.4.

- La fonction $\begin{cases} \mathbb{R}_+ & \rightarrow \mathbb{R}_+ \\ x & \mapsto x^2 \end{cases}$ est bijective, de réciproque $\begin{cases} \mathbb{R}_+ & \rightarrow \mathbb{R}_+ \\ x & \mapsto \sqrt{x} \end{cases}$,
- La fonction $\begin{cases} \mathbb{R} & \rightarrow \mathbb{R}_+^* \\ x & \mapsto e^x \end{cases}$ est bijective, de réciproque $\begin{cases} \mathbb{R}_+^* & \rightarrow \mathbb{R} \\ x & \mapsto \ln x \end{cases}$.

Remarque 4.0.5. La courbe de f^{-1} est le symétrique de la courbe de f par rapport à la droite d'équation $y = x$.

Proposition 4.0.6.

- Si f est bijective, alors f^{-1} l'est également, et $(f^{-1})^{-1} = f$.
- Soit $g : F \rightarrow E$. On a l'équivalence : (f et g sont bijectives et $g = f^{-1}$) \iff ($g \circ f = \text{Id}_E$ et $f \circ g = \text{Id}_F$).

Remarque 4.0.7. On voit que les propriétés d'injectivité, de surjectivité et de bijectivité dépendent notamment de E et F . Ainsi, la fonction $\begin{cases} E & \rightarrow f(E) \\ x & \mapsto f(x) \end{cases}$ est surjective par définition de $f(E)$. De même, avec un choix judicieux de $A \subset E$, la fonction $\begin{cases} A & \rightarrow F \\ x & \mapsto f(x) \end{cases}$ peut être injective. On appelle cette fonction *restriction de f à A* , notée $f|_A$.

Le choix de A peut être guidé par le théorème suivant :

Théorème 4.0.8 (Théorème de la bijection continue). Soit I un intervalle de \mathbb{R} , soit $f : I \rightarrow \mathbb{R}$. Si f est continue et strictement monotone sur I , alors f réalise une bijection de I dans l'intervalle $J = f(I)$, et f^{-1} est continue et strictement monotone sur J .

Remarque 4.0.9. Dans ce théorème, le fait que J soit un intervalle est une conséquence immédiate du théorème des valeurs intermédiaires. **Le théorème de la bijection monotone n'est cependant pas un corollaire du TVI.**

Exemple 4.0.10. La fonction $\begin{cases} [0, \pi] & \rightarrow [-1, 1] \\ x & \mapsto \cos(x) \end{cases}$ est continue sur $[0, \pi]$, strictement décroissante sur $[0, \pi]$ et surjective puisque $\cos([0, \pi]) = [-1, 1]$, donc est bijective. La fonction réciproque est la fonction *arc cosinus*, définie par :

$$\arccos : \begin{cases} [-1, 1] & \rightarrow [0, \pi] \\ y & \mapsto \text{l'unique } x \in [0, \pi] \text{ t.q. } \cos(x) = y. \end{cases}$$

Proposition 4.0.11. Si f est bijective, $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$.

Démonstration. Comme $f \circ f^{-1} = \text{Id}$, on a $f' \times f' \circ f^{-1} = 1$, donc $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$. \square

Chapitre 4

Nombres complexes

1 Généralités

L'apparition des *nombres complexes* est relativement récente dans l'histoire des mathématiques : ils sont utilisés pour la première fois en 1535 par Jérôme Cardan et par Tartaglia, qui proposent une méthode de résolution des équations polynomiales de degré 3 en s'autorisant à utiliser au cours du calcul des nombres de carré négatif, qu'ils appellent « nombres impossibles ». Leur méthode est généralisée dès 1540 par Ludovico Ferrari pour résoudre les équations polynomiales de degré 4. Les nombres complexes sont ensuite progressivement assimilés par les mathématiciens : Descartes introduit ainsi la terminologie de « nombres imaginaires » en 1637 et Euler introduit la notation « i » en 1777.

1.1 L'ensemble \mathbb{C}

Définition 1.1.1. On définit le nombre i comme étant une solution de l'équation $x^2 + 1 = 0$.

Remarque 1.1.2. Le nombre i a donc pour propriété fondamentale de vérifier l'égalité : $i^2 = -1$. L'autre solution de l'équation $x^2 + 1 = 0$ est le nombre $-i$.

L'utilisation de ce nouveau nombre a deux conséquences immédiates :

Proposition 1.1.3. Soit a un nombre réel. L'équation $x^2 = a$ admet deux solutions (confondues si $a = 0$) :

- Si $a \geq 0$: $x_{1,2} = \pm\sqrt{a}$,
- Si $a \leq 0$: $x_{1,2} = \pm i\sqrt{-a}$.

Proposition 1.1.4. Soit (a, b, c) dans \mathbb{R}^3 avec $a \neq 0$. On note $\Delta = b^2 - 4ac$. L'équation $ax^2 + bx + c = 0$ admet deux solutions (confondues si $\Delta = 0$) :

- Si $\Delta \geq 0$: $x_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a}$,
- Si $\Delta \leq 0$: $x_{1,2} = \frac{-b \pm i\sqrt{-\Delta}}{2a}$.

On voit, dans le cas $\Delta < 0$, que les solutions trouvées sont de la forme « (nombre réel) $+ i \times$ (nombre réel) ». On est donc amené à considérer un nouvel ensemble de nombres :

Définition 1.1.5. L'ensemble des *nombres complexes*, noté \mathbb{C} , est l'ensemble :

$$\mathbb{C} = \{x + iy \mid (x, y) \in \mathbb{R}^2\},$$

muni des opérations suivantes : $\forall (x, y, x', y') \in \mathbb{R}^4$,

- $(x + iy) + (x' + iy') = (x + x') + i(y + y')$,

- $(x + iy) \times (x' + iy') = (xx' - yy') + i(xy' + x'y)$.

Remarque 1.1.6. L'ensemble \mathbb{R} des nombres réels est un sous-ensemble de \mathbb{C} : $\mathbb{R} = \{x + i \times 0 \mid x \in \mathbb{R}\}$. On note $i\mathbb{R}$ le sous-ensemble de \mathbb{C} : $i\mathbb{R} = \{0 + iy \mid y \in \mathbb{R}\}$, appelé *ensemble des nombres imaginaires purs*.

Proposition 1.1.7. Soit $z = x + iy$ dans \mathbb{C} , avec $(x, y) \in \mathbb{R}^2$. Les réels x et y sont uniquement déterminés par z . On appelle x la partie réelle de z , y la partie imaginaire de z , et on note $x = \operatorname{Re}(z)$ et $y = \operatorname{Im}(z)$. La forme $x + iy$ est appelée forme algébrique de z .

Proposition 1.1.8 (Caractérisation des réels et des imaginaires purs par les parties réelle et imaginaire).

Soit z dans \mathbb{C} .

1. $(z \in \mathbb{R}) \iff (\operatorname{Im}(z) = 0)$,
2. $(z \in i\mathbb{R}) \iff (\operatorname{Re}(z) = 0)$.

Proposition 1.1.9 (Propriétés des parties réelle et imaginaire).

1. $\forall (z, z') \in \mathbb{C}^2$, $\operatorname{Re}(z + z') = \operatorname{Re}(z) + \operatorname{Re}(z')$ et $\operatorname{Im}(z + z') = \operatorname{Im}(z) + \operatorname{Im}(z')$,
2. $\forall (\lambda, z) \in \mathbb{R} \times \mathbb{C}$, $\operatorname{Re}(\lambda z) = \lambda \operatorname{Re}(z)$ et $\operatorname{Im}(\lambda z) = \lambda \operatorname{Im}(z)$,
3. $\forall (z, z') \in \mathbb{C}^2$, $(z = z') \iff (\operatorname{Re}(z) = \operatorname{Re}(z') \text{ et } \operatorname{Im}(z) = \operatorname{Im}(z'))$.

Proposition 1.1.10. Soit z dans \mathbb{C}^* . Il existe un unique w dans \mathbb{C}^* tel que $zw = 1$. Le nombre w est appelé l'inverse de z , noté $w = \frac{1}{z}$.

Démonstration. On raisonne par analyse-synthèse. Notons $z = x + iy$ sous forme algébrique, soit $w = a + ib$ tel que $zw = 1$. Alors par identification des parties réelle et imaginaire :

$$\begin{cases} ax - by = 1 \\ ay + bx = 0 \end{cases} \text{ donc } \begin{cases} axy - by^2 = y \\ axy + bx^2 = 0 \end{cases} \text{ donc } \begin{cases} b = -\frac{y}{x^2 + y^2} \\ a = \frac{x}{x^2 + y^2} \end{cases} \text{ donc } w = \frac{x - iy}{x^2 + y^2},$$

où $x^2 + y^2 \neq 0$ puisque $z \neq 0$. Réciproquement, $w = \frac{x - iy}{x^2 + y^2}$ vérifie bien $zw = 1$. \square

On voit que l'inverse de $z = x + iy \neq 0$ fait intervenir les quantités $x - iy$ et $x^2 + y^2$, auxquelles on va s'intéresser maintenant :

1.2 Conjugaison complexe

Définition 1.2.1. Soit z dans \mathbb{C} , de forme algébrique $z = x + iy$. Le *conjugué complexe de z* , noté \bar{z} , est le nombre complexe $\bar{z} = x - iy$.

Proposition 1.2.2 (Opérations sur le conjugué).

1. $\forall (z, z') \in \mathbb{C}^2$, $\overline{z + z'} = \bar{z} + \bar{z'}$,
2. $\forall (z, z') \in \mathbb{C}^2$, $\overline{zz'} = \bar{z} \cdot \bar{z'}$,
3. $\forall (\lambda, z) \in \mathbb{R} \times \mathbb{C}$, $\overline{\lambda z} = \lambda \bar{z}$,
4. $\forall z \in \mathbb{C}^*$, $\overline{\frac{1}{z}} = \frac{1}{\bar{z}}$.

Le conjugué permet également d'écrire les parties réelle et imaginaire :

Proposition 1.2.3 (Formules d'Euler).

Soit z dans \mathbb{C} .

1. $\operatorname{Re}(z) = \frac{z + \bar{z}}{2}$,

$$2. \operatorname{Im}(z) = \frac{z - \bar{z}}{2i}.$$

Proposition 1.2.4 (Caractérisation des réels et des imaginaires purs par le conjugué).

Soit z dans \mathbb{C} .

1. $(z \in \mathbb{R}) \iff (z = \bar{z})$,
2. $(z \in i\mathbb{R}) \iff (z = -\bar{z})$.

1.3 Interprétation géométrique

L'ensemble \mathbb{C} s'assimile au plan \mathbb{R}^2 , de la façon suivante :

Proposition 1.3.1. La fonction $\begin{cases} \mathbb{C} & \rightarrow \mathbb{R}^2 \\ x + iy & \mapsto (x, y) \end{cases}$ est bijective. Pour z dans \mathbb{C} , on dit du point M du plan de coordonnées $f(z)$ qu'il est d'affixe z .

Exemple 1.3.2 (Interprétation graphique du conjugué). Soit M le point du plan d'affixe z . Le point M' d'affixe \bar{z} est le symétrique de M par rapport à l'axes des abscisses (Ox).

1.4 Module

Définition 1.4.1. Soit z dans \mathbb{C} , de forme algébrique $z = x + iy$. Le *module de z* , noté $|z|$, est le nombre réel positif $|z| = \sqrt{x^2 + y^2}$.

Remarque 1.4.2 (Interprétation graphique). Comme la distance absolue, dont il partage la notation, est la distance à 0 d'un nombre réel, le module d'un nombre complexe z est la distance à l'origine du point du plan d'affixe z . Plus généralement, la distance entre deux nombres complexes z et z' vaut $|z - z'|$.

On peut à présent réécrire la propriété 1.1.10 sous la forme suivante, à retenir :

Proposition 1.4.3. $\forall z \in \mathbb{C}, z\bar{z} = |z|^2$.

Proposition 1.4.4 (Opérations sur le module).

1. $\forall z \in \mathbb{C}, |z| = 0 \iff z = 0$,
2. (Inégalité triangulaire) $\forall (z, z') \in \mathbb{C}^2, |z + z'| \leq |z| + |z'|$,
3. $\forall (z, z') \in \mathbb{C}^2, |zz'| = |z| \cdot |z'|$,
4. $\forall (\lambda, z) \in \mathbb{R}_+ \times \mathbb{C}, |\lambda z| = \lambda |z|$,
5. $\forall z \in \mathbb{C}, |\bar{z}| = |z|$,
6. $\forall z \in \mathbb{C}^*, \left| \frac{1}{z} \right| = \frac{1}{|z|}$.

Preuve de l'inégalité triangulaire. Soit (z, z') dans \mathbb{C}^2 . Alors :

$$|z + z'|^2 = (z + z')\overline{(z + z')} = (z + z')(\bar{z} + \bar{z}') = |z|^2 + |z'|^2 + 2\operatorname{Re}(z\bar{z}'),$$

d'où $|z + z'| \leq |z| + |z'| \iff |z + z'|^2 \leq (|z| + |z'|)^2 \iff \operatorname{Re}(z\bar{z}') \leq |z||z'| = |z||\bar{z}'| = |z\bar{z}'|$. Cette dernière inégalité étant vraie, l'inégalité triangulaire l'est aussi. \square

Remarque 1.4.5 (Cas d'égalité dans l'inégalité triangulaire). Dans la preuve ci-dessus, il y a égalité si et seulement si $\operatorname{Re}(z\bar{z}') = |z\bar{z}'|$, c'est-à-dire lorsque $z\bar{z}' \in \mathbb{R}_+$, c'est-à-dire lorsque $z = \frac{\lambda}{z'} = \frac{\lambda}{|z'|}z'$: autrement dit, il y a égalité dans l'inégalité triangulaire si et seulement si $z = kz'$ où $k \in \mathbb{R}_+$.

Remarque 1.4.6 (Interprétation graphique de l'inégalité triangulaire). Soient M , M' et M'' les points d'affixes respectives z , z' et $z + z'$. Ces points sont liés par l'égalité vectorielle $\overrightarrow{OM''} = \overrightarrow{OM} + \overrightarrow{OM'}$, soit $\overrightarrow{OM'} = \overrightarrow{MM''}$. L'inégalité triangulaire s'écrit alors : $\|\overrightarrow{OM''}\| \leq \|\overrightarrow{OM}\| + \|\overrightarrow{MM''}\|$: le chemin le plus court de O à M'' est la ligne droite.

Il y a égalité dans cette inégalité si et seulement si O , M et M' sont alignés.

Proposition 1.4.7 (Caractérisation des réels et des imaginaires purs par le module).

Soit z dans \mathbb{C} .

1. $(z \in \mathbb{R}) \iff (z = \pm|z|)$,
2. $(z \in i\mathbb{R}) \iff (z = \pm i|z|)$.

L'interprétation du module comme distance entre les nombres complexes permet de considérer les ensembles suivants :

Proposition 1.4.8. Soient z_0 dans \mathbb{C} et $r \in \mathbb{R}_+$. Soit M_0 le point du plan d'affixe z_0 .

- $C(z_0, r) = \{z \in \mathbb{C} \mid |z - z_0| = r\}$ correspond dans le plan au cercle de centre M_0 et de rayon r ,
- $D(z_0, r) = \{z \in \mathbb{C} \mid |z - z_0| \leq r\}$ correspond dans le plan au disque de centre M_0 et de rayon r .

1.5 Équations du second degré

Comme on l'a vu en début de chapitre, la motivation première d'utilisation des nombres complexes est la résolution d'équations polynomiales. Le cas des équations du second degré à coefficients complexes est une généralisation directe du cas réel :

Proposition 1.5.1. Soit Z un nombre complexe. L'équation $z^2 = Z$ admet deux solutions (confondues si $Z = 0$).

Proposition 1.5.2. Soit (a, b, c) dans \mathbb{C}^3 avec $a \neq 0$. On note $\Delta = b^2 - 4ac$. Soit δ une racine carrée de Δ . L'équation $az^2 + bz + c = 0$ admet deux solutions (confondues si $\Delta = 0$) :

$$z_{1,2} = \frac{-b \pm \delta}{2a}.$$

Exemple 1.5.3. $z^2 - (1 + 2i)z + 1 - i = 0$.

Remarque 1.5.4. On réserve la notation $\sqrt{}$ aux nombres réels positifs : il n'y a pas de choix « canonique » de racine carrée dans \mathbb{C} . De même, on rappelle que la fonction \ln ne prend en argument que des réels strictement positifs.

Proposition 1.5.5. Avec les notations de la proposition 1.5.2, on a :

1. $z_1 + z_2 = -\frac{b}{a}$,
2. $z_1 z_2 = \frac{c}{a}$,
3. $az^2 + bz + c = a(z - z_1)(z - z_2)$.

Théorème 1.5.6 (Théorème de d'Alembert-Gauss). Toute fonction polynomiale à coefficients complexes de degré $n \geq 1$ admet une racine dans \mathbb{C} .

Ce théorème, dit *théorème fondamental de l'algèbre*, est faux dans \mathbb{R} : la fonction $x^2 + 1$ n'a par exemple pas de racine réelle. On dit que \mathbb{C} est *algébriquement clos*.

2 Nombres complexes et trigonométrie

2.1 Le cercle unité

Définition 2.1.1. On appelle *cercle unité*, et on note \mathbb{U} , le cercle $C(0, 1) = \{z \in \mathbb{C} \mid |z| = 1\}$.

Définition 2.1.2. Soit $z \in \mathbb{U}$. Notons M_0 le point d'affixe 1 et M le point d'affixe z . On appelle *mesure de l'angle* $\widehat{M_0OM}$ la longueur de l'arc de cercle M_0M , comptée algébriquement, dans le sens *trigonométrique* (c'est-à-dire antihoraire). Le point et l'affixe correspondant à une mesure d'angle (ou, par abus de langage, un angle) $\theta \in \mathbb{R}$ sont respectivement notés M_θ et z_θ . L'unité de la mesure d'angle est le *radian*.

Définition 2.1.3. Soit $\theta \in \mathbb{R}$. On note respectivement $\cos(\theta)$ et $\sin(\theta)$ l'abscisse et l'ordonnée de M_θ .

Soit D_θ^+ la demi-droite du plan issue de O et d'angle θ . Si $\theta \neq \frac{\pi}{2} + k\pi$, où $k \in \mathbb{Z}$, on note :

$\varepsilon = \begin{cases} 1 & \text{si } (\theta \bmod 2\pi) \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\\ -1 & \text{si } (\theta \bmod 2\pi) \in \left]\frac{\pi}{2}, \frac{3\pi}{2}\right[\end{cases}$, puis $\tan(\theta) = \varepsilon y$ où y est l'ordonnée de l'unique point d'intersection de D_θ^+ avec la droite d'équation $x = \varepsilon 1$.

Les *fonctions trigonométriques sinus, cosinus et tangente* ainsi définies sur \mathbb{R} ont les premières propriétés suivantes, directement issues de leur définition :

Proposition 2.1.4. Pour tout θ dans \mathbb{R} :

1. $\cos^2(\theta) + \sin^2(\theta) = 1$,
2. $\cos(\theta + 2\pi) = \cos(\theta)$ et $\sin(\theta + 2\pi) = \sin(\theta)$,
3. $\cos(-\theta) = \cos(\theta)$ et $\sin(-\theta) = -\sin(\theta)$,
4. $\cos(\theta + \pi) = -\cos(\theta)$ et $\sin(\theta + \pi) = -\sin(\theta)$,
5. $\cos(\pi - \theta) = -\cos(\theta)$ et $\sin(\pi - \theta) = \sin(\theta)$,
6. $\cos\left(\theta + \frac{\pi}{2}\right) = -\sin(\theta)$ et $\sin\left(\theta + \frac{\pi}{2}\right) = \cos(\theta)$,
7. $\cos\left(\frac{\pi}{2} - \theta\right) = \sin(\theta)$ et $\sin\left(\frac{\pi}{2} - \theta\right) = \cos(\theta)$.

Proposition 2.1.5. Pour tout $\theta \neq \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$,

1. $\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$,
2. $\tan(-\theta) = -\tan(\theta)$,
3. $\tan(\theta + \pi) = \tan(\theta)$,
4. $\tan\left(\frac{\pi}{2} - \theta\right) = \frac{1}{\tan(\theta)}$.

On retiendra les valeurs remarquables suivantes des fonctions trigonométriques :

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos \theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
$\sin \theta$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\tan \theta$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	non défini

On notera également les équivalences :

Proposition 2.1.6. $\forall (a, b) \in \mathbb{R}^2$,

- $(\cos a = \cos b) \iff (a = \pm b \pmod{2\pi})$,
- $(\sin a = \sin b) \iff (a = b \text{ ou } \pi - b \pmod{2\pi})$,
- $(\tan a = \tan b) \iff (a = b \text{ ou } \pi + b \pmod{2\pi}) \iff (a = b \pmod{\pi})$,

On a de plus l'égalité :

Proposition 2.1.7. $\forall (a, b) \in \mathbb{R}^2$, $\cos(a - b) = \cos(a) \cos(b) + \sin(a) \sin(b)$.

Démonstration. Soit (a, b) dans \mathbb{R}^2 . Alors $\overrightarrow{OM}_a = (\cos a, \sin a)$, $\overrightarrow{OM}_b = (\cos b, \sin b)$, et $\overrightarrow{OM}_a \cdot \overrightarrow{OM}_b = \|\overrightarrow{OM}_a\| \|\overrightarrow{OM}_b\| \cos(a - b) = \cos(a - b)$, d'où l'égalité. \square

Avec les propriétés précédentes, on en déduit immédiatement :

Proposition 2.1.8. Pour tout (a, b) dans \mathbb{R}^2 :

1. $\cos(a + b) = \cos(a) \cos(b) - \sin(a) \sin(b)$,
2. $\sin(a + b) = \sin(a) \cos(b) + \sin(b) \cos(a)$,
3. $\sin(a - b) = \sin(a) \cos(b) - \sin(b) \cos(a)$,
4. Si $a, b, a + b \neq \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$, $\tan(a + b) = \frac{\tan(a) + \tan(b)}{1 - \tan(a) \tan(b)}$,
5. Si $a, b, a - b \neq \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$, $\tan(a - b) = \frac{\tan(a) - \tan(b)}{1 + \tan(a) \tan(b)}$.

En particulier, pour $a = b$ (on supprime les parenthèses pour alléger les formules) :

Proposition 2.1.9. Pour tout a dans \mathbb{R} :

1. $\cos 2a = \cos^2 a - \sin^2 a = 2\cos^2 a - 1 = 1 - 2\sin^2 a$,
2. $\sin 2a = 2\sin a \cos a$,
3. Si $a, 2a \neq \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$, $\tan 2a = \frac{2\tan a}{1 - \tan^2 a}$.

Enfin, on déduit des définitions géométriques du sinus et du cosinus les limites suivantes :

Proposition 2.1.10.

1. $\frac{1 - \cos x}{x} \xrightarrow{x \rightarrow 0} 0$,
2. $\frac{\sin x}{x} \xrightarrow{x \rightarrow 0} 1$,

Démonstration. Graphiquement, on a, pour x proche de 0 :

- $\sin x \leq x$ et $1 - \cos x \geq 0$,
- $\sin^2(x) + (1 - \cos x)^2 \leq x^2$, donc $1 - \cos x \leq \frac{x^2}{2}$,
- $\sin x + (1 - \cos x) \geq x$, donc $\sin x \geq x - \frac{x^2}{2}$,

d'où, par encadrement, les limites indiquées. \square

ce qui entraîne :

Proposition 2.1.11. Les fonctions \cos et \sin sont dérивables sur \mathbb{R} , et :

$$\cos' = -\sin, \quad \sin' = \cos.$$

2.2 L'exponentielle complexe

On considère la fonction $f : \begin{cases} \mathbb{R} & \rightarrow \mathbb{C} \\ \theta & \mapsto \cos(\theta) + i \sin(\theta) \end{cases}$. Cette fonction est 2π -périodique et de classe C^∞ sur \mathbb{R} puisque cos et sin le sont. De plus :

Proposition 2.2.1. *La restriction $]-\pi, \pi] \rightarrow \mathbb{U}$ de f est bijective.*

La fonction f a en outre les propriétés suivantes :

Proposition 2.2.2.

1. $f(0) = 1$,
2. $\forall k \in \mathbb{N}, f^{(k)}(\theta) = i^k f(\theta)$,
3. $\forall (a, b) \in \mathbb{R}^2, f(a + b) = f(a)f(b)$.

Les propriétés de la fonction f sont exactement celles d'une fonction exponentielle. On note donc :

Définition 2.2.3. $\forall \theta \in \mathbb{R}, e^{i\theta} = f(\theta) = \cos \theta + i \sin \theta$.

Remarque 2.2.4. En particulier, $e^{i0} = e^{i2\pi} = 1$, $e^{i\pi} = -1$ et $e^{i\frac{\pi}{2}} = i$.

Cette *exponentielle complexe* hérite des propriétés de la fonction f :

Proposition 2.2.5.

1. $\forall (\theta, k) \in \mathbb{R} \times \mathbb{Z}, e^{i(\theta+2k\pi)} = e^{i\theta}$,
2. $\forall (a, b) \in \mathbb{R}^2, e^{i(a+b)} = e^{ia}e^{ib}$,
3. $\forall (a, b) \in \mathbb{R}^2, e^{i(a-b)} = \frac{e^{ia}}{e^{ib}}$,
4. $\forall \theta \in \mathbb{R}, \overline{e^{i\theta}} = e^{-i\theta}$,
5. $\forall \theta \in \mathbb{R}, |e^{i\theta}| = 1$,
6. (Formules d'Euler) $\forall \theta \in \mathbb{R}, \cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$ et $\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$.

En revanche, l'exponentielle complexe ne se comporte pas « simplement » vis-à-vis de la somme. On a néanmoins les formules suivantes :

Proposition 2.2.6 (Méthode de l'angle moitié).

- $\forall \theta \in \mathbb{R}, 1 + e^{i\theta} = e^{i\frac{\theta}{2}} \left(e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}} \right) = 2 \cos \left(\frac{\theta}{2} \right) e^{i\frac{\theta}{2}}$,
- $\forall (a, b) \in \mathbb{R}^2, e^{ia} + e^{ib} = e^{ia} \left(1 + e^{i(b-a)} \right) = 2 \cos \left(\frac{a-b}{2} \right) e^{i\frac{a+b}{2}}$.

La partie réelle de cette dernière égalité donne la formule de factorisation :

$$\forall (a, b) \in \mathbb{R}^2, \cos a + \cos b = 2 \cos \left(\frac{a-b}{2} \right) \cos \left(\frac{a+b}{2} \right).$$

Exemple 2.2.7. Résolution de l'équation $\cos(3x) + \cos(5x) = 0$.

Plus généralement, l'exponentielle complexe permet de (re)trouver plus facilement nombre de formules de trigonométrie, notamment via les formules d'Euler et la formule de linéarisation suivante, dite *de Moivre* :

Proposition 2.2.8 (Formule de Moivre). $\forall \theta \in \mathbb{R}, (\cos \theta + i \sin \theta)^n = e^{in\theta} = \cos n\theta + i \sin n\theta$.

L'exponentielle complexe offre également une nouvelle écriture des nombres complexes, dite *forme exponentielle* :

Proposition 2.2.9. Pour tout z dans \mathbb{C}^* , il existe un unique $r > 0$ et un unique θ modulo 2π tels que $z = re^{i\theta}$.

Démonstration. Soit z dans \mathbb{C}^* . On pose $r = |z|$, alors $\left|\frac{z}{r}\right| = 1$ donc $z \in \mathbb{U}$. Il existe donc un unique θ modulo 2π tel que $\frac{z}{r} = e^{i\theta}$, donc $z = re^{i\theta}$.

Réiproquement, si $re^{i\theta} = r'e^{i\theta'}$, alors, par égalité des modules, $r = r'$, donc $e^{i\theta} = e^{i\theta'}$, donc $\cos(\theta) = \cos(\theta')$ et $\sin(\theta) = \sin(\theta')$, donc $\theta = \theta' \pmod{2\pi}$. \square

Définition 2.2.10. Si $z \in \mathbb{C}^*$ a pour écriture exponentielle $z = re^{i\theta}$, on dit que θ est un *argument de z* , noté $\arg(z)$.

Proposition 2.2.11 (Propriétés de l'argument). Pour tout (z, z') dans $(\mathbb{C}^*)^2$,

1. $\arg(zz') = \arg(z) + \arg(z')$,
2. $\arg\left(\frac{z}{z'}\right) = \arg(z) - \arg(z')$,
3. $\arg(\bar{z}) = -\arg(z)$.

Enfin, on généralise la fonction exponentielle à tout le plan complexe :

Définition 2.2.12. Soit z dans \mathbb{C} d'écriture algébrique $z = x + iy$. On appelle *exponentielle* de z , et on note e^z , le nombre complexe $e^z = e^x e^{iy}$.

Proposition 2.2.13. Pour tout (z, z') dans \mathbb{C}^2 ,

1. $e^{z+z'} = e^z e^{z'}$,
2. $e^{z-z'} = \frac{e^z}{e^{z'}}$,
3. $(e^z = e^{z'}) \iff (\operatorname{Re}(z) = \operatorname{Re}(z') \text{ et } (\operatorname{Im}(z) = \operatorname{Im}(z') \pmod{2\pi}))$.

2.3 Racines de l'unité

L'écriture exponentielle des nombres complexes permet une écriture simple des racines : Soit n dans \mathbb{N}^* .

Définition 2.3.1. On appelle *racines n èmes de 1*, ou *de l'unité*, les solutions complexes de l'équation $z^n = 1$. L'ensemble des racines n èmes de l'unité est noté \mathbb{U}_n .

Exemple 2.3.2. Ainsi, $\mathbb{U}_2 = \{-1, 1\}$ et $\mathbb{U}_4 = \{-1, 1, -i, i\}$. On notera de plus que pour tout n , $\mathbb{U}_n \subset \mathbb{U}$.

Remarque 2.3.3. Dans le cas réel, l'étude de la fonction $x \mapsto x^n$ permet de constater qu'il y a une seule racine réelle de l'unité (1) lorsque n est impair, deux (-1 et 1) lorsque n est pair.

Proposition 2.3.4. $\mathbb{U}_n = \left\{ e^{i\frac{2k\pi}{n}} \mid k \in \llbracket 0, n-1 \rrbracket \right\}$.

Démonstration. Soit z dans \mathbb{U}_n . Comme $z \neq 0$ (car $0^n \neq 1$), soit $z = re^{i\theta}$ l'écriture exponentielle de z . Alors :

$$\begin{aligned} z^n = 1 &\iff \left(re^{i\theta}\right)^n = e^{i0} \\ &\iff r^n e^{in\theta} = e^{i0} \\ &\iff r^n = 1 \quad \text{et} \quad n\theta = 0 \pmod{2\pi} \\ &\iff r = 1 \quad (\text{car } r > 0) \quad \text{et} \quad \theta = 0 \pmod{\frac{2\pi}{n}} \quad \text{d'après la proposition 2.2.9} \\ &\iff z = e^{i\frac{2k\pi}{n}}, \quad k \in \mathbb{Z}. \end{aligned}$$

Enfin, la fonction $\theta \mapsto e^{i\theta}$ étant 2π -périodique, la fonction $k \mapsto e^{i\frac{2k\pi}{n}}$ est n -périodique, d'où l'écriture de \mathbb{U}_n . \square

Exemple 2.3.5. Ainsi, $\mathbb{U}_3 = \left\{ 1, e^{i\frac{2\pi}{3}}, e^{i\frac{4\pi}{3}} \right\} = \left\{ 1, -\frac{1}{2} + i\frac{\sqrt{3}}{2}, -\frac{1}{2} - i\frac{\sqrt{3}}{2} \right\} = \{1, j, j^2\}$ où $j = e^{i\frac{2\pi}{3}}$.

Remarque 2.3.6. Pour n dans \mathbb{N}^* fixé, on note usuellement $\omega = e^{i\frac{2\pi}{n}}$ la « première racine $n^{\text{ème}}$ de l’unité non-triviale ». On a ainsi $\mathbb{U}_n = \{\omega^k \mid k \in \llbracket 0, n-1 \rrbracket\}$. Notons que $\forall k \in \llbracket 0, n-1 \rrbracket$, $\overline{\omega^k} = \omega^{-k} = \omega^{n-k} \in \mathbb{U}_n$. De plus :

$$\sum_{z \in \mathbb{U}_n} z = \sum_{k=0}^{n-1} \omega^k = \frac{1 - \omega^n}{1 - \omega} = 0.$$

Remarque 2.3.7 (Interprétation graphique). Les points du plan d’affixe les racines $n^{\text{èmes}}$ de l’unité sont les sommets du polygone régulier à n sommets inscrit dans le cercle unité et passant par le point d’affixe 1. Le barycentre de ces points est l’origine du plan.

2.4 Racines $n^{\text{èmes}}$

Généralisons les remarques précédentes : soient n dans \mathbb{N}^* et a dans \mathbb{C} .

Définition 2.4.1. On appelle *racines $n^{\text{èmes}}$ de a* les solutions de l’équation $z^n = a$.

Proposition 2.4.2. *La seule racine $n^{\text{ème}}$ de 0 est 0. Si $a \neq 0$, notons $a = re^{i\theta}$. On a :*

$$z^n = a \iff z = r^{\frac{1}{n}} e^{i\frac{\theta+2k\pi}{n}} = r^{\frac{1}{n}} e^{i\frac{\theta}{n}} \omega^k, \quad k \in \llbracket 0, n-1 \rrbracket.$$

Remarque 2.4.3. En particulier, on notera que : $\forall (z, z') \in \mathbb{C}^2, z^n = z'^n \iff z = z' \omega^k, k \in \llbracket 0, n-1 \rrbracket$.

3 Géométrie du plan

On revient ici sur quelques aspects du lien étroit entre nombres complexes et géométrie.

3.1 Alignement et orthogonalité

Soient A, B, C trois points distincts du plan, d’affixes respectives z_A, z_B, z_C .

Proposition 3.1.1.

1. *Le vecteur \overrightarrow{AB} a pour affixe $z_B - z_A$. Sa norme vaut donc $|z_B - z_A|$ et son angle avec (Ox) vaut $\arg(z_B - z_A)$.*
2. *L’angle $BAC = (\overrightarrow{AB}, \overrightarrow{AC})$ est égal à $\arg(z_C - z_A) - \arg(z_B - z_A) = \arg\left(\frac{z_C - z_A}{z_B - z_A}\right)$.*

Par conséquent :

Proposition 3.1.2.

1. *(A, B, C sont alignés) $\iff \left(\arg\left(\frac{z_C - z_A}{z_B - z_A}\right) = 0 \pmod{\pi}\right) \iff \left(\frac{z_C - z_A}{z_B - z_A} \in \mathbb{R}\right)$,*
2. *((AB) et (AC) sont perpendiculaires en A) $\iff \left(\arg\left(\frac{z_C - z_A}{z_B - z_A}\right) = \frac{\pi}{2} \pmod{\pi}\right) \iff \left(\frac{z_C - z_A}{z_B - z_A} \in i\mathbb{R}\right)$.*

3.2 Transformations du plan

Définition 3.2.1. Une *transformation du plan* est une bijection $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, ou $f : \mathbb{C} \rightarrow \mathbb{C}$.

Exemple 3.2.2. La fonction $\begin{cases} \mathbb{C} & \rightarrow \mathbb{C} \\ z & \mapsto \bar{z} \end{cases}$ est une transformation du plan. Elle correspond géométriquement à la symétrie d’axe (Ox) .

On s’intéresse ici à deux familles de transformations du plan : les transformations $z \mapsto z + a$, où $a \in \mathbb{C}$, et les transformations $z \mapsto az$, où $a \in \mathbb{C}^*$.

- Pour $a = x + iy \in \mathbb{C}$, la transformation $z \mapsto z + a$ correspond à la *translation* de vecteur (x, y) . Sa réciproque est la translation de vecteur opposé $(-x, -y)$: $z \mapsto z - a$.
- Pour $a = re^{i\theta} \in \mathbb{C}^*$, la transformation $z \mapsto az$ correspond :
 - si $a = r$: à l'*homothétie* de centre O et de rapport r . Sa réciproque est l'homothétie de centre O et de rapport inverse $\frac{1}{r}$: $z \mapsto \frac{z}{r}$,
 - si $a = e^{i\theta}$: à la *rotation* de centre O et d'angle θ . Sa réciproque est la rotation de centre O et d'angle opposé $-\theta$: $z \mapsto e^{-i\theta}z$,
 - si $a = re^{i\theta}$: à la *similitude* de centre O , de rapport r et d'angle θ , c'est-à-dire à la composée de l'homothétie de rapport r et de la rotation d'angle θ . Sa réciproque est la similitude de centre O , de rapport $\frac{1}{r}$ et d'angle $-\theta$: $z \mapsto \frac{1}{r}e^{-i\theta}z$.

Chapitre 5

Fonctions usuelles

1 Logarithmes, exponentielles, puissances

1.1 Logarithmes

Le mot « logarithme » est un néologisme dû à John Napier en 1614, formé à partir des mots λόγος (discours, raison...) et ἀριθμός (nombre). Il s'agit d'une *mise en rapport* d'une progression géométrique (b^n) de base b avec la progression arithmétique (n) de ses puissances. Multiplier deux termes de la progression géométrique revient alors à additionner deux termes de la progression arithmétique !

Une *fonction logarithme* vérifie donc :

$$\forall x, y, \quad \log(xy) = \log(x) + \log(y).$$

En particulier, $\log(1) = 0$. En fixant y et en notant $f = \log'$, on obtient alors : $\forall x, f(x) = \frac{f(1)}{x}$. On définit alors :

Définition 1.1.1. La fonction *logarithme népérien*, notée \ln et définie sur \mathbb{R}_+^* , est la primitive de la fonction inverse ($x \mapsto \frac{1}{x}$) qui s'annule en 1.

Remarque 1.1.2. Toute fonction logarithme est donc un multiple de la fonction \ln .

Par construction, la fonction \ln a les propriétés remarquables suivantes :

Proposition 1.1.3.

1. $\ln(1) = 0$,
2. $\forall (x, y) \in (\mathbb{R}_+^*)^2, \ln(xy) = \ln(x) + \ln(y)$,
3. $\forall x \in \mathbb{R}_+^*, \ln\left(\frac{1}{x}\right) = -\ln(x)$,
4. $\forall (x, n) \in \mathbb{R}_+^* \times \mathbb{Z}, \ln(x^n) = n \ln(x)$.

Étude de la fonction \ln : Comme $\ln' > 0$, la fonction \ln est strictement croissante sur \mathbb{R}_+^* . Comme $\ln'' < 0$, la fonction \ln est concave sur \mathbb{R}_+^* . De plus, $\forall n \in \mathbb{Z}, \ln(2^n) = n \ln(2) \xrightarrow[n \rightarrow \pm\infty]{} \pm\infty$, donc $\ln(x) \xrightarrow[x \rightarrow 0]{} -\infty$ et $\ln(x) \xrightarrow[x \rightarrow +\infty]{} +\infty$.

On retiendra également les formules suivantes :

Proposition 1.1.4.

1. $\forall x > -1, \ln(1+x) \leq x$,
2. $\frac{\ln(1+x)}{x} \xrightarrow[x \rightarrow 0]{} 1$.

Démonstration.

1. La fonction \ln étant concave, elle est en-dessous de sa tangente en 1, d'équation $y = \ln'(1)(x - 1) + \ln(1) = x - 1$, c'est-à-dire : $\forall x > 0$, $\ln(x) \leq x - 1$. On applique ensuite le changement de variables $t = x - 1$.
2. Le terme de gauche est le taux d'accroissement de $f : x \mapsto \ln(1 + x)$ en 0. Sa limite vaut donc $f'(0) = 1$.

□

D'après le théorème de la bijection monotone, la fonction \ln réalise donc une bijection de \mathbb{R}_+^* dans \mathbb{R} . On note en particulier e l'unique antécédent de 1 par \ln : $\ln(e) = 1$. On peut calculer que $e \simeq 2,71828$.

On a alors : $\forall n \in \mathbb{Z}$, $\ln(e^n) = n \ln(e) = n$. La fonction \ln est donc le *logarithme de base e*. On peut alors généraliser : comme le logarithme de base b doit vérifier : $\forall n \in \mathbb{Z}$, $\log(b^n) = n$, on a :

Définition 1.1.5. Soit $b > 0$ différent de 1. La fonction *logarithme de base b*, notée \log_b , est la fonction \log_b :

$$\begin{cases} \mathbb{R}_+^* & \rightarrow \mathbb{R} \\ x & \mapsto \frac{\ln(x)}{\ln(b)} \end{cases} .$$

et donc, par construction :

Proposition 1.1.6.

1. $\log_b(1) = 0$,
2. $\forall n \in \mathbb{Z}$, $\log_b(b^n) = n$,
3. $\forall (a, b, n) \in (\mathbb{R}_+^*)^2 \times \mathbb{Z}$, $a = b^n \iff n = \log_b(a)$.

Démonstration. Exercice. □

Remarque 1.1.7. Les logarithmes les plus utilisés sont les logarithmes de base 2 (logarithme binaire) et 10 (logarithme décimal, parfois simplement noté log).

1.2 Exponentielles

D'après la section précédente, la fonction $\ln : \mathbb{R}_+^* \rightarrow \mathbb{R}$ est bijective. On définit donc :

Définition 1.2.1. La fonction *exponentielle*, notée \exp , définie sur \mathbb{R} , est la bijection réciproque de la fonction \ln .

La fonction \exp a les propriétés suivantes, directement héritées de la fonction \ln :

Proposition 1.2.2.

1. $\forall x \in \mathbb{R}$, $\exp(x) > 0$,
2. $\exp(0) = 1$,
3. $\forall (x, y) \in \mathbb{R}^2$, $\exp(x + y) = \exp(x) \exp(y)$,
4. $\forall x \in \mathbb{R}$, $\exp(-x) = \frac{1}{\exp(x)}$,
5. $\forall (x, n) \in \mathbb{R} \times \mathbb{Z}$, $(\exp(x))^n = \exp(nx)$.

La fonction exponentielle est de plus dérivable car réciproque d'une fonction dérivable, et on a la propriété fondamentale :

Proposition 1.2.3. $\forall x \in \mathbb{R}$, $\exp'(x) = \exp(x)$.

Démonstration. D'après la formule de la dérivée de la réciproque : $\forall x \in \mathbb{R}$, $\exp'(x) = \frac{1}{\ln'(\exp(x))} = \exp(x)$. □

Étude de la fonction \exp : D'après le théorème de la bijection monotone, comme \ln est strictement croissante sur \mathbb{R}_+^* , \exp est croissante sur \mathbb{R} . Comme \ln est concave, \exp est convexe. Enfin, $\exp(x) \xrightarrow{x \rightarrow -\infty} 0$ et $\exp(x) \xrightarrow{x \rightarrow +\infty} +\infty$.

On retiendra également les formules suivantes :

Proposition 1.2.4.

1. $\forall x \in \mathbb{R}$, $\exp(x) \geq 1 + x$,
2. $\frac{\exp(x) - 1}{x} \xrightarrow{x \rightarrow 0} 1$.

Démonstration.

1. La fonction \exp étant convexe, elle est en-dessus de sa tangente en 0, d'équation $y = \exp'(0)x + \exp(0) = x + 1$.
2. Le terme de gauche est le taux d'accroissement de \exp en 0. Sa limite vaut donc $\exp'(0) = 1$.

□

1.3 Puissances

Dans cette partie, on définit progressivement la notion de *puissance réelle* (c'est-à-dire x^α où $\alpha \in \mathbb{R}$). Les *puissances entières* sont définies naturellement :

Définition 1.3.1.

$\forall x \in \mathbb{R}$,

- $x^0 = 1$,
- $\forall n \in \mathbb{N}$, $x^{n+1} = x^n \times x$.

et : $\forall x \in \mathbb{R}^*$, $\forall n \in \mathbb{N}$, $x^{-n} = \frac{1}{x^n}$.

Pour tout n dans \mathbb{Z} , on note $p_n : x \mapsto x^n$, la fonction *puissance n* , définie sur \mathbb{R} si $n \geq 0$ et sur \mathbb{R}^* si $n < 0$.

Proposition 1.3.2. Pour tout n dans \mathbb{N} , la fonction p_n est de même parité que n , de classe C^∞ sur \mathbb{R} , et strictement croissante sur \mathbb{R}_+ .

Démonstration. Exercice. □

On rappelle les opérations suivantes sur les puissances :

Proposition 1.3.3.

1. $\forall (x, y, n) \in \mathbb{R}^2 \times \mathbb{N}$, $(xy)^n = x^n y^n$,
2. $\forall (x, p, q) \in \mathbb{R} \times \mathbb{N}^2$, $x^p x^q = x^{p+q}$,
3. $\forall (x, p, q) \in \mathbb{R} \times \mathbb{N}^2$, $(x^p)^q = x^{pq}$.

Remarque 1.3.4. Attention, $x^{p^q} = x^{(pq)} \neq (x^p)^q$!

On déduit de la proposition 1.3.2 la propriété suivante :

Proposition 1.3.5. Pour tout n dans \mathbb{N}^* , la fonction p_n réalise une bijection de \mathbb{R}_+ dans lui-même. On appelle fonction racine $n^{\text{ème}}$, notée $x \mapsto \sqrt[n]{x}$, la bijection réciproque.

On introduit alors la notation suivante, **valable uniquement sur \mathbb{R}_+** :

Définition 1.3.6. Pour tout (x, n) dans $\mathbb{R}_+ \times \mathbb{N}^*$, on note $x^{\frac{1}{n}} = \sqrt[n]{x}$.

Remarque 1.3.7. La motivation de cette notation est sa compatibilité avec les règles de calcul sur les puissances : ainsi, $\left(x^{\frac{1}{n}}\right)^n = x^{\frac{n}{n}} = x^1 = x$, donc $x \mapsto x^{\frac{1}{n}}$ est bien la réciproque de p_n .

On peut alors étendre la notation puissance à l'ensemble des rationnels :

Définition 1.3.8. Pour tout (x, p, q) dans $\mathbb{R}_+ \times \mathbb{Z} \times \mathbb{N}^*$, on note $x^{\frac{p}{q}} = \sqrt[q]{x^p}$.

À ce stade, on constate l'identité suivante :

Proposition 1.3.9. $\forall (x, r) \in \mathbb{R}_+^* \times \mathbb{Q}$, $x^r = \exp(r \ln(x))$.

Démonstration. Soient $x \in \mathbb{R}_+^*$ et $r = \frac{p}{q} \in \mathbb{Q}$. Alors $(x^r)^q = x^p$, donc $q \ln(x^r) = p \ln(x)$, donc $\ln(x^r) = \frac{p}{q} \ln(x) = r \ln(x)$, et donc $x^r = \exp(r \ln(x))$. \square

On étend finalement à l'ensemble des puissances réelles, **uniquement sur \mathbb{R}_+^*** :

Définition 1.3.10. $\forall (x, \alpha) \in \mathbb{R}_+^* \times \mathbb{R}$, $x^\alpha = \exp(\alpha \ln(x))$.

Les règles de calcul sur les puissances réelles sont alors les mêmes que pour les puissances entières :

Proposition 1.3.11.

1. $\forall (x, y, \alpha) \in (\mathbb{R}_+^*)^2 \times \mathbb{R}$, $(xy)^\alpha = x^\alpha y^\alpha$,
2. $\forall (x, \alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{R}^2$, $x^\alpha x^\beta = x^{\alpha+\beta}$,
3. $\forall (x, \alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{R}^2$, $(x^\alpha)^\beta = x^{\alpha\beta}$.

Remarque 1.3.12. En particulier, pour tout $a \in \mathbb{R}$, $e^a = \exp(a \ln(e)) = \exp(a)$. On notera alors indifféremment \exp ou e^{\cdot} la fonction exponentielle.

1.4 Croissances comparées

On a vu que les fonctions logarithmes, exponentielles et puissances sont croissantes tendent toutes vers l'infini (elles divergent vers $+\infty$) en $+\infty$. Mais leurs *vitesses de divergence* diffèrent, et on peut les comparer ainsi :

Proposition 1.4.1. $\forall a > 0$, $\frac{\ln(x)}{x^a} \xrightarrow{x \rightarrow +\infty} 0$.

Démonstration. On traite d'abord le cas $a = 1$: la fonction $f : x \mapsto \frac{\ln(x)}{x}$ est dérivable sur \mathbb{R}_+^* , et : $\forall x \in \mathbb{R}_+^*$, $f'(x) = \frac{1 - \ln(x)}{x^2}$, donc f est décroissante sur $[e, +\infty[$, et par ailleurs positive sur cet intervalle. Elle admet donc une limite $l \geq 0$ en $+\infty$.

De plus : $\forall x \in \mathbb{R}_+^*$, $f(2x) = \frac{\ln(2x)}{2x} = \frac{\ln(2)}{2x} + \frac{\ln(x)}{2x} \xrightarrow{x \rightarrow +\infty} \frac{1}{2}l$, donc $l = \frac{1}{2}l$, donc $l = 0$.

Dans le cas général, on fait le changement de variable $t = x^a$. Alors $\frac{\ln(x)}{x^a} = \frac{\ln(t^{\frac{1}{a}})}{t} = \frac{1}{a} \frac{\ln(t)}{t} \xrightarrow{x \rightarrow +\infty} 0$ puisque $t \xrightarrow{x \rightarrow +\infty} +\infty$. \square

Remarque 1.4.2. On dit que *les puissances l'emportent sur le logarithme*.

On en déduit les limites suivantes :

Proposition 1.4.3. $\forall a > 0$,

1. $x^a \ln(x) \xrightarrow{x \rightarrow 0} 0$,
2. $x^a e^{-x} \xrightarrow{x \rightarrow +\infty} 0$,
3. $x^{-a} e^x \xrightarrow{x \rightarrow +\infty} +\infty$.

Remarque 1.4.4. On dit que *l'exponentielle l'emporte sur les puissances*.

2 Fonctions trigonométriques

On a vu dans les chapitres précédents que les restrictions suivantes des fonctions trigonométriques sont bijectives :

- $\cos : [0, \pi] \rightarrow [-1, 1]$,
- $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1]$,
- $\tan : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\rightarrow \mathbb{R}$,

de bijections réciproques :

- $\arccos : [-1, 1] \rightarrow [0, \pi]$,
- $\arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$,
- $\arctan : \mathbb{R} \rightarrow \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.

Par définition de ces fonctions, on a directement :

Proposition 2.0.1.

- $\forall x \in [-1, 1], \sin(\arccos x) = \sqrt{1 - x^2}$,
- $\forall x \in [-1, 1], \cos(\arcsin x) = \sqrt{1 - x^2}$.

Proposition 2.0.2. *Les fonctions trigonométriques réciproques sont dérивables, respectivement sur $] -1, 1[$ et sur \mathbb{R} , de dérivées respectives :*

- $\forall x \in] -1, 1[, \arccos'(x) = -\frac{1}{\sqrt{1 - x^2}}$,
- $\forall x \in] -1, 1[, \arcsin'(x) = \frac{1}{\sqrt{1 - x^2}}$,
- $\forall x \in \mathbb{R}, \arctan'(x) = \frac{1}{1 + x^2}$.

Démonstration. On applique la formule de la dérivée de la réciproque :

- Soit $x \in] -1, 1[$. $\arccos'(x) = \frac{1}{\cos'(\arccos(x))} = -\frac{1}{\sin(\arccos(x))} = -\frac{1}{\sqrt{1 - x^2}}$.
- De même pour \arcsin .
- Soit $x \in \mathbb{R}$. $\arctan'(x) = \frac{1}{\tan'(\arctan(x))} = \frac{1}{1 + \tan^2(\arctan(x))} = \frac{1}{1 + x^2}$.

□

On en déduit les formules suivantes :

Proposition 2.0.3.

1. $\forall x \in [-1, 1], \arccos(x) + \arcsin(x) = \frac{\pi}{2}$,
2. $\forall x \in \mathbb{R}^*, \arctan(x) + \arctan\left(\frac{1}{x}\right) = \text{sgn}(x)\frac{\pi}{2}$, où $\text{sgn}(x) = 1$ si $x > 0$, -1 si $x < 0$.

Démonstration.

1. D'après la propriété précédente, $(\arccos + \arcsin)' = 0$ sur $] -1, 1[$, donc la fonction $\arccos + \arcsin$ est constante sur $[-1, 1]$, égale à $\arccos(0) + \arcsin(0) = \frac{\pi}{2}$.
2. Soit $f : x \mapsto \arctan(x) + \arctan\left(\frac{1}{x}\right)$, définie sur \mathbb{R}^* . Cette fonction est dérivable, et : $\forall x \in \mathbb{R}^*, f'(x) = 0$, donc f est constante sur tout *intervalle* inclus dans \mathbb{R}^* . Donc sur \mathbb{R}_+^* , f est constante égale à $f(1) = 2\arctan(1) = \frac{\pi}{2}$, et sur \mathbb{R}_-^* , f est constante égale à $f(-1) = 2\arctan(-1) = -\frac{\pi}{2}$.

□

On retiendra également les formules suivantes :

Proposition 2.0.4.

1. $\frac{\sin(x)}{x} \xrightarrow{x \rightarrow 0} 1$,
2. $\frac{\cos(x) - 1}{x} \xrightarrow{x \rightarrow 0} 0$,
3. $\frac{1 - \cos(x)}{x^2} \xrightarrow{x \rightarrow 0} \frac{1}{2}$.

Démonstration. Les deux premières formules correspondent aux taux d'accroissement respectifs en 0 des fonctions \sin et \cos , de limites respectives $\sin'(0) = 1$ et $\cos'(0) = 0$. Pour la troisième formule :

$$\frac{1 - \cos(x)}{x^2} = \frac{2 \sin^2(\frac{x}{2})}{x^2} = \frac{1}{2} \left(\frac{\sin(\frac{x}{2})}{\frac{x}{2}} \right)^2 \xrightarrow{x \rightarrow 0} \frac{1}{2}.$$

ainsi que les inégalités :

Proposition 2.0.5.

1. $\forall x \geq 0, \sin(x) \leq x$ et $\arcsin(x) \geq x$,
2. $\forall x \geq 0, \tan(x) \geq x$ et $\arctan(x) \leq x$.

3 Fonctions hyperboliques

On rappelle les définitions des fonctions hyperboliques, déjà croisées dans le chapitre 2 comme parties paire et impaire de la fonction exponentielle :

Définition 3.0.1.

La fonction *cosinus hyperbolique*, notée ch , est la fonction définie sur \mathbb{R} par : $\forall x \in \mathbb{R}, \text{ch}(x) = \frac{e^x + e^{-x}}{2}$.

La fonction *sinus hyperbolique*, notée sh , est la fonction définie sur \mathbb{R} par : $\forall x \in \mathbb{R}, \text{sh}(x) = \frac{e^x - e^{-x}}{2}$.

Étude des fonctions hyperboliques : Les fonctions ch et sh sont respectivement paire et impaire. Ces deux fonctions sont dérivables, avec $\text{ch}' = \text{sh}$ et $\text{sh}' = \text{ch}$. Étant de plus strictement croissantes sur \mathbb{R}_+ avec $\text{ch}(\mathbb{R}_+) = [1, +\infty[$ et $\text{sh}(\mathbb{R}_+) = \mathbb{R}_+$, les restrictions suivantes sont bijectives :

- $\text{ch} : \mathbb{R}_+ \rightarrow [1, +\infty[$,
- $\text{sh} : \mathbb{R} \rightarrow \mathbb{R}$,

de bijections réciproques :

- $\text{argch} : [1, +\infty[\rightarrow \mathbb{R}_+$,
- $\text{argsh} : \mathbb{R} \rightarrow \mathbb{R}$.

Remarque 3.0.2. Ces fonctions s'expriment explicitement :

$$\forall x \in \mathbb{R}, \text{argsh}(x) = \ln \left(x + \sqrt{x^2 + 1} \right) \quad \text{et} \quad \forall x \geq 1, \text{argch}(x) = \ln \left(x + \sqrt{x^2 - 1} \right).$$

Proposition 3.0.3. *Les fonctions hyperboliques réciproques sont dérivables, respectivement sur $]1, +\infty[$ et sur \mathbb{R} , de dérivées respectives :*

- $\forall x \in]1, +\infty[, \text{argch}'(x) = \frac{1}{\sqrt{x^2 - 1}}$,

- $\forall x \in \mathbb{R}, \arg \operatorname{sh}'(x) = \frac{1}{\sqrt{1+x^2}}.$

On retiendra également les formules suivantes, à mettre en regard des formules de trigonométrie vues au chapitre 3 :

Proposition 3.0.4.

1. $\forall x \in \mathbb{R}, \operatorname{ch}^2(x) - \operatorname{sh}^2(x) = 1,$
2. $\forall (x, y) \in \mathbb{R}^2, \operatorname{ch}(x+y) = \operatorname{ch}(x)\operatorname{ch}(y) + \operatorname{sh}(x)\operatorname{sh}(y),$
3. $\forall (x, y) \in \mathbb{R}^2, \operatorname{sh}(x+y) = \operatorname{sh}(x)\operatorname{ch}(y) + \operatorname{ch}(x)\operatorname{sh}(y).$

Proposition 3.0.5.

1. $\frac{\operatorname{sh}(x)}{x} \xrightarrow{x \rightarrow 0} 1,$
2. $\frac{\operatorname{ch}(x) - 1}{x} \xrightarrow{x \rightarrow 0} 0,$
3. $\frac{\operatorname{ch}(x) - 1}{x^2} \xrightarrow{x \rightarrow 0} \frac{1}{2}.$

Chapitre 6

Ensembles et applications

On développe dans ce chapitre la notion d'*ensemble*, dont on a vu qu'elle permet de définir tous les autres objets mathématiques (suites, fonctions, espaces, probabilités...), ainsi que celle d'*application*, qui permet de relier des ensembles entre eux.

1 Ensembles

Le concept d'*ensemble* permet de définir toutes les autres notions vues cette année. Autrement dit, la théorie des ensembles est utilisée comme *fondement* des mathématiques. Ce point de vue est aujourd'hui très majoritaire dans la communauté mathématique, même si d'autres fondements (par exemple la théorie des *topos*) sont proposés pour lever certains paradoxes inhérents à la théorie des ensembles (paradoxes de Russell, de Banach-Tarski...).

1.1 Définition

Définition 1.1.1. Un *ensemble* est une collection d'objets, appelés *éléments*. On dit d'un élément x d'un ensemble E qu'il *appartient* à E , et on note $x \in E$.

Le nombre d'éléments que contient un ensemble E est appelé son *cardinal*, et noté $\text{Card}(E)$. On dit que E est un ensemble *fini* si son cardinal est fini.

L'ensemble ne contenant aucun élément est noté \emptyset , et appelé l'*ensemble vide*. Il est de cardinal 0.

Exemple 1.1.2.

- Les nombres 1, 5 et 34 sont des éléments de l'ensemble des entiers naturels, noté \mathbb{N} . Le cardinal de \mathbb{N} est infini.
- L'ensemble collectionnant les fonctions \cos , \sin et \ln est l'ensemble $\{\cos, \sin, \ln\}$. Cet ensemble est de cardinal 3.

Remarque 1.1.3.

- Lorsqu'il ne porte pas de nom spécifique, un ensemble se note entre accolades : $E = \{\dots\}$.
- Dans un ensemble, les éléments ne sont ni ordonnés, ni répétés. Ainsi, $\{2, 1\} = \{1, 2\} = \{1, 2, 2, 1\}$.
- Un ensemble peut être défini implicitement (écriture *en compréhension*) (par exemple : $\mathbb{U}_5 = \{z \in \mathbb{C} \mid z^5 = 1\}$) ou explicitement (écriture *en extension*) ($\mathbb{U}_5 = \{1, e^{\frac{2i\pi}{5}}, e^{\frac{4i\pi}{5}}, e^{\frac{6i\pi}{5}}, e^{\frac{8i\pi}{5}}\}$).

Définition 1.1.4. Un *sous-ensemble* (ou une *partie*) A d'un ensemble B est un ensemble dont tous les éléments sont également des éléments de B . On dit alors que A est *inclus dans* B , et on note $A \subset B$.

Exemple 1.1.5.

- L'ensemble $\{1, 5, 34\}$ est un sous-ensemble de \mathbb{N} .
- L'ensemble vide est un sous-ensemble de n'importe quel ensemble.

Remarque 1.1.6. Attention à bien distinguer les notions d'élément et de sous-ensemble. Par exemple, $\sqrt{2}$ est un **élément** de \mathbb{R} , et on note $\sqrt{2} \in \mathbb{R}$. Par contre, $\{\sqrt{2}\}$ est un **sous-ensemble** de \mathbb{R} , et on note $\{\sqrt{2}\} \subset \mathbb{R}$. (Un ensemble ne contenant qu'un élément, de cardinal 1 donc, est par ailleurs appelé un *singleton*).

Proposition 1.1.7. Soient A et B deux ensembles. Alors $A = B \Leftrightarrow (A \subset B \text{ et } B \subset A)$.

Définition 1.1.8. Soit E un ensemble. L'*ensemble des parties* de E est l'ensemble des sous-ensembles de E , noté $\mathcal{P}(E)$.

Exemple 1.1.9. $\mathcal{P}(\{1, 4, 5\}) = \{\emptyset, \{1\}, \{4\}, \{5\}, \{1, 4\}, \{1, 5\}, \{4, 5\}, \{1, 4, 5\}\}$.

Proposition 1.1.10. Si E est un ensemble fini, alors $\mathcal{P}(E)$ l'est aussi et $\text{Card}(\mathcal{P}(E)) = 2^{\text{Card}(E)}$.

Démonstration. Construire un sous-ensemble de E consiste à faire le choix, pour chaque élément x de E , de le prendre ou non. Il y a donc 2 possibilités pour chaque élément, donc $2^{\text{Card}(E)}$ possibilités en tout. \square

1.2 Opérations

Définition 1.2.1. Soit E un ensemble et A et B deux sous-ensembles de E .

- Le *complémentaire* de A dans E , noté $E \setminus A$ ou \bar{A} , est l'ensemble des éléments de E qui n'appartiennent pas à A : $\bar{A} = \{x \in E \mid x \notin A\}$.
- L'*intersection* de A et B , notée $A \cap B$, est l'ensemble des éléments de E qui appartiennent à la fois à A et à B : $A \cap B = \{x \in E \mid x \in A \text{ et } x \in B\}$.
- L'*union* de A et B , notée $A \cup B$, est l'ensemble des éléments de E qui appartiennent à A ou à B : $A \cup B = \{x \in E \mid x \in A \text{ ou } x \in B\}$.

On peut constater à ce stade une correspondance entre le langage des ensembles et celui des assertions :

Ensembles	Assertions
$x \in \bar{A}$	$\neg(x \in A)$
$x \in A \cap B$	$(x \in A) \wedge (x \in B)$
$x \in A \cup B$	$(x \in A) \vee (x \in B)$
$A \subset B$	$(x \in A) \Rightarrow (x \in B)$
$A = B$	$(x \in A) \Leftrightarrow (x \in B)$

On obtient ainsi directement les propriétés suivantes, héritées des propriétés des assertions :

Proposition 1.2.2. Soient A , B et C des sous-ensembles d'un ensemble E .

- *Lois de De Morgan* :

$$\begin{aligned} \overline{A \cap B} &= \bar{A} \cup \bar{B}, \\ \overline{A \cup B} &= \bar{A} \cap \bar{B}. \end{aligned}$$

- *Associativité* :

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

- *Distributivité* :

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

Les notions d'intersection et d'union se généralisent avec les notations suivantes, semblables aux notations somme et produit vues dans le premier chapitre :

Définition 1.2.3. Soient I un ensemble et $(A_i)_{i \in I}$ une famille de sous-ensembles d'un ensemble E indexée par I . On note $\bigcap_{i \in I} A_i$ l'intersection de tous les A_i , et $\bigcup_{i \in I} A_i$ leur union.

Exemple 1.2.4. $\bigcap_{n \in \mathbb{N}^*} \left[-\frac{1}{n}, \frac{1}{n} \right] = \{0\}.$

On peut alors considérer la notion suivante :

Définition 1.2.5. Soient I un ensemble et $(A_i)_{i \in I}$ une famille de sous-ensembles d'un ensemble E indexée par I . On dit que $(A_i)_{i \in I}$ est une *partition* de E si :

- $\forall i \in I, A_i \neq \emptyset,$
- $\forall i \neq j, A_i \cap A_j = \emptyset,$
- $\bigcup_{i \in I} A_i = E.$

Exemple 1.2.6. Notons $2\mathbb{N}$ l'ensemble des entiers naturels pairs et $2\mathbb{N} + 1$ l'ensemble des entiers naturels impairs. Alors la famille $(2\mathbb{N}, 2\mathbb{N} + 1)$ est une partition de \mathbb{N} .

1.3 Produit cartésien

Définition 1.3.1. Soient E et F deux ensembles. Le *produit cartésien* de E par F est l'ensemble

$$E \times F = \{(x, y) \mid x \in E, y \in F\}.$$

On note alors, pour $n \in \mathbb{N}$: $E^n = E \times \dots \times E$ (avec n facteurs).

Remarque 1.3.2. Ici, (x, y) désigne un *couple* d'éléments, c'est-à-dire une famille *ordonnée* de deux éléments. Le couple (y, x) est donc distinct du couple (x, y) (sauf si $x = y$). Par conséquent, le produit $F \times E$ est distinct de $E \times F$.

Exemple 1.3.3. L'ensemble $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ s'identifie au plan muni d'un repère cartésien. Le sous-ensemble $\mathbb{R} \times [0, 1]$ de \mathbb{R}^2 désigne alors une bande horizontale comprise entre les droites d'équation $y = 0$ et $y = 1$, tandis que le sous-ensemble $[0, 1] \times \mathbb{R}$ désigne une bande verticale comprise entre les droites d'équation $x = 0$ et $x = 1$.

Proposition 1.3.4. Si E et F sont des ensembles finis, alors $E \times F$ l'est aussi et $\text{Card}(E \times F) = \text{Card}(E) \times \text{Card}(F)$. De même, $\text{Card}(E^n) = (\text{Card}(E))^n$.

Démonstration. Construire un élément (x, y) de $E \times F$ consiste à choisir un élément x de E et un élément y de F . Il y a $\text{Card}(E)$ possibilités pour x et $\text{Card}(F)$ possibilités pour y , donc $\text{Card}(E) \times \text{Card}(F)$ possibilités en tout. De même pour E^n par une récurrence évidente. \square

2 Applications

Comme on l'a vu pour les fonctions réelles, la notion de produit cartésien de deux ensembles permet de mettre les éléments de ces ensembles en correspondance :

2.1 Définition

Définition 2.1.1. Soient E et F des ensembles. Une *application* f de E vers F est la donnée d'une partie Γ de $E \times F$ telle que : $\forall x \in E, \exists! y \in F, (x, y) \in \Gamma$. On note alors $y = f(x)$ et $f : \begin{cases} E & \rightarrow F \\ x & \mapsto f(x) \end{cases}$.

Exemple 2.1.2.

- $f : \begin{cases} \mathbb{C}^n & \rightarrow \mathbb{C} \\ (z_1, \dots, z_n) & \mapsto \sum_{k=1}^n z_k \end{cases},$
- $D : \begin{cases} \mathbb{C}^1(\mathbb{R}, \mathbb{R}) & \rightarrow \mathbb{C}^0(\mathbb{R}, \mathbb{R}) \\ f & \mapsto f' \end{cases}$ (opérateur de dérivation),

- Pour $a \in \mathbb{R}$, $\text{ev}_a : \begin{cases} \mathcal{F}(\mathbb{R}, \mathbb{R}) & \rightarrow \mathbb{R} \\ f & \mapsto f(a) \end{cases}$ (application évaluation en a).

On retrouve dans ce contexte une bonne partie du vocabulaire des fonctions réelles :

Définition 2.1.3. Soit $f : E \rightarrow F$ une application.

- Soit (x, y) dans $E \times F$. Si $y = f(x)$, on dit que y est l'*image* de x par f et que x est un **antécédent** de y par f ,
- E est l'*ensemble de départ* de f , F est son *ensemble d'arrivée*, $f(E) = \{f(x) \mid x \in E\}$ est son *ensemble image*, et Γ est son *graphe*.

Enfin, on note $\mathcal{F}(E, F)$ (ou F^E) l'ensemble des applications de E vers F .

Cette dernière notation se justifie par la propriété suivante :

Proposition 2.1.4. Si E et F sont des ensembles finis, alors $\mathcal{F}(E, F)$ l'est aussi et $\text{Card}(\mathcal{F}(E, F)) = (\text{Card}(F))^{\text{Card}(E)}$.

Démonstration. Construire une application $f : E \rightarrow F$ consiste à choisir, pour chaque x dans E , un élément y de F . Il y a $\text{Card}(F)$ possibilités pour chaque x , donc $(\text{Card}(F))^{\text{Card}(E)}$ possibilités en tout. \square

On peut en particulier définir les applications et les types d'applications suivants :

- l'*application identité* $\text{id}_E : \begin{cases} E & \rightarrow E \\ x & \mapsto x, \end{cases}$,
- pour $A \subset E$, l'*application indicatrice* $1_A : \begin{cases} E & \rightarrow \{0, 1\} \\ x & \mapsto 1 \text{ si } x \in A, 0 \text{ sinon} \end{cases}$,
- les *suites* $u : \begin{cases} \mathbb{N} & \rightarrow \mathbb{R} \\ n & \mapsto u(n) = u_n \end{cases}$,
- et plus généralement les *familles indexées* par un ensemble I , $a : \begin{cases} I & \rightarrow E \\ i & \mapsto a_i \end{cases}$.

Définition 2.1.5. Soit E et F des ensembles, $f : E \rightarrow F$ une application, A un sous-ensemble de E et B un sous-ensemble de F . On appelle :

- *image directe* de A par f l'ensemble $f(A) = \{f(x) \in F \mid x \in A\}$,
- *image réciproque* de B par f l'ensemble $f^{-1}(B) = \{x \in E \mid f(x) \in B\}$.

Remarque 2.1.6. L'écriture $f^{-1}(B)$ ci-dessus est une notation, qui ne présuppose pas l'existence d'une application réciproque f^{-1} . L'existence d'une telle application dépend des propriétés de f décrites ci-après.

2.2 Injections, surjections, bijections

Comme pour les fonctions réelles, on s'intéresse au comportement des applications vis-à-vis de leurs antécédents :

Définition 2.2.1. Soit $f : E \rightarrow F$ une application.

- Si tout élément y de F a **au plus** un antécédent par f , on dit que f est *injective*.
- Si tout élément y de F a **au moins** un antécédent par f , on dit que f est *surjective*.
- Si f est à la fois injective et surjective, on dit que f est *bijective*.

Dans ce dernier cas, tout élément y de F a **exactement** un antécédent par $f : \forall y \in F, \exists! x \in E, (x, y) \in \Gamma_f$. L'application f admet donc une *application réciproque* $f^{-1} : F \rightarrow E$, définie par : $\forall y \in F, f^{-1}(y)$ est l'unique x dans E tel que $y = f(x)$.

Exemples 2.2.2.

- $\begin{cases} \mathbb{C} & \rightarrow \mathbb{R} \\ z & \mapsto \text{Re}(z) \end{cases}$ est surjective, non injective,

- $\begin{cases} \mathbb{R}^* \times [0, 2\pi[& \rightarrow \mathbb{R}^2 \\ (r, \theta) & \mapsto (r \cos \theta, r \sin \theta) \end{cases}$ est injective, non surjective,
- $\begin{cases} \mathbb{N} & \rightarrow \mathbb{Z} \\ n & \mapsto k \text{ si } n = 2k, -k \text{ si } n = 2k - 1 \end{cases}$ est bijective.

Ces propriétés ont des conséquences sur les cardinaux des ensembles :

Proposition 2.2.3. *Soient E et F deux ensembles finis, et soit $f : E \rightarrow F$.*

- *Si f est injective, alors $\text{Card}(E) \leq \text{Card}(F)$,*
- *Si f est surjective, alors $\text{Card}(E) \geq \text{Card}(F)$,*
- *Si f est bijective, alors $\text{Card}(E) = \text{Card}(F)$.*

Démonstration. Notons qu'on a toujours $\text{Card}(f(E)) \leq \text{Card}(E)$ (car $\forall y \in f(E), \exists! x \in E, y = f(x)$) et $\text{Card}(f(E)) \leq \text{Card}(F)$ (car $f(E) \subset F$). Alors :

- Supposons f injective. Alors $\text{Card}(f(E)) = \text{Card}(E)$ (car $x \neq y \Rightarrow f(x) \neq f(y)$), donc $\text{Card}(E) \leq \text{Card}(F)$.
- Supposons f surjective. Alors $f(E) = F$, donc $\text{Card}(F) \leq \text{Card}(E)$.
- La dernière assertion est conséquence directe des deux premières.

□

A contrario, les cardinaux des ensembles ont des conséquences sur les applications :

Proposition 2.2.4. *Soient E et F deux ensembles finis avec $\text{Card}(E) = \text{Card}(F)$, et soit $f : E \rightarrow F$. Alors : $(f \text{ est injective}) \Leftrightarrow (f \text{ est surjective}) \Leftrightarrow (f \text{ est bijective})$.*

Enfin, dans le cas d'une application composée, on a les propriétés suivantes :

Proposition 2.2.5. *Soient $u : E \rightarrow F$ et $v : F \rightarrow G$ deux applications. On a :*

1. *Si u et v sont injectives, alors $v \circ u$ est injective,*
2. *Si u et v sont surjectives, alors $v \circ u$ est surjective,*
3. *Si $v \circ u$ est surjective, alors v est surjective,*
4. *Si $v \circ u$ est injective, alors u est injective,*
5. *Si u et v sont bijectives, alors $v \circ u$ est bijective, et $(v \circ u)^{-1} = u^{-1} \circ v^{-1}$.*

Chapitre 7

Arithmétique

Dans ce chapitre, les seuls nombres « autorisés » sont les nombres entiers. Les équations $2x - 3 = 0$ et $x^2 = 2$ n'ont donc pas de solution. Cette considération est essentielle dans l'étude de la divisibilité : le nombre 7 est premier en arithmétique, pourtant $7 = \frac{7}{2} \times 2\dots$

1 Nombres premiers et divisibilité

Définition 1.0.1. Soit (d, n) dans \mathbb{Z}^2 . On dit que d divise n , ou que d est un *diviseur* de n , ou encore que n est un *multiple* de d , et on note $d|n$, s'il existe q dans \mathbb{Z} tel que $n = dq$. On dit alors que q est le *quotient* de n par d .

Exemple 1.0.2.

- Les diviseurs de 12 sont 1, 2, 3, 4, 6, 12 et leurs opposés.
Les premiers multiples positifs de 12 sont 0, 12, 24, 36, 48\dots
- Pour tout entier n , 1 est un diviseur de n (n est un multiple de 1) puisque $n = 1 \times n$. Les seuls diviseurs de 1 sont 1 et -1 .
- Pour tout entier n , 0 est un multiple de n (n est un diviseur de 0) puisque $0 = n \times 0$. Le seul multiple de 0 est 0.

Proposition 1.0.3. Soit (d, m, n) dans \mathbb{Z}^3 .

1. Si $d|m$ et $d|n$, alors, pour tout (a, b) dans \mathbb{Z}^2 , $d|am + bn$.
2. Si $d|m$ et $m|n$, alors $d|n$.

Démonstration.

1. Si $m = kd$ et $n = ld$, alors $am + bn = d(ak + bl)$, donc $d|am + bn$.
2. Si $m = kd$ et $n = pm$, alors $n = (pk)d$, donc $d|n$.

□

Remarque 1.0.4.

- $\forall (d, n) \in \mathbb{Z}^2$, $d|n \Leftrightarrow -d|n$. On ne mentionne donc en général que les diviseurs positifs d'un entier.
- Tout entier est au moins divisible par 1 et par lui-même (et par leurs opposés) ; ce sont ses diviseurs *triviaux*.
Les entiers n'ayant aucun autre diviseur sont particulièrement intéressants.

Définition 1.0.5. Un *nombre premier* est un entier naturel ayant exactement deux diviseurs positifs distincts. On note \mathcal{P} l'ensemble des nombres premiers.

Cette définition correspond aux entiers naturels n'ayant pas de diviseur non trivial, **hormis** 1, qui n'a qu'un diviseur positif. Les premiers nombres premiers sont :

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47\dots

Proposition 1.0.6. *Tout entier est divisible par un nombre premier.*

Théorème 1.0.7. *L'ensemble \mathcal{P} des nombres premiers est infini.*

Démonstration. On raisonne par l'absurde : supposons que \mathcal{P} soit fini, de cardinal n . Notons $\mathcal{P} = \{p_1, \dots, p_n\}$. Soit $N = p_1 \times p_2 \times \dots \times p_n + 1$. D'après le théorème 1.0.9, il existe $i \in \llbracket 1, n \rrbracket$ tel que p_i divise N . Or p_i divise $p_1 \times p_2 \times \dots \times p_n$, donc d'après la proposition 1.0.3, p_i divise $N - p_1 \times p_2 \times \dots \times p_n = 1$. Comme aucun diviseur de 1 n'est premier, cela est absurde. Donc \mathcal{P} est infini. \square

Un entier $n \geq 2$ est non premier si et seulement s'il admet des diviseurs *non triviaux*, c'est-à-dire qu'il existe $(n_1, n_2) \in \llbracket 2, n-1 \rrbracket$ tel que $n = n_1 n_2$. En supposant par exemple $n_1 \leq n_2$, on voit que $n \geq n_1^2$, donc $n_1 \leq \sqrt{n}$. On en déduit la propriété suivante :

Proposition 1.0.8. *Un entier n est premier si et seulement s'il n'admet aucun diviseur dans $[2, \sqrt{n}] \cap \mathcal{P}$.*

On en déduit un algorithme de détermination des nombres premiers inférieurs à $N \in \mathbb{N}$, dit du *crible d'Ératosthène* : on considère l'ensemble $\llbracket 2, N \rrbracket$, et on le parcourt dans l'ordre croissant jusqu'à \sqrt{N} , en lui enlevant à chaque étape les multiples stricts du premier nombre rencontré. Par exemple, à la première étape, on enlève tous les nombres pairs strictement supérieurs à 2 ; puis à la deuxième étape, tous les multiples de 3 strictement supérieurs à 3. Les nombres restants à la fin de l'algorithme sont les nombres premiers inférieurs à N .

Les nombres premiers sont les « atomes » des entiers par rapport à la multiplication, comme le formalise le théorème suivant.

Théorème 1.0.9 (Théorème fondamental de l'arithmétique). *Soit $n \geq 2$ entier. Il existe un unique $r \in \mathbb{N}$, puis $(p_1, \dots, p_r) \in \mathcal{P}^r$ et $(\alpha_1, \dots, \alpha_r) \in (\mathbb{N}^*)^r$ uniques à l'ordre près, tels que*

$$n = \prod_{i=1}^r p_i^{\alpha_i}.$$

Ce produit est appelé *décomposition en facteurs premiers de n* , et est donc unique à l'ordre des facteurs près.

Exemple 1.0.10.

- Comme $28 = 4 \times 7 = 2 \times 2 \times 7$, la décomposition en facteurs premiers de 28 est $2^2 \times 7^1$.
- De même, la décomposition en facteurs premiers de 2022 est $2^1 \times 3^1 \times 337^1$, et celle de 2023 est $7^1 \times 17^2$.

Les nombres premiers permettent par exemple de caractériser les diviseurs d'un entier donné :

Proposition 1.0.11. *Soit $n \geq 2$, et soit $n = \prod_{i=1}^r p_i^{\alpha_i}$ sa décomposition en facteurs premiers. Les diviseurs positifs*

de n sont alors les $n = \prod_{i=1}^r p_i^{\beta_i}$, où : $\forall i \in \llbracket 1, r \rrbracket$, $0 \leq \beta_i \leq \alpha_i$.

Les nombres premiers sont essentiels en arithmétique, entre autres. Ils ne sont pas encore complètement compris ; on sait par exemple que :

Remarque 1.0.12.

- Le plus grand nombre premier connu en novembre 2025 (découvert le 12 octobre 2024) est $2^{136\,279\,841} - 1$. C'est un nombre premier de la forme $2^k - 1$, dit *de Mersenne*.
- Les nombres premiers ont tendance à se raréfier à mesure que l'on avance dans \mathbb{N} . En notant $\pi(N)$ le nombre de nombres premiers inférieurs à N , le *théorème des nombres premiers* affirme que $\pi(N) \simeq \frac{N}{\ln(N)}$.
- La *répartition des nombres premiers*, c'est-à-dire l'existence d'une formule donnant directement le $n^{\text{ème}}$ nombre premier en fonction de n , est un problème ouvert (la *formule de Minác et Willans* (1995) utilise pour cela le *théorème de Wilson*, mais est impraticable). La résolution de l'*hypothèse de Riemann* (l'un des *problèmes du millénaire*) serait un pas vers une telle formule.

2 PGCD et PPCM

Définition 2.0.1. Soit (m, n) dans \mathbb{N}^2 .

- Le *PGCD* de m et n , noté $PGCD(m, n)$ ou $m \wedge n$, est le plus grand diviseur commun à m et n .
- Le *PPCM* de m et n , noté $PPCM(m, n)$ ou $m \vee n$, est le plus petit multiple commun à m et n .

Le PGCD et le PPCM peuvent directement être calculés à partir des décompositions en nombres premiers de m et n . On utilise pour cela la notion suivante :

Définition 2.0.2. Soit (m, n) dans $(\mathbb{N}^*)^2$. Notons \mathcal{P}_m et \mathcal{P}_n les facteurs premiers de m et n respectivement, puis $\mathcal{P}_{(m,n)} = \mathcal{P}_m \cup \mathcal{P}_n$. On appelle *décompositions communes à m et n* les décompositions :

$$m = \prod_{i=1}^r p_i^{\alpha_i} \quad \text{et} \quad n = \prod_{i=1}^r p_i^{\beta_i},$$

où $r \in \mathbb{N}$, $(p_1, \dots, p_r) \in \mathcal{P}_{(m,n)}$, $(\alpha_1, \dots, \alpha_r) \in \mathbb{N}^r$ et $(\beta_1, \dots, \beta_r) \in \mathbb{N}^r$.

Exemple 2.0.3. Les décompositions communes à $12 = 2^2 \times 3$ et $56 = 2^3 \times 7$ sont $12 = 2^2 \times 3^1 \times 7^0$ et $56 = 2^3 \times 3^0 \times 7^1$.

Proposition 2.0.4. Soit (m, n) dans $(\mathbb{N}^*)^2$, et soient $m = \prod_{i=1}^r p_i^{\alpha_i}$ et $n = \prod_{i=1}^r p_i^{\beta_i}$ leurs décompositions en facteurs premiers communes. Alors

$$m \wedge n = \prod_{i=1}^r p_i^{\min(\alpha_i, \beta_i)} \quad \text{et} \quad m \vee n = \prod_{i=1}^r p_i^{\max(\alpha_i, \beta_i)}.$$

Exemple 2.0.5. $12 \wedge 56 = 2^2 \times 3^0 \times 7^0 = 4$ et $12 \vee 56 = 2^3 \times 3^1 \times 7^1 = 168$.

Corollaire 2.0.6. $\forall (m, n) \in (\mathbb{N}^*)^2$, $(m \wedge n) \times (m \vee n) = m \times n$.

Démonstration.

$$(m \wedge n) \times (m \vee n) = \prod_{i=1}^r p_i^{\min(\alpha_i, \beta_i)} \times \prod_{i=1}^r p_i^{\max(\alpha_i, \beta_i)} = \prod_{i=1}^r p_i^{\min(\alpha_i, \beta_i) + \max(\alpha_i, \beta_i)} = \prod_{i=1}^r p_i^{\alpha_i + \beta_i} = m \times n. \quad \square$$

Définition 2.0.7. Si $m \wedge n = 1$, on dit que m et n sont *premiers entre eux*.

Remarque 2.0.8.

- Deux nombres premiers entre eux n'ont donc aucun facteur premier commun.
- De manière équivalente, la fraction $\frac{m}{n}$ est irréductible.

3 Division euclidienne

Proposition 3.0.1. Soit (a, b) dans $\mathbb{Z} \times \mathbb{N}^*$. Il existe un unique (q, r) dans \mathbb{Z}^2 tel que $a = bq + r$ et $r \in \llbracket 0, b - 1 \rrbracket$. On dit que q est le *quotient* de la division euclidienne de a par b et que r est le *reste* de cette division.

Exemple 3.0.2. $56 = 12 \times 4 + 8$, donc 4 est le quotient et 8 le reste de la division euclidienne de 56 par 12.

Démonstration. Existence : Considérons l'ensemble $E = \{a - bk \mid k \in \mathbb{Z}\} \cap \mathbb{N}$. C'est une partie de \mathbb{N} non-vide, donc elle admet un plus petit élément $r = a - bq$. On a bien $r \in \llbracket 0, b - 1 \rrbracket$ (car sinon $r - b \in E$).

Unicité : S'il existe (q_1, r_1) et (q_2, r_2) solutions, alors $bq_1 + r_1 = bq_2 + r_2$ donc $b(q_1 - q_2) = r_2 - r_1$, où $|r_2 - r_1| \leq b - 1$ et $b(q_1 - q_2)$ est un multiple de b , donc $b(q_1 - q_2) = r_2 - r_1 = 0$, donc $(q_1, r_1) = (q_2, r_2)$. \square

La notion de division euclidienne fournit un algorithme rapide de calcul du PGCD, grâce à la propriété suivante :

Proposition 3.0.3. Soit (a, b) dans $(\mathbb{N}^*)^2$ et $a = bq + r$ la division euclidienne de a par b . Alors $a \wedge b = b \wedge r$.

Démonstration. Soit d dans \mathbb{N}^* . Si $d|a$ et $d|b$, alors $d|a - bq = r$. Réciproquement, si $d|b$ et $d|r$, alors $d|bq + r = a$. Donc les diviseurs communs à a et b sont exactement les diviseurs communs à b et r , et en particulier $\text{PGCD}(a, b) = \text{PGCD}(b, r)$. \square

Par conséquent, pour déterminer $a \wedge b$, on peut appliquer l'algorithme suivant, dit *d'Euclide* :

1. On fait la division euclidienne de a par b : $a = bq + r$. Si $r = 0$, alors $a \wedge b = b$.
2. Sinon, on fait la division euclidienne de b par r : $b = rq + r'$. Si $r' = 0$, alors $a \wedge b = b \wedge r = r$.
3. Sinon, on fait la division euclidienne de r par r' : $r = r'q + r''$. Si $r'' = 0$, alors $a \wedge b = b \wedge r = r \wedge r' = r'$.
4. etc. L'algorithme s'arrête quand un reste nul est rencontré, ce qui arrive en temps fini puisque la suite des restes est une suite strictement décroissante d'entiers naturels (c'est le *principe de descente infinie* de Fermat). $a \wedge b$ est alors le *dernier reste non nul* rencontré.

Exemple 3.0.4. Pour déterminer $5742 \wedge 1320$:

- $5742 = 1320 \times 4 + 462$,
- $1320 = 462 \times 2 + 396$,
- $462 = 396 \times 1 + \mathbf{66}$,
- $396 = 66 \times 6 + 0$,

donc $5742 \wedge 1320 = 66$.

Chapitre 8

Primitives

1 Primitives et intégrales

1.1 Primitives

Définition 1.1.1. Soit I un intervalle de \mathbb{R} . Soit $f : I \rightarrow \mathbb{R}$ une fonction. On dit qu'une fonction $F : I \rightarrow \mathbb{R}$ est une *primitive* de f si F est dérivable sur I et si $F' = f$.

Exemple 1.1.2.

- Les fonctions $x \mapsto \frac{x^3}{3}$ et $\frac{x^3}{3} - 15$ sont des primitives de $x \mapsto x^2$ sur \mathbb{R} ,
- La fonction $x \mapsto e^{7x}$ a entre autres pour primitives $x \mapsto \frac{e^{7x}}{7} - 2$, $x \mapsto \frac{e^{7x}}{7} + \pi$ et $x \mapsto \frac{e^{7x}}{7} + 3\sqrt{2}$.

Remarque 1.1.3.

- La *primitivation* d'une fonction, opération qui consiste à en déterminer une primitive, est donc l'opération inverse de sa *dérivation*. Ce lien entre primitives et dérivées est central dans ce chapitre.
- Contrairement à la dérivée, une fonction n'admet pas de primitive unique : on observe notamment que si F est une primitive de f , alors, pour $c \in \mathbb{R}$, $F + c$ est également une primitive de f . C'est en fait une équivalence :

Proposition 1.1.4. Soient $f : I \rightarrow \mathbb{R}$ et soit $F : I \rightarrow \mathbb{R}$ une primitive de f . L'ensemble des primitives de f est alors :

$$\{F + c \mid c \in \mathbb{R}\}.$$

Démonstration. Soit $G : I \rightarrow \mathbb{R}$ dérivable. Alors :

$$(F' = G' = f) \Leftrightarrow (F' = f \text{ et } (G - F)' = 0) \Leftrightarrow (F' = f \text{ et } G - F = c, c \in \mathbb{R}).$$

□

Cela assure que si l'on connaît une primitive, on connaît *toutes* les autres : elles ne diffèrent que par une constante.

Remarque 1.1.5. C'est faux si I n'est pas un intervalle : par exemple, la fonction $\begin{cases} \mathbb{R}^* \rightarrow \mathbb{R} \\ x \mapsto 1 \end{cases}$ a pour primitives $F : \begin{cases} \mathbb{R}^* \rightarrow \mathbb{R} \\ x \mapsto x \end{cases}$ et $G : \begin{cases} \mathbb{R}^* \rightarrow \mathbb{R} \\ x \mapsto \begin{cases} x + 2 & \text{si } x > 0 \\ x - 5 & \text{si } x < 0 \end{cases} \end{cases}$, et $\forall c \in \mathbb{R}, F \neq G + c$.

Comme pour les dérivées, les primitives des fonctions usuelles sont à connaître ; elles sont indiquées dans le tableau récapitulatif joint à ce cours.

On a vu qu'il existe des fonctions continues non dérивables. De même, on peut se demander s'il existe des fonctions continues non *primitivables*, c'est-à-dire n'admettant pas de primitive.

Réponse :

- *Oui*, si l'on veut une primitive *usuelle*, c'est-à-dire écrite à l'aide de fonctions usuelles et d'opérations standard (addition, produit, composition...). Par exemple, la fonction $x \mapsto e^{-x^2}$ n'admet pas de telle primitive.
- *Non*, si l'on s'autorise à exprimer les primitives à l'aide d'*intégrales*.

1.2 Intégrales

On présente ici une introduction aux intégrales, qui suffira à notre utilisation. La notion d'intégrale sera définie plus rigoureusement dans un chapitre ultérieur.

On considère un segment $[a, b]$ de \mathbb{R} .

Définition 1.2.1. Soit $f : [a, b] \rightarrow \mathbb{R}$. On dit que f est une fonction *en escalier* s'il existe $a = x_0 < x_1 < \dots < x_n = b$ et $(y_0, \dots, y_{n-1}) \in \mathbb{R}^n$ tels que : $\forall k \in \llbracket 0, n-1 \rrbracket$, $\forall x \in]x_k, x_{k+1}[$, $f(x) = y_k$.

L'intégrale de f est alors le **nombre** :

$$\int_a^b f = \sum_{k=0}^{n-1} y_k (x_{k+1} - x_k).$$

Remarque 1.2.2. L'intégrale de f correspond donc à l'aire entre le graphe de f et l'axe des abscisses (Ox), comptée *algébriquement*, c'est-à-dire positivement si le graphe est au-dessus de l'axe, et négativement s'il est en-dessous.

Cette notion peut être généralisée à toute fonction continue sur $[a, b]$; c'est l'objet de la proposition suivante, admise pour l'instant :

Proposition 1.2.3. Soit $f : [a, b] \rightarrow \mathbb{R}$ continue. Il existe une suite de fonctions en escalier (f_n) qui approche f : $f_n \xrightarrow{n \rightarrow +\infty} f$. La suite $\left(\int_a^b f_n \right)$ est alors convergente, on note $\int_a^b f$, ou $\int_a^b f(x) dx$, sa limite.

Remarque 1.2.4.

- La notation « $\int_a^b f(x) dx$ » est à mettre en regard de la formule 1.2.1 : le $f(x)$ correspond à l'image y_k , le dx à la différence $x_{k+1} - x_k$ (c'est un Δx infinitésimal, appelé *différentielle*), et le symbole \int_a^b correspond à la somme $\sum_{k=0}^{n-1}$, comme le manifeste le fait que ces deux symboles soient des S majuscule (le premier en ancien français, le second en grec).
- L'intégrale d'une fonction continue correspond alors bien à l'aire algébrique entre son graphe et (Ox). De cette interprétation graphique découlent les principales propriétés de l'intégrale :

Proposition 1.2.5. Soient $f : [a, b] \rightarrow \mathbb{R}$ et $g : [a, b] \rightarrow \mathbb{R}$ deux fonctions continues, et soit $(\lambda, \mu) \in \mathbb{R}^2$. Alors :

1. $\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g$ (linéarité de l'intégrale),
2. $\forall c \in [a, b]$, $\int_a^b f = \int_a^c f + \int_c^b f$ (Relation de Chasles),
3. $\int_b^a f = - \int_a^b f$,
4. Si $f \geq 0$, alors $\int_a^b f \geq 0$ et $\left(\int_a^b f = 0 \Rightarrow f = 0 \right)$ (positivité de l'intégrale),
5. Si $f \leq g$, alors $\int_a^b f \leq \int_a^b g$ et $\left(\int_a^b f = \int_a^b g \Rightarrow f = g \right)$ (croissance de l'intégrale).

Enfin, on pourra remarquer le cas des fonctions paires et des fonctions impaires :

Proposition 1.2.6. Soit $f : [-a, a] \rightarrow \mathbb{R}$.

- Si f est paire, alors $\int_{-a}^a f = 2 \int_0^a f$,
- Si f est impaire, alors $\int_{-a}^a f = 0$.

1.3 Théorème fondamental de l'analyse

On apporte ici notre réponse à la question de la primitivabilité des fonctions continues : toute fonction continue admet une primitive d'après le théorème ci-dessous, dû à Newton et Leibniz, appelé *théorème fondamental de l'analyse*.

Théorème 1.3.1 (Théorème fondamental de l'analyse). Soit I un intervalle de \mathbb{R} , soit $a \in I$ et soit $f : I \rightarrow \mathbb{R}$ continue. Alors la fonction

$$F : \begin{cases} I & \rightarrow \mathbb{R} \\ x & \mapsto \int_a^x f \end{cases}$$

est dérivable sur I et $F' = f$. C'est donc une primitive de f .

Démonstration. Soit $(x, x_0) \in I^2$. Alors

$$\begin{aligned} F(x) - F(x_0) &= \int_{a_x}^x f - \int_{a_{x_0}}^{x_0} f \\ &= \int_{a_x}^x f + \int_{x_0}^x f \\ &= \int_{x_0}^x f \quad \text{d'après la relation de Chasles.} \end{aligned}$$

Or, lorsque x est proche de x_0 , $\int_{x_0}^x f$ est proche de l'aire du rectangle de base $x - x_0$ et de hauteur $f(x_0)$, soit $f(x_0)(x - x_0)$. Donc $\frac{F(x) - F(x_0)}{x - x_0} \xrightarrow{x \rightarrow x_0} f(x_0)$. Donc F est dérivable en x_0 , et $F'(x_0) = f(x_0)$. \square

Remarque 1.3.2. D'après la proposition 1.1.4, toutes les primitives de f sont donc de la forme $x \mapsto \int_a^x f + c$, $(a, c) \in I \times \mathbb{R}$. On note $\int f$ l'ensemble des primitives de f .

2 Calcul de primitives et d'intégrales

Calculer une intégrale consiste à trouver le nombre correspondant à ladite intégrale. En particulier, *calculer* une primitive $x \mapsto \int_a^x f + c$ consiste à trouver une expression de ladite primitive qui n'utilise pas d'intégrale.

Contrairement à la dérivation, il n'existe pas de formule systématique de calcul de primitive. Outre l'identification directe, quelques méthodes existent¹, dont les principales sont présentées ci-dessous :

1. En voici une représentation : <https://xkcd.com/2117/>

2.1 Calcul direct

Les seules primitives usuelles que l'on peut déterminer directement sont celles du tableau ci-joint, et celles des fonctions suivantes :

Proposition 2.1.1. *Soient $u : I \rightarrow I$ dérivable et $f : I \rightarrow \mathbb{R}$ de primitive F , alors $u' \cdot f \circ u$ a pour primitive $F \circ u$.*

Exemple 2.1.2. $x \mapsto \frac{e^x}{1 + (e^x)^2}$ a pour primitive $x \mapsto \arctan(e^x)$.

Dans le cas où une primitive usuelle est connue, la valeur de l'intégrale est obtenue grâce à la propriété suivante :

Proposition 2.1.3. *Soit $f : [a, b] \rightarrow \mathbb{R}$ continue, et soit $F : [a, b] \rightarrow \mathbb{R}$ une primitive de f . Alors*

$$\int_a^b f = F(b) - F(a).$$

Démonstration. D'après le théorème fondamental de l'analyse, il existe $(x_0, c) \in I \times \mathbb{R}$ tel que $F : x \mapsto \int_{x_0}^x f + c$. Alors, d'après la relation de Chasles : $F(b) - F(a) = \left(\int_{x_0}^b f + c \right) - \left(\int_{x_0}^a f + c \right) = \int_{x_0}^b f + \int_a^{x_0} f = \int_a^b f$. \square

Remarque 2.1.4. On note usuellement $[F(x)]_a^b = F(b) - F(a)$.

Exemple 2.1.5. $\int_1^3 x dx = \left[\frac{x^2}{2} \right]_1^3 = \frac{3^2}{2} - \frac{1^2}{2} = 4$.

2.2 Cas des fonctions trigonométriques

Dans certains cas, le calcul de l'intégrale d'une fonction trigonométrique peut être résolu :

- Par linéarisation, lorsqu'on cherche à intégrer un produit de fonctions trigonométriques :

Exemple 2.2.1. $\int_0^{\frac{\pi}{2}} \cos^2 x dx = \int_0^{\frac{\pi}{2}} \left(\frac{1}{2} + \frac{1}{2} \cos(2x) \right) dx$.

- Par utilisation de la forme exponentielle, lorsqu'on est face à un produit de fonctions trigonométriques et d'exponentielles :

Exemple 2.2.2. $\int_0^{\pi} e^x \cos(3x) dx = \operatorname{Re} \left(\int_0^{\pi} e^x e^{i3x} dx \right)$.

2.3 Cas des fractions rationnelles

Le calcul de l'intégrale d'une fraction rationnelle passe en général par la division euclidienne du numérateur par le dénominateur. La primitivation directe, la primitivation des inverses des fonctions polynomiales de degré 2 (présentée à la fin du chapitre) et le théorème de factorisation des polynômes (plus tard dans l'année), associé à la décomposition en éléments simples, assurent le calcul de toute fraction rationnelle.

Exemple 2.3.1. $\int \frac{x}{1+x} dx, \int \frac{x^4}{(x-4)(x+2)} dx$

2.4 Intégration par parties

Comme pour la dérivation, la primitive d'un produit de fonctions *n'est pas* le produit des primitives ! On se souvient de la formule de Leibniz : étant donné deux fonctions $f, g : [a, b] \rightarrow \mathbb{R}$ dérivables :

$$(fg)' = f'g + fg'.$$

Si les fonctions f' et g' sont continues, les intégrales des fonctions de l'égalité ci-dessus sont définies et sont égales :

$$\int_a^b (fg)' = \int_a^b (f'g + fg') = \int_a^b f'g + \int_a^b fg'.$$

Or la fonction $(fg)'$ a pour primitive la fonction fg , donc :

$$[f(x)g(x)]_a^b = \int_a^b f'g + \int_a^b fg'.$$

Cette dernière égalité est connue sous le nom d'*intégration par parties*, valable pour des fonctions f et g dérivables et de dérivées continues, c'est-à-dire *de classe C^1* (cf. le chapitre 2). Elle est utilisée sous la forme suivante :

Théorème 2.4.1 (Intégration par parties). *Soient f et g deux fonctions dans $C^1([a, b], \mathbb{R})$. Alors :*

$$\int_a^b f'g = [f(x)g(x)]_a^b - \int_a^b fg'.$$

Sous cette forme, c'est une formule de *transformation* de l'intégrale d'un produit. Il ne faut pas l'utiliser systématiquement, mais elle peut être très utile lorsque le produit à intégrer est composé :

- d'une fonction « peu coûteuse » à primitiver, c'est-à-dire dont la primitive n'est pas trop compliquée ; c'est le « f' » initial. Par exemple \exp , \cos , \sin ...
- d'une fonction « intéressante » à dériver, c'est-à-dire dont la dérivée est plus simple que la fonction de départ ; c'est le « g » initial. Par exemple $x \mapsto x^n$, \ln ...

Exemples 2.4.2. $\int_0^1 xe^x dx$, $\int \ln x dx$, $\int_0^{\frac{\pi}{2}} x^2 \sin x dx$.

2.5 Changement de variable

La transformation d'une intégrale par *changement de variable* permet de simplifier la primitivation d'une composée. Plus précisément : étant donné deux fonctions u et F définies sur $[a, b]$ et $[u(a), u(b)]$ respectivement, on a, avec la notation d'évaluation entre crochets vue plus haut :

$$[F(x)]_{u(a)}^{u(b)} = [F \circ u(t)]_a^b.$$

Si la fonction F est une primitive d'une fonction f , on peut alors écrire :

$$\int_{u(a)}^{u(b)} f(x) dx = \int_a^b (F \circ u)'(t) dt = \int_a^b f(u(t))u'(t) dt.$$

On a ainsi appliqué à l'intégrale $\int_a^b f(x) dx$ le *changement de variable* $x = u(t)$. On peut ainsi transformer l'intégrale en une intégrale plus simple. Par exemple, dans l'intégrale $\int_0^1 \frac{e^{2x}}{1+e^x} dx$, le changement de variables $x = \ln t$ (c'est-à-dire $t = e^x$) est tout indiqué.

Remarque 2.5.1. Il s'agit en fait d'une simple réécriture de la proposition de primitivation directe : $\int u' \times f \circ u = F \circ u$.

Théorème 2.5.2 (Changement de variable). *Soient u dans $C^1([a, b], \mathbb{R})$ et f dans $C^0([u(a), u(b)], \mathbb{R})$. Alors :*

$$\int_{u(a)}^{u(b)} f(x) dx = \int_a^b f(u(t))u'(t) dt.$$

Remarque 2.5.3. Il faut prendre garde à transformer tous les éléments de l'intégrale :

- $f(x)$ devient $f(u(t))$,
- dx devient, puisque $\frac{dx}{dt} = \frac{du(t)}{dt} = u'(t) : dx = u'(t)dt$,
- $\int_a^b = \int_{x=a}^{x=b} = \int_{u(t)=u(c)}^{u(t)=u(d)} \text{devient } \int_{t=c}^{t=d}$.

Exemple 2.5.4. $\int_0^1 \frac{e^{2x}}{1+e^x} dx, \int_0^1 \sqrt{1-x^2} dx.$

2.6 Cas des fonctions $x \mapsto \frac{1}{ax^2 + bx + c}$

Lorsqu'on cherche à primitiver $x \mapsto \frac{1}{ax^2 + bx + c}$ ($a > 0$), trois cas se présentent selon la valeur de $\Delta = b^2 - 4ac$:

- Si $\Delta > 0$, alors $ax^2 + bx + c = a(x - r_1)(x - r_2)$ où r_1 et r_2 sont les deux racines réelles distinctes du trinôme. Alors :

$$\frac{1}{ax^2 + bx + c} = \frac{1}{a(x - r_1)(x - r_2)} = \frac{1}{a(r_2 - r_1)} \left(\frac{1}{x - r_2} - \frac{1}{x - r_1} \right),$$

donc

$$\int \frac{dx}{ax^2 + bx + c} = \frac{1}{a(r_2 - r_1)} \ln \left| \frac{x - r_2}{x - r_1} \right|.$$

- Si $\Delta = 0$, alors $ax^2 + bx + c = a(x - r)^2$, donc

$$\int \frac{dx}{ax^2 + bx + c} = \frac{1}{a} \int \frac{dx}{(x - r)^2} = -\frac{1}{a(x - r)}.$$

- Si $\Delta < 0$, alors on met sous forme canonique : $ax^2 + bx + c = a \left((x + p)^2 + q^2 \right)$ (avec $p = \frac{b}{2a}$ et $q = \sqrt{-\frac{\Delta}{4a^2}}$), donc :

$$\frac{1}{ax^2 + bx + c} = \frac{1}{a} \frac{1}{(x + p)^2 + q^2} = \frac{1}{aq^2} \frac{1}{1 + \left(\frac{x+p}{q} \right)^2},$$

d'où :

$$\int \frac{dx}{ax^2 + bx + c} = \frac{1}{aq} \arctan \left(\frac{x+p}{q} \right).$$

Exemple 2.6.1. $\int_{-1}^1 \frac{x-1}{x^2+2x+5} dx$

Chapitre 9

Équations différentielles

1 Généralités

Définition 1.0.1. Une *équation différentielle* est une équation reliant une fonction $I \rightarrow \mathbb{R}$ inconnue (généralement notée y) à ses dérivées successives. Cette équation peut être accompagnée de conditions sur les valeurs prises par la fonction et ses dérivées en un point, appelées *conditions initiales*.

Exemple 1.0.2. $y'' = -y$, $\begin{cases} y' = y \\ y(0) = 1 \end{cases}$, $y' = 1 + y^2$, $\forall t \in \mathbb{R}_+^*, y'(t) = \frac{y(t)}{t}$.

Pour simplifier l'écriture, on écrira la dernière équation ci-dessus sous la forme : $y' = \frac{y}{t}$. Il est important de garder à l'esprit que la formulation exacte est la première !

Il existe une grande variété d'équations différentielles, et la résolution de la plupart d'entre elles reste un problème ouvert. C'est en quelque sorte une généralisation de la primitivation, opération dont on a vu la complexité. On s'intéressera uniquement dans ce chapitre à la résolution de certaines équations différentielles *linéaires*.

Définition 1.0.3. Une équation différentielle *linéaire* d'ordre $n \in \mathbb{N}$ et d'inconnue $y : I \rightarrow \mathbb{R}$ est une équation différentielle de la forme :

$$a_n y^{(n)} + \dots + a_1 y' + a_0 y = b,$$

où $a_0, a_1, \dots, a_n, b : I \rightarrow \mathbb{R}$ sont des fonctions, où a_n ne s'annule pas sur I .

On va traiter dans ce chapitre le cas $n = 1$ dans une certaine généralité, ainsi que le cas $n = 2$ dans des cas particuliers. Pour cela, on va intensivement utiliser la propriété vue dans le chapitre précédent dans le cadre des systèmes linéaires :

Définition 1.0.4. L'*équation homogène* associée à l'équation ci-dessus est l'équation différentielle « sans second membre » d'inconnue $y_h : I \rightarrow \mathbb{R}$:

$$a_n y_h^{(n)} + \dots + a_1 y'_h + a_0 y_h = 0.$$

Proposition 1.0.5. Soit S l'ensemble des solutions de l'équation différentielle considérée. Soit y_p une solution particulière de l'équation, et soit S_h l'ensemble des solutions de l'équation homogène associée. Alors :

$$S = \{y_p + y_h \mid y_h \in S_h\}.$$

Démonstration. Comme y_p est une solution de l'équation, on a :

$$a_n y_p^{(n)} + \dots + a_1 y'_p + a_0 y_p = b,$$

donc :

$$\begin{aligned}
 y \in S &\Leftrightarrow a_n y^{(n)} + \dots + a_1 y' + a_0 y = b \\
 &\Leftrightarrow a_n(y - y_p)^{(n)} + \dots + a_1(y - y_p)' + a_0(y - y_p) = 0 \\
 &\Leftrightarrow (y - y_p) \in S_h.
 \end{aligned}$$

□

Exemple 1.0.6. $y'' + y = 2t + 3$.

Remarque 1.0.7. La conséquence directe de cette proposition est que, pour résoudre une équation différentielle linéaire, il suffit d'en connaître **une** solution et de savoir résoudre l'équation homogène associée. Cette idée est au cœur des méthodes de résolution présentées ci-après.

2 Équations différentielles linéaires d'ordre 1

On s'intéresse dans cette partie aux équations différentielles linéaires d'ordre 1, et plus précisément aux équations de la forme :

$$y' + ay = b,$$

d'inconnue $y : I \rightarrow \mathbb{R}$, où $a, b : I \rightarrow \mathbb{R}$ sont supposées *continues*.

2.1 À coefficients constants

Dans un premier temps, on suppose a et b constantes, ou ce qui revient au même : $a, b \in \mathbb{R}$.

On constate alors que :

- si a est non-nul, alors une solution de l'équation est la fonction constante $y_p = \frac{b}{a}$,
- si a est nul, alors une solution est la fonction $y_p : t \mapsto bt$.

Il ne reste plus qu'à résoudre l'équation homogène :

Proposition 2.1.1. *L'ensemble des solutions S_h de l'équation $y'_h + ay_h = 0$ est égal à :*

$$S_h = \{y_h : t \mapsto \lambda e^{-at} \mid \lambda \in \mathbb{R}\}.$$

Démonstration. La fonction $f : t \mapsto e^{at}$ ne s'annulant pas sur I , on a l'équivalence :

$$\begin{aligned}
 \forall t \in I, y'_h(t) + ay_h(t) = 0 &\iff \forall t \in I, y'_h(t)e^{at} + ay_h(t)e^{at} = 0, \\
 &\iff \forall t \in I, (fy_h)'(t) = 0, \\
 &\iff \exists \lambda \in \mathbb{R}, \forall t \in I, (fy_h)(t) = \lambda \quad \text{car } I \text{ est un intervalle,} \\
 &\iff \exists \lambda \in \mathbb{R}, \forall t \in I, y_h(t) = \lambda e^{-at}.
 \end{aligned}$$

□

Par conséquent, l'ensemble des solutions de l'équation globale est égal à :

- si a est non-nul, alors $S = \left\{ y : t \mapsto \lambda e^{-at} + \frac{b}{a} \mid \lambda \in \mathbb{R} \right\}$,
- si a est nul, alors $S = \{y : t \mapsto \lambda + bt \mid \lambda \in \mathbb{R}\}$.

2.2 À coefficients non constants

On peut généraliser l'analyse ci-dessus au cas où a et b sont des fonctions non constantes. On commence par l'équation homogène :

Équation homogène

Proposition 2.2.1. Soit A une primitive de a sur I . L'ensemble des solutions S_h de l'équation $y'_h + ay_h = 0$ est égal à :

$$S_h = \{y_h : t \mapsto \lambda e^{-A(t)} \mid \lambda \in \mathbb{R}\}.$$

Démonstration. La fonction e^A ne s'annulant pas sur I , on a l'équivalence :

$$\begin{aligned} y'_h + ay_h = 0 &\iff y'_h e^A + ay_h e^A = 0, \\ &\iff (y_h e^A)' = 0, \\ &\iff \exists \lambda \in \mathbb{R}, y_h e^A = \lambda \quad \text{car } I \text{ est un intervalle,} \\ &\iff \exists \lambda \in \mathbb{R}, y_h = \lambda e^{-A}. \end{aligned}$$

□

Exemple 2.2.2. $y' - \frac{y}{t} = t$.

Solution particulière

Une fois l'équation homogène résolue, et si une solution de l'équation globale n'est pas identifiable directement, on utilisera la proposition suivante, appelée *méthode de variation de la constante* :

Proposition 2.2.3. Il existe une solution de la forme $y_p : t \mapsto \lambda(t)e^{-A(t)}$, où A est une primitive de a sur I .

Démonstration. Soit une telle fonction y_p . On a alors :

$$\forall t \in I, y'_p(t) = \lambda'(t)e^{-A(t)} - a(t)\lambda(t)e^{-A(t)},$$

donc

$$\forall t \in I, y'_p(t) + ay_p(t) = \lambda'(t)e^{-A(t)},$$

donc y_p est une solution de l'équation si et seulement si

$$\forall t \in I, \lambda'(t)e^{-A(t)} = b(t).$$

En pratique, on pourra directement écrire cette dernière ligne. On a alors :

$$\forall t \in I, \lambda'(t) = b(t)e^{A(t)},$$

c'est-à-dire que λ est une primitive de be^A . Soit alors $t_0 \in I$, on a vu dans le chapitre 7 que :

$$\lambda : t \mapsto \int_{t_0}^t b(x)e^{A(x)}dx = \int_{t_0}^t be^A$$

est une primitivation possible, donc :

$$y_p : t \mapsto e^{-A(t)} \int_{t_0}^t b(x)e^{A(x)}dx$$

est une solution particulière de l'équation. On peut l'écrire de manière plus concise : $y_p = e^{-A} \int_{t_0}^t be^A$. □

Finalement, l'ensemble des solutions S de l'équation $y' + ay = b$ est donné, en choisissant $t_0 \in I$, par :

$$S = \left\{ y : t \mapsto e^{-A(t)} \left(\lambda + \int_{t_0}^t b(x)e^{A(x)}dx \right) \mid \lambda \in \mathbb{R} \right\}.$$

Mentionnons également le *principe de superposition* des solutions d'une équation différentielle linéaire, qui permet notamment de traiter séparément les éventuelles différentes parties du second membre :

Proposition 2.2.4. Soient $a, b_1, b_2 : I \rightarrow \mathbb{R}$ continues.

Si $y_{p,1}$ est solution de l'équation $y' + ay = b_1$,

et si $y_{p,2}$ est solution de l'équation $y' + ay = b_2$,

alors $y_p = y_{p,1} + y_{p,2}$ est solution de l'équation $y' + ay = b_1 + b_2$.

Problème de Cauchy

En pratique, une équation différentielle représente souvent l'équation d'*évolution* d'une quantité au sein d'un système (par exemple de la position d'un objet soumis à des contraintes). Elle s'accompagne alors en général d'informations sur l'*état* préalable du système (dans l'exemple, la position initiale de l'objet). L'ensemble de ces informations forme un *problème de Cauchy* :

Définition 2.2.5 (Problème de Cauchy d'ordre 1). Soient $a, b : I \rightarrow \mathbb{R}$ continues, et soit $(t_0, y_0) \in I \times \mathbb{R}$. Le système

$$\begin{cases} y' + ay = b \\ y(t_0) = y_0 \end{cases}$$

est le *problème de Cauchy* d'équation $y' + ay = b$ et de *condition initiale* $y(t_0) = y_0$.

On peut se convaincre que la donnée de l'état initial d'un système et de sa loi d'évolution *décrit* le système, c'est-à-dire permet d'en déterminer l'état à tout moment. C'est l'objet du théorème suivant :

Théorème 2.2.6 (Théorème de Cauchy-Lipschitz à l'ordre 1). *Tout problème de Cauchy d'ordre 1 a une et une seule solution.*

Démonstration. On prend les notations de la définition ci-dessus. On a vu que l'équation a pour solutions les fonctions :

$$y : t \mapsto e^{-A(t)} \left(\lambda + \int_{t_0}^t b e^A \right),$$

où λ parcourt \mathbb{R} . La condition initiale appliquée à ces fonctions s'écrit : $y_0 = e^{-A(t_0)} \times \lambda$, c'est-à-dire $\lambda = y_0 e^{A(t_0)}$. Le problème de Cauchy a donc bien pour unique solution :

$$y : t \mapsto y_0 e^{-(A(t)-A(t_0))} + e^{-A(t)} \int_{t_0}^t b e^A.$$

□

3 Équations différentielles linéaires d'ordre 2

On s'intéresse dans cette partie aux équations différentielles linéaires d'ordre 2 à coefficients constants (et réels), que l'on notera :

$$ay'' + by' + cy = f,$$

d'inconnue $y : I \rightarrow \mathbb{R}$, où $(a, b, c) \in \mathbb{R}^3$ avec $a \neq 0$ et où $f : I \rightarrow \mathbb{R}$ est continue.

3.1 Équation homogène

Comme précédemment, on commence par résoudre l'équation homogène

$$ay_h'' + by_h' + cy_h = 0.$$

On commence par observer que, pour $(\lambda, r) \in \mathbb{R}^2$, la fonction $t \mapsto \lambda e^{rt}$ est solution de l'équation homogène si et seulement si

$$ar^2 + br + c = 0.$$

Cette équation est appelée *équation caractéristique* de l'équation différentielle (car elle va nous permettre d'en *caractériser* les solutions). Notre expérience des équations du second degré nous permet de distinguer deux cas :

Proposition 3.1.1. *Notons $\Delta = b^2 - 4ac$ le discriminant de l'équation caractéristique.*

- Si $\Delta > 0$, notons r_1, r_2 les deux racines distinctes de l'équation caractéristique. L'ensemble des solutions de l'équation homogène est alors égal à :

$$S_h = \{y_h : t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t} \mid (\lambda, \mu) \in \mathbb{R}^2\}.$$

- Si $\Delta = 0$, notons r_0 la racine double de l'équation caractéristique. L'ensemble des solutions de l'équation homogène est alors égal à :

$$S_h = \{y_h : t \mapsto (\lambda t + \mu) e^{r_0 t} \mid (\lambda, \mu) \in \mathbb{R}^2\}.$$

- Si $\Delta < 0$, notons $r_1, r_2 = R \pm i\omega$ les deux racines complexes conjuguées de l'équation caractéristique. L'ensemble des solutions de l'équation homogène est alors égal à :

$$S_h = \{y_h : t \mapsto e^{Rt} (\lambda \cos(\omega t) + \mu \sin(\omega t)) \mid (\lambda, \mu) \in \mathbb{R}^2\}.$$

Démonstration.

- Si $\Delta = 0$: soient $y : I \rightarrow \mathbb{R}$ et $z : t \mapsto y(t) e^{-r_0 t}$. La fonction y est alors solution de l'équation homogène si et seulement si z est solution de l'équation différentielle

$$z'' = 0,$$

d'où la solution annoncée.

- Si $\Delta \neq 0$: on procède de la même manière, en posant $y : I \rightarrow \mathbb{R}$ et $z = y' - r_2 y$. La fonction y est alors solution de l'équation homogène si et seulement si z est solution de l'équation différentielle

$$z' - r_1 z = 0.$$

D'après notre étude des équations d'ordre 1, z est dans ce cas de la forme $z : t \mapsto C e^{r_1 t}$, $C \in \mathbb{R}$, c'est-à-dire :

$$y' - r_2 y = C e^{r_1 t},$$

que l'on sait également résoudre, et on trouve la solution annoncée.

Dans le cas $\Delta < 0$, comme les coefficients sont réels, les racines r_1 et r_2 sont complexes conjuguées. Notons-les $r_{1,2} = R \pm i\omega$, on a :

$$\begin{aligned} \forall t \in I, y_h(t) &= \lambda e^{(R+i\omega)t} + \mu e^{(R-i\omega)t} \\ &= e^{Rt} (\lambda e^{i\omega t} + \mu e^{-i\omega t}) \\ &= e^{Rt} ((\lambda + \mu) \cos(\omega t) + i(\lambda - \mu) \sin(\omega t)) \\ &= e^{Rt} (\tilde{\lambda} \cos(\omega t) + \tilde{\mu} \sin(\omega t)), \end{aligned}$$

où l'on note $\tilde{\lambda} = \lambda + \mu$ et $\tilde{\mu} = i(\lambda - \mu)$.

□

3.2 Solution particulière

On va se limiter dans cette partie à donner une méthode de détermination de solution particulière dans le cas où le second membre f est de la forme $f : t \mapsto \alpha e^{\beta t}$, où $(\alpha, \beta) \in \mathbb{R}^2$. On peut noter qu'outre les fonctions exponentielles, cela couvre également le cas des fonctions trigonométriques et hyperboliques.

Comme pour l'équation homogène, la forme de la solution particulière va dépendre des racines de l'équation caractéristique, qui est toujours :

$$P(r) = ar^2 + br + c = 0,$$

de discriminant $\Delta = b^2 - 4ac$.

Proposition 3.2.1.

- Si $P(\beta) \neq 0$, alors $y_p : t \mapsto \frac{\alpha}{P(\beta)} e^{\beta t}$ est solution,
- Si $P(\beta) = 0$ et $\Delta \neq 0$, alors $y_p : t \mapsto \frac{\alpha}{P'(\beta)} t e^{\beta t}$ est solution,
- Si $P(\beta) = 0$ et $\Delta = 0$, alors $y_p : t \mapsto \frac{\alpha}{P''(\beta)} t^2 e^{\beta t}$ est solution.

Remarque 3.2.2.

- Si le second membre f est une fonction polynomiale de degré n , il est toujours possible de trouver une solution particulière polynomiale de degré n .
- Une *méthode de variation des constantes* λ et μ est applicable dans le cas général, avec un raisonnement similaire au cas des équations d'ordre 1. Nous ne détaillerons pas cette méthode ici.

De même qu'à l'ordre 1, on peut faire appel au *principe de superposition* des solutions :

Proposition 3.2.3.

Soient $f_1, f_2 : I \rightarrow \mathbb{R}$ continues.

Si $y_{p,1}$ est solution de l'équation $ay'' + by' + cy = f_1$,
et si $y_{p,2}$ est solution de l'équation $ay'' + by' + cy = f_2$,
alors $y_p = y_{p,1} + y_{p,2}$ est solution de l'équation $ay'' + by' + cy = f_1 + f_2$.

Problème de Cauchy

Un problème de Cauchy associé à une équation d'ordre 2 nécessite *deux* conditions initiales. On peut l'interpréter de la façon suivante : une équation d'ordre 2 régit l'*accélération* d'un objet. La détermination de son état à tout temps nécessite donc de connaître la fois sa *position* et sa *vitesse* initiales.

Définition 3.2.4 (Problème de Cauchy d'ordre 2). Soient $(a, b, c) \in \mathbb{R}^* \times \mathbb{R}^2$, $f : I \rightarrow \mathbb{R}$ continue, et soit $(t_0, y_0, v_0) \in I \times \mathbb{R}^2$. Le système

$$\begin{cases} ay'' + by' + cy = f \\ y(t_0) = y_0 \\ y'(t_0) = v_0 \end{cases}$$

est le *problème de Cauchy* d'équation $ay'' + by' + cy = f$ et de *conditions initiales* $y(t_0) = y_0$ et $y'(t_0) = v_0$.

Le théorème de Cauchy-Lipschitz s'applique alors :

Théorème 3.2.5 (Théorème de Cauchy-Lipschitz à l'ordre 2). *Tout problème de Cauchy d'ordre 2 a une et une seule solution.*

Démonstration. On prend les notations de la définition ci-dessus. Notons y_p une solution particulière de l'équation, et Δ le discriminant de l'équation caractéristique. Alors :

- Si $\Delta \neq 0$, notons r_1, r_2 les deux racines distinctes de l'équation caractéristique. Les solutions de l'équation sont les :

$$y : t \mapsto \lambda e^{r_1 t} + \mu e^{r_2 t} + y_p(t),$$

où (λ, μ) parcourt \mathbb{R}^2 . Les conditions initiales appliquées à ces fonctions s'écrivent :

$$\begin{cases} \lambda e^{r_1 t_0} + \mu e^{r_2 t_0} + y_p(t_0) = y_0 \\ \lambda r_1 e^{r_1 t_0} + \mu r_2 e^{r_2 t_0} + y'_p(t_0) = v_0 \end{cases},$$

c'est-à-dire

$$\begin{cases} e^{r_1 t_0} \lambda + e^{r_2 t_0} \mu = y_0 - y_p(t_0) \\ r_1 e^{r_1 t_0} \lambda + r_2 e^{r_2 t_0} \mu = v_0 - y'_p(t_0) \end{cases},$$

où l'on reconnaît un système linéaire de 2 équations à 2 inconnues. Son déterminant est égal à $e^{r_1 t_0} \times r_2 e^{r_2 t_0} - r_1 e^{r_1 t_0} \times e^{r_2 t_0} = (r_2 - r_1) e^{(r_1 + r_2)t_0} \neq 0$: c'est donc un système de Cramer, dont on sait qu'il admet une et une seule solution.

- Si $\Delta = 0$, notons r_0 la racine double de l'équation caractéristique. Les solutions de l'équation sont les :

$$y : t \mapsto (\lambda t + \mu) e^{r_0 t} + y_p(t),$$

où (λ, μ) parcourt \mathbb{R}^2 . Les conditions initiales appliquées à ces fonctions s'écrivent :

$$\begin{cases} (\lambda t_0 + \mu) e^{r_0 t_0} + y_p(t_0) = y_0 \\ \lambda e^{r_0 t_0} + r_0 (\lambda t_0 + \mu) e^{r_0 t_0} + y'_p(t_0) = v_0 \end{cases},$$

c'est-à-dire

$$\begin{cases} t_0 e^{r_0 t_0} \lambda + e^{r_0 t_0} \mu = y_0 - y_p(t_0) \\ (1 + r_0 t_0) e^{r_0 t_0} \lambda + r_0 e^{r_0 t_0} \mu = v_0 - y'_p(t_0) \end{cases},$$

où l'on reconnaît un système linéaire de 2 équations à 2 inconnues. Son déterminant est égal à $-e^{2r_0 t_0} \neq 0$: c'est donc un système de Cramer, dont on sait qu'il admet une et une seule solution.

□

Chapitre 10

Calcul matriciel et systèmes linéaires

On s'intéresse dans ce chapitre aux matrices, qui sont au départ un outil de manipulation des systèmes linéaires. Nous allons progressivement nous détacher de ce point de vue, pour étudier les matrices comme un objet mathématique à part entière.

1 Systèmes linéaires

Dans tout le chapitre, on note $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , et n, p deux entiers naturels non nuls.

1.1 Définitions

Définition 1.1.1. Une *équation linéaire* à p inconnues x_1, \dots, x_p est une équation de la forme

$$a_1x_1 + a_2x_2 + \dots + a_px_p = b,$$

où $(a_1, \dots, a_p) \in \mathbb{K}^p$ sont les *coefficients* de l'équation et $b \in \mathbb{K}$ en est le *second membre*. Les p -uplets (x_1, \dots, x_p) qui satisfont l'équation en sont les *solutions*.

Exemple 1.1.2.

- $2x + 3y + 4z = 1$
- $z + it = 2$
- Contre-exemples : $x^2 + y^2 = 1$, $xy + 2x - y = 1$

Observation 1.1.3.

- Pour $p = 2$, l'ensemble des solutions d'une équation linéaire est une droite de \mathbb{R}^2 (sauf si $a_1 = a_2 = 0$, auquel cas l'ensemble des solutions est \mathbb{R}^2 si $b = 0$, \emptyset sinon).
- Pour $p = 3$, l'ensemble des solutions est un plan de \mathbb{R}^3 (sauf si $a_1 = a_2 = a_3 = 0$, auquel cas l'ensemble des solutions est \mathbb{R}^3 si $b = 0$, \emptyset sinon).

Définition 1.1.4. Un *système linéaire* de n équations à p inconnues x_1, \dots, x_p est un système de n équations linéaires de la forme :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p &= b_2 \\ \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p &= b_n \end{cases},$$

où : $\forall i \in \llbracket 1, n \rrbracket$, $\forall j \in \llbracket 1, p \rrbracket$, $a_{ij} \in \mathbb{K}$ et $(b_1, \dots, b_n) \in \mathbb{K}^n$.

Pour tout i dans $\llbracket 1, n \rrbracket$, l'équation linéaire

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p = b_i$$

est appelée *i^{ème} ligne* du système, et sera notée L_i .

Exemple 1.1.5.
$$\begin{cases} 3x - 4y = 1 \\ 2x - 5y = 2 \end{cases}.$$

Observation 1.1.6.

- Pour $n = 2$ et $p = 2$, l'ensemble des solutions d'une équation linéaire est l'intersection de deux droites de \mathbb{R}^2 , donc est :
 - un point si les deux droites ne sont pas parallèles,
 - vide si les deux droites sont parallèles et non confondues,
 - une droite si les deux droites sont parallèles et confondues,
 - sauf si tous les coefficients d'une ligne sont nuls, auquel cas on est ramené au cas $n = 1$.
- Pour $n = 2$ et $p = 3$, l'ensemble des solutions d'une équation linéaire est l'intersection de deux plans de \mathbb{R}^3 , donc est :
 - une droite si les deux plans ne sont pas parallèles,
 - vide si les deux plans sont parallèles et non confondues,
 - un plan si les deux plans sont parallèles et confondues,
 - sauf si tous les coefficients d'une ligne sont nuls, auquel cas on est ramené au cas $n = 1$.

1.2 Opérations élémentaires

Définition 1.2.1. Les *opérations élémentaires* sur les lignes d'un système linéaire (ou d'une matrice) sont :

1. l'*échange* de deux lignes : $L_i \leftrightarrow L_j$,
2. l'*ajout* d'une ligne à une autre : $L_i \leftarrow L_i + \lambda L_j$, où $\lambda \in \mathbb{K}$,
3. la *multiplication* d'une ligne par un scalaire : $L_i \leftarrow \lambda L_i$, où $\lambda \in \mathbb{K}^*$.

Proposition 1.2.2. *Effectuer des opérations élémentaires sur un système n'en modifie pas l'ensemble des solutions. On dit que l'ensemble des solutions du système est invariant par opérations élémentaires.*

Définition 1.2.3. On dit que deux systèmes linéaires sont *équivalents* si l'on peut passer de l'un à l'autre par un nombre fini d'opérations élémentaires.

De même, on dit que deux matrices sont *équivalentes par lignes* si l'on peut passer de l'une à l'autre par un nombre fini d'opérations élémentaires.

Remarque 1.2.4. Par construction, deux systèmes sont équivalents si et seulement si leurs matrices respectives sont équivalentes par lignes.

Cette notion définit une relation d'équivalence sur l'ensemble des systèmes linéaires (et des matrices) !

On peut alors reformuler la proposition précédente comme suit :

Théorème 1.2.5. *Deux systèmes linéaires équivalents ont même ensemble de solutions.*

1.3 Algorithme du pivot de Gauss-Jordan

Définition 1.3.1. Une matrice est dite *échelonnée* par lignes si le premier coefficient non nul de chaque ligne, appelé *pivot*, est situé strictement à gauche du premier coefficient non nul des lignes suivantes.

Une matrice échelonnée par lignes est dite *réduite* si ses pivots sont égaux à 1 et sont les seuls coefficients non nuls de leur colonne.

Théorème 1.3.2 (Théorème de l'algorithme du pivot de Gauss-Jordan). *Toute matrice est équivalente à une unique matrice échelonnée réduite par lignes.*

Démonstration. L'unicité est admise. La preuve de l'existence consiste à décrire un algorithme, dit *du pivot*, dû à C. F. Gauss et C. Jordan. Étant donné une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$ (ou un système linéaire) non nulle, on lui applique les opérations élémentaires suivantes :

1. On identifie la première colonne non nulle (c'est-à-dire contenant un coefficient non nul). Notons j l'indice de cette colonne, puis i l'indice du premier coefficient non nul dans cette colonne. On a donc $a_{ij} \neq 0$. On échange alors les lignes L_1 et L_i :

$$L_i \leftrightarrow L_1,$$

Le coefficient a_{1j} est alors le premier coefficient non nul de la matrice.

2. On normalise la première ligne :

$$L_1 \leftarrow \frac{L_1}{a_{1j}},$$

Le coefficient a_{1j} est alors égal à 1.

3. On annule les coefficients situés sous $a_{1j} = 1$:

$$\forall i \in \llbracket 2, n \rrbracket, L_i \leftarrow L_i - a_{ij}L_1,$$

Le coefficient $a_{1j} = 1$ est alors un pivot de la matrice.

4. On réitère les étapes 1. à 3. sur la sous-matrice allant des colonnes $j + 1$ à p et des lignes 2 à n ,
On obtient alors une matrice échelonnée.
5. On revient à la matrice globale, où l'on annule les coefficients situés au-dessus des pivots a_{ij} :

$$\forall k \in \llbracket 1, i - 1 \rrbracket, L_k \leftarrow L_k - a_{kj}L_i.$$

On obtient alors une matrice échelonnée réduite.

Cet algorithme est *valide* : il donne le résultat attendu (il est donc *correct*), en un nombre fini d'étapes (il *termine*). Il est en effet correct par construction, et termine car à chaque itération des étapes 1. à 3., la taille de la matrice considérée diminue strictement. \square

Remarque 1.3.3.

- Cet algorithme s'applique de manière strictement similaire à un système linéaire. On considère qu'un système linéaire mis sous forme échelonnée réduite est *résolu* : sous cette forme, toutes les informations souhaitées sur l'ensemble des solutions sont facilement accessibles.
- Cet algorithme *ne dépend pas* du second membre du système : il ne prend en compte que les coefficients.

1.4 Résolution

On considère dans cette partie un système mis sous forme échelonnée réduite.

Définition 1.4.1.

- Le *rang* d'un système (ou d'une matrice) est le nombre de ses pivots,
- Les inconnues x_j correspondant aux colonnes contenant des pivots sont appelées *inconnues principales*, ou (par abus) *pivots*, du système,
- Les inconnues x_j non principales sont appelées *inconnues secondaires*, ou *paramètres*, du système.
- Les éventuelles lignes du système situées en-dessous du dernier pivot sont appelées *conditions de compatibilité* du système.

Remarque 1.4.2.

- Les coefficients des éventuelles conditions de compatibilité du système sont tous nuls. Cela est compatible avec l'existence d'une solution si et seulement si les seconds membres correspondants sont également nuls.

- L'ensemble des solutions du système est donné par l'ensemble des p -uplets (x_1, \dots, x_p) où les x_j principaux sont écrits en fonction des x_j secondaires, ces derniers parcourant \mathbb{K} .

Proposition 1.4.3. *Notons r le rang du système considéré. Comme avant, n désigne le nombre de lignes et p le nombre d'inconnues. Alors :*

- $r \leq \min(n, p)$,
 - *Le nombre de paramètres est égal à $p - r$,*
 - *Le nombre de conditions de compatibilité est égal à $n - r$,*
 - *S'il existe une condition de compatibilité qui n'est pas vérifiée, alors l'ensemble des solutions du système est l'ensemble vide (on dit que le système est incompatible).*
- Si toutes les conditions de compatibilité sont vérifiées (on dit que le système est compatible), alors l'ensemble des solutions est non vide, et sa nature géométrique est caractérisée par le nombre de paramètres.*

Exemple 1.4.4. Supposons le système compatible.

Si $p - r = 0$, alors l'ensemble de ses solutions est un point de \mathbb{R}^p .

Si $p - r = 1$, c'est une droite de \mathbb{R}^p .

Si $p - r = 2$, c'est un plan de \mathbb{R}^p .

Etc.

Définition 1.4.5. Si $n = p = r$, on dit que le système est *de Cramer*.

Proposition 1.4.6. *D'après l'étude ci-dessus, un système de Cramer a une et une seule solution.*

Proposition 1.4.7. *Pour $n = p = 2$, le système*

$$\begin{cases} ax + by = r \\ cx + dy = s \end{cases}$$

est de Cramer si et seulement si $ad - bc \neq 0$. On appelle $ad - bc$ le déterminant d'un tel système.

1.5 Vers les matrices

Terminons cette partie par la constatation suivante : le second membre n'est pas intervenu dans l'algorithme du pivot, et n'intervient dans la forme de l'ensemble des solutions que via les conditions de compatibilité. Seule la partie *homogène* du système (=hors second membre) guide le raisonnement :

Définition 1.5.1. Le *système homogène* associé à un système linéaire est le système linéaire de mêmes coefficients, mais « sans second membre » :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p = 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{np}x_p = 0 \end{cases}.$$

La méthode du pivot, qui ne fait intervenir que les coefficients du système, justifie ensuite de ne plus considérer que la *matrice*, c'est-à-dire le tableau, des coefficients :

Définition 1.5.2. La *matrice* associée à un système linéaire est le tableau de ses coefficients :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}.$$

C'est une matrice à n lignes et p colonnes ; on note $M_{n,p}(\mathbb{K})$ l'ensemble de ces matrices.

La *matrice augmentée* est la matrice faisant apparaître le second membre :

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1p} & b_1 \\ a_{21} & a_{22} & \dots & a_{2p} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} & b_n \end{array} \right).$$

Exemple 1.5.3. La matrice associée au système $\begin{cases} 3x - 4y = 1 \\ 2x - 5y = 2 \end{cases}$ est $\begin{pmatrix} 3 & -4 \\ 2 & -5 \end{pmatrix}$. Sa matrice augmentée est $\left(\begin{array}{cc|c} 3 & -4 & 1 \\ 2 & -5 & 2 \end{array} \right)$.

Avec ce formalisme, on retrouve un résultat familier :

Proposition 1.5.4. Soit S l'ensemble des solutions du système $AX = B$, soit X_0 une solution particulière du système, et soit S_h l'ensemble des solutions du système homogène associé. On a alors la relation :

$$S = \{X_0 + Y \mid Y \in S_h\}.$$

Démonstration. $X \in S \Leftrightarrow AX = B = AX_0 \Leftrightarrow A(X - X_0) = 0 \Leftrightarrow X - X_0 \in S_h$. \square

2 Ensembles de matrices

2.1 Définition

Définition 2.1.1. Une *matrice* A à n lignes et p colonnes, ou de *taille* (n, p) , est un tableau de nombres :

$$A = \left(\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{array} \right),$$

où $\forall i \in \llbracket 1, n \rrbracket$, $\forall j \in \llbracket 1, p \rrbracket$, $a_{ij} \in \mathbb{K}$. Les nombres a_{ij} sont les *coefficients* de la matrice A . On note alors $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$. On note $M_{n,p}(\mathbb{K})$ l'ensemble des matrices à n lignes et p colonnes. Si $n = p$, on note plus simplement $M_n(\mathbb{K})$.

Exemple 2.1.2.

- La matrice $\begin{pmatrix} 1 & 2 & -1 \\ 4 & 0 & 3 \end{pmatrix}$ est un élément de $M_{2,3}(\mathbb{R})$.
- La matrice $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ est dans $M_2(\mathbb{R})$. C'est la *matrice nulle* de $M_2(\mathbb{R})$, qu'on pourra noter 0_2 .

2.2 Opérations matricielles

Dans le cadre des systèmes linéaires, on a effectué des additions de lignes et des multiplications de lignes par un scalaire (c'est-à-dire par un nombre). On peut facilement étendre ces opérations à l'ensemble des coefficients d'une matrice, ce qui permet de définir deux premières opérations :

Définition 2.2.1.

1. Soient A et B dans $M_{n,p}(\mathbb{K})$. La *somme* de A et B est la matrice de $M_{n,p}(\mathbb{K})$, notée $A + B$, dont chaque coefficient est la somme des coefficients correspondants de A et de B . Autrement dit :

$$A + B = (a_{ij} + b_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}.$$

2. Soit A dans $M_{n,p}(\mathbb{K})$ et soit λ dans \mathbb{K} . Le *produit* de A par λ est la matrice de $M_{n,p}(\mathbb{K})$, notée λA , dont chaque coefficient est le produit du coefficient correspondant de A par λ . Autrement dit :

$$\lambda A = (\lambda a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}.$$

Remarque 2.2.2.

- On notera que l'addition de deux matrices est définie *si et seulement si elles ont la même taille*.
- En particulier, pour toute matrice A dans $M_{n,p}(\mathbb{K})$, $A + 0_{n,p} = 0_{n,p} + A = A$. De plus, $0 \cdot A = 0_{n,p}$.
- L'addition de deux matrices est une opération *commutative* : $\forall (A, B) \in (M_{n,p}(\mathbb{K}))^2$, $A + B = B + A$.

Ces deux opérations sont rassemblées par la notion suivante :

Définition 2.2.3. Soient A et B dans $M_{n,p}(\mathbb{K})$, et soient λ et μ dans \mathbb{K} . La *combinaison linéaire* de A et B par λ et μ est la matrice suivante :

$$\lambda A + \mu B = (\lambda a_{ij} + \mu b_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}.$$

On va maintenant définir le *produit de deux matrices*. Pour cela, on revient à la notion d'*opérations sur les lignes* vue dans la section précédente. Ces opérations, notées jusqu'ici sous la forme $L_i \leftarrow \sum_j \lambda_{ij} L_j$, peuvent elles-mêmes être codées par la matrice $(\lambda_{ij})_{i,j}$, dont l'action sera représentée comme un *produit à gauche* sur la matrice considérée.

Exemple 2.2.4. Considérons la matrice $\begin{pmatrix} 1 & 2 & -1 \\ 4 & 0 & 3 \end{pmatrix}$. Les opérations $L_1 \leftarrow 2L_1 + 3L_2$ et $L_2 \leftarrow -L_2$ appliquées à cette matrice produisent la matrice $\begin{pmatrix} 14 & 4 & 7 \\ -4 & 0 & -3 \end{pmatrix}$. Cette opération sera dorénavant notée

$$\begin{pmatrix} 2 & 3 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -1 \\ 4 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 4 & 7 \\ -4 & 0 & -3 \end{pmatrix}.$$

Remarque 2.2.5.

- On observe que ce produit est défini si et seulement si à chaque ligne L_j de la matrice de droite correspond une colonne $(\lambda_{ij})_{1 \leq i \leq n}$ dans la matrice de gauche ; c'est-à-dire si *le nombre de lignes de la matrice de droite est égal au nombre de colonnes de la matrice de gauche*.
- Dans le cas le plus élémentaire, le produit d'une matrice à 1 colonne (ou *matrice colonne*) par une matrice à 1 ligne (ou *matrice ligne*) donne :

$$(\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_n) \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = (\lambda_1 c_1 + \lambda_2 c_2 + \dots + \lambda_n c_n) = \left(\sum_{k=1}^n \lambda_k c_k \right).$$

Définition 2.2.6. Soient A dans $M_{n,p}(\mathbb{K})$ et B dans $M_{p,q}(\mathbb{K})$. Le *produit* de B par A est la matrice de $M_{n,q}(\mathbb{K})$, notée AB , dont chaque coefficient est le produit de la colonne correspondante de B par la ligne correspondante de A . Autrement dit :

$$AB = \left((a_{i1} \quad a_{i2} \quad \dots \quad a_{ip}) \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}} = \left(\sum_{k=1}^p a_{ik} b_{kj} \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq q}}.$$

Remarque 2.2.7.

- Comme on l'a vu, le produit AB est défini si et seulement si le nombre de colonnes de A est égal au nombre de lignes de B .
- Le produit matriciel est une opération *non commutative* : la plupart du temps, $AB \neq BA$,
- Le produit matriciel est une opération *non intègre* : $AB = 0 \Rightarrow A = 0$ ou $B = 0$ (mais la réciproque est vraie),
- Le produit matriciel est une opération *associative* : $A(BC) = (AB)C$,
- Le produit matriciel est une opération *distributive* par rapport à l'addition : $A(B + C) = AB + AC$.

Remarquons enfin que l'ensemble d'opérations sur les lignes qui laisse une matrice $A \in M_{n,p}(\mathbb{K})$ invariante, à savoir : $\forall i \in \llbracket 1, n \rrbracket$, $L_i \leftarrow L_i$, correspond au produit à gauche par la matrice :

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in M_n(\mathbb{K}),$$

appelée pour cette raison la *matrice identité* de taille n . On peut d'autre part se convaincre que l'ensemble des opérations sur les lignes codées par la matrice A , appliquées à la matrice identité de taille p , redonne la matrice A . En résumé :

$$\forall A \in M_{n,p}(\mathbb{K}), \quad I_n A = A I_p = A.$$

La matrice identité est donc le « 1 » (ou *élément neutre*) du produit matriciel.

3 Matrices carrées

Outre les opérations décrites ci-dessus, deux opérations peuvent être définies pour les matrices carrées : l'*exponentiation*, ou mise à une puissance, entière, et l'*inversion*.

3.1 Puissances

Étant donné les contraintes mentionnées sur le produit matriciel, une matrice peut être multipliée par elle-même si et seulement si elle est *carrée*, c'est-à-dire lorsque $n = p$. Dans ce cas :

Définition 3.1.1. Soient A dans $M_n(\mathbb{K})$ et k dans \mathbb{N} . La *puissance $k^{\text{ème}}$* de A est la matrice de $M_n(\mathbb{K})$, notée A^k , définie récursivement par : $A^0 = I_n$, et, si $k \in \mathbb{N}^*$: $A^k = A \cdot A^{k-1}$.

Le calcul de A^k est difficile en général. Deux cas font exception :

Définition 3.1.2. Soit $A = (a_{ij})$ dans $M_n(\mathbb{K})$.

1. La matrice A est dite *diagonale* si $\forall i \neq j$, $a_{ij} = 0$.
2. La matrice A est dite *triangulaire supérieure* (resp. *inférieure*) si $\forall i > j$ (resp. $i < j$), $a_{ij} = 0$.
La matrice A est dite *strictement triangulaire supérieure* (resp. *inférieure*) si $\forall i \geq j$ (resp. $i \leq j$), $a_{ij} = 0$.

Proposition 3.1.3.

1. Soit $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n$. On considère la matrice diagonale D de coefficients $\lambda_1, \lambda_2, \dots, \lambda_n$:

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Les puissances k èmes de D sont données par :

$$\forall k \in \mathbb{N}, \quad D^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix}.$$

2. Toute matrice T strictement triangulaire supérieure (ou inférieure) est nilpotente d'ordre n , c'est-à-dire que $T^n = 0_n$.

Remarque 3.1.4. Plus généralement, une matrice A est dite *nilpotente d'ordre* $k \in \mathbb{N}$ si $A^k = 0_n$. Toutes les puissances supérieures de A sont alors triviales.

Il est donc pertinent, pour déterminer les puissances d'une matrice A donnée, de chercher à l'écrire comme somme d'une matrice diagonale et d'une matrice nilpotente : $A = D + N$. Cette décomposition, dite *de Jordan*, existe toujours mais ne sera pas traitée dans ce chapitre. Mentionnons toutefois que cette décomposition n'est intéressante que si elle permet de se ramener au calcul des puissances de D et N . Or, le produit matriciel n'étant pas commutatif, on a par exemple :

$$A^2 = (D + N)^2 = D^2 + DN + ND + N^2,$$

puis

$$A^3 = D^3 + D^2N + DND + ND^2 + DN^2 + NDN + N^2D + N^3\dots$$

Le cas traitable est donc celui où les deux matrices *commutent* ; on retrouve alors une formule bien connue :

Proposition 3.1.5 (Formule du binôme de Newton matricielle). *Soient A et B dans $M_n(\mathbb{K})$ et soit k dans \mathbb{N} . Si A et B commutent, alors :*

$$(A + B)^k = \sum_{j=0}^k \binom{k}{j} A^j B^{k-j}.$$

Démonstration. La preuve suit la même démarche que dans le cas numérique : on raisonne par récurrence. La formule est vérifiée pour $k = 0$; soit $k \in \mathbb{N}^*$, supposons-la vraie jusqu'au rang $k - 1$. Alors :

$$\begin{aligned} (A + B)^k &= (A + B)(A + B)^{k-1} \\ &= (A + B) \sum_{j=0}^{k-1} \binom{k-1}{j} A^j B^{k-1-j} \quad \text{par hypothèse de récurrence,} \\ &= \sum_{j=0}^{k-1} \binom{k-1}{j} A^{j+1} B^{k-1-j} + \sum_{j=0}^{k-1} \binom{k-1}{j} A^j B^{k-j} \quad \text{car } A \text{ et } B \text{ commutent,} \\ &= \sum_{i=1}^k \binom{k-1}{i-1} A^i B^{k-i} + \sum_{j=0}^{k-1} \binom{k-1}{j} A^j B^{k-j} \quad \text{par changement de variables } i = j + 1, \\ &= A^k + \sum_{j=1}^{k-1} \left(\binom{k-1}{j-1} + \binom{k-1}{j} \right) A^j B^{k-j} + B^k \quad \text{en regroupant les termes,} \\ &= \sum_{j=0}^k \binom{k}{j} A^j B^{k-j} \quad \text{d'après la formule du triangle de Pascal.} \end{aligned}$$

□

Remarque 3.1.6. Cette formule est en particulier applicable dès lors que l'une des matrices est *scalaire*, c'est-à-dire de la forme λI_n avec $\lambda \in \mathbb{K}$. Ces matrices ont en effet la propriété de commuter avec toute matrice de $M_n(\mathbb{K})$ (elles en forment le *centre*).

3.2 Inversion

Après avoir défini l'addition, la soustraction et la multiplication matricielles, intéressons-nous à la division. On sait que l'*inverse* d'un nombre non nul $a \in \mathbb{K}^*$ est le nombre b tel que $ab = 1$, et on note $b = \frac{1}{a} = a^{-1}$. De la même manière :

Définition 3.2.1. Soit A dans $M_n(\mathbb{K})$. On dit que A est *inversible* s'il existe B dans $M_n(\mathbb{K})$ telle que

$$AB = BA = I_n.$$

On note alors $B = A^{-1}$. On appelle $GL_n(\mathbb{K})$ (pour *groupe linéaire*) l'ensemble des matrices inverses de taille n .

Proposition 3.2.2. Si une matrice A est inversible, alors son inverse est unique.

Démonstration. Supposons qu'il existe B et B' deux matrices inverses de A . Alors :

$$B = BI_n = B(AB') = (BA)B' = I_nB' = B'.$$

□

Exemple 3.2.3.

- La matrice $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ a pour inverse la matrice $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.
- La matrice $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ n'est pas inversible.

Proposition 3.2.4. Soient A et B dans $M_n(\mathbb{K})$.

1. Si A et B sont inversibles, alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$,
2. Si A est inversible, alors A^{-1} est inversible et $(A^{-1})^{-1} = A$.

Démonstration.

1. Supposons A et B inversibles. Alors $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = I_n$, et de même $(AB)(B^{-1}A^{-1}) = I_n$, d'où le résultat.
2. Supposons A inversible. Alors $AA^{-1} = A^{-1}A = I_n$, d'où le résultat.

□

La détermination de l'inversibilité et le calcul éventuel de A^{-1} sont difficiles en général. On note le cas des matrices diagonales :

Proposition 3.2.5. Soit $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n$. La matrice diagonale D de coefficients $\lambda_1, \lambda_2, \dots, \lambda_n$ est inversible si et seulement si : $\forall j \in \llbracket 1, n \rrbracket$, $\lambda_j \neq 0$, et alors :

$$D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n} \end{pmatrix}.$$

Dans le cas général, on va établir une méthode d'inversion en revenant à notre approche des matrices héritée des systèmes linéaires. Remarquons pour commencer que l'inversion est d'une grande utilité pour la résolution d'un système linéaire. En effet, considérons un système linéaire mis sous forme matricielle : $AX = B$. Si A est inversible, alors le système est résolu par l'écriture : $X = A^{-1}B$.

4 Opérations élémentaires

4.1 Matrices d'opérations élémentaires

Les *opérations élémentaires* vues pour les systèmes linéaires peuvent, avec les remarques précédentes, s'écrire matriciellement.

On commence par définir les *matrices élémentaires* :

Définition 4.1.1. Soit (i, j) dans $\llbracket 1, n \rrbracket^2$. La *matrice élémentaire* $E_{ij} = (e_{kl})_{\substack{1 \leq k \leq n \\ 1 \leq l \leq n}}$ est la matrice dont tous les coefficients sont nuls sauf e_{ij} , qui vaut 1. Autrement dit :

$$E_{ij} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & & \ddots & 1 & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

On peut à présent définir les *matrices d'opérations élémentaires* :

Définition 4.1.2. Soit (i, j) dans $\llbracket 1, n \rrbracket^2$.

1. Soit λ dans \mathbb{K} . La *matrice de transvection* $T_{ij}(\lambda)$ est la matrice correspondant à l'opération élémentaire $L_i \leftarrow L_i + \lambda L_j$, c'est-à-dire :

$$T_{ij}(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & & & & \vdots \\ 0 & \vdots & & 1 & & \lambda & & 0 \\ \vdots & \vdots & & & \ddots & & & \vdots \\ 0 & \vdots & & & & 1 & & 0 \\ \vdots & \vdots & & & & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & \cdots & 1 \end{pmatrix} = I_n + \lambda E_{ij}.$$

2. Soit λ dans \mathbb{K}^* . La *matrice de dilatation* $D_i(\lambda)$ est la matrice correspondant à l'opération élémentaire $L_i \leftarrow \lambda L_i$, c'est-à-dire :

$$D_i(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & & \vdots \\ 0 & \vdots & & \lambda & & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} = I_n + (\lambda - 1) E_{ii}.$$

3. La *matrice de permutation* P_{ij} est la matrice correspondant à l'opération élémentaire $L_i \leftrightarrow L_j$, c'est-à-dire :

$$P_{ij} = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & & & & \vdots \\ 0 & \vdots & & 0 & & 1 & & 0 \\ \vdots & \vdots & & & \ddots & & & \vdots \\ 0 & \vdots & & 1 & & 0 & & 0 \\ \vdots & \vdots & & & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} = I_n + E_{ij} + E_{ji} - E_{ii} - E_{jj}.$$

Proposition 4.1.3. *Les matrices d'opérations élémentaires sont inversibles.*

Démonstration. Il suffit de considérer à chaque fois l'opération élémentaire inverse. On vérifie ainsi que :

1. La matrice inverse de $T_{ij}(\lambda)$ est la matrice correspondant à l'opération $L_i \leftarrow L_i - \lambda L_j$, c'est-à-dire que :

$$T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda),$$

2. La matrice inverse de $D_i(\lambda)$ est la matrice correspondant à l'opération $L_i \leftarrow \frac{L_i}{\lambda}$, c'est-à-dire :

$$D_i(\lambda)^{-1} = D_i\left(\frac{1}{\lambda}\right),$$

3. La matrice inverse de P_{ij} est la matrice correspondant à l'opération $L_i \leftrightarrow L_j$, c'est-à-dire :

$$P_{ij}^{-1} = P_{ij}.$$

□

4.2 Méthode du pivot

D'après notre définition du produit matriciel, le produit d'une matrice $A \in M_n(\mathbb{K})$ à gauche par une matrice d'opération élémentaire a pour résultat la matrice, équivalente par lignes à A , obtenue à partir de A par cette opération élémentaire. Il en découle directement la version matricielle du théorème de Gauss-Jordan :

Théorème 4.2.1 (Théorème de Gauss-Jordan matriciel). *Soit A dans $M_n(\mathbb{K})$. Il existe une unique matrice échelonnée réduite $R \in M_n(\mathbb{K})$, et il existe une matrice E produit de matrices d'opérations élémentaires, telles que :*

$$EA = R.$$

Remarque 4.2.2. Notons que l'unicité n'est garantie que pour la matrice R : comme on l'a vu pour les systèmes linéaires, la suite des opérations élémentaires permettant de passer de A à R n'est quant à elle pas unique.

Exemple 4.2.3. La matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ est équivalente par lignes à la matrice échelonnée réduite $R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, en appliquant les opérations élémentaires successives :

- $L_2 \leftarrow L_2 - L_1$,
- $L_2 \leftarrow -\frac{L_2}{2}$,
- $L_1 \leftarrow L_1 - L_2$,

ce qui s'écrit matriciellement :

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

soit, après multiplication des matrices d'opérations élémentaires entre elles :

$$\underbrace{\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_E \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_A = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_R.$$

Il est alors fondamental d'observer que la matrice R est égale à la matrice identité si et seulement si la matrice A (ou le système linéaire correspondant) possède n pivots, c'est-à-dire est de rang n , c'est-à-dire est de Cramer. On a en fait le résultat :

Proposition 4.2.4. *Soit A dans $M_n(\mathbb{K})$. La matrice A est inversible si et seulement si elle est de rang n .*

Dans ce cas, l'algorithme du pivot de Gauss-Jordan nous fournit une suite d'opérations élémentaires, c'est-à-dire une matrice E , telle que :

$$EA = I_n, \quad \text{c'est-à-dire : } A^{-1} = E.$$

L'algorithme du pivot de Gauss-Jordan nous fournit ainsi une méthode de détermination de l'inversibilité et de calcul éventuel de l'inverse d'une matrice. Couramment appelée *méthode du pivot*, elle peut se présenter ainsi : on écrit côté à côté la matrice A à étudier et la matrice identité :

$$(A \mid I_n),$$

puis on applique l'algorithme de Gauss-Jordan à la matrice A , en effectuant *en parallèle* les *mêmes* opérations élémentaires sur la matrice I_n . On obtient alors l'écriture

$$(EA \mid EI_n), \quad \text{c'est-à-dire } (R \mid E).$$

La matrice A est alors inversible si et seulement si $R = I_n$, et on a alors $A^{-1} = E$.

Avec cette méthode, on détermine par exemple la formule explicite d'inversion des matrices de taille 2 :

Proposition 4.2.5. *Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{K})$. La matrice A est inversible si et seulement si son déterminant est non nul : $ad - bc \neq 0$, et alors :*

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Enfin, récapitulons le lien entre matrices inversibles et systèmes linéaires :

Proposition 4.2.6. *Soient $A \in M_n(\mathbb{K})$ et $X \in \mathbb{K}^n$. Les assertions suivantes sont équivalentes :*

1. *A est inversible,*
2. *$A \sim_L I_n$,*
3. *Si $AX = 0$, alors $X = 0$,*
4. *Pour tout $B \in \mathbb{K}^n$, le système $AX = B$ admet une unique solution,*
5. *Pour tout $B \in \mathbb{K}^n$, le système $AX = B$ admet une solution.*

5 Transposition et symétrie

Les opérations matricielles étudiées jusqu'ici portent sur les *lignes* des matrices. Il est en fait possible de les effectuer, et plus généralement de raisonner, sur les *colonnes*. Il suffit pour cela de changer les colonnes en lignes !

Définition 5.0.1. Soit $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \in M_{n,p}(\mathbb{K})$. La matrice *transposée* de A , notée A^\top ou encore ${}^t A$, est la matrice $A^\top = (b_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} \in M_{p,n}(\mathbb{K})$ telle que : $\forall i \in \llbracket 1, p \rrbracket$, $\forall j \in \llbracket 1, n \rrbracket$, $b_{ij} = a_{ji}$.

Remarque 5.0.2. Cette opération consiste donc à échanger les lignes et les colonnes. Elle revient à effectuer une « symétrie » sur les coefficients de la matrice par rapport à sa diagonale.

Exemple 5.0.3. ${}^t \begin{pmatrix} 1 & 2 & -1 \\ 4 & \mathbf{0} & 3 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 2 & \mathbf{0} \\ -1 & 3 \end{pmatrix}$.

Proposition 5.0.4 (Propriétés de la transposition). *Soient $A \in M_{n,p}(\mathbb{K})$, $B \in M_{q,r}(\mathbb{K})$ et $(\lambda, \mu) \in \mathbb{K}^2$.*

1. *Si $n = q$ et $p = r$, alors $(\lambda A + \mu B)^\top = \lambda A^\top + \mu B^\top$,*
2. *Si $p = q$, alors $(AB)^\top = B^\top A^\top \in M_{r,n}(\mathbb{K})$,*
3. *Si $n = p$ et si A est inversible, alors A^\top est inversible et $(A^\top)^{-1} = (A^{-1})^\top$.*

Démonstration.

1. La preuve est immédiate.
2. Notons $AB = (c_{ij})$, $(AB)^\top = (d_{ij})$, $A^\top = (e_{ij})$ et $B^\top = (f_{ij})$. Alors :

$$\forall i \in \llbracket 1, r \rrbracket, \forall j \in \llbracket 1, n \rrbracket, \quad d_{ij} = c_{ji} = \sum_{k=1}^p a_{jk} b_{ki} = \sum_{k=1}^p f_{ik} e_{kj},$$

d'où le résultat.

3. D'après l'assertion précédente : $AA^{-1} = A^{-1}A = I_n \Rightarrow (A^{-1})^\top A^\top = A^\top (A^{-1})^\top = I_n$, d'où le résultat.

□

L'assertion 2. de la proposition ci-dessus entraîne que, pour appliquer aux colonnes les transformations obtenues sur les lignes comme un produit à gauche, il suffit d'effectuer un produit à droite :

- La transvection $C_i \leftarrow C_i + \lambda C_j$ est un produit à droite par $T_{ij}(\lambda)^\top = T_{ji}(\lambda)$,
- La dilatation $C_i \leftarrow \lambda C_i$ est un produit à droite par $D_i(\lambda)^\top = D_i(\lambda)$,
- La permutation $C_i \leftrightarrow C_j$ est un produit à droite par $P_{ij}^\top = P_{ij}$.

La méthode du pivot peut donc être mise en œuvre en considérant les transformations sur les colonnes ! **Attention** toutefois à ne pas mélanger opérations sur les lignes et opérations sur les colonnes, puisqu'on obtiendrait alors une écriture de la forme :

$$(E_L A E_C \mid E_L E_C),$$

où E_L correspond aux opérations sur les lignes et E_C aux opérations sur les colonnes ; or $E_L A E_C = I_n \Rightarrow A = E_L^{-1} E_C^{-1} \Rightarrow A^{-1} = E_C E_L \neq E_L E_C$.

Notons que ce raisonnement amène au résultat suivant, qui n'a rien d'intuitif :

Proposition 5.0.5. *Le rang d'une matrice est égal au rang de sa transposée.*

Exemple 5.0.6. Le système $\begin{cases} x + 2y - z = 0 \\ 4x + 3z = 0 \end{cases}$ a donc le même nombre de pivots que le système $\begin{cases} x + 4y = 0 \\ 2x = 0 \\ -x + 3y = 0 \end{cases}$.

Terminons ce chapitre par quelques notions sur les matrices *symétriques* :

Définition 5.0.7. Soit A dans $M_n(\mathbb{K})$. La matrice A est dite :

- *symétrique* si $A^\top = A$,
- *antisymétrique* si $A^\top = -A$.

Proposition 5.0.8.

1. *Les matrices antisymétriques sont de diagonale nulle,*
2. *Si A est une matrice à la fois symétrique et antisymétrique, alors $A = 0_n$,*
3. *Toute matrice se décompose de façon unique en somme d'une matrice symétrique et antisymétrique.*

Démonstration.

1. Soit $A = (a_{ij})$ antisymétrique. Alors : $\forall i \in \llbracket 1, n \rrbracket$, $a_{ii} = -a_{ii}$, d'où $a_{ii} = 0$.
2. Soit A symétrique et antisymétrique. Alors $A^\top = A = -A$, d'où $A = 0_n$.
3. Soit $A \in M_n(\mathbb{K})$. Notons $S = \frac{A + A^\top}{2}$ et $T = \frac{A - A^\top}{2}$, alors S et T conviennent. L'unicité est due à l'assertion 2.

□

Remarque 5.0.9. Cette dernière propriété rappelle la décomposition d'une fonction comme somme de sa partie paire et de sa partie impaire. C'est un premier pas vers l'utilisation ultérieure que nous ferons des matrices, qui représenteront... des applications.

Chapitre 11

Suites numériques

On revient en détail dans ce chapitre sur la notion de *suite*, déjà rencontrée sporadiquement dans les chapitres précédents.

1 Généralités

Définition 1.0.1. Une *suite numérique* est une application $u : \mathbb{N} \rightarrow \mathbb{K}$ où $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Par convention, on note :

$$\forall n \in \mathbb{N}, u_n := u(n).$$

De même, la suite u est généralement notée $(u_n)_{n \in \mathbb{N}}$, $(u_n)_{n \geq 0}$, ou plus simplement (u_n) . Si n désigne un entier naturel quelconque, u_n est appelé le *terme général* de la suite (u_n) .

Plus généralement, pour $n_0 \in \mathbb{N}$, une application $u : \mathbb{N} \setminus \{0, \dots, n_0\} \rightarrow \mathbb{K}$ est également considérée comme une suite. On note alors $u = (u_n)_{n \geq n_0}$.

Dans tout ce chapitre, sauf mention contraire, on supposera que les suites considérées sont *réelles*, c'est-à-dire que : $\forall n \in \mathbb{N}, u_n \in \mathbb{R}$.

Remarque 1.0.2. Une suite peut être définie de façon :

- *explicite*, par exemple : $\forall n \in \mathbb{N}, u_n = n^2 + 1$,
- *implicite*, par exemple : $\forall n \in \mathbb{N}, u_n = \begin{cases} 0 & \text{si } n \text{ est pair} \\ 1 & \text{si } n \text{ est impair} \end{cases}$,
- *par récurrence*, par exemple : $u_0 = 1$, et $\forall n \in \mathbb{N}, u_{n+1} = 2u_n$.

Définition 1.0.3 (Vocabulaire sur les suites). Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle. Cette suite est dite :

- *croissante* si : $\forall n \in \mathbb{N}, u_{n+1} \geq u_n$,
- *décroissante* si : $\forall n \in \mathbb{N}, u_{n+1} \leq u_n$,
- *monotone* si elle est croissante *ou* décroissante,
- *constante* si : $\forall n \in \mathbb{N}, u_{n+1} = u_n$,
- *stationnaire* si elle est constante *à partir d'un certain rang* : $\exists N \in \mathbb{N}, \forall n \geq N, u_{n+1} = u_n$,
- *périodique* si : $\exists T \in \mathbb{N}^*, \forall n \in \mathbb{N}, u_{n+T} = u_n$,
- *majorée* si : $\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \leq M$,
- *minorée* si : $\exists m \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \geq m$,
- *bornée* si elle est majorée *et* minorée, ou de manière équivalente si : $\exists C \in \mathbb{R}, \forall n \in \mathbb{N}, |u_n| \leq C$.

Représentation graphique : Une suite réelle peut être représentée dans le plan comme une fonction, ou directement sur l'axe réel.

Étudier une suite consiste, comme pour une fonction, à en déterminer les propriétés, notamment parmi celles ci-dessus. Si la suite n'est pas définie explicitement, cela peut commencer par en déterminer le terme général.

2 Suites récurrentes

On s'intéresse dans cette partie à l'étude des suites définies par récurrence. On y retrouvera des méthodes rencontrées dans le cadre des équations différentielles !

2.1 Suites récurrentes linéaires d'ordre 1

Il s'agit des suites de la forme :

$$u_0 \in \mathbb{R}, \text{ et } \forall n \in \mathbb{N}, u_{n+1} = au_n + b,$$

où $(a, b) \in \mathbb{R}^2$. Dans le cas $a = 1$, ce sont les *suites arithmétiques*, déjà rencontrées au chapitre 1. On a alors :

$$\forall n \in \mathbb{N}, \quad u_n = u_0 + nb, \quad \text{et} \quad \sum_{k=0}^n u_k = (n+1)u_0 + \frac{n(n+1)}{2}b.$$

Les suites arithmétiques sont croissantes si $b \geq 0$, décroissantes si $b \leq 0$, constantes si $b = 0$.

Dans le cas $a \neq 1$ et $b = 0$, ce sont les *suites géométriques*. On a alors :

$$\forall n \in \mathbb{N}, \quad u_n = a^n u_0, \quad \text{et} \quad \sum_{k=0}^n u_k = u_0 \frac{1 - a^{n+1}}{1 - a}.$$

Les suites géométriques sont bornées si et seulement si $-1 \leq a \leq 1$.

Dans le cas $a \neq 1$ et $b \neq 0$, ce sont les *suites arithmético-géométriques*. Pour trouver le terme général de ces suites, on commence par trouver une *solution particulière* : la suite constante de terme général $u_n = k$ vérifie la formule de récurrence si et seulement si :

$$k = ak + b, \quad \text{c'est-à-dire} \quad k = \frac{b}{1 - a}.$$

On soustrait alors la solution particulière pour se ramener à une formule *homogène* :

$$\left\{ \begin{array}{rcl} \forall n \in \mathbb{N}, & u_{n+1} &= au_n + b \\ & k &= ak + b \end{array} \right. \Rightarrow \forall n \in \mathbb{N}, u_{n+1} - k = a(u_n - k).$$

La suite $(u_n - k)$ est donc géométrique, d'où : $\forall n \in \mathbb{N}, u_n - k = a^n(u_0 - k)$, et finalement :

$$\forall n \in \mathbb{N}, u_n = a^n \left(u_0 - \frac{b}{1 - a} \right) + \frac{b}{1 - a}.$$

Les suites arithmético-géométriques sont donc bornées si et seulement si $-1 \leq a \leq 1$.

2.2 Suites récurrentes linéaires d'ordre 2

On considère les suites de la forme :

$$(u_0, u_1) \in \mathbb{R}^2, \text{ et } \forall n \in \mathbb{N}, u_{n+2} = au_{n+1} + bu_n,$$

où $(a, b) \in \mathbb{R}^2$. L'*équation caractéristique* associée à une telle suite est l'équation du second degré :

$$r^2 = ar + b.$$

Notons Δ le discriminant de cette équation. Alors :

Proposition 2.2.1.

- Si $\Delta \neq 0$, notons r_1 et r_2 les deux racines distinctes de l'équation caractéristique. Les suites satisfaisant la formule de récurrence sont les suites de la forme :

$$\forall n \in \mathbb{N}, u_n = \lambda r_1^n + \mu r_2^n, \quad \text{où } (\lambda, \mu) \in \mathbb{R}^2,$$

- Si $\Delta = 0$, notons r_0 la racine double de l'équation caractéristique. Les suites satisfaisant la formule de récurrence sont les suites de la forme :

$$\forall n \in \mathbb{N}, u_n = (\lambda n + \mu) r_0^n, \quad \text{où } (\lambda, \mu) \in \mathbb{R}^2.$$

2.3 Suites récurrentes générales d'ordre 1

Soit $f : I \rightarrow \mathbb{R}$ une fonction définie sur un intervalle I . On considère les suites de la forme :

$$u_0 \in I, \text{ et } \forall n \in \mathbb{N}, u_{n+1} = f(u_n).$$

Proposition 2.3.1. Si $f(I) \subset I$, alors (u_n) est bien définie.

On dit alors que I est stable par f .

Exemple 2.3.2.

- La suite donnée par $\begin{cases} u_0 = 0 \\ \forall n \in \mathbb{N}, u_{n+1} = \sqrt{1 + u_n} \end{cases}$ est bien définie,
- La suite donnée par $\begin{cases} u_0 = 3 \\ \forall n \in \mathbb{N}, u_{n+1} = \sqrt{u_n - 2} \end{cases}$ ne l'est pas.

On suppose désormais que I est stable par f .

Représentation graphique : Une telle suite peut être représentée dans le plan à partir de u_0 , du graphe de f et de la droite $y = x$.

Exemple 2.3.3. $\begin{cases} u_0 \in [0, \frac{\pi}{2}] \\ \forall n \in \mathbb{N}, u_{n+1} = \cos(u_n) \end{cases}$

Remarque 2.3.4. On dit que dans le premier exemple, la suite (u_n) a une structure *en escalier*; dans le second, (u_n) a une structure *en spirale*.

Proposition 2.3.5.

- Si f est croissante, alors la suite (u_n) est monotone, et son sens de variation dépend du signe de $x \mapsto f(x) - x$,
- Si f est décroissante, alors les sous-suites (u_{2n}) et (u_{2n+1}) sont monotones de sens opposé.

L'observation de ces suites montrent que leurs valeurs tendent à se concentrer autour de valeurs "limites" à mesure que n augmente : on dit que ces suites *convergent*.

3 Suites convergentes

3.1 Définition

On dit d'une suite réelle (u_n) qu'elle *converge* vers un réel l si les termes de la suite *se rapprochent de l lorsque n augmente*. Plus précisément, la suite (u_n) converge vers l si le terme général u_n est *aussi proche que l'on veut de l à partir d'un rang n assez grand*. Précisons ces termes :

- Soient $l \in \mathbb{R}$ et $\varepsilon > 0$. On dit d'un réel x qu'il est ε -proche de l si $|x - l| \leq \varepsilon$, c'est-à-dire : $l - \varepsilon \leq x \leq l + \varepsilon$.
- On a déjà vu d'autre part que la locution « à partir d'un certain rang » s'écrit : « $\exists N \in \mathbb{N}, \forall n \geq N$ ».

On peut donc établir la définition :

Définition 3.1.1. Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle et soit $l \in \mathbb{R}$. On dit que la suite (u_n) est *convergente* vers l , ou *converge* vers l , si pour tout ε dans \mathbb{R}_+^* , la suite (u_n) est ε -proche de l à partir d'un certain rang, c'est-à-dire :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |u_n - l| \leq \varepsilon.$$

Remarque 3.1.2. Graphiquement, les valeurs de la suite (u_n) sont « piégées » dans un « tube » de rayon ε à partir du rang N .

Exemple 3.1.3. La suite $\left(\frac{1}{n}\right)_{n \in \mathbb{N}^*}$ converge vers 0.

Remarque 3.1.4.

- Il est fondamental, dans la définition ci-dessus, de prendre $\varepsilon > 0$ et non $\varepsilon \geq 0$ (cf TD). Les autres inégalités peuvent par contre être strictes ou larges sans problème.
- Notons que la suite (u_n) converge vers l si et seulement si la suite $(|u_n - l|)$ converge vers 0.
- Une suite non convergente est dite *divergente*. C'est en particulier le cas des suites qui « tendent vers l'infini », dont nous reparlerons plus bas.

Proposition 3.1.5 (Unicité de la limite). *Si une suite (u_n) converge vers $l \in \mathbb{R}$, alors l est unique. On dit que l est la limite de la suite (u_n) .*

Démonstration. Supposons par l'absurde que (u_n) converge à la fois vers l et vers $l' \neq l$. Posons alors $\varepsilon = \frac{|l - l'|}{3} > 0$. Il existe alors $N_1 \in \mathbb{N}$ tel que : $\forall n \geq N_1, |u_n - l| \leq \varepsilon$, et de même, il existe $N_2 \in \mathbb{N}$ tel que : $\forall n \geq N_2, |u_n - l'| \leq \varepsilon$. Posons alors $N = \max(N_1, N_2)$. On a :

$$\forall n \geq N, |l - l'| = |l - u_n + u_n - l'| \leq |l - u_n| + |u_n - l'| \leq 2\varepsilon = \frac{2}{3}|l - l'|,$$

par inégalité triangulaire, donc $1 \leq \frac{2}{3}$, ce qui est absurde. Donc $l = l'$. \square

La notion de suite convergente vient s'ajouter à notre vocabulaire de description des suites. Elle est liée aux notions précédentes, avec par exemple la propriété suivante :

Proposition 3.1.6. *Toute suite convergente est bornée.*

Démonstration. Soit (u_n) une suite convergente vers $l \in \mathbb{R}$. Prenons $\varepsilon = 5$. Soit $N \in \mathbb{N}$ tel que : $\forall n \geq N, |u_n - l| \leq 5$, c'est-à-dire $l - 5 \leq u_n \leq l + 5$. À partir du rang N , la suite (u_n) est donc bornée par $C_2 = \max(|l - 5|, |l + 5|)$. Avant le rang N , la suite (u_n) est bornée par $C_1 = \max(|u_0|, |u_1|, \dots, |u_N|)$. La suite (u_n) est donc bornée par $C = \max(C_1, C_2)$. \square

Et plus spécifiquement :

Proposition 3.1.7. *Si une suite (u_n) converge vers $l \neq 0$, alors les termes de la suite sont non nuls à partir d'un certain rang.*

Démonstration. Posons $\varepsilon = \frac{|l|}{2} > 0$. Soit $N \in \mathbb{N}$ tel que : $\forall n \geq N, |u_n - l| \leq \varepsilon$, alors : $\forall n \geq N, u_n \in \left[\frac{l}{2}, \frac{3l}{2}\right]$ si $l > 0$, $\left[\frac{3l}{2}, \frac{l}{2}\right]$ si $l < 0$. Dans tous les cas : $\forall n \geq N, u_n \neq 0$. \square

3.2 Opérations sur les limites

Proposition 3.2.1. Soient (u_n) et (v_n) deux suites convergeant vers $l \in \mathbb{R}$ et $l' \in \mathbb{R}$ respectivement. Alors :

1. La suite $(u_n + v_n)$ est convergente, de limite $l + l'$,
2. Pour tout $\lambda \in \mathbb{R}$, la suite (λu_n) est convergente, de limite λl ,
3. La suite $(u_n v_n)$ est convergente, de limite ll' ,
4. Si $l' \neq 0$, alors la suite $\left(\frac{u_n}{v_n}\right)$ est bien définie à partir d'un certain rang et est convergente, de limite $\frac{l}{l'}$.

Remarque 3.2.2. Plus généralement, avec les notations ci-dessus, si $f : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction **continue en l** , alors la suite $(f(u_n))$ est convergente, de limite $f(l)$.

Proposition 3.2.3 (Limites et inégalités). Soient (u_n) et (v_n) deux suites convergeant vers $l \in \mathbb{R}$ et $l' \in \mathbb{R}$ respectivement. Alors :

1. Si $\forall n \in \mathbb{N}, u_n \leq v_n$, alors $l \leq l'$,
2. Si $\forall n \in \mathbb{N}, u_n < v_n$, alors $l < l'$.

Remarque 3.2.4. Il faut bien sûr prendre garde au fait que dans la seconde assertion, l'inégalité **ne reste pas toujours** stricte lorsque l'on « passe à la limite ». C'est par exemple le cas pour les suites de termes généraux $u_n = 1 - \frac{1}{n}$ et $v_n = 1 + \frac{1}{n}$.

En pratique, on notera le fait qu'une suite converge vers $l \in \mathbb{R}$ comme suit :

$$u_n \xrightarrow[n \rightarrow +\infty]{} l.$$

On utilisera la même notation pour certaines suites divergentes :

Définition 3.2.5.

- On dit qu'une suite (u_n) *tend vers $+\infty$* , et on note $u_n \xrightarrow[n \rightarrow +\infty]{} +\infty$, si : $\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \geq N, u_n \geq A$.
- On dit qu'une suite (u_n) *tend vers $-\infty$* , et on note $u_n \xrightarrow[n \rightarrow +\infty]{} -\infty$, si : $\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \geq N, u_n \leq A$.

Les règles d'opérations sur les limites vues ci-dessus s'appliquent également à ces suites, mais celles-ci peuvent donner lieu à des *formes indéterminées*, c'est-à-dire des opérations dont le résultat peut varier selon le cas considéré :

$$\infty - \infty, \quad \infty \times 0, \quad \frac{\infty}{\infty}, \quad 1^\infty \dots$$

Exemple 3.2.6.

- $u_n = \sqrt{n+1} - \sqrt{n}$,
 $v_n = e^{n+1} - e^n$,
- $u_n = n - n^2$,
 $v_n = n^2 - n$,
- $u_n = \left(1 + \frac{1}{n}\right)^n$.

3.3 Résultats d'existence d'une limite

La détermination de la convergence d'une suite, puis le calcul éventuel de sa limite, peuvent être difficiles. Les résultats ci-après permettent de traiter certains de ces cas.

Théorème 3.3.1 (Théorème d'encadrement (ou des gendarmes)). Soient (u_n) , (v_n) et (w_n) trois suites telles que :

$$\forall n \in \mathbb{N}, u_n \leq v_n \leq w_n.$$

Si (u_n) et (w_n) convergent vers la même limite $l \in \mathbb{R}$, alors la suite (v_n) est convergente, de limite l également.

Démonstration. Soit $\varepsilon > 0$. Soit $N_1 \in \mathbb{N}$ tel que : $\forall n \geq N_1$, $|u_n - l| \leq \varepsilon$ et soit $N_2 \in \mathbb{N}$ tel que : $\forall n \geq N_1$, $|w_n - l| \leq \varepsilon$. Notons $N = \max(N_1, N_2)$, on a alors :

$$\forall n \geq N, \quad l - \varepsilon \leq u_n \leq v_n \leq w_n \leq l + \varepsilon,$$

et donc : $\forall n \geq N$, $|v_n - l| \leq \varepsilon$. Donc (v_n) converge vers l . \square

Exemple 3.3.2. La suite de terme général $v_n = \frac{\sin n}{n}$ converge vers 0.

Un résultat similaire peut être établi pour les suites qui tendent vers l'infini :

Théorème 3.3.3 (Théorème de divergence par minoration (ou majoration)). *Soient (u_n) et (v_n) deux suites telles que :*

$$\forall n \in \mathbb{N}, \quad u_n \leq v_n.$$

Si (u_n) tend vers $+\infty$, alors (v_n) tend vers $+\infty$.

Si (v_n) tend vers $-\infty$, alors (u_n) tend vers $-\infty$.

Démonstration. On va montrer la première assertion (la démonstration de la seconde est similaire). Soit $A \in \mathbb{R}$. Soit $N \in \mathbb{N}$ tel que : $\forall n \geq N$, $u_n \geq A$, alors : $\forall n \geq N$, $v_n \geq u_n \geq A$. Donc (v_n) tend vers $+\infty$. \square

Exemple 3.3.4. La suite de terme général $v_n = n + (-1)^n$ tend vers $+\infty$.

Théorème 3.3.5 (Théorème de la convergence monotone (ou de la limite monotone)).

- *Toute suite croissante et majorée est convergente.*
- *Toute suite décroissante et minorée est convergente.*

Exemple 3.3.6. La suite de terme général $u_n = \sum_{k=1}^n \frac{1}{k^2}$ est convergente.

La détermination de la limite de cette suite est l'objet du problème de Bâle, résolu par Euler en 1741. Elle est égale à $\frac{\pi^2}{6}$.

La démonstration du théorème de la convergence monotone fait appel de façon fine à la structure de \mathbb{R} , et plus précisément au résultat suivant que nous admettrons :

Proposition 3.3.7.

- *Soit A une partie de \mathbb{R} non vide et majorée. L'ensemble des majorants de A admet un plus petit élément, appelé borne supérieure de A et noté $\sup A$.*
- *Soit A une partie de \mathbb{R} non vide et minorée. L'ensemble des minorants de A admet un plus grand élément, appelé borne inférieure de A et noté $\inf A$.*

Exemple 3.3.8.

- Pour $A = [-5, 4[$: $\inf A = -5$, $\sup A = 4$,
- Pour $A = \left\{ \frac{1}{n} \mid n \in \mathbb{N}^* \right\}$: $\inf A = 0$, $\sup A = 1$.

Pour montrer le théorème de la convergence monotone, nous aurons plus particulièrement besoin de la propriété suivante :

Proposition 3.3.9. *Soit A une partie de \mathbb{R} non vide et majorée, de borne supérieure l . Alors : $\forall \varepsilon > 0$, $\exists a \in A$, $l - \varepsilon \leq a \leq l$.*

Démonstration. Par l'absurde : Supposons qu'il existe $\varepsilon > 0$ tel que : $\forall a \in A$, $|l - a| > \varepsilon$. Comme l est un majorant de A , on a donc : $\forall a \in A$, $a < l - \varepsilon$. Donc $l - \varepsilon$ est un majorant de A , ce qui contredit le fait que l soit la borne supérieure de A . \square

Démonstration du théorème de la convergence monotone. Soit (u_n) une suite croissante majorée. Posons :

$$A = \{u_n \mid n \in \mathbb{N}\}.$$

La partie A de \mathbb{R} est non vide et majorée, donc admet une borne supérieure l . Soit $\varepsilon > 0$. D'après la proposition ci-dessus, il existe un élément de A , c'est-à-dire un entier N , tel que $|u_N - l| \leq \varepsilon$. Comme la suite (u_n) est croissante et majorée par l , on a donc :

$$\forall n \geq N, \quad l - \varepsilon \leq u_N \leq u_n \leq l,$$

et donc : $\forall n \geq N, |u_n - l| \leq \varepsilon$. Donc (u_n) converge vers l . \square

Ce dernier théorème nous permet également d'accéder au résultat suivant :

Théorème 3.3.10 (Théorème des suites adjacentes). *Soient (u_n) et (v_n) deux suites telles que :*

1. *la suite (u_n) est croissante,*
2. *la suite (v_n) est décroissante,*
3. $u_n - v_n \xrightarrow{n \rightarrow +\infty} 0$.

Alors les suites (u_n) et (v_n) sont convergentes, et ont la même limite.

Démonstration. D'après les hypothèses, la suite $(u_n - v_n)$ est croissante et converge vers 0. Par conséquent : $\forall n \in \mathbb{N}, u_n - v_n \leq 0$, et plus précisément :

$$\forall n \in \mathbb{N}, \quad u_0 \leq u_n \leq v_n \leq v_0.$$

En particulier, la suite (u_n) est donc majorée par v_0 , et croissante, donc convergente vers $l \in \mathbb{R}$ d'après le théorème précédent. De même la suite (v_n) est minorée par u_0 et décroissante, donc convergente vers $l' \in \mathbb{R}$.

De plus, $u_n - v_n \xrightarrow{n \rightarrow +\infty} l - l'$, donc, d'après la troisième hypothèse, $l = l'$. \square

Exemple 3.3.11. Les suites de termes généraux $u_n = \sum_{k=1}^{2n} \frac{(-1)^{k+1}}{k}$ et $v_n = u_n + \frac{1}{2n+1}$ sont adjacentes.

Leur limite commune est égale à $\ln 2$.

4 Suites extraites

Étant donné une suite (u_n) , on peut en *extraire* une *sous-suite*, comme entre ensembles et sous-ensembles. On demande toutefois que l'**ordre** des termes extraits soit conservé. On peut formaliser cela ainsi :

Définition 4.0.1. Une *extraction* de \mathbb{N} est une application $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante.

Définition 4.0.2. Soit $(u_n)_{n \in \mathbb{N}}$ une suite et φ une extraction de \mathbb{N} . La *suite extraite* de (u_n) par φ est la suite $(u_{\varphi(n)})_{n \in \mathbb{N}}$. On dit également que c'est une *sous-suite* de la suite (u_n) .

Exemple 4.0.3. La suite extraite de $\left(\frac{(-1)^n}{n}\right)_{n \in \mathbb{N}^*}$ par $\varphi : n \rightarrow 2n$ est la suite $\left(\frac{1}{2n}\right)_{n \in \mathbb{N}^*}$.

L'intérêt d'extraire des sous-suites consiste à faire émerger des structures sous-jacentes dans une suite donnée. Ainsi, la suite de terme général $u_n = (-1)^n$ est constituée des sous-suites constantes $(u_{2n} = 1)$ et $(u_{2n+1} = -1)$.

La plupart des propriétés d'une suite restent valables pour ses sous-suites :

Proposition 4.0.4. *Si une suite (u_n) est croissante (respectivement décroissante, constante, stationnaire, majorée, minorée, bornée, convergente), alors toutes ses sous-suites $(u_{\varphi(n)})$ sont croissantes (respectivement décroissantes, constantes, stationnaires, majorées, minorées, bornées, convergentes).*

La périodicité est une exception notable : si une suite (u_n) est périodique, alors ses sous-suites ne sont pas forcément périodiques. Par exemple, la suite de terme général $u_n = \sin\left(\frac{n\pi}{3}\right)$ est 6-périodique, mais sa sous-suite (u_{n^2}) ne l'est pas.

Inversement, si une sous-suite $(u_{\varphi(n)})$ d'une suite (u_n) a une propriété P , cela n'indique **pas** que la suite (u_n) vérifie P . Par exemple, la suite définie par : $\forall n \in \mathbb{N}, u_n = \begin{cases} n & \text{si } n \text{ est pair} \\ \frac{1}{n} & \text{si } n \text{ est impair} \end{cases}$ a une sous-suite (u_{2n+1}) convergente vers 0, mais la suite (u_n) n'est pas convergente.

On peut cependant énoncer les résultats suivants :

Proposition 4.0.5. *Si une suite (u_n) admet deux sous-suites $(u_{\varphi(n)})$ et $(u_{\psi(n)})$ convergentes vers l et $l' \neq l$ respectivement, alors (u_n) diverge.*

Proposition 4.0.6. *Si une suite (u_n) est monotone et admet une sous-suite $(u_{\varphi(n)})$ qui tend vers $l \in \mathbb{R} \cup \{\pm\infty\}$, alors (u_n) tend vers l .*

Exemple 4.0.7. La suite de terme général $H_n = \sum_{k=1}^n \frac{1}{k}$ tend vers $+\infty$.

Enfin, mentionnons que la propriété 3.1.6 admet la "réciproque" suivante :

Théorème 4.0.8 (Théorème de Bolzano-Weierstrass). *Toute suite bornée admet une sous-suite convergente.*

Exemple 4.0.9. La suite $(\cos n)$ admet une sous-suite convergente.

5 Cas des suites complexes

Nous avons traité dans ce chapitre du cas des suites réelles. Un certain nombre des notions et résultats rencontrés peuvent être étendus aux suites de nombres complexes. Le seul obstacle est en fait l'absence de relation d'ordre sur \mathbb{C} , qui rend caducs les énoncés faisant intervenir une comparaison des termes de la suite : notion de suite croissante ou décroissante majorée ou minorée, tendant vers l'infini, théorème d'encadrement, de convergence monotone, des suites adjacentes...

Les notions de suite bornée et convergente sont cependant conservées : il suffit d'y remplacer la valeur absolue par le module, dont on rappelle qu'ils correspondent tous deux à la distance du nombre à 0.

Exemple 5.0.1. La suite $\left(\frac{e^{in}}{n^2}\right)$ converge vers 0.

Une suite complexe (u_n) peut être étudiée via sa décomposition en partie réelle et partie imaginaire : notons, pour tout $n \in \mathbb{N}$, $u_n = x_n + iy_n$ cette décomposition. On a par exemple le résultat suivant :

Proposition 5.0.2. *La suite (u_n) est convergente si et seulement si les suites (x_n) et (y_n) sont convergentes.*

Démonstration. Soit $l = a + ib$ dans \mathbb{C} , on a :

$$\forall n \in \mathbb{N}, \quad |u_n - l|^2 = |x_n - a|^2 + |y_n - b|^2,$$

donc si $|x_n - a|$ et $|y_n - b|$ tendent vers 0, alors $|u_n - l|$ aussi, et inversement : $\forall n \in \mathbb{N}, |x_n - a|^2 \leq |u_n - l|^2$, donc si $|u_n - l|$ tend vers 0, alors $|x_n - a|$ aussi, et de même pour $|y_n - b|$. \square

6 Relations de comparaison

Revenons aux suites réelles, et terminons ce chapitre par l'introduction des *notations de Landau*, qui permettent de décrire et de comparer entre elles les *vitesses de croissance* des suites.

Dans toute cette partie, on considère deux suites (u_n) et (v_n) , et **on suppose que (v_n) ne s'annule pas**.

6.1 Équivalents

Définition 6.1.1. On dit que (u_n) est *équivalente* à (v_n) , et on note $u_n \sim v_n$ (« équivalent à »), si la suite $\left(\frac{u_n}{v_n}\right)$ converge vers 1.

Remarque 6.1.2.

- Dire que $u_n \sim v_n$ signifie donc intuitivement que « la vitesse de croissance de (u_n) est *la même* que celle de (v_n) ».
- Par définition, **on ne peut pas écrire** $u_n \sim 0$.

Exemple 6.1.3.

- $3n^2 - 2n + 7 \sim 3n^2$,
- $(-1)^n \sim (-1)^n + \frac{1}{n}$.

Proposition 6.1.4.

1. $u_n \sim l \in \mathbb{R}^*$ si et seulement si (u_n) converge vers l .
2. Si $u_n \sim v_n$ et si $v_n \rightarrow l \in \overline{\mathbb{R}}$, alors $u_n \rightarrow l$.

Proposition 6.1.5. Soient (u_n) , (v_n) , (w_n) et (y_n) des suites.

1. Si $u_n \sim w_n$ et $v_n \sim y_n$, alors $u_n v_n \sim w_n y_n$,
2. Si $u_n \sim v_n$ et si ces suites ne s'annulent pas, alors $\frac{1}{u_n} \sim \frac{1}{v_n}$,
3. Si $u_n \sim v_n$ et $v_n \sim w_n$, alors $u_n \sim w_n$,

Remarque 6.1.6.

- La relation \sim est réflexive, symétrique et transitive : c'est une *relation d'équivalence*, comme son nom l'indique.
- **On ne peut pas sommer des équivalents.** Ainsi, pour $u_n = \frac{1}{n} + \frac{1}{n^2}$, $v_n = \frac{1}{n} + \frac{1}{n^3}$ et $w_n = -\frac{1}{n}$, on a $u_n \sim v_n$ et $w_n \sim w_n$, mais $u_n + w_n \not\sim v_n + w_n$.

On retiendra les équivalents usuels suivants :

Proposition 6.1.7. Si $u_n \rightarrow 0$ et ne s'annule pas, alors :

- $\sin u_n \sim u_n$,
- $\tan u_n \sim u_n$,
- $\ln(1 + u_n) \sim u_n$,
- $e^{u_n} - 1 \sim u_n$,
- $\forall \alpha \in \mathbb{R}^*$, $(1 + u_n)^\alpha - 1 \sim \alpha u_n$,
- $1 - \cos(u_n) \sim \frac{u_n^2}{2}$.

Mentionnons enfin la mirifique *formule de Stirling* : $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$.

6.2 Petits- o

Définition 6.2.1. On dit que (u_n) est *négligeable* devant (v_n) , et on note $u_n = o(v_n)$ (« petit o de »), si la suite $\left(\frac{u_n}{v_n}\right)$ converge vers 0.

Remarque 6.2.2.

- Dire que $u_n = o(v_n)$ signifie donc intuitivement que « la vitesse de croissance de (u_n) est *strictement inférieure* à celle de (v_n) ».
- Par définition, **on ne peut pas écrire** $u_n = o(0)$.

Exemple 6.2.3.

- $3n^2 - 2n + 7 = o(n^5)$,
- $\ln n = o(e^n)$.

Proposition 6.2.4. Soient (u_n) , (v_n) , (w_n) et (y_n) des suites.

1. Si $u_n = o(w_n)$ et $v_n = o(w_n)$, alors $u_n + v_n = o(w_n)$,
2. Si $u_n = o(w_n)$ et $v_n = o(y_n)$, alors $u_n v_n = o(w_n y_n)$,
3. Si $u_n = o(v_n)$ et $v_n = o(w_n)$, alors $u_n = o(w_n)$,
4. $u_n = o(1)$ si et seulement si (u_n) converge vers 0.

Remarque 6.2.5. La relation o est transitive, mais n'est ni réflexive, ni symétrique ni antisymétrique.

Proposition 6.2.6.

1. $(u_n \sim v_n) \Leftrightarrow (u_n - v_n = o(v_n))$,
2. $(u_n = o(v_n)) \Leftrightarrow (u_n + v_n \sim v_n)$,
3. Si $u_n \sim v_n$, alors (u_n) et (v_n) sont de même signe à partir d'un certain rang,
4. Si $u_n \sim v_n$ et que (v_n) tend vers $l \in \overline{\mathbb{R}}$, alors (u_n) tend vers l ,
5. Si (u_n) et (v_n) convergent vers un réel l **non nul**, alors $u_n \sim v_n$.

Contre-exemple si $l = 0$: $u_n = \frac{1}{n}$ et $v_n = \frac{1}{n^2}$.

6.3 Grands- O

Définition 6.3.1. On dit que (u_n) est *de l'ordre de* (v_n) , et on note $u_n = O(v_n)$ (« grand O de »), si la suite $\left(\frac{u_n}{v_n}\right)$ est bornée.

Remarque 6.3.2.

- Dire que $u_n = O(v_n)$ signifie donc intuitivement que « la vitesse de croissance de (u_n) est *inférieure ou égale* à celle de (v_n) ».
- Par définition, **on ne peut pas écrire** $u_n = O(0)$.

Exemple 6.3.3.

- $3n^2 - 2n + 7 = O(n^2)$,
- $\ln n = O(e^n)$.

Proposition 6.3.4. Soient (u_n) , (v_n) , (w_n) et (y_n) des suites.

1. Si $u_n = O(w_n)$ et $v_n = O(w_n)$, alors $u_n + v_n = O(w_n)$,
2. Si $u_n = O(w_n)$ et $v_n = O(y_n)$, alors $u_n v_n = O(w_n y_n)$,
3. Si $u_n = O(v_n)$ et $v_n = O(w_n)$, alors $u_n = O(w_n)$,
4. $u_n = O(1)$ si et seulement si (u_n) est bornée.

Remarque 6.3.5. La relation O est réflexive et transitive, mais n'est ni symétrique ni antisymétrique.

Chapitre 12

Polynômes

On revient en détail dans ce chapitre sur la notion de *polynôme*, déjà rencontrée sporadiquement dans les chapitres précédents. On note dans tout le chapitre $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

1 L'ensemble $\mathbb{K}[X]$

1.1 Définitions

Définition 1.1.1. Un *polynôme* à coefficients dans \mathbb{K} est une expression de la forme

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

où $n \in \mathbb{N}$, $(a_n, a_{n-1}, \dots, a_0) \in K^{n+1}$ sont les *coefficients* du polynôme, et X est l'*indéterminée*. On note $\mathbb{K}[X]$ l'ensemble des polynômes à coefficients dans \mathbb{K} .

Exemple 1.1.2.

- $X^2 + 3X + 2 \in \mathbb{R}[X]$,
- $-2X^3 + iX - \sqrt{2} \in \mathbb{C}[X]$,
- $0 \in \mathbb{K}[X]$.

Remarque 1.1.3.

- On a bien sûr $\mathbb{R}[X] \subset \mathbb{C}[X]$.
- Lorsqu'il n'y a pas de confusion possible, on notera souvent $P = P(X)$.
- Dans la définition ci-dessus, X n'est pas une *inconnue* ni une *variable* : cela voudrait dire que X appartient à un ensemble bien identifié. En tant qu'*indéterminée*, c'est simplement une notation permettant d'écrire le polynôme. Autrement dit, on peut assimiler le polynôme à l'ensemble de ses coefficients $(a_n, a_{n-1}, \dots, a_0)$.
- Cette approche permet une étude plus générale que celle de la *fonction polynomiale*

$$f : x \mapsto a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

définie sur \mathbb{R} ou sur \mathbb{C} . Comme on va le voir, l'*indéterminée* X d'un polynôme peut a priori représenter un élément de n'importe quel ensemble sur lequel il existe une addition, une multiplication, et la notion de multiplication par un *scalaire* (=un élément de \mathbb{K}) : X peut donc désigner une matrice, une fonction...

Définition 1.1.4. Avec les notations de la définition précédente, si a_n est non nul, alors n est appelé le *degré* du polynôme P , et a_n est son *coefficent dominant*. On note $n = \deg P$. Par convention, $\deg 0 = -\infty$.

Pour tout n dans \mathbb{N} , on note $\mathbb{K}_n[X]$ l'ensemble des polynômes à coefficients dans \mathbb{K} de degré inférieur ou égal à n .

Exemple 1.1.5.

- $\mathbb{K}_2[X] = \{aX^2 + bX + c \mid (a, b, c) \in \mathbb{K}^3\}$,
- $\mathbb{K}_0[X] \simeq \mathbb{K}$ est l'ensemble des *polynômes constants*,
- On a ainsi une suite d'ensembles *croissante par inclusion* :

$$\mathbb{K}_0[X] \subset \mathbb{K}_1[X] \subset \dots \subset \mathbb{K}_n[X] \subset \mathbb{K}_{n+1}[X] \subset \dots \subset \mathbb{K}[X].$$

Définition 1.1.6. Soit P dans $\mathbb{K}[X]$, avec les notations précédentes. Soit A une \mathbb{K} -algèbre, c'est-à-dire un ensemble dont les éléments peuvent être additionnés et multipliés entre eux, et multipliés par des éléments de \mathbb{K} . Soit b dans A . L'*évaluation de P en b* est l'élément de A , noté $P(b)$, défini par :

$$P(b) = a_n b^n + a_{n-1} b^{n-1} + \dots + a_1 b + a_0 b^0.$$

Exemple 1.1.7.

- L'évaluation de $P = X^4 - 3$ en i est le nombre $P(i) = i^4 - 3 = -2$.
- L'évaluation de $P = X^2 + 1$ en la matrice $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ est la matrice $P(M) = M^2 + I_2 = 0_2$.
- L'évaluation de $P = X^3$ en la fonction $f : x \mapsto \frac{1}{x}$ est la fonction $P(f) : x \mapsto \frac{1}{x^3}$.

1.2 Opérations

On définit ici formellement les règles de manipulation des polynômes, héritées de celles pratiquées sur les fonctions polynomiales.

Mentionnons tout d'abord que, d'après leur définition, deux polynômes sont égaux si et seulement si leurs coefficients sont égaux :

Proposition 1.2.1. Soient $P = \sum_{k=0}^n a_k X^k$ et $Q = \sum_{k=0}^m b_k X^k$ dans $\mathbb{K}[X]$. On a l'équivalence :

$$(P = Q) \Leftrightarrow (n = m \quad \text{et} \quad \forall k \in \llbracket 0, n \rrbracket, a_k = b_k).$$

Les opérations sur $\mathbb{K}[X]$ sont définies comme suit :

Définition 1.2.2. Soient $P = \sum_{k=0}^n a_k X^k$ et $Q = \sum_{k=0}^m b_k X^k$ dans $\mathbb{K}[X]$. On suppose que $n \geq m$. Alors :

1. $P + Q = \sum_{k=0}^m (a_k + b_k) X^k + \sum_{k=m+1}^n a_k X^k$,
2. $\forall \lambda \in \mathbb{K}, \lambda P = \sum_{k=0}^n \lambda a_k X^k$,
3. $PQ = P \times Q = \sum_{k=0}^{m+n} c_k X^k$, où : $\forall k \in \llbracket 0, m+n \rrbracket, c_k = \sum_{i+j=k} a_i b_j = \sum_{i=0}^k a_i b_{k-i}$,
4. $P^0 = 1$ et $\forall k \in \mathbb{N}^*, P^k = P \times P^{k-1}$,
5. $P \circ Q = \sum_{k=0}^n a_k Q^k$.

On observe en particulier le comportement du degré vis-à-vis de ces opérations :

Proposition 1.2.3. Soient P et Q dans $\mathbb{K}[X]$. Alors :

1. $\deg(P + Q) \leq \max(\deg P, \deg Q)$,
2. $\forall \lambda \in \mathbb{K}^*, \deg(\lambda P) = \deg P$,
3. $\deg(PQ) = \deg P + \deg Q$,
4. $\forall k \in \mathbb{N}^*, \deg P^k = k \deg P$,
5. $\deg(P \circ Q) = \deg P \deg Q$.

Remarque 1.2.4. En particulier, on remarque que pour tout $n \in \mathbb{N}$, l'ensemble $\mathbb{K}_n[X]$ est *stable par combinaison linéaire* : $\forall (P, Q) \in (\mathbb{K}_n[X])^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \lambda P + \mu Q \in \mathbb{K}_n[X]$.

2 Arithmétique dans $\mathbb{K}[X]$

Les opérations sur $\mathbb{K}[X]$ décrites ci-dessus, ainsi que la structure apportée à cet ensemble par la notion de degré (c'est un *préordre*), permettent d'y définir une division euclidienne, comme on va le voir maintenant.

2.1 Divisibilité

Définition 2.1.1. Soient A et B dans $\mathbb{K}[X]$. On dit que B divise A , ou que A est un *multiple* de B , et on note $B|A$, s'il existe Q dans $\mathbb{K}[X]$ tel que $A = BQ$.

Exemple 2.1.2.

- $X^2 + X + 1 \mid X^3 - 1$,
- $\forall P \in \mathbb{K}[X], P \mid 0$,
- $\forall \lambda \in \mathbb{K}^*, \forall P \in \mathbb{K}[X], \lambda \mid P$.

En particulier, on observe que :

Proposition 2.1.3. Soient A et B dans $\mathbb{K}[X]$. Si B divise A et si $A \neq 0$, alors $\deg B \leq \deg A$.

Cette observation va nous permettre de parler de *division euclidienne* sur $\mathbb{K}[X]$. Remarquons par ailleurs les similitudes et les différences avec certains résultats vus en arithmétique sur \mathbb{Z} :

Proposition 2.1.4.

- Soient A, B et D dans $\mathbb{K}[X]$. On a l'équivalence :

$$(D|A \text{ et } D|B) \Leftrightarrow (\forall (P, Q) \in (\mathbb{K}[X])^2, D|PA + QB).$$

- Soient A et B dans $\mathbb{K}[X]$. On a l'implication :

$$(B|A \text{ et } A|B) \Rightarrow (\exists \lambda \in \mathbb{K}^*, A = \lambda B).$$

Remarque 2.1.5. Cette dernière propriété est une manifestation du fait que le *groupe des unités* (c'est-à-dire l'ensemble des diviseurs de 1) de $\mathbb{K}[X]$ est \mathbb{K}^* , alors que celui de \mathbb{Z} est réduit à $\{-1, 1\}$.

2.2 Division euclidienne

Théorème 2.2.1. (*Division euclidienne dans $\mathbb{K}[X]$*)

Soient A et B dans $\mathbb{K}[X]$, avec $B \neq 0$. Il existe un unique (Q, R) dans $(\mathbb{K}[X])^2$ tel que $\begin{cases} A = BQ + R, \\ \deg R < \deg B. \end{cases}$

Démonstration. Cette preuve est une adaptation de celle vue dans \mathbb{Z} :

- Commençons par l'existence. Si $\deg A < \deg B$, alors $(0, A)$ convient. Si $B|A, R = 0$ convient. Sinon, notons $n = \deg A, m = \deg B \leq n$. Considérons l'ensemble :

$$E = \{\deg(A - BQ) \mid Q \in \mathbb{K}[X]\}.$$

Cet ensemble est une partie de \mathbb{N} non-vide (car $n \in E$), donc admet un plus petit élément r , et il existe $Q_0 \in \mathbb{K}[X]$ tel que $\deg(R_0 := A - BQ_0) = r$.

Si $r < m$, (Q_0, R_0) convient. Sinon, notons b_m et c_r les coefficients dominants respectifs de B et R_0 , puis posons $Q_1 = Q_0 + \frac{c_r}{b_m} X^{r-m}$. On observe alors que $\deg(R_1 := A - BQ_1) < \deg R_0$, ce qui est absurde par définition de R_0 . Donc $r < m$.

- Montrons l'unicité. Supposons qu'il existe (Q_1, R_1) et (Q_2, R_2) dans $(\mathbb{K}[X])^2$ qui conviennent. Alors $A = BQ_1 + R_1 = BQ_2 + R_2$, donc $B(Q_1 - Q_2) = R_2 - R_1$. Or $\deg(R_2 - R_1) < \deg B$ et $\deg(B(Q_1 - Q_2)) = \deg B + \deg(Q_1 - Q_2)$, donc $\deg(Q_1 - Q_2) < 0$, donc $Q_1 - Q_2 = 0$, donc $R_2 - R_1 = 0$, donc $(Q_1, R_1) = (Q_2, R_2)$.

□

3 Déivation

Comme pour les autres opérations sur $\mathbb{K}[X]$, on définit la déivation sur $\mathbb{K}[X]$ par analogie avec celle que l'on connaît sur les fonctions polynomiales :

Définition 3.0.1. Soit $P = \sum_{k=0}^n a_k X^k$ dans $\mathbb{K}[X]$. Le *polynôme dérivé* de P , noté P' , est le polynôme :

$$P' = \sum_{k=1}^n k a_k X^{k-1} = \sum_{k=0}^{n-1} (k+1) a_{k+1} X^k.$$

Exemple 3.0.2. Le polynôme dérivé de $P = X^3 - 2X + 4$ est le polynôme $P' = 3X^2 - 2$.

On retrouve les propriétés et opérations déjà connues pour les fonctions polynomiales :

Proposition 3.0.3. Soient P et Q dans $\mathbb{K}[X]$. Alors :

1. Si $\deg P \geq 1$, alors $\deg P' = \deg P - 1$,
2. $P' = 0 \Leftrightarrow P \in \mathbb{K}_0[X]$,
3. $\forall \lambda \in \mathbb{K}, (\lambda P)' = \lambda P'$,
4. $(P + Q)' = P' + Q'$,
5. $(PQ)' = P'Q + PQ'$,
6. $(P \circ Q)' = Q' \times (P' \circ Q)$.

Définition 3.0.4 (Dérivées d'ordre supérieur). Soit P dans $\mathbb{K}[X]$. Pour tout k dans \mathbb{N} , le polynôme de P dérivé k fois, noté $P^{(k)}$, est défini par $P^{(0)} = P$ et : $\forall k \in \mathbb{N}^*, P^{(k)} = (P^{(k-1)})' = (P')^{(k-1)}$.

Proposition 3.0.5 (Formule de Leibniz). Soient P et Q dans $\mathbb{K}[X]$. Pour tout k dans \mathbb{N} :

$$(PQ)^{(k)} = \sum_{j=0}^k \binom{k}{j} P^{(j)} Q^{(k-j)}.$$

Proposition 3.0.6. Si P est de degré n , de coefficient dominant a_n , alors :

$$\forall k \leq n, P^{(k)}(0) = k! a_k, \quad P^{(n)} = n! a_n, \quad \text{et} \quad \forall k > n, P^{(k)} = 0.$$

Proposition 3.0.7 (Formule de Taylor). *Soit P dans $\mathbb{K}[X]$ de degré n , et soit x_0 dans \mathbb{K} . Alors :*

$$P = \sum_{k=0}^n \frac{P^{(k)}(x_0)}{k!} (X - x_0)^k.$$

Démonstration. Si $n = 0$, alors P est constant et on a bien $P = P(x_0)$.

Si $n \geq 1$, supposons la formule de Taylor vérifiée sur $\mathbb{K}_{n-1}[X]$. Elle est alors en particulier vérifiée pour P' :

$$P' = \sum_{k=0}^{n-1} \frac{(P')^{(k)}(x_0)}{k!} (X - x_0)^k = \sum_{k=0}^{n-1} \frac{P^{(k+1)}(x_0)}{k!} (X - x_0)^k.$$

Posons d'autre part $Q = \sum_{k=0}^n \frac{P^{(k)}(x_0)}{k!} (X - x_0)^k$. On a :

$$Q' = \sum_{k=1}^n \frac{P^{(k)}(x_0)}{(k-1)!} (X - x_0)^{k-1},$$

donc $Q' = P'$, donc $(Q - P)' = 0$, donc $Q - P$ est un polynôme constant. Or $(Q - P)(x_0) = 0$, donc $Q = P$. La formule est donc vérifiée sur $\mathbb{K}_n[X]$. Par récurrence, elle est donc vérifiée sur $\mathbb{K}[X]$. \square

4 Racines et décomposition

Après les considérations précédentes, il est temps d'observer les polynômes sous l'angle qui fait leur spécificité et leur richesse : l'étude de leurs racines.

4.1 Factorisation

Définition 4.1.1. Soient P dans $\mathbb{K}[X]$ et r dans \mathbb{K} . On dit que r est une *racine* de P si $P(r) = 0$.

Exemple 4.1.2.

- $P = X^2 + 1$ n'a pas de racine réelle, mais a deux racines complexes : i et $-i$.
- Si $P = aX + b$ est de degré 1 (c'est-à-dire que $a \neq 0$), alors $r = -\frac{b}{a}$ est l'unique racine de P dans \mathbb{K} .
- Soit $P \in \mathbb{R}[X]$. Si $z_0 \in \mathbb{C}$ est racine de P , alors $\overline{z_0}$ également.

Proposition 4.1.3 (Lemme de factorisation). *Soit P dans $\mathbb{K}[X]$ admettant une racine $r \in \mathbb{K}$. Alors $X - r \mid P$.*

Démonstration. Notons $P = (X - r)Q + R$ la division euclidienne de P par $X - r$. Comme $\deg R < \deg(X - r) = 1$, R est un polynôme constant. De plus, cette égalité évaluée en r s'écrit $0 = R(r)$, donc $R = 0$. Donc $X - r$ divise P . \square

Ce lemme se généralise :

Proposition 4.1.4. *Soit p dans \mathbb{N}^* et P dans $\mathbb{K}[X]$ admettant p racines distinctes $(r_1, \dots, r_p) \in \mathbb{K}^p$. Alors $\prod_{i=1}^p (X - r_i) \mid P$.*

Démonstration. Si $p = 1$, c'est le lemme précédent. Si $p > 1$, supposons cette propriété vérifiée pour $p - 1$ racines.

Alors $\prod_{i=1}^{p-1} (X - r_i) \mid P$; soit donc $Q \in \mathbb{K}[X]$ tel que $P = Q \prod_{i=1}^{p-1} (X - r_i)$. L'évaluation en r_p de cette égalité s'écrit $0 = Q(r_p) \prod_{i=1}^{p-1} (r_p - r_i)$. Or les r_i sont distincts, donc $Q(r_p) = 0$, et donc $X - r_p \mid Q$. Donc $\prod_{i=1}^p (X - r_i) \mid P$, et la propriété est vérifiée pour p racines. Donc par récurrence, elle est vérifiée pour tout p . \square

Une conséquence directe de ce résultat est la limite du nombre de racines :

Proposition 4.1.5. *Soit P dans $\mathbb{K}[X]$ de degré $n \in \mathbb{N}$ admettant p racines distinctes. Alors $p \leq n$.*

Corollaire 4.1.6. Si P dans $\mathbb{K}_n[X]$ admet $n + 1$ racines distinctes, alors $P = 0$.

Que se passe-t-il si les racines ne sont pas toutes distinctes ? On fait alors appel à la notion de *multiplicité* :

Définition 4.1.7. Soient P dans $\mathbb{K}[X]$, r dans \mathbb{K} et m dans \mathbb{N}^* . On dit que r est une *racine de multiplicité m* de P si : $\forall k < m, P^{(k)}(r) = 0$, et $P^{(m)}(r) \neq 0$.

Une racine de multiplicité 1 est dite *simple* ; une racine de multiplicité 2 est dite *double*, etc.

Le lemme de factorisation énoncé ci-dessus s'écrit dans ce contexte :

Proposition 4.1.8. *Soient P dans $\mathbb{K}[X]$, r dans \mathbb{K} et m dans \mathbb{N}^* . Alors r est une racine de multiplicité m de P si et seulement si $(X - r)^m \mid P$ et $(X - r)^{m+1} \nmid P$.*

Démonstration.

\Leftarrow : Supposons tout d'abord que $(X - r)^m \mid P$ et $(X - r)^{m+1} \nmid P$. Soit Q tel que $P = (X - r)^m Q$, r n'est alors pas racine de Q . D'après la formule de Leibniz :

$$\forall k \in \mathbb{N}, P^{(k)}(r) = \sum_{j=0}^k \binom{k}{j} ((X - r)^m)^{(j)}(r) Q^{(k-j)}(r).$$

Or, pour tout $j \in \mathbb{N}$, on a : $((X - r)^m)^{(j)} = \frac{m!}{(m - j)!} (X - r)^{m-j}$ si $j \leq m$, 0 sinon. Donc $((X - r)^m)^{(j)}(r) = m!$ si $j = m$, 0 sinon, et donc $P^{(k)}(r) = m! Q(r) = 0$ si $k < m$, $\neq 0$ si $k = m$. Donc r est une racine de multiplicité m de P .

\Rightarrow : Réciproquement, supposons que r est une racine de multiplicité m de P . Alors, d'après la formule de Taylor :

$$P = \sum_{k=0}^n \frac{P^{(k)}(r)}{k!} (X - r)^k = \sum_{k=m}^n \frac{P^{(k)}(r)}{k!} (X - r)^k = (X - r)^m Q,$$

où $Q = \sum_{k=m}^n \frac{P^{(k)}(r)}{k!} (X - r)^{k-m}$ vérifie $Q(r) = \frac{P^{(m)}(r)}{m!} \neq 0$. Donc $(X - r)^m \mid P$ et $(X - r)^{m+1} \nmid P$. \square

On arrive finalement au résultat de factorisation suivant, qui se démontre par récurrence comme son homologue précédent :

Proposition 4.1.9. *Soit p dans \mathbb{N}^* et P dans $\mathbb{K}[X]$ admettant p racines distinctes $(r_1, \dots, r_p) \in \mathbb{K}^p$ de multiplicités respectives $(m_1, \dots, m_p) \in (\mathbb{N}^*)^p$. Alors $\prod_{i=1}^p (X - r_i)^{m_i} \mid P$.*

On en déduit comme précédemment :

Proposition 4.1.10. *Soit P dans $\mathbb{K}[X]$ de degré $n \in \mathbb{N}$ admettant p racines distinctes de multiplicités respectives $(m_1, \dots, m_p) \in (\mathbb{N}^*)^p$. Alors $\sum_{i=1}^p m_i \leq n$.*

4.2 Décomposition en facteurs irréductibles

Il est légitime, après l'étude ci-dessus, de se demander si les racines d'un polynôme P donné décrivent entièrement P . On introduit la notion suivante :

Définition 4.2.1. Un polynôme $P \in \mathbb{K}[X]$ de degré $n \in \mathbb{N}$ est dit *scindé* si, avec les notations précédentes, $\sum_{i=1}^p m_i = n$. On a alors :

$$P = a_n \prod_{i=1}^p (X - r_i)^{m_i},$$

où $a_n \in \mathbb{K}^*$ est le coefficient dominant de P .

On dit que P est *scindé à racines simples* si toutes les racines sont simples (c'est-à-dire : $\forall i \in \llbracket 1, p \rrbracket$, $m_i = 1$).

Par convention, les polynômes constants sont scindés.

La question est alors de savoir quels polynômes sont scindés. Dans \mathbb{C} , une réponse radicale (si l'on peut dire) est apportée par l'énoncé suivant :

Théorème 4.2.2 (Théorème de d'Alembert-Gauss, ou *théorème fondamental de l'algèbre*). *Tout polynôme non constant de $\mathbb{C}[X]$ admet une racine dans \mathbb{C} .*

Corollaire 4.2.3. Tout polynôme de $\mathbb{C}[X]$ est scindé dans $\mathbb{C}[X]$.

Ces énoncés sont faux dans $\mathbb{R}[X]$: par exemple, pour tout n dans \mathbb{N} , le polynôme $X^{2n} + 1$ n'a pas de racine réelle. On peut néanmoins en chercher les *facteurs irréductibles* :

Définition 4.2.4. Un polynôme $P \in \mathbb{K}[X]$ est dit *irréductible* dans $\mathbb{K}[X]$ si :

$$\forall (A, B) \in (\mathbb{K}[X])^2, (P = AB) \Rightarrow (A \text{ ou } B \text{ est constant}).$$

Proposition 4.2.5.

1. *Les polynômes irréductibles dans $\mathbb{C}[X]$ sont les éléments de $\mathbb{C}_1[X]$.*
2. *Les polynômes irréductibles dans $\mathbb{R}[X]$ sont les éléments de $\mathbb{R}_1[X]$ et les polynômes de degré 2 de discriminant strictement négatif.*

On aboutit ainsi au résultat suivant, version polynomiale de la décomposition des entiers en facteurs premiers :

Théorème 4.2.6 (Décomposition en facteurs irréductibles).

1. *Dans $\mathbb{C}[X]$: Soit $P \in \mathbb{C}[X]$ non nul. Il existe un unique $\alpha \in \mathbb{C}^*$, un unique $p \in \mathbb{N}$, et $(z_1, \dots, z_p) \in \mathbb{C}^p$, $(m_1, \dots, m_p) \in (\mathbb{N}^*)^p$ uniques à l'ordre près, tels que*

$$P = \alpha \prod_{i=1}^p (X - z_i)^{m_i}.$$

2. *Dans $\mathbb{R}[X]$: Soit $P \in \mathbb{R}[X]$ non nul. Il existe un unique $\alpha \in \mathbb{R}^*$, un unique $(p, q) \in \mathbb{N}^2$, et $(r_1, \dots, r_p) \in \mathbb{R}^p$, $(m_1, \dots, m_p) \in (\mathbb{N}^*)^p$, $(b_1, \dots, b_q) \in \mathbb{R}^q$, $(c_1, \dots, c_q) \in \mathbb{R}^q$, $(n_1, \dots, n_q) \in (\mathbb{N}^*)^q$ uniques à l'ordre près, tels que*

$$P = \alpha \prod_{i=1}^p (X - r_i)^{m_i} \prod_{j=1}^q (X^2 + b_j X + c_j)^{n_j},$$

avec : $\forall j \in \llbracket 1, q \rrbracket$, $b_j^2 - 4c_j < 0$.

4.3 Relations coefficients-racines

L'aboutissement ci-dessus ne doit pas faire oublier que nous n'avons rien dit dans ce chapitre sur le *calcul* des racines ! Déterminer les racines d'un polynôme P à partir de ses coefficients et des opérations élémentaires, c'est-à-dire *résoudre par radicaux l'équation polynomiale* $P(x) = 0$ d'inconnue $x \in \mathbb{K}$, fut un problème ouvert pendant très longtemps : si la résolution des équations de degré 1 et 2 fut maîtrisée dès l'Antiquité, il fallut attendre le XVI^{ème} siècle et les travaux de Tartaglia, Cardan et Ferrari pour parvenir à résoudre les équations de degré 3 et 4, grâce à l'introduction des nombres complexes. L'impossibilité de résoudre par radicaux l'ensemble des équations de degré 5 fut démontrée, suite aux travaux de Lagrange et de Ruffini, par Abel en 1826, précédant de quelques années la caractérisation par Galois des équations résolubles par radicaux de degré quelconque, en 1832.

Il ne s'agit pas dans cette partie de présenter ces derniers résultats, dont la technicité dépasse notre programme. Nous nous contenterons de relever les relations coefficients-racines suivantes :

Proposition 4.3.1. Soit $P = \sum_{k=0}^n a_k X^k$ dans $\mathbb{C}[X]$ de degré $n \in \mathbb{N}$. Notons (r_1, \dots, r_n) ses racines complexes, comptées avec leur multiplicité. Alors :

$$1. \sigma_1 = \sum_{k=1}^n r_k = -\frac{a_{n-1}}{a_n},$$

$$2. \sigma_n = \prod_{k=1}^n r_k = (-1)^n \frac{a_0}{a_n}.$$

Exemple 4.3.2. Pour $P = X^n - 1$, on retrouve les formules déjà connues sur les racines de l'unité.

Chapitre 13

Limites et continuité

On revient dans ce chapitre sur la notion de *fonction continue*, à la lumière du formalisme introduit pour les limites des suites numériques.

Dans tout le chapitre, sauf indication contraire, on considère un intervalle I et une fonction $f : I \rightarrow \mathbb{R}$.

1 Limites et fonctions

1.1 Définitions

L'étude des suites convergentes, et la formalisation de leurs propriétés, peut être adaptée à l'étude des fonctions. De même qu'on a quantifié l'écriture « $u_n \xrightarrow[n \rightarrow +\infty]{} l$ », il s'agit de donner un sens à l'écriture « $f(x) \xrightarrow[x \rightarrow a]{} l$ ».

À la différence des suites, où la variable n tend toujours vers $+\infty$, la variable x d'une fonction peut tendre vers n'importe quel élément de I , et même de l'*adhérence* de I .

Définition 1.1.1. On appelle *adhérence* de l'intervalle I l'ensemble \bar{I} défini par :

1. si $I = \emptyset$, alors $\bar{I} = \emptyset$,
2. si $I =]a, b[$, $[a, b[$, $]a, b]$ ou $[a, b]$, alors $\bar{I} = [a, b]$,
3. si $I =]a, +\infty[$ ou $[a, +\infty[$, alors $\bar{I} = [a, +\infty] = [a, +\infty[\cup \{+\infty\}$,
4. si $I =]-\infty, b[$ ou $]-\infty, b]$, alors $\bar{I} = [-\infty, b] = \{-\infty\} \cup]-\infty, b]$,
5. si $I = \mathbb{R}$, alors $\bar{I} = [-\infty, +\infty] = \mathbb{R} \cup \{-\infty, +\infty\}$.

Soient alors a dans \bar{I} et l dans $\bar{\mathbb{R}}$. Comme pour les suites, dire que « $f(x)$ tend vers l quand x tend vers a » signifie que $f(x)$ est aussi proche que l'on veut de l à condition que x soit suffisamment proche de a .

Si a et l sont dans \mathbb{R} , cela veut dire que pour tout $\varepsilon > 0$, $f(x)$ est ε -proche de l à condition que x soit δ -proche de a pour un certain $\delta > 0$. Si a ou l est infini, on se ramène de même aux quantifications vues pour les suites qui tendent vers l'infini.

Cela donne donc :

Définition 1.1.2. Soient a dans \bar{I} et l dans $\bar{\mathbb{R}}$. On dit que $f(x)$ tend vers l quand x tend vers a , ou que f admet l pour *limite* en a , et on note $f(x) \xrightarrow[x \rightarrow a]{} l$, si :

1. si $a \in \mathbb{R}$ et $l \in \mathbb{R}$:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in I, (|x - a| \leq \delta) \Rightarrow (|f(x) - l| \leq \varepsilon).$$

2. si $a \in \mathbb{R}$ et $l = +\infty$ (resp. $-\infty$) :

$$\forall A \in \mathbb{R}, \exists \delta > 0, \forall x \in I, (|x - a| \leq \delta) \Rightarrow (f(x) \geq A \text{ (resp. } \leq A)).$$

3. si $a = +\infty$ (resp. $-\infty$) et $l \in \mathbb{R}$:

$$\forall \varepsilon > 0, \exists C \in \mathbb{R}, \forall x \in I, (x \geq C \text{ (resp. } \leq C)) \Rightarrow (|f(x) - l| \leq \varepsilon).$$

4. si $a = +\infty$ (resp. $-\infty$) et $l = +\infty$ (resp. $-\infty$) :

$$\forall \varepsilon > 0, \exists C \in \mathbb{R}, \forall x \in I, (x \geq C \text{ (resp. } \leq C)) \Rightarrow (f(x) \geq A \text{ (resp. } \leq A)) .$$

Plus synthétiquement, on peut rassembler ces différentes écritures en faisant appel à la notion de *voisinage* :

Définition 1.1.3. Soit $a \in \overline{\mathbb{R}}$. On appelle *voisinage* de a tout intervalle V contenant :

1. si $a \in \mathbb{R}$: $[a - \delta, a + \delta]$ pour un certain $\delta > 0$,
2. si $a = +\infty$: $[C, +\infty[$ pour un certain $C \in \mathbb{R}$,
3. si $a = -\infty$: $] - \infty, C]$ pour un certain $C \in \mathbb{R}$.

Les définitions ci-dessus peuvent alors être rassemblées sous la forme :

Définition 1.1.4. Soient a dans \overline{I} et l dans $\overline{\mathbb{R}}$. On dit que $f(x)$ tend vers l quand x tend vers a , et on note $f(x) \xrightarrow{x \rightarrow a} l$, si :

$$\forall W \text{ voisinage de } l, \exists V \text{ voisinage de } a, \forall x \in V, (x \in V) \Rightarrow (f(x) \in W) .$$

Constatons enfin que la définition de la limite a pour conséquence immédiate le résultat suivant :

Proposition 1.1.5. Si f est définie en $a \in \mathbb{R}$ et si $f(x) \xrightarrow{x \rightarrow a} l$, alors $l = f(a)$.

Remarque 1.1.6. La réciproque de cette dernière propriété est fausse : si f est définie en a , alors f n'admet pas nécessairement $f(a)$ pour limite en a . Cf par exemple $f = 1_{\mathbb{R}_+}$ en 0. Les fonctions qui satisfont cette réciproque sont qualifiées de *continues*, elles seront étudiées en détail dans la seconde partie de ce chapitre.

Ceci étant posé, il est légitime de se demander si les résultats observés dans le cadre des suites sont conservés pour les fonctions.

1.2 Ce qui ne change pas par rapport aux suites

Les propriétés suivantes, héritées des suites, restent vérifiées :

Proposition 1.2.1 (Unicité de la limite). Soient a dans \overline{I} et (l, l') dans $\overline{\mathbb{R}}^2$ tels que $f(x) \xrightarrow{x \rightarrow a} l$ et $f(x) \xrightarrow{x \rightarrow a} l'$. Alors $l = l'$.

Démonstration. On traite le cas où a et l sont réels. Rédiger les autres cas est un bon exercice d'entraînement !

Supposons $l \neq l'$. Posons $\varepsilon = \frac{|l - l'|}{3}$. Soient $\delta_1 > 0$ et $\delta_2 > 0$ tels que : $\forall x \in I, (|x - a| \leq \delta_1) \Rightarrow (|f(x) - l| \leq \varepsilon)$ et $(|x - a| \leq \delta_2) \Rightarrow (|f(x) - l'| \leq \varepsilon)$. Notons $\delta = \min(\delta_1, \delta_2)$, on a alors :

$$\forall x \in I, (|x - a| \leq \delta) \Rightarrow \left(|l - l'| \leq |l - f(x)| + |f(x) - l'| \leq 2\varepsilon = \frac{2}{3}|l - l'| \right) ,$$

ce qui est absurde. Donc $l = l'$. □

Théorème 1.2.2 (Théorème d'encadrement). Soient a dans \overline{I} , l dans \mathbb{R} et $g, h : I \rightarrow \mathbb{R}$ tels que : $\forall x \in I, g(x) \leq f(x) \leq h(x)$, $g(x) \xrightarrow{x \rightarrow a} l$ et $h(x) \xrightarrow{x \rightarrow a} l$. Alors $f(x) \xrightarrow{x \rightarrow a} l$.

Démonstration. On traite le cas où a est réel. Rédiger les autres cas est un bon exercice d'entraînement !

Soit $\varepsilon > 0$. Soient $\delta_1 > 0$ et $\delta_2 > 0$ tels que : $\forall x \in I, (|x - a| \leq \delta_1) \Rightarrow (|g(x) - l| \leq \varepsilon)$ et $(|x - a| \leq \delta_2) \Rightarrow (|h(x) - l| \leq \varepsilon)$. Notons $\delta = \min(\delta_1, \delta_2)$, on a alors :

$$\forall x \in I, (|x - a| \leq \delta) \Rightarrow (l - \varepsilon \leq g(x) \leq f(x) \leq h(x) \leq l + \varepsilon) ,$$

d'où le résultat. □

On conserve de même les théorèmes de minoration et de majoration :

Théorème 1.2.3 (Théorèmes de minoration et de majoration). *Soit a dans \bar{I} .*

1. *Soit $g : I \rightarrow \mathbb{R}$ telle que : $\forall x \in I, g(x) \leq f(x)$ et $g(x) \xrightarrow{x \rightarrow a} +\infty$. Alors $f(x) \xrightarrow{x \rightarrow a} +\infty$.*
2. *Soit $h : I \rightarrow \mathbb{R}$ telle que : $\forall x \in I, f(x) \leq h(x)$ et $h(x) \xrightarrow{x \rightarrow a} -\infty$. Alors $f(x) \xrightarrow{x \rightarrow a} -\infty$.*

Et les relations entre limites et égalités sont conservées :

Proposition 1.2.4. *Soient $f, g : I \rightarrow \mathbb{R}$ telles que $f(x) \xrightarrow{x \rightarrow a} l$ et $g(x) \xrightarrow{x \rightarrow a} l'$. Alors :*

1. *Si $\forall x \in I, f(x) \leq g(x)$, alors $l \leq l'$,*
2. *Si $\forall x \in I, f(x) < g(x)$, alors $l \leq l'$.*

Les opérations sur les limites restent valables, aux formes indéterminées près :

Proposition 1.2.5. *Soient a dans \bar{I} et (l, l') dans $\bar{\mathbb{R}}^2$. Soient $f, g : I \rightarrow \mathbb{R}$ telles que $f(x) \xrightarrow{x \rightarrow a} l$ et $g(x) \xrightarrow{x \rightarrow a} l'$. Alors :*

1. $f(x) + g(x) \xrightarrow{x \rightarrow a} l + l'$,
2. $\forall \lambda \in \mathbb{R}, \lambda f(x) \xrightarrow{x \rightarrow a} \lambda l$,
3. $f(x)g(x) \xrightarrow{x \rightarrow a} ll'$,
4. *Si $l' \neq 0$, alors il existe un voisinage de a sur lequel g ne s'annule pas, et $\frac{f(x)}{g(x)} \xrightarrow{x \rightarrow a} \frac{l}{l'}$.*

Pour lever certaines formes indéterminées, on renvoie à l'étude des *croissances comparées* vues dans le chapitre consacré aux fonctions usuelles.

Enfin, les *relations de comparaison* peuvent être utilisées dans le cadre des fonctions, avec les mêmes règles que pour les suites :

Définition 1.2.6. Soient $f, g : I \rightarrow \mathbb{R}$ et a dans \bar{I} , tels que g **ne s'annule pas** au voisinage de a . On dit que :

1. *f est dominée par g en a , et on note $f(x) = \underset{x \rightarrow a}{O}(g(x))$, si il existe un voisinage de a sur lequel $\frac{f}{g}$ est bornée,*
2. *f est négligeable devant g en a , et on note $f(x) = \underset{x \rightarrow a}{o}(g(x))$, si $\frac{f(x)}{g(x)} \xrightarrow{x \rightarrow a} 0$,*
3. *f est équivalente à g en a , et on note $f(x) \underset{x \rightarrow a}{\sim} g(x)$, si $\frac{f(x)}{g(x)} \xrightarrow{x \rightarrow a} 1$.*

Exemple 1.2.7 (Équivalents usuels).

- $\sin x \underset{x \rightarrow 0}{\sim} x$,
- $\ln(1 + x) \underset{x \rightarrow 0}{\sim} x$,
- $e^x - 1 \underset{x \rightarrow 0}{\sim} x$,
- $(1 + x)^\alpha - 1 \underset{x \rightarrow 0}{\sim} \alpha x$,
- $1 - \cos x \underset{x \rightarrow 0}{\sim} \frac{x^2}{2}$,
- $\tan x \underset{x \rightarrow 0}{\sim} x$.

Soulignons enfin la *compatibilité* entre les notions de limite d'une suite et d'une fonction :

Proposition 1.2.8 (Caractérisation séquentielle de la limite d'une fonction). *f a pour limite $l \in \bar{\mathbb{R}}$ en $a \in \bar{I}$ si et seulement si, pour toute suite (u_n) de limite a , $(f(u_n))$ a pour limite l .*

1.3 Ce qui change (un peu) par rapport aux suites

Certains résultats obtenus pour les suites doivent être adaptés aux fonctions ; ceci est essentiellement dû aux différences de structure entre \mathbb{N} et \mathbb{R} . Ainsi, alors que « toute suite convergente est bornée », on a maintenant :

Proposition 1.3.1. *Soient a dans \bar{I} et l dans \mathbb{R} . Si $f(x) \xrightarrow[x \rightarrow a]{} l$, alors il existe un voisinage de a sur lequel f est bornée.*

Le théorème de la limite monotone s'écrit quant à lui :

Théorème 1.3.2 (Théorème de la limite monotone). *Soit (a, b) dans $\bar{\mathbb{R}}^2$ avec $a < b$. Si $f :]a, b[\rightarrow \mathbb{R}$ est croissante et majorée (ou décroissante et minorée), alors f admet une limite finie à gauche en b .*

À première vue, c'est le même résultat que pour les suites. Pourtant, la notion de limite est ici contrainte : b est au bord de I , et dans l'écriture « $f(x) \xrightarrow[x \rightarrow b]{} l$ », seuls les $x < b$ sont considérés. Si $b \neq \pm\infty$, on parle alors de *limite à droite* (resp. *à gauche*) :

Définition 1.3.3. Soient a dans $\bar{I} \cap \mathbb{R}$ et l dans $\bar{\mathbb{R}}$.

1. On dit que f admet l pour *limite à gauche* en a , et on note $f(x) \xrightarrow[x \rightarrow a^-]{} l$ ou $f(x) \xrightarrow[x \rightarrow a, x < a]{} l$, si :

$$\forall W \text{ voisinage de } l, \exists \delta > 0, \forall x \in I, (a - \delta \leq x < a) \Rightarrow (f(x) \in W).$$

2. On dit que f admet l pour *limite à droite* en a , et on note $f(x) \xrightarrow[x \rightarrow a^+]{} l$ ou $f(x) \xrightarrow[x \rightarrow a, x > a]{} l$, si :

$$\forall W \text{ voisinage de } l, \exists \delta > 0, \forall x \in I, (a < x \leq a + \delta) \Rightarrow (f(x) \in W).$$

3. On dit que f admet l pour *limite épointée* en a , et on note $f(x) \xrightarrow[x \rightarrow a, x \neq a]{} l$, si :

$$\forall W \text{ voisinage de } l, \exists \delta > 0, \forall x \in I, (|x - a| \leq \delta \text{ et } x \neq a) \Rightarrow (f(x) \in W).$$

Proposition 1.3.4. *Soient a dans $\bar{I} \cap \mathbb{R}$ et l dans $\bar{\mathbb{R}}$. f admet l pour limite épointée en a si et seulement si f admet l pour limite à droite et à gauche en a .*

2 Continuité en un point

Définition 2.0.1. Soit a dans $I \cap \mathbb{R}$. On dit que f est *continue* en a si f admet une limite épointée en a , et si cette limite est égale à $f(a)$.

De même que pour la limite, on peut parler de continuité à gauche ou à droite :

Définition 2.0.2. Soit a dans $I \cap \mathbb{R}$. On dit que f est *continue à gauche*, resp. *à droite*, en a si $f(x) \xrightarrow[x \rightarrow a^-]{} f(a)$, resp. $f(x) \xrightarrow[x \rightarrow a^+]{} f(a)$.

Proposition 2.0.3. *Soit a dans $I \cap \mathbb{R}$. f est continue en a si et seulement si f est continue à droite et à gauche en a .*

Inversement, dans le cas où a est un réel au bord de I et où f n'est pas définie en a , on peut alors parfois prolonger f par continuité en a :

Définition 2.0.4. Soit a dans $(\bar{I} \setminus I) \cap \mathbb{R}$. On dit que f est *prolongeable par continuité à gauche* (resp. *à droite*) en a si f admet une limite finie l à gauche (resp. à droite) en a . On appelle alors *prolongement par continuité de f en a* la fonction $\tilde{f} : I \cup \{a\} \rightarrow \mathbb{R}$ définie par :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in I \\ l & \text{si } x = a \end{cases}.$$

Exemples 2.0.5. Les fonctions $x \mapsto x^x$ et $x \mapsto (1+x)^{\frac{1}{x}}$ sont prolongeables par continuité en 0.

Notons la conservation de la continuité par opérations élémentaires :

Proposition 2.0.6. Soient $f, g : I \rightarrow \mathbb{R}$ et $a \in I$. Alors :

1. si f et g sont continues en a , lorsqu'elles sont définies, $f + g$, λf , fg , $\frac{f}{g}$ sont continues en a ,
2. si $g(I) \subset I$, si g est continue en a et si f est continue en $g(a)$, alors $f \circ g$ est continue en a .

Enfin, la compatibilité déjà mentionnée avec le langage des suites nous donne :

Proposition 2.0.7 (Caractérisation séquentielle de la continuité). f est continue en $a \in I$ si et seulement si, pour toute suite (u_n) convergente vers a , $(f(u_n))$ est convergente vers $f(a)$.

Remarque 2.0.8. Cette propriété est surtout utile dans sa version négative : pour montrer qu'une fonction n'est pas continue en a , il suffit de trouver une suite convergente vers a dont l'image par f ne converge pas vers $f(a)$. Pour montrer qu'une fonction n'est pas prolongeable par continuité en a , il suffit de trouver deux suites convergentes vers a dont les images par f convergent vers des limites différentes (ou divergent).

Exemple 2.0.9. $f : x \mapsto \cos\left(\frac{1}{x}\right)$ n'est pas prolongeable par continuité en 0.

3 Continuité sur un intervalle

Définition 3.0.1. Une fonction $f : I \rightarrow \mathbb{R}$ est dite *continue* si, pour tout a dans I , f est continue en a .

Pour un intervalle $J \subset I$, f est dite *continue sur J* si la restriction $f|_J$ est continue.

On note $C(I, \mathbb{R})$ l'ensemble des fonctions $f : I \rightarrow \mathbb{R}$ continues.

Cette définition est l'aboutissement du formalisme développé depuis la notion de suite convergente : elle permet de donner un sens précis à la *continuité de f* , auparavant décrite, et intuitivement comprise, comme le fait de pouvoir tracer le graphe de f sans lever le crayon.

La définition entraîne directement la conservation de la continuité, déjà observée pour la continuité ponctuelle :

Proposition 3.0.2. Soient $f, g : I \rightarrow \mathbb{R}$ continues. Alors :

1. lorsqu'elles sont définies, $f + g$, λf , fg , $\frac{f}{g}$ sont continues,
2. si $g(I) \subset I$, si g est continue et si f est continue, alors $f \circ g$ est continue.

Une propriété fondamentale des fonctions continues est le *théorème des valeurs intermédiaires* :

Théorème 3.0.3 (Théorème des valeurs intermédiaires). Soit $f : [a, b] \rightarrow \mathbb{R}$ continue. Alors :

$$\forall y \in [f(a), f(b)], \exists c \in [a, b], f(c) = y.$$

Démonstration. On traite le cas où $f(a) \leq f(b)$ (la démonstration dans le cas contraire est similaire). On pose $a_0 = a$ et $b_0 = b$. Soit y dans $[f(a), f(b)]$, on a donc : $f(a_0) \leq y \leq f(b_0)$. Soit $m_0 = \frac{a_0 + b_0}{2}$:

- Si $f(m_0) \leq y$, alors on pose $a_1 = m_0$ et $b_1 = b_0$,
- Si $f(m_0) > y$, alors on pose $a_1 = a_0$ et $b_1 = m_0$.

Dans tous les cas, on a donc encore : $f(a_1) \leq y \leq f(b_1)$. On définit de la même manière a_2 et b_2 , puis, par récurrence, deux suites (a_n) et (b_n) telles que :

$$\forall n \in \mathbb{N}, \quad f(a_n) \leq y \leq f(b_n) \quad \text{et} \quad b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}.$$

Les suites (a_n) et (b_n) sont respectivement croissante et décroissante, et : $\forall n \in \mathbb{N}, b_n - a_n = \frac{b_0 - a_0}{2^n} \xrightarrow[n \rightarrow +\infty]{} 0$. Elles sont donc adjacentes, et convergent donc vers un même réel c . En particulier, $a_0 \leq c \leq b_0$, donc $c \in [a, b]$; et, comme f est continue, $f(a_n) \xrightarrow[n \rightarrow +\infty]{} f(c)$ et $f(b_n) \xrightarrow[n \rightarrow +\infty]{} f(c)$, donc, par passage à la limite : $f(c) \leq y \leq f(c)$, c'est-à-dire que $f(c) = y$. □

Exemples 3.0.4.

- Si $f : [a, b] \rightarrow \mathbb{R}$ continue vérifie $f(a)f(b) \leq 0$, alors f s'annule sur $[a, b]$,
- Toute fonction polynomiale de degré impair s'annule sur \mathbb{R} .

Une autre manière d'écrire ce théorème met en évidence la compatibilité des fonctions continues avec la structure de \mathbb{R} :

Corollaire 3.0.5. Si $f : I \rightarrow \mathbb{R}$ est continue, alors $f(I)$ est un intervalle.

Démonstration. Soient $r \leq s$ dans $f(I)$, puis $y \in [r, s]$. Soient a, b dans I tels que $r = f(a)$ et $s = f(b)$, alors d'après le TVI, il existe c dans $[a, b]$ tel que $y = f(c)$, donc $y \in f(I)$. Donc $f(I)$ est un intervalle. □

Remarque 3.0.6. Par conséquent, si $f : [a, b] \rightarrow \mathbb{R}$ est continue et croissante (resp. décroissante), alors $f([a, b]) = [f(a), f(b)]$ (resp. $[f(b), f(a)]$).

Ce n'est pas aussi simple si f n'est ni croissante, ni décroissante (cf. par exemple \sin sur $[0, \pi]$). On peut néanmoins utiliser le résultat suivant :

Théorème 3.0.7 (Théorème de Weierstrass (ou des bornes atteintes)). *Soit $f : [a, b] \rightarrow \mathbb{R}$ continue. Alors f est bornée et atteint ses bornes :*

$$\exists(m, M) \in \mathbb{R}^2, \forall x \in [a, b], m \leq f(x) \leq M,$$

et

$$\exists(c, d) \in [a, b]^2, f(c) = m \text{ et } f(d) = M.$$

Autrement dit : $f([a, b]) = [m, M]$.

Enfin, attention au fait que le TVI n'assure pas l'*unicité* de l'antécédent c de y par f . On a besoin pour cela du contexte suivant :

Théorème 3.0.8 (Théorème de la bijection monotone continue). *Soit $f : I \rightarrow \mathbb{R}$ continue et strictement monotone. Alors f réalise une bijection de I dans l'intervalle $J = f(I)$, et sa bijection réciproque f^{-1} est continue et strictement monotone, de même sens que f .*

Démonstration. On traite le cas où f est strictement croissante (la démonstration dans le cas où f est strictement décroissante est similaire).

- D'après le théorème des valeurs intermédiaires, $J = f(I)$ est un intervalle et, par construction, $f : I \rightarrow J$ est surjective. Montrons que f est injective : soient x_1 et x_2 dans I tels que $f(x_1) = f(x_2)$. Si $x_1 < x_2$, alors, comme f est strictement croissante, $f(x_1) < f(x_2)$, ce qui est exclu ; de même si $x_1 > x_2$, donc $x_1 = x_2$. Donc f est injective, donc bijective.
- Soient $y_1 < y_2$ dans J . Soient $x_1 = f^{-1}(y_1)$ et $x_2 = f^{-1}(y_2)$. Alors $f(x_1) < f(x_2)$ donc, comme f est strictement croissante, $x_1 < x_2$, c'est-à-dire $f^{-1}(y_1) < f^{-1}(y_2)$. Donc f^{-1} est strictement croissante.
- Il reste à montrer que $f^{-1} : J \rightarrow I$ est continue. Soit b dans J , montrons que f^{-1} est continue en b : Soit $\varepsilon > 0$. On cherche $\delta > 0$ tel que : $\forall y \in J, |y - b| \leq \delta \Rightarrow |f^{-1}(y) - f^{-1}(b)| \leq \varepsilon$. Notons $a = f^{-1}(b)$. Comme f est croissante, on a :

$$|f^{-1}(y) - a| \leq \varepsilon \Leftrightarrow (a - \varepsilon \leq f^{-1}(y) \leq a + \varepsilon) \Leftrightarrow (f(a - \varepsilon) \leq y \leq f(a + \varepsilon)) \Leftrightarrow y \in [f(a - \varepsilon), f(a + \varepsilon)].$$

Or, comme f est strictement croissante, $b \in]f(a - \varepsilon), f(a + \varepsilon)[$, donc il existe $\delta > 0$ tel que $[b - \delta, b + \delta] \subset [f(a - \varepsilon), f(a + \varepsilon)]$. Plus précisément $\delta = \min(f(a + \varepsilon) - b, b - f(a - \varepsilon)) > 0$ convient :

$$\forall y \in J, |y - b| \leq \delta \Rightarrow (f(a - \varepsilon) \leq y \leq f(a + \varepsilon)) \Rightarrow |f^{-1}(y) - a| \leq \varepsilon.$$

Donc f^{-1} est continue en b , donc f^{-1} est continue.

□

4 Fonctions complexes

Qu'en est-il des fonctions à valeurs complexes, c'est-à-dire des fonctions $f : I \rightarrow \mathbb{C}$ (où $I \subset \mathbb{R}$ est encore un intervalle) ?

Comme pour les suites, les définitions et résultats traités ci-dessus restent valables, *dès lors que leur énoncé ou leur démonstration ne fait pas intervenir de relation d'ordre sur les valeurs de f* . Ainsi :

- Restent valables : la notion de limite *finie* $l \in \mathbb{C}$ de f en $a \in \bar{I}$, l'unicité de cette limite, les relations de comparaison, les notions de continuité et de prolongement par continuité, leur conservation par opérations élémentaires, le caractère borné d'une fonction continue sur un segment (c'est-à-dire la « moitié » du théorème de Weierstrass),
- Ne sont plus valables : la notion de limite infinie $l = \pm\infty$, les théorèmes d'encadrement, de minoration et de majoration, le théorème de la limite monotone, le théorème des valeurs intermédiaires, le théorème de la bijection monotone continue.

Exemple 4.0.1. La fonction $f : t \mapsto e^{it}$ n'obéit pas au TVI : elle est continue sur \mathbb{R} et vérifie $f(0) = 1$ et $f(\pi) = -1$, mais ne s'annule pas sur $[0, \pi]$.

Chapitre 14

Dérivabilité

On revient dans ce chapitre sur la notion de *fonction dérivable*, à la lumière du formalisme introduit pour les suites numériques et pour les fonctions continues.

Dans tout le chapitre, sauf indication contraire, on considère un intervalle I et une fonction $f : I \rightarrow \mathbb{R}$.

1 Nombre dérivé, fonction dérivée

1.1 Dérivabilité

L'étude des fonctions continues effectuée précédemment s'étend naturellement à l'étude de la dérivabilité :

Définition 1.1.1. Soit a dans I . La fonction *taux d'accroissement* de f en a , notée $\tau_{a,f}$, est la fonction :

$$\tau_{a,f} : \begin{cases} I \setminus \{a\} & \rightarrow \mathbb{R} \\ x & \mapsto \frac{f(x) - f(a)}{x - a} \end{cases} .$$

Définition 1.1.2. On dit que la fonction f est *dérivable* en $a \in I$ si la fonction $\tau_{a,f}$ est prolongeable par continuité en a ; c'est-à-dire s'il existe $l \in \mathbb{R}$ tel que $\tau_{a,f}(x) \xrightarrow[x \rightarrow a]{} l$. On dit que l est le *nombre dérivé* de f en a , noté $f'(a)$.

Exemple 1.1.3. $f : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto \sin(x) \end{cases}$, $g : \begin{cases} \mathbb{R}_+ & \rightarrow \mathbb{R} \\ x & \mapsto \sqrt{x} \end{cases}$.

Définition 1.1.4. On dit que la fonction f est *dérivable à gauche* (resp. à droite) en $a \in I$ si la fonction $\tau_{a,f}$ est prolongeable par continuité à gauche (resp. à droite) en a . On note en général $f'_g(a)$ (resp. $f'_d(a)$) le nombre dérivé à gauche (resp. à droite) de f en a .

Exemple 1.1.5. $f : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto |x| \end{cases}$.

Définition 1.1.6. On dit que f admet un *développement limité à l'ordre 1* en $a \in I$, et on note « f est DL_1 en a » s'il existe $(\alpha_0, \alpha_1) \in \mathbb{R}^2$ tel que :

$$f(x) = \alpha_0 + \alpha_1(x - a) + o_{x \rightarrow a}(x - a).$$

L'égalité suivante est une simple réécriture de la définition du taux d'accroissement :

Proposition 1.1.7. $\forall x \in I \setminus \{a\}$, $f(x) = f(a) + (x - a)\tau_{a,f}(x)$.

Elle a plusieurs conséquences :

Proposition 1.1.8.

1. Si f est dérivable en a , alors f est continue en a .
2. Si f est dérivable en a , alors la courbe de f admet une tangente en a , d'équation cartésienne $y = f(a) + f'(a)(x - a)$.
3. La fonction f est drivable en a si et seulement si f est DL_1 en a .
- 4.

Démonstration.

1. $\forall x \in I \setminus \{a\}$, $f(x) = f(a) + (x - a)\tau_{a,f}(x)$, donc si f est dérivable en a , alors $f(x) \xrightarrow{x \rightarrow a} f(a)$.
2. Immédiat
3. • Supposons f DL_1 en a . Alors $f(x) \xrightarrow{x \rightarrow a} \alpha_0 = f(a)$, donc f est continue en a . Par conséquent, $\tau_{a,f}(x) = \alpha_1 + \underset{x \rightarrow a}{o}(1) \xrightarrow{x \rightarrow a} \alpha_1$, donc f est dérivable en a et $f'(a) = \alpha_1$.
• Réciproquement, supposons f dérivable en a . Alors $\tau_{a,f}(x) = f'(a) + \underset{x \rightarrow a}{o}(1)$, donc $f(x) = f(a) + f'(a)(x - a) + \underset{x \rightarrow a}{o}(x - a)$. Donc f est DL_1 en a .

□

Définition 1.1.9. On dit que f est *dérivable* (ou *dérivable sur I*), si f est dérivable en a pour tout $a \in I$. La fonction dérivée de f , notée f' , est alors la fonction $f' : \begin{cases} I & \rightarrow \mathbb{R} \\ x & \mapsto f'(x) \end{cases}$.

1.2 Opérations sur les dérivées

On rappelle les opérations suivantes :

Proposition 1.2.1. Soient $f, g : I \rightarrow \mathbb{R}$ dérivables en $a \in I$ et $\lambda \in \mathbb{R}$. Alors :

1. $f + g$ est dérivable en a et $(f + g)'(a) = f'(a) + g'(a)$,
2. λf est dérivable en a et $(\lambda f)'(a) = \lambda f'(a)$,
3. fg est dérivable en a et $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$ (Formule de Leibniz),
4. si g ne s'annule pas, alors $\frac{1}{g}$ est dérivable en a et $\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g(a)^2}$,
5. si $g(I) \subset I$ et si f est dérivable en $g(a)$, alors $f \circ g$ est dérivable en a et $(f \circ g)'(a) = g'(a)f'(g(a))$,
6. si $f : I \rightarrow J$ est bijective, si f est dérivable en $f^{-1}(a)$ et si $f'(f^{-1}(a)) \neq 0$, alors f^{-1} est dérivable en a et $(f^{-1})'(a) = \frac{1}{f'(f^{-1}(a))}$.

2 Propriétés des fonctions dérivables

Commençons par revenir sur les propriétés connues des fonctions dérivables. Nous allons en tirer de nouveaux énoncés, issus des théorèmes vus plus haut sur les fonctions continues.

2.1 Dérivées et extrema

Définition 2.1.1. Soit a dans I . On dit que f admet un *maximum* (resp. *minimum*) local en a si :

$$\exists \delta > 0, \forall x \in I, |x - a| \leq \delta \Rightarrow f(x) \leq f(a) \text{ (resp. } f(x) \geq f(a)).$$

On dit que f admet un *extremum local* en a si f admet un maximum local ou un minimum local en a .

Remarque 2.1.2. Attention : à la différence d'une *borne supérieure*, un *maximum* est par définition une valeur atteinte par la fonction.

Définition 2.1.3. On dit que $a \in I$ est à l'intérieur de I , et on note $a \in \overset{\circ}{I}$, s'il existe $\delta > 0$ tel que $]a - \delta, a + \delta[\subset I$.

Proposition 2.1.4. Si f est dérivable en $a \in \overset{\circ}{I}$, et si f admet un extremum local en a , alors $f'(a) = 0$.

Démonstration. Cas d'un maximum : $\forall x < a, \tau_{a,f}(x) \geq 0$ et $\forall x > a, \tau_{a,f}(x) \leq 0$, donc $f'(a) = \lim_{x \rightarrow a} \tau_{a,f}(x) = 0$. \square

Remarque 2.1.5. La réciproque est fausse : l'annulation de f' ne signale pas toujours un extremum ! Par exemple, la fonction $f : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto x^3 \end{cases}$ n'a pas extremum en 0.

Un autre lien entre dérivation et extrema est le théorème suivant, fondamental, dont les conséquences vont occuper la plus grande partie de ce chapitre :

Théorème 2.1.6 (Théorème de Rolle). Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$ telle que $f(a) = f(b)$. Il existe alors $c \in]a, b[$ tel que $f'(c) = 0$.

Démonstration. Comme f est continue sur $[a, b]$, d'après le théorème de Weierstrass, f est bornée sur $[a, b]$ et atteint ses bornes. Il existe donc un minimum m et un maximum M de f sur $[a, b]$, et $(x_0, x_1) \in [a, b]^2$ tel que $f(x_0) = m$ et $f(x_1) = M$. Par conséquent :

- si $m = M$, alors f est constante sur $[a, b]$, donc : $\forall c \in]a, b[, f'(c) = 0$,
- si $m \neq M$, alors, comme $f(a) = f(b)$, l'un de ces deux extrema est atteint à l'intérieur de $[a, b] : x_0 \in]a, b[$ ou $x_1 \in]a, b[$. Si c'est x_0 , alors d'après la propriété précédente, $f'(x_0) = 0$, donc $c = x_0$ convient. De même si c'est x_1 .

\square

Exemple 2.1.7. Si un polynôme P est scindé sur \mathbb{R} , alors P' aussi.

Ce résultat se généralise au non moins fondamental théorème suivant :

Théorème 2.1.8 (Théorème des accroissements finis). Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$. Il existe alors $c \in]a, b[$ tel que $f'(c) = \frac{f(b) - f(a)}{b - a}$.

Démonstration. Notons $\alpha = \frac{f(b) - f(a)}{b - a}$. On applique le théorème de Rolle à la fonction $g : x \mapsto f(x) - \alpha(x - a)$. \square

Application 2.1.9. Une première utilisation de cette propriété est la détermination d'encadrements numériques : $\sqrt{1,02}$, $\ln(0,97)$...

2.2 Conséquences du théorème des accroissements finis

Le théorème des accroissements finis donne accès à un nombre considérable de résultats sur les fonctions dérivables, à commencer par le lien entre signe de la dérivée et sens de variation de la fonction :

Dérivées et monotonie

Proposition 2.2.1. Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$. La fonction f est croissante (resp. décroissante) si et seulement si sa dérivée f' est positive (resp. négative).

Démonstration. On traite le cas où f est croissante (le cas f décroissante est similaire).

- Supposons f croissante sur $[a, b]$. Soit $c \in [a, b]$, alors : $\forall x \neq c, \tau_{c,f}(x) \geq 0$, donc par passage à la limite : $f'(c) \geq 0$.

- Réciproquement, supposons f' positive sur $[a, b]$. Soient $x < y$ dans $[a, b]$. D'après le théorème des accroissements finis, il existe $c \in]x, y[$ tel que $f(x) - f(y) = f'(c)(x - y)$, donc $f(x) - f(y) \leq 0$, donc $f(x) \leq f(y)$. Donc f est croissante sur $[a, b]$.

□

De façon similaire, on a :

Proposition 2.2.2. Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$. Si f' est strictement positive (resp. strictement négative), alors f est strictement croissante (resp. strictement décroissante).

Remarque 2.2.3. Attention, la réciproque est cette fois fausse ! La fonction $x \mapsto x^3$ est ainsi strictement croissante sur \mathbb{R} , mais sa dérivée n'est pas strictement positive sur \mathbb{R} . On a néanmoins :

Proposition 2.2.4. Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$. La fonction f est strictement croissante (resp. décroissante) si et seulement si sa dérivée f' est positive (resp. négative) et s'il n'existe pas d'intervalle non réduit à un point sur lequel f' est nulle.

Inégalité des accroissements finis

Une conséquence directe du théorème des accroissements finis est la suivante :

Proposition 2.2.5 (Inégalité des accroissements finis). Soit $f : I \rightarrow \mathbb{R}$ dérivable. Si f' est bornée par C , alors :

$$\forall (x, y) \in I^2, |f(x) - f(y)| \leq C|x - y|.$$

Exemple 2.2.6. Cette formule permet par exemple de majorer l'erreur dans l'approximation numérique de $\ln(1 - \varepsilon)$ pour $\varepsilon > 0$...

Cette inégalité appelle à considérer les fonctions suivantes :

Définition 2.2.7. Soit $f : I \rightarrow \mathbb{R}$, et soit $k \in \mathbb{R}_+$. La fonction f est dite k -lipschitzienne si :

$$\forall (x, y) \in I^2, |f(x) - f(y)| \leq k|x - y|.$$

Exemple 2.2.8. Les fonctions \sin et $x \mapsto |x|$ sont 1-lipschitziennes. La fonction $x \mapsto x^2$ ne l'est pas.

Proposition 2.2.9. Toute fonction lipschitzienne est continue.

Démonstration. Soit a dans I . Alors : $\forall x \in I, |f(x) - f(a)| \leq k|x - a| \xrightarrow{x \rightarrow a} 0$, donc $f(x) \xrightarrow{x \rightarrow a} f(a)$. Donc f est continue en a . □

Remarques 2.2.10.

- L'exemple de la fonction valeur absolue montre qu'une fonction lipschitzienne n'est pas forcément dérivable.
- Graphiquement, les courbes des fonctions lipschitziennes sont celles dont les cordes sont de pente bornée.

Si f est k -lipschitzienne avec $k < 1$, on dit que f est *contractante*.

Théorème de la limite de la dérivée

Comme pour les fonctions continues, la question de la dérivalibilité du *prolongement* d'une fonction se pose. Par définition, il faut vérifier si le taux d'accroissement admet une limite au point considéré.

Exemple 2.2.11. La fonction $x \mapsto \frac{\sin x}{x}$ est prolongeable par continuité en 0, de prolongement dérivable en 0. Ce résultat est une conséquence de la célèbre :

Lemme 2.2.12 (Règle de L'Hôpital). Si f et g sont dérivables en a , avec $f(a) = g(a) = 0$, alors : $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$.

Le théorème suivant nous simplifie la tâche :

Théorème 2.2.13 (Théorème de limite de la dérivée). Soit $a \in I$. Soit $f : I \rightarrow \mathbb{R}$ continue sur I et dérivable sur $I \setminus \{a\}$. S'il existe $l \in \mathbb{R}$ tel que $f'(x) \xrightarrow{x \rightarrow a} l$, alors $\tau_{a,f}(x) \xrightarrow{x \rightarrow a} l$.

Dans ce cas, la fonction f est donc dérivable en a , et sa dérivée f' est continue en a .

Démonstration. Soit $x \in I \setminus \{a\}$. D'après le théorème des accroissements finis, il existe $c \in]a, x[$ (ou $]x, a[$) tel que $\tau_{a,f}(x) = f'(c)$. Or $c \xrightarrow{x \rightarrow a} a$, donc $f'(c) \xrightarrow{x \rightarrow a} l$, d'où le résultat. \square

Remarques 2.2.14.

- Attention, la réciproque est fausse : par exemple, la fonction $f : x \mapsto \begin{cases} x^2 \sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$ est dérivable en 0, mais $f'(x) \xrightarrow{x \rightarrow 0} f'(0)$.
- D'après ce théorème, il suffit que f' admette la même limite à gauche et à droite en a pour que f soit dérivable en a !
- Par contre, si f' a des limites distinctes à gauche et à droite en a , alors $\tau_{a,f}$ également ; ce qui suffit à conclure à la non-dérivabilité de f en a . C'est le cas pour $f : x \mapsto e^x$ si $x < 0$, x^2 si $x \geq 0$.

Le théorème de limite de la dérivée peut se lire de la façon suivante : lorsqu'une fonction dérivée admet une limite en un point, cette limite est nécessairement égale à la valeur en ce point ; propriété également vérifiée par les fonctions continues. De même, les fonctions dérivées satisfont le théorème des valeurs intermédiaires (résultat connu sous le nom de *théorème de Darboux*).

Pour autant, une fonction dérivée n'est pas forcément continue : on est donc amené à *classer* les fonctions par *régularité*. On dit ainsi d'une fonction qui satisfait le théorème de limite de la dérivée, c'est-à-dire dont la dérivée est continue, qu'elle est *de classe C^1* .

3 Fonctions de classe C^n

Définition 3.0.1. Soit $n \in \mathbb{N}$. On définit par récurrence : $f^{(0)} = f$ et, pour $n \in \mathbb{N}^*$, si $f^{(n-1)}$ est dérivable, alors $f^{(n)} = (f^{(n-1)})'$. On dit alors que f est *n fois dérivable*, et on note $f \in D^n(I, \mathbb{R})$.

Définition 3.0.2. On dit que f est *de classe C^n* , et on note $f \in C^n(I, \mathbb{R})$, si f est n fois dérivable et si $f^{(n)}$ est continue. Si f est de classe C^n pour tout $n \in \mathbb{N}$, on dit que f est *de classe C^∞* .

Proposition 3.0.3 (Opérations sur les dérivées $n^{\text{èmes}}$). Soit $n \in \mathbb{N}$. Soient $f, g : I \rightarrow \mathbb{R}$ de classe C^n et $\lambda \in \mathbb{R}$. Alors :

1. $f + g$ est de classe C^n et $(f + g)^{(n)} = f^{(n)} + g^{(n)}$,
2. λf est de classe C^n et $(\lambda f)^{(n)} = \lambda f^{(n)}$,
3. fg est de classe C^n et $(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}$ (Formule de Leibniz),
4. si g ne s'annule pas, alors $\frac{1}{g}$ est de classe C^n ,
5. si $g(I) \subset I$, alors $f \circ g$ est de classe C^n ,
6. si $f : I \rightarrow J$ est bijective et si f' ne s'annule pas, alors f^{-1} est de classe C^n .

Remarque 3.0.4. La démonstration de la formule de Leibniz ci-dessus est similaire à celle vue pour les polynômes, elle-même similaire à celle du binôme de Newton vue pour les nombres et les matrices...

Une formule du même type, quoique bien plus compliquée, existe pour $(f \circ g)^{(n)}$, c'est la *formule de Faà di Bruno*.

4 Fonctions convexes

On a vu qu'une fonction dérivable est croissante si et seulement si sa dérivée est positive. On peut s'intéresser de même aux fonctions *dont la croissance est croissante* :

Définition 4.0.1. Une fonction $f : I \rightarrow \mathbb{R}$ est dite *convexe* si :

$$\forall (x, y) \in I^2, \forall \lambda \in [0, 1], \quad f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y).$$

Si $-f$ est convexe, on dit que f est *concave*.

Exemple 4.0.2. Les fonctions \exp et $x \mapsto x^2$ sont convexes sur \mathbb{R} . Les fonctions \ln et $x \mapsto \sqrt{x}$ sont concaves.

Proposition 4.0.3. Soit $a \in I$. La fonction f est convexe si et seulement si $\tau_{a,f}$ est croissante.

Démonstration. Soient $a < x < y \in I$, il existe alors $\lambda \in [0, 1]$ tel que $x = (1 - \lambda)a + \lambda y$. Donc :

$$f(x) \leq (1 - \lambda)f(a) + \lambda f(y) \Leftrightarrow f(x) - f(a) \leq \lambda(f(y) - f(a)) \Leftrightarrow \tau_{a,f}(x) \leq \tau_{a,f}(y).$$

□

Cette caractérisation permet d'interpréter graphiquement la notion de convexité :

Proposition 4.0.4. La fonction f est convexe si et seulement si elle est *en-dessous de ses cordes*.

Démonstration. Soient $a < b \in I$. La corde à f entre a et b est $g : x \mapsto \tau_{a,f}(b)(x - a) + f(a)$. Notons $h = f - g$. On a :

$$h(x) \geq 0 \Leftrightarrow \tau_{a,f}(x) \geq \tau_{a,f}(b) \text{ si } x \geq a, \quad \tau_{a,f}(x) \leq \tau_{a,f}(b) \text{ si } x \leq a.$$

□

Par conséquent, si f est dérivable, la fonction f est convexe si et seulement si elle est *au-dessus de ses tangentes*. On a de cette façon les *inégalités de convexité* :

- $\forall x > -1, \ln(1 + x) \leq x$,
- $\forall x \in \left[0, \frac{\pi}{2}\right], \frac{2}{\pi}x \leq \sin(x) \leq x$.

La convexité est alors liée à la dérivée par :

Proposition 4.0.5. Soit $f : I \rightarrow \mathbb{R}$ dérivable. La fonction f est convexe si et seulement si f' est croissante.

Démonstration. Soient $x < y \in I$. Si, pour tout $a \in I$, $\tau_{a,f}$ est croissante, alors :

$$f'(x) \leq \tau_{x,f}(y) = \tau_{y,f}(x) \leq f'(y),$$

donc f' est croissante.

Réiproquement, si f' est croissante, soit $a \in]x, y[$. D'après le théorème des accroissements finis, il existe $c_1 \in]x, a[$ et $c_2 \in]a, y[$ tels que :

$$\tau_{a,f}(x) = f'(c_1) \leq f'(c_2) = \tau_{a,f}(y).$$

Donc $\tau_{a,f}$ est croissante, donc f est convexe. □

c'est-à-dire :

Corollaire 4.0.6. Soit $f : I \rightarrow \mathbb{R}$ deux fois dérivable. La fonction f est convexe si et seulement si f'' est positive.

5 Fonctions complexes dérivables

Comme d'habitude, on termine par un point sur les fonctions à valeurs complexes $f : I \rightarrow \mathbb{C}$. Notons tout d'abord l'équivalence :

Proposition 5.0.1. *La fonction f est dérivable si et seulement si $\operatorname{Re} f$ et $\operatorname{Im} f$ sont dérивables.*

Plus généralement, ce qui a été dit ci-dessus reste vrai à *condition que* n'interviennent pas dans les définitions, énoncés ou démonstrations, de relation d'ordre sur des nombres complexes (c'est-à-dire sur des valeurs prises par f). Par conséquent, restent valables :

- les définitions du taux d'accroissement, du nombre dérivé, de la fonction dérivée, des dérivées d'ordre supérieur,
- les opérations sur les dérivées et sur les dérivées d'ordre supérieur,
- l'inégalité des accroissements finis, les définitions et propriétés des fonctions lipschitziennes,
- le théorème de la limite de la dérivée (que l'on démontre à l'aide de l'inégalité des accroissements finis).

Remarque 5.0.2. La validité de l'inégalité des accroissements finis dans \mathbb{C} ne va pas de soi, et ne découle en particulier pas directement de la propriété précédente. Nous la démontrerons ultérieurement.

Ne sont plus valables :

- le théorème de Rolle : par exemple, $x \mapsto e^{ix}$ est dérivable et 2π -périodique, mais sa dérivée ne s'annule pas,
- le théorème des accroissements finis,
- les liens entre dérivation et monotonie.

Chapitre 15

Espaces vectoriels

Ce chapitre est consacré aux premières notions d'*algèbre linéaire*, partie des mathématiques qui formalise l'étude des systèmes d'équations linéaires vue précédemment. On note dans tout le chapitre $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

1 Espaces vectoriels

1.1 Définition

Considérons un ensemble non-vide E , muni d'une opération interne $E \times E \rightarrow E$, appelée *addition* et notée $+$, et d'une opération externe $\mathbb{K} \times E \rightarrow E$, appelée *multiplication* et notée \cdot :

$$\begin{array}{rcl} E \times E & \rightarrow & E \\ (u, v) & \mapsto & u + v \end{array} \quad \text{et} \quad \begin{array}{rcl} \mathbb{K} \times E & \rightarrow & E \\ (\lambda, u) & \mapsto & \lambda \cdot u \end{array}.$$

Définition 1.1.1. On dit que $(E, +, \cdot)$ est un *espace vectoriel* sur \mathbb{K} , ou un \mathbb{K} -*espace vectoriel*, si :

1. $+$ est *associative* : $\forall (u, v, w) \in E^3$, $(u + v) + w = u + (v + w)$,
2. E contient un *élément neutre* pour $+$, noté 0_E , tel que : $\forall u \in E$, $u + 0_E = 0_E + u = u$,
3. tout vecteur de E a un *symétrique* par $+$: $\forall u \in E$, $\exists v \in E$, $u + v = v + u = 0_E$,
4. $+$ est *commutative* : $\forall (u, v) \in E^2$, $u + v = v + u$,
5. \cdot est *associative* par rapport à la multiplication de \mathbb{K} : $\forall (\lambda, \mu) \in \mathbb{K}^2$, $\forall u \in E$, $\lambda \cdot (\mu \cdot u) = (\lambda \mu) \cdot u$,
6. L'élément neutre 1 de la multiplication de \mathbb{K} vérifie : $\forall u \in E$, $1 \cdot u = u$,
7. \cdot est *distributive par rapport à l'addition de E* : $\forall \lambda \in \mathbb{K}$, $\forall (u, v) \in E^2$, $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$,
8. \cdot est *distributive par rapport à l'addition de \mathbb{K}* : $\forall (\lambda, \mu) \in \mathbb{K}^2$, $\forall u \in E$, $(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$.

Les éléments de E sont alors appelés *vecteurs*, et les éléments de \mathbb{K} sont appelés *scalaires*. Le vecteur 0_E est appelé *vecteur nul*, et le symétrique d'un vecteur u est noté $-u$.

Remarque 1.1.2. Les conditions sur l'addition de E font de $(E, +)$ un *groupe commutatif*.

Les conditions sur $+$ et \cdot permettent d'assurer un certain nombre d'opérations :

Proposition 1.1.3. Soient E un \mathbb{K} -espace vectoriel, $u \in E$ un vecteur, $\lambda \in \mathbb{K}$ un scalaire. Alors :

1. $\forall n \in \mathbb{N}^*$, $n \cdot u = \underbrace{u + \dots + u}_{n \text{ fois}}$,
2. $0 \cdot u = 0_E$,
3. $\lambda \cdot 0_E = 0_E$,
4. $(-1) \cdot u = -u$,
5. $\lambda \cdot u = 0_E \Leftrightarrow (\lambda = 0 \text{ ou } u = 0_E)$.

1.2 Exemples fondamentaux

Les ensembles $\mathbb{K}^n, \mathbb{K}^{\mathbb{N}}, \mathcal{F}(I, \mathbb{K}), M_{n,p}(\mathbb{K}), \mathbb{K}[X]$ sont des \mathbb{K} -espaces vectoriels pour les opérations usuelles.

Remarque 1.2.1. Si $\mathbb{K} = \mathbb{C}$, ce sont également des \mathbb{R} -espaces vectoriels ! Lorsqu'il n'y a pas de confusion possible, on parlera plus simplement d'*espace vectoriel*.

On peut également les combiner :

Proposition 1.2.2. *Soient E et F des espaces vectoriels. Alors $E \times F$ est un espace vectoriel.*

1.3 Combinaison linéaire

Un espace vectoriel est fondamentalement un espace dont on sait *additionner* les éléments (les vecteurs) entre eux, et les *multiplier* par des scalaires. On nomme donc le fait de *combiner* ces opérations :

Définition 1.3.1. Soit $n \in \mathbb{N}$, soient u_1, \dots, u_n des vecteurs de E . La *combinaison linéaire* des u_i de *coefficients* $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ est le vecteur $u \in E$ suivant :

$$u = \lambda_1 \cdot u_1 + \dots + \lambda_n \cdot u_n.$$

2 Sous-espaces vectoriels

Les opérations $+$ et \cdot sur E , le vecteur nul, la notion de combinaison linéaire, donnent une *structure* à E . Celle-ci se précise lorsqu'on s'intéresse à ses *sous-espaces vectoriels*.

2.1 Définition

Définition 2.1.1. Soient $(E, +, \cdot)$ un \mathbb{K} -espace vectoriel et F un sous-ensemble de E . Notons $+_F$ et \cdot_F respectivement les restrictions de $+$ et \cdot à F . On dit que F est un *sous-espace vectoriel* de E si $(F, +_F, \cdot_F)$ est un espace vectoriel.

Remarque 2.1.2. Cette définition sous-tend essentiellement qu'un sous-espace vectoriel F est *stable par $+$ et par \cdot* (c'est-à-dire *stable par combinaison linéaire*). En effet, si les opérations $+_F$ et \cdot_F sont bien définies, alors elles vérifient les 8 hypothèses de la définition 1.1.1., puisque leurs prolongements respectifs $+$ et \cdot les vérifient sur tout E .

En particulier, le vecteur nul de F est 0_E .

Étant donné un espace vectoriel E bien identifié, il est donc relativement simple de vérifier si $F \subset E$ est un sous-espace vectoriel :

Proposition 2.1.3. *Soient $(E, +, \cdot)$ un \mathbb{K} -espace vectoriel et F un sous-ensemble de E . Alors :*

$$(F \text{ est un sous-espace vectoriel de } E) \Leftrightarrow (F \text{ est non-vide et } \forall (u, v) \in F^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \lambda \cdot u + \mu \cdot v \in F).$$

Exemple 2.1.4.

- Les sous-espaces vectoriels de \mathbb{R}^2 sont $\{0_{\mathbb{R}^2}\}$, les droites passant par l'origine et \mathbb{R}^2 ,
- Les sous-espaces vectoriels de \mathbb{R}^3 sont $\{0_{\mathbb{R}^3}\}$, les droites passant par l'origine, les plans passant par l'origine et \mathbb{R}^3 ,
- $\forall n \in \mathbb{N}$, $\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$,
- L'ensemble des solutions d'un système linéaire homogène ou d'une équation différentielle linéaire homogène est un sous-espace vectoriel.

2.2 Sous-espaces engendrés

Du point de vue des vecteurs, un sous-espace vectoriel est un ensemble de combinaisons linéaires :

Définition 2.2.1. Soient E un \mathbb{K} -espace vectoriel et (u_1, \dots, u_n) des vecteurs de E . On note $\text{Vect}(u_1, \dots, u_n)$ l'ensemble des combinaisons linéaires des u_i :

$$\text{Vect}(u_1, \dots, u_n) = \{\lambda_1 \cdot u_1 + \dots + \lambda_n \cdot u_n \mid (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n\}.$$

Proposition 2.2.2. Avec les notations ci-dessus, $\text{Vect}(u_1, \dots, u_n)$ est un sous-espace vectoriel de E .

Proposition 2.2.3. Soit F un sous-espace vectoriel de E et soient v_1, \dots, v_k des vecteurs de E . Alors :

$$(v_1, \dots, v_k \in F) \Leftrightarrow (\text{Vect}(v_1, \dots, v_k) \subset F).$$

2.3 Opérations sur les sous-espaces vectoriels

Soit E un \mathbb{K} -espace vectoriel. Les sous-espaces vectoriels de E sont avant tout des *sous-ensembles* de E : comment se comportent-ils vis-à-vis des opérations vues sur les sous-ensembles ?

Proposition 2.3.1. Soient F et G deux sous-espaces vectoriels de E . Alors leur intersection $F \cap G$ est un sous-espace vectoriel de E .

Démonstration. Soient u, v dans $F \cap G$ et $\lambda, \mu \in \mathbb{K}$. Comme u et v sont dans F , $\lambda \cdot u + \mu \cdot v$ est dans F ; et de même, comme u et v sont dans G , $\lambda \cdot u + \mu \cdot v$ est dans G . Donc $\lambda \cdot u + \mu \cdot v$ est dans $F \cap G$. Donc $F \cap G$ est un sous-espace vectoriel de E . \square

Remarque 2.3.2. À l'inverse, l'union $F \cup G$ de deux sous-espaces vectoriels n'est en général *pas* un sous-espace vectoriel ! L'opération adaptée est en fait la *somme* :

Définition 2.3.3. Soient F et G deux sous-espaces vectoriels de E . Leur *somme*, notée $F + G$, est le sous-ensemble de E suivant :

$$F + G = \{u + v \mid u \in F, v \in G\}.$$

Proposition 2.3.4. C'est un sous-espace vectoriel de E .

Démonstration. Soient u, v dans $F + G$ et $\lambda, \mu \in \mathbb{K}$. Soient u_1, v_1 dans F et u_2, v_2 dans G tels que $u = u_1 + u_2$ et $v = v_1 + v_2$. Alors $\lambda \cdot u + \mu \cdot v = (\lambda \cdot u_1 + \mu \cdot v_1) + (\lambda \cdot u_2 + \mu \cdot v_2)$. Donc $\lambda \cdot u + \mu \cdot v$ est dans $F + G$. Donc $F + G$ est un sous-espace vectoriel de E . \square

Définition 2.3.5. Soient F et G deux sous-espaces vectoriels de E . Soit u un vecteur de $F + G$. La décomposition $u = v + w$, où $v \in F$ et $w \in G$, est appelée *décomposition de u selon F et G* .

Si tout vecteur u de $F + G$ admet une unique décomposition selon F et G , on dit que F et G sont en *somme directe*. On note alors $F \oplus G$ leur somme.

Proposition 2.3.6. Deux sous-espaces vectoriels F et G de E sont en somme directe si et seulement si $F \cap G = \{0_E\}$.

Remarque 2.3.7. Si F et G sont en somme directe, tout vecteur de E admet donc *au plus* une décomposition selon F et G . Inversement, si $F + G = E$, alors tout vecteur de E admet *au moins* une décomposition selon F et G . Lorsque ces deux conditions sont réunies, tout vecteur de E admet donc *exactement* une décomposition selon F et G :

Définition 2.3.8. Deux sous-espaces vectoriels F et G de E en somme directe et de somme E sont dits *supplémentaires*. On a alors $F \oplus G = E$.

Remarque 2.3.9. Attention, un sous-espace vectoriel F peut admettre plusieurs supplémentaires : ainsi $F = \text{Vect}((1, 0))$ a pour supplémentaire $G_1 = \text{Vect}((0, 1))$ dans \mathbb{R}^2 , mais aussi $G_2 = \text{Vect}((1, 1))$.

3 Familles de vecteurs

Les notions introduites ci-dessus donnent une idée de la structure *ensembliste* des espaces vectoriels. Relisons-les maintenant du point de vue des *vecteurs*.

Soit E un \mathbb{K} -espace vectoriel, et soit $\mathcal{F} = (u_1, \dots, u_n)$ une famille de vecteurs de E . À quelle condition un vecteur $u \in E$ admet-il une décomposition selon \mathcal{F} ? Est-elle unique?

3.1 Familles génératrices

Définition 3.1.1. Soit F un sous-espace vectoriel de E . La famille \mathcal{F} est *génératrice de F* si $\text{Vect}(\mathcal{F}) = F$.

Cette notion nous permet de compléter notre réponse :

Proposition 3.1.2. Si la famille \mathcal{F} est génératrice de E , alors tout vecteur de E a au moins une décomposition selon \mathcal{F} .

Comparativement aux sous-espaces vectoriels, on a cette fois :

Proposition 3.1.3. Soient F et G deux sous-espaces vectoriels de E . Soient $\mathcal{F} = (u_1, \dots, u_k)$ une famille génératrice de F et $\mathcal{G} = (v_1, \dots, v_l)$ une famille génératrice de G . Si $F + G = E$, alors la famille $\mathcal{F} \cup \mathcal{G}$ est génératrice de E .

Démonstration. Soit $w \in E$. Comme $F + G = E$, il existe $u \in F$ et $v \in G$ tels que $w = u + v$. Comme \mathcal{F} est génératrice de F , il existe $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tels que $u = \lambda_1 \cdot u_1 + \dots + \lambda_k \cdot u_k$ et, de même, il existe $\mu_1, \dots, \mu_l \in \mathbb{K}$ tels que $v = \mu_1 \cdot v_1 + \dots + \mu_l \cdot v_l$. Donc :

$$w = \lambda_1 \cdot u_1 + \dots + \lambda_k \cdot u_k + \mu_1 \cdot v_1 + \dots + \mu_l \cdot v_l,$$

donc $\mathcal{F} \cup \mathcal{G}$ est génératrice de E . □

Enfin, mentionnons que :

Proposition 3.1.4. Toute sur-famille d'une famille génératrice de E est génératrice de E .

3.2 Familles libres

Par linéarité, on constate que l'unicité de la décomposition d'un vecteur quelconque est équivalente à l'unicité de la décomposition du vecteur nul. On définit donc :

Définition 3.2.1. La famille \mathcal{F} est *libre* si :

$$\forall (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \quad (\lambda_1 \cdot u_1 + \dots + \lambda_n \cdot u_n = 0_E) \Rightarrow (\lambda_1 = \dots = \lambda_n = 0).$$

Une famille qui n'est pas libre est dite *liée*.

Remarque 3.2.2.

- Une famille est donc *liée* s'il existe une *relation* entre les vecteurs : $\lambda_1 \cdot u_1 + \dots + \lambda_n \cdot u_n = 0_E$, non-triviale, c'est-à-dire telle que les λ_i sont non tous nuls.
- Une famille de deux vecteurs (u_1, u_2) est liée si et seulement si u_1 et u_2 sont colinéaires (c'est-à-dire proportionnels).
- Lorsqu'une famille de trois vecteurs (u_1, u_2, u_3) est liée, on dit que les vecteurs sont *coplanaires*.

Exemple 3.2.3. Toute famille de polynôme échelonnée en degrés est libre.

Démonstration. Soit (P_1, \dots, P_k) une famille de polynômes, telle que $\deg P_1 < \deg P_2 < \dots < \deg P_k$. Soient $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tels que $\lambda_1 P_1 + \dots + \lambda_k P_k = 0$. Si $\lambda_k \neq 0$, alors $\deg \lambda_k P_k = \deg P_k$, et : $\forall j < k, \deg \lambda_j P_j < \deg P_k$, donc $\deg \lambda_1 P_1 + \dots + \lambda_k P_k = \deg P_k$, ce qui est impossible. Donc $\lambda_k = 0$. Donc $\lambda_1 P_1 + \dots + \lambda_{k-1} P_{k-1} = 0$, et par une récurrence immédiate, $\lambda_{k-1} = \dots = \lambda_1 = 0$. Donc la famille (P_1, \dots, P_k) est libre. \square

Cette notion nous permet de répondre en partie à notre question :

Proposition 3.2.4. *Si la famille \mathcal{F} est libre, alors tout vecteur de $\text{Vect}(\mathcal{F})$ a une unique décomposition selon \mathcal{F} .*

En particulier, elle répond à la notion de sous-espaces en somme directe :

Proposition 3.2.5. *Soient F et G deux sous-espaces vectoriels de E . Soient $\mathcal{F} = (u_1, \dots, u_k)$ une famille libre de vecteurs de F et $\mathcal{G} = (v_1, \dots, v_l)$ une famille libre de vecteurs de G . Si F et G sont en somme directe, alors la famille $\mathcal{F} \cup \mathcal{G}$ est libre.*

Démonstration. Soient $\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_l \in \mathbb{K}$ tels que $\lambda_1 \cdot u_1 + \dots + \lambda_k \cdot u_k + \mu_1 \cdot v_1 + \dots + \mu_l \cdot v_l = 0_E$. Alors :

$$\lambda_1 \cdot u_1 + \dots + \lambda_k \cdot u_k = -\mu_1 \cdot v_1 - \dots - \mu_l \cdot v_l \in F \cap G,$$

donc, puisque F et G sont en somme directe :

$$\lambda_1 \cdot u_1 + \dots + \lambda_k \cdot u_k = -\mu_1 \cdot v_1 - \dots - \mu_l \cdot v_l = 0_E,$$

donc, puisque \mathcal{F} et \mathcal{G} sont libres,

$$\lambda_1 = \dots = \lambda_k = \mu_1 = \dots = \mu_l = 0.$$

Donc $\mathcal{F} \cup \mathcal{G}$ est libre. \square

Enfin, mentionnons que :

Proposition 3.2.6. *Toute sous-famille d'une famille libre est libre.*

3.3 Bases

Si l'on rassemble les notions précédentes, on obtient :

Définition 3.3.1. Soit F un sous-espace vectoriel de E . La famille \mathcal{F} est une *base de F* si elle est à la fois libre et génératrice de F .

Proposition 3.3.2. *Si la famille \mathcal{F} est une base de E , alors tout vecteur de E a exactement une décomposition selon \mathcal{F} .*

Définition 3.3.3. Soient $\mathcal{F} = (u_1, \dots, u_n)$ une base de E , $u \in E$, et $u = \lambda_1 \cdot u_1 + \dots + \lambda_n \cdot u_n$ la décomposition de u selon \mathcal{F} . Les λ_i sont appelés les *coordonnées de u selon \mathcal{F}* . On note :

$$u = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{F}}.$$

Proposition 3.3.4. *Soient F et G deux sous-espaces vectoriels de E . Soient $\mathcal{F} = (u_1, \dots, u_k)$ une base de F et $\mathcal{G} = (v_1, \dots, v_l)$ une base de G . Si F et G sont supplémentaires, alors la famille $\mathcal{F} \cup \mathcal{G}$ est une base de E .*

Démonstration. Comme \mathcal{F} et \mathcal{G} sont libres et que F et G sont en somme directe, $\mathcal{F} \cup \mathcal{G}$ est libre. Comme \mathcal{F} est génératrice de F , que \mathcal{G} est génératrice de G et que $F + G = E$, $\mathcal{F} \cup \mathcal{G}$ est génératrice de E . Donc $\mathcal{F} \cup \mathcal{G}$ est une base de E . \square

Exemple 3.3.5.

- Notons $e_1 = (1, 0)$ et $e_2 = (0, 1)$. La famille (e_1, e_2) est une base de \mathbb{R}^2 , appelée *base canonique de \mathbb{R}^2* .
- Notons $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ et $e_3 = (0, 0, 1)$. La famille (e_1, e_2, e_3) est une base de \mathbb{R}^3 , appelée *base canonique de \mathbb{R}^3* .
- Et de même pour tout \mathbb{R}^n , où $n \in \mathbb{N}$...

Chapitre 16

Applications linéaires

Suite au chapitre précédent consacré aux espaces vectoriels, on s'intéresse aux applications reliant ces espaces. On note dans tout le chapitre $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

1 Généralités

Soient E et F deux \mathbb{K} -espaces vectoriels, et soit $f : E \rightarrow F$ une application de E vers F .

1.1 Définitions

L'application f est dite *linéaire* si elle est *compatible avec les structures d'espaces vectoriels* de E et F , c'est-à-dire que l'image d'une combinaison linéaire est la combinaison linéaire des images :

Définition 1.1.1. On dit que f est linéaire si :

$$\forall (u, v) \in E^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \quad f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

On dit également que f est un *morphisme* d'espaces vectoriels.

Remarque 1.1.2. En particulier, si f est linéaire, alors $f(0_E) = 0_F$.

Proposition 1.1.3. Soit $f : E \rightarrow F$ linéaire. Si E' est un sous-espace vectoriel de E , alors $f(E')$ est un sous-espace vectoriel de F .

Définition 1.1.4. On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E vers F .

Si $F = E$, on note cet ensemble $\mathcal{L}(E)$. Les applications linéaires $f : E \rightarrow E$ sont appelées les *endomorphismes* de E .

Si $F = \mathbb{K}$, on dit que f est une *forme linéaire*.

1.2 Opérations sur les applications linéaires

Définition 1.2.1. Soient f, g dans $\mathcal{L}(E, F)$ et λ dans \mathbb{K} :

1. L'application $f + g : E \rightarrow F$ est définie par : $\forall u \in E, (f + g)(u) = f(u) + g(u)$,
2. L'application $\lambda f : E \rightarrow F$ est définie par : $\forall u \in E, (\lambda f)(u) = \lambda f(u)$.

Proposition 1.2.2. Les applications $f + g$ et λf sont linéaires.

Par conséquent, $\mathcal{L}(E, F)$ est un \mathbb{K} -espace vectoriel, de vecteur nul l'application nulle $0_{\mathcal{L}(E, F)} : \begin{cases} E & \rightarrow F \\ u & \mapsto 0_F \end{cases}$.

On retrouve également les opérations de composition et de restriction, héritées du chapitre sur les applications :

Proposition 1.2.3. Soient E , F et G des \mathbb{K} -espaces vectoriels.

1. Si $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$, alors $g \circ f \in \mathcal{L}(E, G)$.
2. Si $f \in \mathcal{L}(E, F)$ et si E_1 est un sous-espace vectoriel de E , alors $f|_{E_1} \in \mathcal{L}(E_1, F)$.

1.3 Image et noyau

Lors de l'étude des applications entre ensembles, nous avons introduit les notions d'*injectivité* et de *surjectivité*. Dans le cas des applications linéaires, on peut développer des outils spécifiques pour savoir si une application linéaire donnée est injective ou surjective.

Ainsi, soit $f \in \mathcal{L}(E, F)$: on sait que f est injective si et seulement si :

$$\forall (u, v) \in E^2, \quad (f(u) = f(v)) \Rightarrow (u = v).$$

Or, dans le cas des applications linéaires :

$$(f(u) = f(v)) \Leftrightarrow (f(u - v) = 0_F) \quad \text{et} \quad (u = v) \Leftrightarrow (u - v = 0_E),$$

donc f est injective si et seulement si :

$$\forall w \in E, \quad (f(w) = 0_F) \Rightarrow (w = 0_E),$$

autrement dit si l'ensemble des antécédents de 0_F par f est réduit à $\{0_E\}$:

$$\{w \in E \mid f(w) = 0_F\} = \{0_E\}.$$

On qualifie cet ensemble :

Définition 1.3.1. Soit $f \in \mathcal{L}(E, F)$. Le *noyau* de f , noté $\text{Ker } f$ (de « kernel », noyau en anglais) est l'ensemble des antécédents de 0_F par f :

$$\text{Ker } f = \{w \in E \mid f(w) = 0_F\}.$$

Proposition 1.3.2. L'ensemble $\text{Ker } f$ est un sous-espace vectoriel de E .

On a montré ci-dessus la caractérisation :

Proposition 1.3.3. L'application linéaire f est injective si et seulement si $\text{Ker } f = \{0_E\}$.

De même, on sait que la surjectivité de f dépend de l'ensemble $\text{Im } f = \{f(u) \mid u \in E\}$:

Proposition 1.3.4. L'application linéaire f est surjective si et seulement si $\text{Im } f = F$.

Proposition 1.3.5. L'ensemble $\text{Im } f$ est un sous-espace vectoriel de F .

Remarque 1.3.6. Si f est une forme linéaire, alors $\text{Im } f = \{0\}$ ou \mathbb{K} .

La linéarité de f permet de simplifier la détermination de l'ensemble $\text{Im } f$. En effet, soit $\beta = (u_1, \dots, u_n)$ une famille génératrice de E ; soit u dans E et $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ tel que $u = \lambda_1 u_1 + \dots + \lambda_n u_n$, alors, par linéarité :

$$f(u) = \lambda_1 f(u_1) + \dots + \lambda_n f(u_n),$$

c'est-à-dire que :

Proposition 1.3.7. L'application linéaire f est entièrement déterminée par $f(\beta) = (f(u_1), \dots, f(u_n))$.

En particulier :

$$\begin{aligned} \text{Im } f &= \{f(u) \mid u \in E\} \\ &= \{\lambda_1 f(u_1) + \dots + \lambda_n f(u_n) \mid (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n\} \\ &= \text{Vect}(f(u_1), \dots, f(u_n)). \end{aligned}$$

Il suffit donc de calculer l'image des vecteurs de β pour déterminer $\text{Im } f$!

Remarque 1.3.8. Si $E = \mathbb{K}^n$ ou $\mathbb{K}_n[X]$, on peut prendre pour β la *base canonique* de E :
$$\left(e_i = \underbrace{(0, \dots, 0)}_{i-1 \text{ zéros}}, \underbrace{1, \underbrace{0, \dots, 0}_{n-i \text{ zéros}}}_{1 \leq i \leq n} \right)$$
 pour \mathbb{K}^n , $(1, X, \dots, X^n)$ pour $\mathbb{K}_n[X]$.

2 Isomorphismes

Maintenant qu'on sait identifier une application linéaire injective ou surjective, on peut s'intéresser aux bijections :

Définition 2.0.1. Une application linéaire $f \in \mathcal{L}(E, F)$ bijective est appelée un *isomorphisme*. On note $\text{Isom}(E, F)$ l'ensemble des isomorphismes de E vers F .

Rappelons que toute bijection $f : E \rightarrow F$ admet une bijection réciproque $f^{-1} : F \rightarrow E$, telle que $f \circ f^{-1} = \text{Id}_F$ et $f^{-1} \circ f = \text{Id}_E$. Notons que Id_E est linéaire. La propriété suivante est remarquable :

Proposition 2.0.2. *La bijection réciproque f^{-1} d'un isomorphisme $f \in \text{Isom}(E, F)$ est une application linéaire. On a donc $f^{-1} \in \text{Isom}(F, E)$.*

Démonstration. Soient u, v dans F et λ, μ dans \mathbb{K} . On a :

$$\begin{aligned} f(\lambda f^{-1}(u) + \mu f^{-1}(v)) &= \lambda f \circ f^{-1}(u) + \mu f \circ f^{-1}(v) \quad \text{par linéarité de } f \\ &= \lambda u + \mu v, \end{aligned}$$

d'où par composition à gauche par f^{-1} :

$$\lambda f^{-1}(u) + \mu f^{-1}(v) = f^{-1}(\lambda u + \mu v),$$

donc f^{-1} est linéaire. □

Définition 2.0.3. Lorsque $E = F$, les isomorphismes sont appelés *automorphismes* de E .

L'ensemble des automorphismes de E est noté $GL(E)$ (« groupe linéaire de E »).

Cette terminologie se justifie par le fait que cet ensemble, muni de la composition \circ , forme un groupe, notion algébrique sur laquelle nous ne nous étendrons pas.

3 Endomorphismes remarquables

Soit E un \mathbb{K} -espace vectoriel. On a parlé ci-dessus de l'endomorphisme Id_E . D'autres types d'endomorphismes sont notables :

3.1 Homothéties et rotations

Définition 3.1.1. Soit $\lambda \in \mathbb{K}$. L'application $h : \begin{cases} E & \rightarrow E \\ u & \mapsto \lambda u \end{cases}$ est un endomorphisme de E , appelé *homothétie* de rapport λ .

Proposition 3.1.2. *Avec les notations ci-dessus : $h \in GL(E) \Leftrightarrow \lambda \neq 0$.*

Définition 3.1.3. Soit $\theta \in \mathbb{R}$. L'application $r_\theta : \begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{R}^2 \\ (x, y) & \mapsto (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta) \end{cases}$ est un endomorphisme de \mathbb{R}^2 , appelé *rotation* d'angle θ .

Proposition 3.1.4. *Avec les notations ci-dessus : $\forall \theta \in \mathbb{R}, r_\theta \in GL(\mathbb{R}^2)$.*

3.2 Projecteurs

Soient F et G deux sous-espaces vectoriels supplémentaires dans E . Soit u dans E . On a vu précédemment qu'il existe une unique décomposition :

$$u = u_F + u_G,$$

où $u_F \in F$ et $u_G \in G$. Ces vecteurs sont les *composantes* de u selon F et G respectivement.

Définition 3.2.1. Le *projecteur sur F parallèlement à G* est l'application $p_{F//G} : \begin{cases} E & \rightarrow E \\ u & \mapsto u_F \end{cases}$.

Proposition 3.2.2. L'application $p_{F//G}$ est linéaire, d'image F et de noyau G .

Proposition 3.2.3. Soit $p \in \mathcal{L}(E)$. L'application p est un projecteur si et seulement si $p \circ p = p$.

Exemple 3.2.4. $f(x, y, z) = (2x - 2z, y, x - z)$, $f(M) = \frac{1}{2}(M + M^\top) \dots$

3.3 Symétries

On considère à nouveau deux sous-espaces vectoriels supplémentaires F et G dans E .

Définition 3.3.1. La *symétrie par rapport à F parallèlement à G* est l'application $s_{F//G} : \begin{cases} E & \rightarrow E \\ u & \mapsto u_F - u_G \end{cases}$.

Proposition 3.3.2. L'application $s_{F//G}$ est un automorphisme de E .

Proposition 3.3.3. $p_{F//G} = \frac{1}{2}(s_{F//G} + \text{Id}_E)$.

Proposition 3.3.4. Soit $s \in \mathcal{L}(E)$. L'application s est une symétrie si et seulement si $s \circ s = \text{Id}_E$.

Exemple 3.3.5. $f(z) = \bar{z}$, $\varphi(f) = (x \mapsto f(-x)) \dots$

Chapitre 17

Développements limités

1 Formules de Taylor

Rappelons la notion de *classe de régularité* d'une fonction :

Définition 1.0.1. Soient I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{R}$ une application. On dit que f est de **classe C^n sur I** si f est n fois dérivable sur I et si sa dérivée $n^{\text{ème}}$ $f^{(n)}$ est continue sur I . On dit que f est de **classe C^∞ sur I** si f est de classe C^n sur I pour tout $n \in \mathbb{N}$.

En particulier pour $n = 0$, f est de classe C^0 sur I si et seulement si f est continue sur I .

Cette notion nous permet de généraliser le théorème des accroissements finis, qui devient :

Théorème 1.0.2 (Formule de Taylor-Lagrange à l'ordre n). Soient $n \in \mathbb{N}$ et $f : [a, b] \rightarrow \mathbb{R}$ une application. On suppose que f est de classe C^n sur $[a, b]$ et que f est $(n+1)$ fois dérivable sur $[a, b]$.

Alors il existe $c \in]a, b[$ tel que

$$f(b) = f(a) + f'(a)(b-a) + f''(a)\frac{(b-a)^2}{2!} + f'''(a)\frac{(b-a)^3}{3!} + \dots + f^{(n)}(a)\frac{(b-a)^n}{n!} + f^{(n+1)}(c)\frac{(b-a)^{n+1}}{(n+1)!},$$

c'est-à-dire :

$$f(b) = \sum_{k=0}^n f^{(k)}(a)\frac{(b-a)^k}{k!} + f^{(n+1)}(c)\frac{(b-a)^{n+1}}{(n+1)!}.$$

Démonstration de la formule de Taylor-Lagrange. La preuve du théorème des accroissements finis se généralise pour obtenir la formule de Taylor-Lagrange :

On pose, $\forall x \in [a, b]$,

$$g(x) = f(x) - \sum_{k=0}^n f^{(k)}(a)\frac{(x-a)^k}{k!} - \lambda\frac{(x-a)^{n+1}}{(n+1)!},$$

où $\lambda \in \mathbb{R}$ est choisi pour avoir $g(b) = 0$. On a alors :

$$g(a) = g'(a) = \dots = g^{(n)}(a) = 0,$$

donc en appliquant n fois le théorème de Rolle : il existe $c \in]a, b[$ tel que $g^{(n+1)}(c) = 0$, c'est-à-dire $\lambda = f^{(n+1)}(c)$.

□

Application : $\ln(1,003) = 0,00299550 + \varepsilon$ avec $0 < \varepsilon < 10^{-8}$ (sans calculatrice !)

Dans le membre de droite de la formule de Taylor-Lagrange, le dernier terme (comportant le c inconnu) peut ainsi être pensé comme une erreur entre $f(b)$ et son approximation $\sum_{k=0}^n f^{(k)}(a)\frac{(b-a)^k}{k!}$. Cette erreur diminue d'autant plus vite quand b tend vers a que l'ordre n est grand. C'est le sens de la formule suivante :

Corollaire 1.0.3 (Formule de Taylor-Young). Soit I un intervalle et $f : I \rightarrow \mathbb{R}$ de classe C^n sur I , alors $\forall x_0 \in I$, on a :

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2!} \\ &\quad + \dots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!} + \underset{x \rightarrow x_0}{o} ((x - x_0)^n). \end{aligned}$$

Démonstration. On applique la formule de Taylor-Lagrange à l'ordre $n - 1$ (comme on ignore si $f^{(n)}$ est dérivable, on ne peut pas l'utiliser à l'ordre n), avec $a = x_0$ et $b = x$. On obtient qu'il existe $c_x \in]x_0, x[$ tel que

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2!} \\ &\quad + \dots + f^{(n-1)}(x_0) \frac{(x - x_0)^{n-1}}{(n-1)!} + f^{(n)}(c_x) \frac{(x - x_0)^n}{n!}. \end{aligned}$$

En écrivant $f^{(n)}(c_x) = f^{(n)}(x_0) + f^{(n)}(c_x) - f^{(n)}(x_0)$, on en déduit

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2!} \\ &\quad + \dots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!} + \frac{f^{(n)}(c_x) - f^{(n)}(x_0)}{n!} (x - x_0)^n. \end{aligned}$$

Le dernier terme est de la forme $r(x)(x - x_0)^n$ avec $r(x) = \frac{f^{(n)}(c_x) - f^{(n)}(x_0)}{n!}$. Comme $f^{(n)}$ est continue, il suffit de voir que $c_x \xrightarrow{x \rightarrow x_0} x_0$ pour en déduire que $r(x) \xrightarrow{x \rightarrow x_0} 0$.

Cela découle évidemment de l'inégalité $|c_x - x_0| < |x - x_0|$. □

De la formule de Taylor-Lagrange découle également l'inégalité suivante, qui donne une majoration de l'erreur :

Corollaire 1.0.4 (Inégalité de Taylor-Lagrange). Soit $f : [a, b] \rightarrow \mathbb{R}$ de classe C^n sur $[a, b]$ et $(n+1)$ fois dérivable sur $]a, b[$. On suppose qu'il existe $M \geq 0$ tel que $\forall x \in]a, b[$, $|f^{(n+1)}(x)| \leq M$. Alors

$$\left| f(b) - \left(\sum_{k=0}^n f^{(k)}(a) \frac{(b-a)^k}{k!} \right) \right| \leq M \frac{|b-a|^{n+1}}{(n+1)!}.$$

Démonstration. On a, d'après la formule de Taylor-Lagrange 1.0.2, l'existence de $c \in]a, b[$ tel que

$$\left| f(b) - \left(\sum_{k=0}^n f^{(k)}(a) \frac{(b-a)^k}{k!} \right) \right| = \left| f^{(n+1)}(c) \frac{(b-a)^{n+1}}{(n+1)!} \right| \leq M \frac{|b-a|^{n+1}}{(n+1)!}.$$

□

Remarque 1.0.5. Les hypothèses de ce corollaire sont satisfaites dès que f est de classe C^{n+1} sur $[a, b]$. En effet, cela signifie que $f^{(n+1)}$ est continue sur $[a, b]$. Le théorème de Weierstrass assure alors que $f^{(n+1)}$ est bornée sur $[a, b]$, c'est-à-dire qu'il existe $M \geq 0$ tel que $\forall x \in [a, b]$, $|f^{(n+1)}(x)| \leq M$.

2 Développements limités

2.1 Définition et premières propriétés

Définition 2.1.1. Soient $f : I \rightarrow \mathbb{R}$ une fonction réelle et $x_0 \in I$. On dit que f admet un **développement limité à l'ordre n en x_0** (on abrégera souvent en DL_n) s'il existe des réels a_0, a_1, \dots, a_n tels que

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + \underset{x \rightarrow x_0}{o} ((x - x_0)^n),$$

c'est-à-dire s'il existe un polynôme P de degré $\leq n$ tel que $f(x) = P(x - x_0) + \underset{x \rightarrow x_0}{o} ((x - x_0)^n)$.

Proposition 2.1.2. *Le DL_n de f en x_0 est unique ; c'est-à-dire que si*

$$\begin{aligned} f(x) &= a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + \underset{x \rightarrow x_0}{o}((x - x_0)^n) \\ \text{et } f(x) &= b_0 + b_1(x - x_0) + b_2(x - x_0)^2 + \cdots + b_n(x - x_0)^n + \underset{x \rightarrow x_0}{o}((x - x_0)^n), \end{aligned}$$

alors $\forall 0 \leq i \leq n$, $a_i = b_i$.

Démonstration. Par l'absurde : si c'est faux, on pose $j = \min\{i \mid a_i \neq b_i\}$. Par soustraction, on a

$$0 = (a_j - b_j)(x - x_0)^j + \underset{x \rightarrow x_0}{o}((x - x_0)^j).$$

En divisant par $(x - x_0)^j$ pour $x \neq x_0$, on obtient $0 = a_j - b_j + \underset{x \rightarrow x_0}{o}(1) \xrightarrow{x \rightarrow x_0} a_j - b_j = 0$. Contradiction. \square

Remarque 2.1.3. On peut montrer en adaptant la preuve ci-dessus que si $f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + \underset{x \rightarrow x_0}{o}((x - x_0)^n)$, alors

$$f(x) \underset{x \rightarrow x_0}{\sim} a_j(x - x_0)^j, \text{ où } j = \min\{i \mid a_i \neq 0\}.$$

Par contre, si tous les termes du DL_n sont nuls, on ne peut pas en déduire d'équivalent de f en x_0 .

Remarques 2.1.4.

- On a les équivalences immédiates :

$$\begin{aligned} f \text{ a un DL}_0 \text{ en } x_0 &\iff \exists a_0 \in \mathbb{R}, f(x) = a_0 + \underset{x \rightarrow x_0}{o}(1), \\ &\iff f(x) \xrightarrow{x \rightarrow x_0} a_0, \\ &\iff f \text{ est continue en } x_0 \text{ et } f(x_0) = a_0. \end{aligned}$$

- De plus :

$$\begin{aligned} f \text{ a un DL}_1 \text{ en } x_0 &\iff \exists a_0, a_1 \in \mathbb{R}, f(x) = a_0 + a_1(x - x_0) + \underset{x \rightarrow x_0}{o}(x - x_0), \\ &\iff f \text{ est dérivable (et donc continue !) en } x_0 \text{ et } f(x_0) = a_0 \text{ et } f'(x_0) = a_1. \end{aligned}$$

- D'après la formule de Taylor-Young 1.0.3, si f est de classe C^n sur un voisinage de x_0 , alors f admet un DL_n en x_0 avec $a_k = \frac{f^{(k)}(x_0)}{k!}$.

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2!} \\ &\quad + \cdots + f^{(n)}(x_0) \frac{(x - x_0)^n}{n!} + \underset{x \rightarrow x_0}{o}((x - x_0)^n). \end{aligned}$$

- Pour $n \geq 2$, il existe des fonctions admettant un DL_n en x_0 mais qui ne sont pas n fois dérivables en x_0 , comme le montre l'exemple suivant avec $n = 2$ et $x_0 = 0$.

Exemple 2.1.5. On considère la fonction définie par

$$f(x) = \begin{cases} x^3 \sin\left(\frac{1}{x^2}\right) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

On a pour $x \neq 0$, $f(x) = x \sin\left(\frac{1}{x^2}\right) x^2 = \underset{x \rightarrow 0}{o}(x^2)$ car $x \sin\left(\frac{1}{x^2}\right)$ tend vers 0 en 0. Donc f admet un DL₂ en 0 avec $a_0 = a_1 = a_2 = 0$.

Néanmoins, on sait que f est dérivable sur $\mathbb{R} \setminus \{0\}$ par les opérations usuelles et on a

$$\forall x \neq 0, f'(x) = 3x^2 \sin\left(\frac{1}{x^2}\right) - 2 \cos\left(\frac{1}{x^2}\right).$$

Quand $x \rightarrow 0$, le premier terme tend vers 0 et le second terme n'a pas de limite. On en déduit que $f'(x)$ n'admet pas de limite quand x tend vers 0, donc que f' n'est pas continue en 0. A fortiori, f' n'est pas dérivable en 0 et donc en particulier, la dérivée seconde en 0 n'existe pas.

2.2 Développements limités usuels

Dans cette section, on donne les développements limités à tout ordre des fonctions usuelles au point $x_0 = 0$. Ces DL sont à connaître **par cœur** !

Proposition 2.2.1. Pour tout $n \in \mathbb{N}$:

1. $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o_{x \rightarrow 0}(x^n)$,
2. $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + o_{x \rightarrow 0}(x^{2n+1})$,
3. $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o_{x \rightarrow 0}(x^{2n+2})$,
4. $\forall \alpha \in \mathbb{R}, (1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \cdots + \frac{\alpha(\alpha-1)\dots(\alpha-(n-1))}{n!} x^n + o_{x \rightarrow 0}(x^n)$,
5. $\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots + x^n + o_{x \rightarrow 0}(x^n)$,
6. $\frac{1}{1+x} = \frac{1}{1-(-x)} = 1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + o_{x \rightarrow 0}(x^n)$,
7. $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n+1} \frac{x^n}{n} + o_{x \rightarrow 0}(x^n)$.

Remarque 2.2.2. On retrouve la formule : $\forall x \in \mathbb{R}, e^{ix} = \cos(x) + i \sin(x)$.

2.3 Opérations sur les développements limités

Proposition 2.3.1. Soient f et g admettant un DL_n en x_0 . Alors :

1. $\forall \lambda, \mu \in \mathbb{R}, \lambda f + \mu g$,
2. fg ,
3. $\forall p \in \mathbb{N}, f^p$,
4. $\frac{f}{g}$ si g ne s'annule pas,
5. $f \circ g$ si f est définie sur $g(I)$,

admettent un DL_n en x_0 .

Exemples 2.3.2.

- $e^x + \frac{2}{1-x} = 3 + 3x + \frac{5}{2}x^2 + \frac{13}{6}x^3 + o_{x \rightarrow 0}(x^3)$,
- $\cos(x)\sqrt{1+x} = 1 + \frac{x}{2} - \frac{5}{8}x^2 + o_{x \rightarrow 0}(x^2)$,
- $\tan(x) = x + \frac{x^3}{3} + \frac{2}{15}x^5 + o_{x \rightarrow 0}(x^5)$,

- $e^{\sin x} = 1 + x + \frac{x^2}{2} + \underset{x \rightarrow 0}{o}(x^3)$,
- $\sqrt{1 + \cos x} = \sqrt{2} - \frac{\sqrt{2}x^2}{8} + \underset{x \rightarrow 0}{o}(x^3)$.

Proposition 2.3.3 (Primitivation de DL). *Soient $f : I \rightarrow \mathbb{R}$ une fonction réelle et $x_0 \in I$. Si*

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + \underset{x \rightarrow x_0}{o}((x - x_0)^n),$$

alors son unique primitive s'annulant en x_0 satisfait

$$F(x) = \int_{x_0}^x f(t)dt = a_0(x - x_0) + a_1 \frac{(x - x_0)^2}{2} + a_2 \frac{(x - x_0)^3}{3} + \cdots + a_n \frac{(x - x_0)^{n+1}}{n+1} + \underset{x \rightarrow x_0}{o}((x - x_0)^{n+1}).$$

Remarque 2.3.4. Attention, cette proposition assure que les DL peuvent se primitiver, mais il n'est pas possible de dériver des DL, comme le montre l'exemple de la fonction $x \mapsto x^3 \sin\left(\frac{1}{x^2}\right)$.

Démonstration. On observe d'abord que la fonction $a_k(x - x_0)^k$ a pour unique primitive s'annulant en x_0 la fonction $a_k \frac{(x - x_0)^{k+1}}{k+1}$.

Il suffit donc de voir que si $g(x) = \underset{x \rightarrow x_0}{o}((x - x_0)^n)$, alors

$$G(x) = \int_{x_0}^x g(t)dt = \underset{x \rightarrow x_0}{o}((x - x_0)^{n+1})$$

On a $g(t) = r(t)(t - x_0)^n$ où $r(t)$ tend vers 0 en 0. Donc pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $|t - x_0| < \delta$ implique $|r(t)| < \varepsilon$.

Fixons un $\varepsilon > 0$ arbitraire, et soit δ associé par ce qui précède. On a alors, $\forall x \in]x_0 - \delta, x_0 + \delta[$,

$$\begin{aligned} |G(x)| &= \left| \int_{x_0}^x g(t)dt \right| = \left| \int_{x_0}^x r(t)(t - x_0)^n dt \right| \\ &\leq \int_{x_0}^x |r(t)| |t - x_0|^n dt \\ &\leq \int_{x_0}^x \varepsilon |x - x_0|^n dt \text{ car } |t - x_0| \leq |x - x_0| < \delta \\ &\leq \varepsilon |x - x_0|^n \int_{x_0}^x 1 dt = \varepsilon |x - x_0|^{n+1}. \end{aligned}$$

Ainsi $\forall \varepsilon > 0$, $\exists \delta > 0$ tel que si $|x - x_0| < \delta$, alors $\left| \frac{G(x)}{(x - x_0)^{n+1}} \right| < \varepsilon$. Cela montre que $\frac{G(x)}{(x - x_0)^{n+1}}$ tend vers 0 quand x tend vers x_0 , c'est-à-dire que $G(x) = \underset{x \rightarrow x_0}{o}((x - x_0)^{n+1})$. \square

Exemple 2.3.5. $\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \underset{x \rightarrow 0}{o}(x^{2n+1})$.

Remarque 2.3.6. On se contente dans bien des exercices de considérer les développements limités en 0. Pour le cas général, on peut toujours s'y ramener en posant $y = x - x_0$ (qui équivaut à $x = x_0 + y$). On a bien $y \rightarrow 0$ si et seulement si $x \rightarrow x_0$. Il faut alors exprimer f en fonction de y , faire le DL en $y \rightarrow 0$ puis revenir à la variable x .

Remarque 2.3.7. En voyant les formules des développements limités à tout ordre des fonctions usuelles, on peu légitimement se demander si la suite des approximations qu'ils fournissent convergent vers la fonction. Plus précisément,

$$\text{si } \forall n \in \mathbb{N}, f(x) = \sum_{k=0}^n a_k x^k + \underset{x \rightarrow 0}{o}(x^n), \text{ a-t-on } f(x) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n a_k x^k ? \quad (17.1)$$

On observe que c'est toujours vrai pour $x = 0$, mais que c'est faux en général pour $x \neq 0$, comme le montre l'exemple de la fonction $x \mapsto e^{-\frac{1}{x^2}}$ si $x \neq 0, 0$ en 0.

Par contre, la réponse à la question (17.1) est oui dans les cas suivants :

- pour tout $x \in \mathbb{R}$ pour les fonctions classiques \exp, \cos, \sin ,
- pour tout $x \in]-1, 1[$ pour les fonctions $(1+x)^\alpha$ où $\alpha \notin \mathbb{N}$ et $\frac{1}{1-x}$,
- pour tout $x \in]-1, 1]$ pour $\ln(1+x)$.

Les fonctions jouissant de cette propriété pour un intervalle $] - R, R[$ sont dites « développables en séries entières », et le plus grand tel R est appelé rayon de convergence de la série $\sum_{k=0}^{+\infty} a_k x^k$. On notera que le contre-exemple $f : x \mapsto e^{-\frac{1}{x^2}}$ est quand même une fonction de classe C^∞ sur \mathbb{R} .

3 Utilisations des DL

Voici quelques exemples d'utilisation des DL. Il y en a bien d'autres !

3.1 Calcul de limites et d'équivalents

On se propose de montrer que $\frac{1}{(\sin x)^2} - \frac{1}{x^2} \xrightarrow[x \rightarrow 0]{} \frac{1}{3}$. Pour cela, on calcule un DL₃ en 0 :

$$\begin{aligned} \frac{1}{(\sin x)^2} - \frac{1}{x^2} &= \frac{1}{\left(x - \frac{x^3}{6} + o_{x \rightarrow 0}(x^4)\right)^2} - \frac{1}{x^2} = \frac{1}{x^2 \left(1 - \frac{x^2}{6} + o_{x \rightarrow 0}(x^3)\right)^2} - \frac{1}{x^2} \\ &= \frac{1}{x^2} \left(\frac{1}{1 - \frac{x^2}{3} + o_{x \rightarrow 0}(x^3)} - 1 \right) = \frac{1}{x^2} \left(1 + \frac{x^2}{3} + o_{x \rightarrow 0}(x^3) - 1 \right) \\ &= \frac{1}{3} + o_{x \rightarrow 0}(x) \xrightarrow[x \rightarrow 0]{} \frac{1}{3}. \end{aligned}$$

Exercice 3.1.1. Déterminer un équivalent en 0 de $\sin \cosh - \sinh \circ \sin$.

3.2 Position d'une courbe par rapport à sa tangente

On se propose de donner la tangente en 0 et localement la position de la courbe par rapport à cette tangente pour la fonction $f(x) = \frac{5x - x^3}{3 + x^2}$.

On rappelle que l'équation de la droite tangente à f en x_0 est $y = f(x_0) + f'(x_0)(x - x_0)$, donnée par un DL₁ en x_0

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o_{x \rightarrow x_0}(x - x_0).$$

Pour déterminer la position de la courbe par rapport à sa tangente, il faut regarder le premier terme de degré ≥ 2 non nul dans les DL de f en x_0 . On fait un DL₃ en 0 :

$$\begin{aligned} f(x) = \frac{5x - x^3}{3 + x^2} &= (5x - x^3) \times \frac{1}{3} \times \frac{1}{1 + \frac{x^2}{3}} = (5x - x^3) \times \frac{1}{3} \times \left(1 - \frac{x^2}{3} + o_{x \rightarrow 0}(x^3)\right) \\ &= \frac{1}{3} \left(5x - \frac{5}{3}x^3 - x^3 + o_{x \rightarrow 0}(x^4)\right) = \frac{5}{3}x - \frac{8}{9}x^3 + o_{x \rightarrow 0}(x^4). \end{aligned}$$

On en déduit que la tangente à f en 0 a pour équation $y = \frac{5}{3}x$. De plus, la différence entre la courbe et la tangente est

$$f(x) - \frac{5}{3}x \underset{x \rightarrow 0}{\sim} -\frac{8}{9}x^3.$$

On en déduit qu'il existe un voisinage de 0 sur lequel $f(x) - \frac{5}{3}x$ a même signe que $-\frac{8}{9}x^3$, soit (cf. la figure 17.1) :

si $x > 0$, alors $f(x) < \frac{5}{3}x$, donc le graphe de f est en dessous de la tangente,

si $x < 0$, alors $f(x) > \frac{5}{3}x$, donc le graphe de f est en dessus de la tangente.

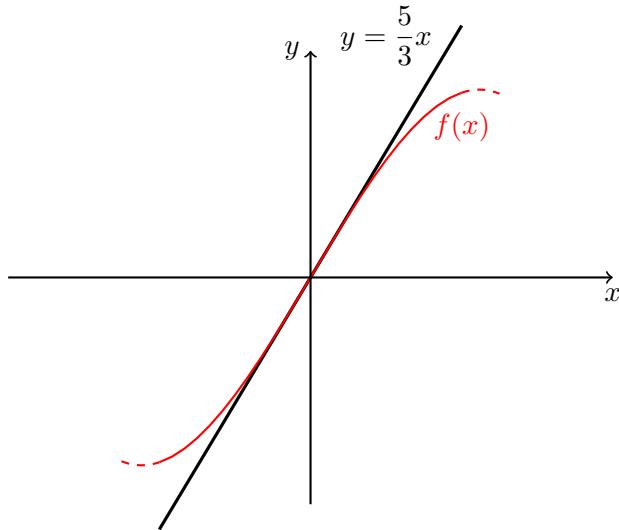


FIGURE 17.1 – Position relative de $f(x) = \frac{5x - x^3}{3 + x^2}$ en rouge par rapport à sa tangente $y = \frac{5}{3}x$ en 0.

3.3 Approximation locale de fonctions réciproques

On se propose de donner des solutions approchées de l'équation $xe^{x^2} = y$ pour x, y proches de 0. Plus précisément, on va calculer un DL₃ de la fonction réciproque de $f(x) = xe^{x^2}$.

Il faut d'abord poser plus précisément le problème. On considère la fonction $f(x) = xe^{x^2}$. Elle est de classe C^∞ par composition et produit de fonctions C^∞ . Sa dérivée est $\forall x \in \mathbb{R}$, $f'(x) = (2x^2 + 1)e^{x^2} > 0$. Donc f est strictement croissante sur \mathbb{R} . De plus, on a par opérations usuelles sur les limites :

$$\lim_{x \rightarrow +\infty} f(x) = +\infty \text{ et } \lim_{x \rightarrow -\infty} f(x) = -\infty.$$

La fonction f étant aussi continue, le théorème de la bijection réciproque vu au premier semestre assure que $f : \mathbb{R} \rightarrow \mathbb{R}$ est bijective, et que sa bijection réciproque $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$, c'est-à-dire l'unique fonction telle que l'on ait

$$\forall x, y \in \mathbb{R}, f(x) = xe^{x^2} = y \iff f^{-1}(y) = x,$$

(c'est-à-dire que $f^{-1}(y)$ est l'unique solution de l'équation $xe^{x^2} = y$) est aussi une bijection croissante continue. De plus, comme f' ne s'annule pas sur \mathbb{R} , la réciproque f^{-1} est aussi dérivable sur \mathbb{R} et on a :

$$\forall y \in \mathbb{R}, (f^{-1})'(y) = \frac{1}{f' \circ f^{-1}(y)}. \tag{17.2}$$

En fait, la fonction f^{-1} est de classe C^∞ sur \mathbb{R} . Pour le justifier, on montre par récurrence que $\forall n \in \mathbb{N}$, f^{-1} est de classe C^n . L'initialisation pour $n = 0$ est vraie d'après le théorème de la bijection réciproque. Supposons par récurrence que f^{-1} est de classe C^n . On sait que f' est de classe C^n (en fait de classe C^∞ car f elle-même est de classe C^∞) et ne s'annule pas sur \mathbb{R} . La formule (17.2) donnant la dérivée de la réciproque assure alors par composition et inversion que $(f^{-1})'$ est de classe C^n , c'est-à-dire que f^{-1} est de classe C^{n+1} , assurant l'hérédité de notre récurrence.

La formule de Taylor-Young 1.0.3 nous assure alors que f^{-1} admet un DL_n à tout ordre $n \in \mathbb{N}$, et ceci en tout point $y_0 \in \mathbb{R}$. Notons a_0, \dots, a_3 les coefficients du DL_3 en 0, c'est-à-dire

$$f^{-1}(y) = a_0 + a_1 y + a_2 y^2 + a_3 y^3 + \underset{y \rightarrow 0}{o}(y^3).$$

On doit maintenant identifier les valeurs numériques de ces coefficients. Pour cela, on utilise l'égalité bien connue $\forall x \in \mathbb{R}$, $x = f^{-1} \circ f(x) = f^{-1}(xe^{x^2})$. Pour $y = xe^{x^2} \xrightarrow{x \rightarrow 0} 0$, on obtient, en notant l'équivalent $x^4 e^{4x^2} \underset{x \rightarrow 0}{\sim} x^4$ et le DL de l'exponentielle $e^{x^2} = 1 + x^2 + \underset{x \rightarrow 0}{o}(x^3)$, que

$$\begin{aligned} x = f^{-1}(xe^{x^2}) &= a_0 + a_1 xe^{x^2} + a_2 x^2 e^{2x^2} + a_3 x^3 e^{3x^2} + \underset{x \rightarrow 0}{o}(x^3 e^{3x^2}) \\ &= a_0 + a_1 x(1 + x^2) + a_2 x^2 + a_3 x^3 + \underset{x \rightarrow 0}{o}(x^3) \\ &= a_0 + a_1 x + a_2 x^2 + (a_1 + a_3)x^3 + \underset{x \rightarrow 0}{o}(x^3). \end{aligned}$$

Par unicité du DL_3 de la fonction x , il vient :

$$\begin{cases} a_0 = 0 \\ a_1 = 1 \\ a_2 = 0 \\ a_1 + a_3 = 0 \end{cases} \iff \begin{cases} a_0 = 0 \\ a_1 = 1 \\ a_2 = 0 \\ a_3 = -1. \end{cases}$$

Finalement :

$$f^{-1}(y) = y - y^3 + \underset{y \rightarrow 0}{o}(y^3).$$

Pour une résolution de l'équation explicite $xe^{x^2} = 0,03 = 3.10^{-2} = y$, nous obtenons une approximation $x_{\text{approx}} = y - y^3 = 3.10^{-2} - 27.10^{-6} = 0,029973$. Une calculatrice donne $x = f^{-1}(y) = 0,029973060571 \dots$

4 Développements asymptotiques

Il est aussi naturel de chercher la position d'une courbe par rapport à une asymptote, ou de vouloir résoudre l'équation $xe^{x^2} = y$ pour x, y très grands, c'est-à-dire au voisinage de $+\infty$. Pour cela, on introduit la notion suivante :

Définition 4.0.1. Soit $x_0 \in \mathbb{R} \cup \{-\infty, +\infty\}$. Supposons que

$$f(x) = u_1(x) + u_2(x) + \dots + u_n(x) + \underset{x \rightarrow x_0}{o}(u_n(x)), \quad (17.3)$$

avec $\forall 1 \leq i \leq n-1$, $u_{i+1}(x) = \underset{x \rightarrow x_0}{o}(u_i(x))$. On dit alors que (17.3) est un **développement asymptotique** à n termes de $f(x)$.

La notion de développement asymptotique généralise celle de développement limité. Certaines fonctions très simples n'ont presque pas de DL, mais ont de bons développements asymptotiques. Par exemple, la fonction $\sin \sqrt{x}$ n'admet pas de DL_1 en 0 (elle n'y est pas dérivable), mais on a :

$$\forall n \in \mathbb{N}, \sin \sqrt{x} = x^{\frac{1}{2}} - \frac{x^{\frac{3}{2}}}{3!} + \frac{x^{\frac{5}{2}}}{5!} - \frac{x^{\frac{7}{2}}}{7!} + \dots + (-1)^n \frac{x^{\frac{2n+1}{2}}}{(2n+1)!} + \underset{x \rightarrow 0}{o}\left(x^{\frac{2n+1}{2}}\right).$$

Définition 4.0.2. S'il existe a, b deux réels tels que $f(x) - (ax + b) \xrightarrow{x \rightarrow +\infty} 0$, alors on dit que la droite d'équation $y = ax + b$ est **asymptote** au graphe de f en $+\infty$.

Application 4.0.3 (Position d'une courbe par rapport à une droite asymptote). On se propose de montrer que la fonction $f(x) = x\sqrt{\frac{x-1}{x+1}}$ admet une droite asymptote en $+\infty$ et de déterminer la position relative de la courbe et cette asymptote.

On va montrer que

$$f(x) = x\sqrt{\frac{x-1}{x+1}} = x - 1 + \frac{1}{2x} + \underset{x \rightarrow +\infty}{o}\left(\frac{1}{x}\right).$$

Ce n'est pas un DL, car $x \rightarrow +\infty$. Pour pouvoir utiliser la puissance de calcul des DL, on pose $y = \frac{1}{x}$. On a $y \rightarrow 0^+$ si et seulement si $x \rightarrow +\infty$. On calcule :

$$\begin{aligned} f(x) &= \frac{1}{y} \sqrt{\frac{\frac{1}{y} - 1}{\frac{1}{y} + 1}} = \frac{1}{y} \sqrt{\frac{1-y}{1+y}} = \frac{1}{y} \sqrt{(1-y)\left(1-y+y^2+\underset{y \rightarrow 0}{o}(y^2)\right)} \\ &= \frac{1}{y} \sqrt{1-y+y^2+\underset{y \rightarrow 0}{o}(y^2)-y+y^2} = \frac{1}{y} \sqrt{1-2y+2y^2+\underset{y \rightarrow 0}{o}(y^2)} \\ &= \frac{1}{y} \left(1 + \frac{1}{2}(-2y+2y^2) - \frac{(-2y)^2}{8} + \underset{y \rightarrow 0}{o}(y^2)\right) = \frac{1}{y} - 1 + \frac{y}{2} + \underset{y \rightarrow 0}{o}(y). \end{aligned}$$

La troisième ligne a été obtenue en utilisant $\sqrt{1+z} = 1 + \frac{z}{2} - \frac{z^2}{8} + \underset{z \rightarrow 0}{o}(z^2)$ avec $z = -2y+2y^2+\underset{y \rightarrow 0}{o}(y^2) \xrightarrow{y \rightarrow 0} 0$.

En revenant à la variable x , on conclut :

$$f(x) = x - 1 + \frac{1}{2x} + \underset{x \rightarrow +\infty}{o}\left(\frac{1}{x}\right).$$

On en déduit que $f(x) - (x-1) = \frac{1}{2x} + \underset{x \rightarrow +\infty}{o}\left(\frac{1}{x}\right) \xrightarrow{x \rightarrow +\infty} 0$, c'est-à-dire que la droite d'équation $y = x-1$ est asymptote au graphe de la fonction f . De plus, comme $\frac{1}{2x} > 0$ quand $x \rightarrow +\infty$, on en déduit qu'il existe un voisinage de $+\infty$ sur lequel le graphe de f est au-dessus de l'asymptote (voir figure 17.2).

Application 4.0.4 (Développement asymptotique de fonctions réciproques au voisinage de $+\infty$). On se propose de donner une bonne estimation de $f^{-1}(y)$ quand $y \rightarrow +\infty$ au moyen de fonctions usuelles.

On commence par déterminer un équivalent en $+\infty$ de $f^{-1}(y)$. On a

$$\begin{aligned} y = xe^{x^2} &\iff \ln y = \ln(x) + x^2 \iff \sqrt{\ln y} = \sqrt{\ln(x) + x^2} \\ &\iff \sqrt{\ln y} = \sqrt{x^2 \left(1 + \frac{\ln(x)}{x^2}\right)} = x\sqrt{1 + \frac{\ln(x)}{x^2}} \underset{x \rightarrow +\infty}{\sim} x = f^{-1}(y). \end{aligned}$$

On en déduit que

$$f^{-1}(y) \underset{y \rightarrow +\infty}{\sim} \sqrt{\ln y}.$$

Cet équivalent est très intéressant en lui-même, mais on voudrait être plus précis, c'est-à-dire estimer l'écart entre ces deux fonctions (l'erreur). On calcule en réinjectant la ligne précédente

$$\begin{aligned} f^{-1}(y) - \sqrt{\ln y} &= x - x\sqrt{1 + \frac{\ln(x)}{x^2}} \\ &= x - x \left(1 + \frac{\ln x}{2x^2} + \underset{x \rightarrow +\infty}{o}\left(\frac{\ln x}{x^2}\right)\right) \\ &= -\frac{\ln x}{2x} + \underset{x \rightarrow +\infty}{o}\left(\frac{\ln x}{x}\right) \underset{x \rightarrow +\infty}{\sim} -\frac{\ln x}{2x}. \end{aligned}$$

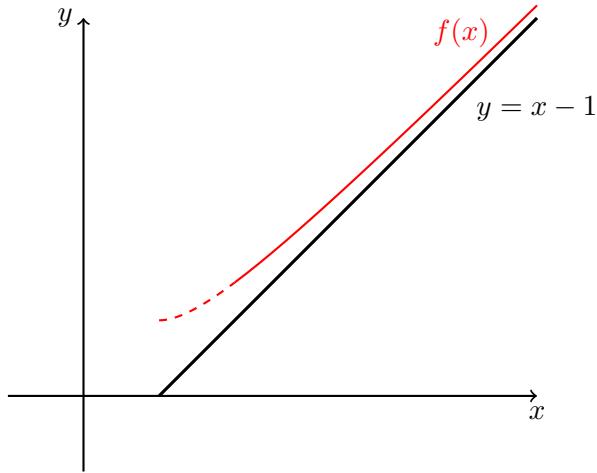


FIGURE 17.2 – Position relative de $f(x) = x\sqrt{\frac{x-1}{x+1}}$ en rouge par rapport à son asymptote $y = x - 1$ en $+\infty$.

à la seconde ligne, on a utilisé $\sqrt{1+z} = 1 + \frac{z}{2} + \underset{z \rightarrow 0}{o}(z)$ avec $z = \frac{\ln x}{x^2} \xrightarrow[x \rightarrow +\infty]{} 0$. Reste à revenir à la variable y dans l'équivalent.

On sait déjà que $x = f^{-1}(y) \underset{y \rightarrow +\infty}{\sim} \sqrt{\ln y}$. Donc

$$\begin{aligned}\ln x &= \ln f^{-1}(y) = \ln \left(\sqrt{\ln y} + \underset{y \rightarrow +\infty}{o}(\sqrt{\ln y}) \right) = \ln \left(\sqrt{\ln y} (1 + \underset{y \rightarrow +\infty}{o}(1)) \right) \\ &= \ln \sqrt{\ln y} + \ln \left(1 + \underset{y \rightarrow +\infty}{o}(1) \right) = \ln \sqrt{\ln y} + \underset{y \rightarrow +\infty}{o}(\ln \sqrt{\ln y}) \underset{y \rightarrow +\infty}{\sim} \ln \sqrt{\ln y}.\end{aligned}$$

(On rappelle qu'on ne peut pas en général composer des équivalents par une fonction, d'où les deux lignes de calcul précédent. C'est en fait possible avec la fonction \ln , pourvu que les équivalents auxquels on l'applique ne tendent pas vers 1) Comme les équivalents s'inversent et se multiplient, on en déduit que

$$f^{-1}(y) - \sqrt{\ln y} = x - \sqrt{\ln f(x)} \underset{x \rightarrow +\infty}{\sim} -\frac{\ln x}{2x} = -\frac{\ln f^{-1}(y)}{2f^{-1}(y)} \underset{y \rightarrow +\infty}{\sim} -\frac{\ln \sqrt{\ln y}}{2\sqrt{\ln y}}.$$

On conclut que :

$$f^{-1}(y) = \sqrt{\ln y} - \frac{\ln \sqrt{\ln y}}{2\sqrt{\ln y}} + \underset{y \rightarrow +\infty}{o}\left(\frac{\ln \sqrt{\ln y}}{2\sqrt{\ln y}}\right).$$

Chapitre 18

Probabilités

On développe dans ce chapitre les premières notions d'une partie des mathématiques que nous avions laissée de côté jusqu'à maintenant : la théorie des probabilités.

1 Expériences aléatoires et univers

La théorie des probabilités, ou plus généralement l'étude du *hasard*, est ancienne, mais sa *mathématisation* est relativement récente : elle est due à Andreï Kolmogorov, en 1933.

Définition 1.0.1. Une expérience *aléatoire* est une expérience dont on ne peut pas connaître le résultat à l'avance.

Exemple 1.0.2. Lancer de dés, jeu de pile ou face...

Définition 1.0.3. Étant donné une expérience aléatoire, on appelle *univers* l'ensemble des résultats possibles (également appelés *issues* ou *observables*).

L'univers est en général noté Ω , et les résultats sont notés ω .

Exemple 1.0.4. L'univers d'un lancer de dé est l'ensemble $\Omega = \{1, 2, 3, 4, 5, 6\}$.

L'univers d'un jeu de pile ou face est l'ensemble $\Omega = \{\text{pile}, \text{face}\}$.

Remarques 1.0.5. • Les résultats sont définis comme les éléments de Ω . Leur nature peut varier selon le contexte : nombres, lettres, assertions...

- On se limitera dans ce chapitre au cas où l'univers est un ensemble *fini*, donc de la forme $\Omega = \{\omega_1, \dots, \omega_n\}$ où $n \in \mathbb{N}$.

Définition 1.0.6. Un *événement* est une partie (c'est-à-dire un sous-ensemble) de Ω . L'ensemble des événements est donc $\mathcal{P}(\Omega)$. Parmi ces événements :

1. L'événement Ω est appelé *événement certain*,
2. L'événement \emptyset est appelé *événement impossible*,
3. Un singleton (de la forme $\{\omega_i\}$) est appelé *événement élémentaire*,
4. Si $A \in \mathcal{P}(\Omega)$, l'événement $\bar{A} = \Omega \setminus A$ est l'*événement contraire* de A ,
5. Deux événements A et B tels que $A \cap B = \emptyset$ sont dits *incompatibles*.

Définition 1.0.7. Un *système complet d'événements* $(A_i)_{i \in I}$ est une partition de Ω , c'est-à-dire :

$$\bigcup_{i \in I} A_i = \Omega \quad \text{et} \quad \forall i \neq j, \quad A_i \cap A_j = \emptyset.$$

On autorise certains A_i à être vides.

Exemple 1.0.8. $\forall A \in \mathcal{P}(\Omega)$, la famille (A, \bar{A}) est un système complet d'événements.

La famille des événements élémentaires est un système complet d'événements.

2 Espaces probabilisés

Définition 2.0.1. Étant donné un univers Ω , une *probabilité* \mathbb{P} sur Ω est une application

$$\mathbb{P} : \begin{cases} \mathcal{P}(\Omega) & \rightarrow [0, 1] \\ A & \mapsto \mathbb{P}(A) \end{cases}$$

telle que

$$\mathbb{P}(\Omega) = 1 \quad \text{et} \quad \forall A, B \text{ incompatibles, } \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

On dit alors que (Ω, \mathbb{P}) est un *espace probabilisé*.

Proposition 2.0.2 (Caractérisation par les probabilités élémentaires). *Soient $\Omega = \{\omega_1, \dots, \omega_n\}$ un univers et $(p_1, \dots, p_n) \in [0, 1]^n$ tel que $p_1 + \dots + p_n = 1$. Il existe alors une unique probabilité \mathbb{P} sur Ω telle que :*

$$\forall i \in \llbracket 1, n \rrbracket, \quad \mathbb{P}(\{\omega_i\}) = p_i.$$

Remarque 2.0.3. Avec ces notations, on a : $\forall A \in \mathcal{P}(\Omega), \mathbb{P}(A) = \sum_{\omega_i \in A} p_i$.

Dans le cas **particulier** où tous les p_i sont égaux, la probabilité est dite *uniforme*. On a alors :

$$\forall i \in \llbracket 1, n \rrbracket, p_i = \frac{1}{n} = \frac{1}{\text{Card}(\Omega)}, \quad \text{et} \quad \forall A \in \mathcal{P}(\Omega), \mathbb{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}.$$

Proposition 2.0.4 (Propriétés des probabilités). *Soit (Ω, \mathbb{P}) un espace probabilisé.*

1. $\forall A \in \mathcal{P}(\Omega), \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$,
2. $\mathbb{P}(\emptyset) = 0$,
3. $\forall (A, B) \in \mathcal{P}(\Omega)^2, A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$ (on dit que \mathbb{P} est *croissante*),
4. $\forall (A, B) \in \mathcal{P}(\Omega)^2, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$,
5. $\forall (A_i)_{i \in I} \in \mathcal{P}(\Omega)^I, \mathbb{P}\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} \mathbb{P}(A_i)$.

Proposition 2.0.5. *Si $(A_i)_{i \in I}$ est un système complet d'événements, alors $\sum_{i \in I} \mathbb{P}(A_i) = 1$.*

3 Probabilités conditionnelles

Considérons l'expérience suivante : on tire, successivement et sans remise, deux boules d'une urne contenant 7 boules blanches et 3 boules noires. Sachant que la première boule est blanche, quelle est la probabilité que la deuxième le soit aussi ?

Lorsqu'une information du type « l'événement $B \subset \mathcal{P}(\Omega)$ est réalisé » est apportée à une expérience aléatoire, l'univers Ω de celle-ci se réduit à B . Si Ω est muni d'une probabilité \mathbb{P} , il faut alors *renormaliser* celle-ci :

Définition 3.0.1. Avec les notations ci-dessus, et à condition que $\mathbb{P}(B) > 0$, la *probabilité conditionnelle* à B est l'application \mathbb{P}_B définie sur $\mathcal{P}(\Omega)$ par :

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Théorème 3.0.2. *C'est une probabilité sur Ω .*

Remarque 3.0.3.

- Si A et B sont incompatibles, alors $\mathbb{P}_B(A) = 0$.

- On note parfois $\mathbb{P}(\cdots | B)$, et on appelle *probabilité sachant B*, l'application \mathbb{P}_B .

Cette notion nous permet en particulier de calculer des probabilités d'intersections d'événements :

Proposition 3.0.4 (Formule des probabilités composées). *Soit (Ω, \mathbb{P}) un espace probabilisé.*

1. *Soient A, B dans $\mathcal{P}(\Omega)$, avec $\mathbb{P}(B) > 0$. Alors :*

$$\mathbb{P}(A \cap B) = \mathbb{P}_B(A) \times \mathbb{P}(B).$$

2. *Plus généralement : soit $(A_i)_{1 \leq i \leq n}$ dans $\mathcal{P}(\Omega)^n$ tel que $\forall k \in \llbracket 1, n \rrbracket$, $\mathbb{P}(A_1 \cap \dots \cap A_k) > 0$. Alors :*

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \dots \times \mathbb{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n).$$

Exemple 3.0.5. Probabilité qu'en tirant une par une les cartes d'un paquet de 32, le premier as sorte au deuxième, puis au troisième tirage.

Remarque 3.0.6. Ce cas de figure peut être représenté par un *arbre de probabilités*, sur les branches duquel sont indiquées les probabilités conditionnelles. La formule des probabilités conditionnelles indique alors que la probabilité d'un chemin dans l'arbre est le produit des probabilités des branches de ce chemin.

Il est également possible d'utiliser les probabilités conditionnelles pour faire des disjonctions de cas :

Proposition 3.0.7 (Formule des probabilités totales). *Soit $(B_i)_{i \in I}$ un système complet d'événements. Soit A un événement, alors :*

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \cap B_i) = \sum_{i \in I} \mathbb{P}_{B_i}(A) \mathbb{P}(B_i).$$

Dans le cas particulier d'un système complet de la forme (B, \bar{B}) , on a donc :

$$\mathbb{P}(A) = \mathbb{P}_B(A) \mathbb{P}(B) + \mathbb{P}_{\bar{B}}(A) \mathbb{P}(\bar{B}).$$

Exemple 3.0.8. Problème de Monty Hall : on décompose la probabilité de l'événement G « le joueur gagne en changeant de porte » selon le système (B, \bar{B}) où B est l'événement « le joueur a choisi la bonne porte du premier coup ».

Remarque 3.0.9. Dans l'arbre des probabilités, la formule des probabilités totales indique que la probabilité d'un événement est la somme des probabilités des chemins qui y mènent.

Enfin, la célèbre formule suivante permet d'échanger les événements dans une probabilité conditionnelle :

Proposition 3.0.10 (Formule de Bayes). *Si $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$, alors :*

$$\mathbb{P}_A(B) = \frac{\mathbb{P}_B(A) \mathbb{P}(B)}{\mathbb{P}(A)}.$$

Exemple 3.0.11. Dans une population, une maladie touche une personne sur 10 000. On dispose d'un test de détection qui, si une personne est malade, est positif dans 99% des cas, et qui, si une personne n'est pas malade, est positif dans 0,1% des cas. Une personne est testée positive. Quelle est la probabilité pour qu'elle soit malade ?

A contrario, la réalisation d'un événement peut ne pas avoir d'influence sur la probabilité d'un autre événement. On parle alors d'*événements indépendants*.

4 Événements indépendants

Si deux événements A et B ont une probabilité non nulle, A est *indépendant de B* si $\mathbb{P}_B(A) = \mathbb{P}(A)$, auquel cas $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$. On a alors $\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \mathbb{P}(B)$, donc B est également indépendant de A !

Plus généralement, on dit donc :

Définition 4.0.1. Soient A et B deux événements d'un univers probabilisé (Ω, \mathbb{P}) . On dit que A et B sont *indépendants* si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Exemple 4.0.2. Lors du tirage d'un dé, on considère $A = \text{« le résultat est } \leq 4 \text{ »}$, $B = \text{« le résultat est pair »}$ et $C = \text{« le résultat est } \leq 3 \text{ »}$.

Remarque 4.0.3. L'indépendance de deux événements... dépend de la probabilité considérée ! En particulier, si A et B sont indépendants pour une probabilité \mathbb{P}_C conditionnelle à un événement C , on parle d'*indépendance conditionnelle* à C .

Exemple 4.0.4. Dans l'usine de montres A , une montre sur 10 est défectueuse. Dans l'usine B , c'est une sur 50. Un horloger reçoit une livraison sans mention de l'usine. La première montre est défectueuse. Quelle est la probabilité que la deuxième le soit ?

Proposition 4.0.5. Soient A et B deux événements. On a :

$$\begin{aligned} (A \text{ et } B \text{ sont indépendants}) &\Leftrightarrow (A \text{ et } \bar{B} \text{ sont indépendants}) \\ &\Leftrightarrow (\bar{A} \text{ et } B \text{ sont indépendants}) \\ &\Leftrightarrow (\bar{A} \text{ et } \bar{B} \text{ sont indépendants}). \end{aligned}$$

La notion suivante généralise la notion d'événements indépendants :

Définition 4.0.6. Soit $(A_i)_{i \in I}$ une famille d'événements. On dit que les événements A_i sont *mutuellement indépendants* si :

$$\forall J \subset I, \quad \mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

Exemple 4.0.7. Dans un jeu de cartes sans figures, les événements A : « on pioche un carreau », B : « on pioche une carte paire », C : « on pioche une carte ≤ 6 » sont mutuellement indépendants.

Remarque 4.0.8. Attention, ce sont bien *toutes les sous-familles* de la famille $(A_i)_{i \in I}$ qui doivent vérifier cette propriété. Considérons par exemple trois lancers de pièce successifs, et notons A l'événement « la première pièce donne pile », B le même événement, et C l'événement « on obtient au moins deux face ». Alors $\mathbb{P}(A \cap B \cap C) = \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$, mais les événements A et B ne sont évidemment pas indépendants.

Proposition 4.0.9. Des événements mutuellement indépendants sont deux à deux indépendants.

Démonstration. Il suffit de considérer les sous-familles à deux éléments. □

Remarque 4.0.10. Attention, la réciproque est fausse ! Considérons par exemple un lancer de deux dés, et notons A l'événement « le premier dé donne un résultat pair », B l'événement « le second dé donne un résultat impair » et C l'événement « un dé donne un résultat pair et l'autre un résultat impair ». Alors $\mathbb{P}(A \cap B) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B)$ et de même pour (A, C) et (B, C) , mais $\mathbb{P}(A \cap B \cap C) = \frac{1}{4} \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

Chapitre 19

Dimension

On revient dans ce chapitre sur les objets vus en algèbre linéaire, en les regardant sous l'angle d'une nouvelle notion : la dimension. Dans tout le chapitre, on notera $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

1 Dimension d'un espace vectoriel

La notion de *dimension* d'un espace peut sembler intuitive : une droite est de dimension 1, un plan est de dimension 2, \mathbb{R}^3 est de dimension 3... Dans le cadre des espaces vectoriels, on va construire cette notion en étudiant la structure des *familles de vecteurs*.

Soit E un \mathbb{K} -espace vectoriel.

1.1 Définition

Définition 1.1.1. On dit que E est de *dimension finie* si E admet une famille (de vecteurs) génératrice *finie* (c'est-à-dire dont le nombre d'éléments est fini). Dans le cas contraire, on dit qu'il est de dimension *infinie*.

Exemple 1.1.2. Les espaces \mathbb{K}^n , $M_{n,p}(\mathbb{K})$, $\mathbb{K}_n[X]$ sont de dimension finie.

L'espace $\mathbb{K}[X]$ est de dimension infinie. On verra qu'il en est de même pour $\mathbb{K}^{\mathbb{N}}$ et $\mathcal{F}(\mathbb{R}, \mathbb{K})$.

On suppose à présent que E est de dimension finie.

On a déjà vu qu'une famille génératrice conserve cette propriété si on lui rajoute des vecteurs ; mais qu'elle peut perdre cette propriété si on lui en enlève. Inversement, une famille libre conserve cette propriété si on lui enlève des vecteurs ; mais elle peut perdre cette propriété si on lui en rajoute. On a en fait :

Proposition 1.1.3. Soit \mathcal{G} une famille génératrice finie de E et soit \mathcal{F} une famille libre de vecteurs de E . Alors \mathcal{F} est finie, et $\text{Card}(\mathcal{F}) \leq \text{Card}(\mathcal{G})$.

Démonstration. Notons $n = \text{Card}(\mathcal{G})$. Une famille de $n + 1$ vecteurs est alors nécessairement liée. En effet, le système linéaire associé à sa liberté, écrit selon les vecteurs de \mathcal{G} , est un système de n équations à $n + 1$ inconnues, donc comporte un paramètre. \square

Remarque 1.1.4. Par conséquent, toute famille libre d'un espace vectoriel de dimension finie est finie. En particulier, $\mathbb{K}^{\mathbb{N}}$ et $\mathcal{F}(\mathbb{R}, \mathbb{K})$ admettent des familles libres infinies, donc sont de dimension infinie.

Admettons temporairement l'existence des bases des espaces vectoriels de dimension finie. On a alors directement :

Théorème 1.1.5. Toutes les bases d'un espace vectoriel de dimension finie ont même cardinal.

Démonstration. Soient β_1 et β_2 deux bases. D'après la première proposition du chapitre, $\text{Card}(\beta_1) \leq \text{Card}(\beta_2)$ et $\text{Card}(\beta_2) \leq \text{Card}(\beta_1)$, donc $\text{Card}(\beta_1) = \text{Card}(\beta_2)$. \square

Remarque 1.1.6. Le cardinal d'une base de E ne dépend donc pas de la base considérée : il ne dépend que de l'espace vectoriel E . On dit que ce nombre est un *invariant* d'espace vectoriel, et c'est ce nombre qui est appelé *dimension* :

Définition 1.1.7. On appelle *dimension* de E le cardinal d'une base de E . C'est un entier naturel, que l'on note $\dim E$.

Exemple 1.1.8. $\dim \mathbb{K}^n = n$, $\dim M_{n,p}(\mathbb{K}) = np$, $\dim \mathbb{K}_n[X] = n + 1$.

En tant que \mathbb{R} -espace vectoriel, $\dim \mathbb{C} = 2$; en tant que \mathbb{C} -espace vectoriel, $\dim \mathbb{C} = 1$. On note alors $\dim_{\mathbb{R}} \mathbb{C} = 2$ et $\dim_{\mathbb{C}} \mathbb{C} = 1$. Plus généralement, s'il y a risque de confusion, on note $\dim_{\mathbb{K}} E$ la dimension du \mathbb{K} -espace vectoriel E .

Définition 1.1.9. Tout espace vectoriel de dimension 1 est appelé une *droite*; tout espace vectoriel de dimension 2 est appelé un *plan*.

Il reste à montrer l'existence des bases que nous avons supposée. Pour cela, examinons la « frontière » entre familles libres et familles génératrices. On a vu que les familles satisfaisant ces deux propriétés sont des *bases* de E . Les théorèmes suivants apportent alors une réponse à notre question :

Théorème 1.1.10 (Théorème de la base incomplète). *Toute famille libre de vecteurs de E peut être complétée en une base de E .*

Démonstration. On considère une famille libre \mathcal{F} et une famille génératrice \mathcal{G} . La complétion libre maximale de \mathcal{F} avec des vecteurs de \mathcal{G} est alors génératrice de $\text{Vect}(\mathcal{G}) = E$. \square

Théorème 1.1.11 (Théorème de la base extraite). *De toute famille génératrice de E peut être extraite une base de E .*

Démonstration. La famille formée par un vecteur non nul de cette famille est libre. On lui applique alors l'algorithme de la base incomplète. \square

Corollaire 1.1.12. Tout espace vectoriel de dimension finie admet une base.

1.2 Propriétés fondamentales

On considère à présent que E est un espace vectoriel de dimension $n \in \mathbb{N}$. De la construction effectuée ci-dessus, on peut tirer les premières propriétés de la notion de dimension, qui en font un outil très efficace en algèbre linéaire :

Proposition 1.2.1. Soit \mathcal{G} une famille génératrice de E . Alors $\text{Card}(\mathcal{G}) \geq n$.

Soit \mathcal{F} une famille libre de E . Alors $\text{Card}(\mathcal{F}) \leq n$.

Théorème 1.2.2. Soit \mathcal{F} une famille de n vecteurs de E . Alors :

$$(\mathcal{F} \text{ est libre}) \Leftrightarrow (\mathcal{F} \text{ est génératrice de } E) \Leftrightarrow (\mathcal{F} \text{ est une base de } E).$$

2 Dimension et sous-espaces vectoriels

Rappelons qu'un sous-espace vectoriel est lui-même un espace vectoriel, donc on peut lui appliquer la notion de dimension.

On considère toujours un espace vectoriel E de dimension $n \in \mathbb{N}$.

Théorème 2.0.1. Soit F un sous-espace vectoriel de E . Alors F est de dimension finie, et $\dim F \leq \dim E$, avec égalité si et seulement si $F = E$.

Démonstration. Soit \mathcal{F} une famille libre maximale de F (c'est aussi une famille libre de E , donc son cardinal est $\leq n$). Elle est alors génératrice de F . \square

Remarque 2.0.2. En particulier, $\dim(\{0_E\}) = 0$.

On a vu qu'étant donnés deux sous-espaces supplémentaires F et G de E , la réunion d'une base de F et d'une base de G est une base de E (dite *adaptée* à F et G). On a donc directement :

Proposition 2.0.3. *Soient F et G deux sous-espaces supplémentaires de E . Alors $\dim E = \dim F + \dim G$.*

Le cas général de la dimension d'une somme de sous-espaces vectoriels est donné par la célèbre formule suivante :

Proposition 2.0.4 (Formule de Grassmann). *Soient F et G deux sous-espaces vectoriels de E . Alors :*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Démonstration. On complète une base de $F \cap G$ en une base de F et une base de G . \square

Remarque 2.0.5. Par conséquent, deux sous-espaces F et G sont supplémentaires dans E si et seulement si ($F \cap G = \{0_E\}$ et $\dim F + \dim G = \dim E$).

3 Rang

Comme dans le premier chapitre d'algèbre linéaire, on reformule les résultats écrits du point de vue des sous-espaces vectoriels (en l'occurrence les propriétés de la dimension), en les adaptant aux familles de vecteurs, puis aux applications linéaires. La dimension est alors appelée *rang*.

3.1 Rang d'une famille de vecteurs

Définition 3.1.1. On appelle *rang* d'une famille de vecteurs \mathcal{F} la dimension du sous-espace vectoriel engendré par \mathcal{F} , et on note :

$$\text{rg}(\mathcal{F}) = \dim(\text{Vect}(\mathcal{F})).$$

Proposition 3.1.2. $(\text{rg}(\mathcal{F}) = \text{Card}(\mathcal{F})) \Leftrightarrow (\mathcal{F} \text{ est libre})$.

3.2 Rang d'une application linéaire

Soient E et F deux espaces vectoriels de dimension finie, et soit $f \in \mathcal{L}(E, F)$ une application linéaire de E dans F . Étant donné $\beta = (e_1, \dots, e_n)$ une base de E , on a vu que f est caractérisée par $f(\beta) = (f(e_1), \dots, f(e_n))$. On appelle alors *rang de f* le rang de la famille $f(\beta)$, c'est-à-dire la dimension de $\text{Vect}(f(e_1), \dots, f(e_n))$. Or cet espace vectoriel est exactement l'espace image de f , donc :

Définition 3.2.1. On appelle *rang de f* la dimension de $\text{Im}(f)$, et on note :

$$\text{rg}(f) = \dim(\text{Im}(f)).$$

De la définition du rang, on tire directement les deux propriétés suivantes :

Proposition 3.2.2. $\text{rg}(f) \leq \min(\dim E, \dim F)$.

Proposition 3.2.3.

- (f est injective) $\Leftrightarrow (\text{rg}(f) = \dim E)$,
- (f est surjective) $\Leftrightarrow (\text{rg}(f) = \dim F)$.

Remarque 3.2.4. Par conséquent, il ne peut exister d'applications linéaires bijectives (*isomorphismes*) entre deux espaces vectoriels que s'ils sont de même dimension ! On dit alors que ces espaces sont *isomorphes*. De plus, et encore d'après la propriété précédente :

Proposition 3.2.5. *Soient E et F deux espaces vectoriels de même dimension $n \in \mathbb{N}$, et $f \in \mathcal{L}(E, F)$. Alors :*

$$(f \text{ est injective}) \Leftrightarrow (f \text{ est surjective}) \Leftrightarrow (f \text{ est bijective}).$$

Considérons l'application linéaire $f : \begin{cases} \mathbb{R}^3 & \rightarrow \mathbb{R}^3 \\ (x, y, z) & \mapsto (2x - y, y - 2z, x - z) \end{cases}$. Alors $\dim \text{Ker}(f) = 1$ et $\text{rg}(f) = 2$. Considérons à présent les modifications suivantes de f :

- Pour $g : \begin{cases} \mathbb{R}^4 & \rightarrow \mathbb{R}^3 \\ (x, y, z, t) & \mapsto (2x - y, y - 2z, x - z) \end{cases}$, on a $\dim \text{Ker}(g) = 2$ et $\text{rg}(g) = 2$,
- Pour $h : \begin{cases} \mathbb{R}^3 & \rightarrow \mathbb{R}^4 \\ (x, y, z) & \mapsto (2x - y, y - 2z, x - z, x - z) \end{cases}$, on a $\dim \text{Ker}(h) = 1$ et $\text{rg}(h) = 2$.

Plus généralement, les dimensions de ces espaces sont liées par la formule suivante, particulièrement utile :

Théorème 3.2.6 (Théorème du rang). *Soient E et F deux espaces vectoriels de dimension finie, et $f \in \mathcal{L}(E, F)$. Alors :*

$$\dim \text{Ker}(f) + \text{rg}(f) = \dim E.$$

Démonstration. Soient \mathcal{F} une base de $\text{Ker}(f)$ et \mathcal{F}' une famille d'antécédents d'une base de $\text{Im}(f)$. Alors $\mathcal{F} \cup \mathcal{F}'$ est une base de E . \square

Remarque 3.2.7.

- Ce théorème, utilisé sous la forme $\text{rg}(f) = \dim E - \dim \text{Ker}(f)$, permet notamment de déterminer directement la dimension de $\text{Im}(f)$, et donc de simplifier l'étude de $\text{Im}(f)$.
- Cette égalité sur les dimensions **n'implique pas** que $\text{Ker}(f)$ et $\text{Im}(f)$ sont supplémentaires - en général, ils ne font même pas partie du même espace vectoriel ! Ces deux espaces **peuvent** être supplémentaires (c'est par exemple le cas pour les projecteurs), mais ce n'est pas systématique : ainsi, pour $f : \begin{cases} \mathbb{R}_3[X] & \rightarrow \mathbb{R}_3[X] \\ P & \mapsto P'' \end{cases}$, on a $\text{Ker}(f) = \text{Im}(f) = \mathbb{R}_1[X]$.

Chapitre 20

Représentation matricielle et déterminants

On décrit dans ce chapitre la représentation matricielle des vecteurs et des applications linéaires, qui, de par sa simplicité et son efficacité, est la plus couramment utilisée.

Ceci nous permettra d'étendre la notion de déterminant, déjà définie pour les matrices de taille 2, à l'ensemble des matrices carrées.

Dans tout le chapitre, sauf mention contraire, on considère un espace vectoriel E de dimension finie n sur $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

1 Représentation matricielle

1.1 Définition

Soient u un vecteur de E et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Rappelons qu'il existe alors un et un seul $(\lambda_1, \dots, \lambda_n)$ dans \mathbb{K}^n tel que :

$$u = \lambda_1 e_1 + \dots + \lambda_n e_n,$$

et que $\lambda_1, \dots, \lambda_n$ sont les *coordonnées* de u . Toute la théorie de la représentation matricielle repose sur cette notion.

Définition 1.1.1. La *représentation matricielle*, ou plus simplement la *matrice*, du vecteur u dans la base \mathcal{B} , est la matrice notée $M_{\mathcal{B}}(u)$ définie par :

$$M_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in M_{n,1}(\mathbb{K}).$$

Remarque 1.1.2. Dans le cas d'un vecteur de \mathbb{K}^n dans la base canonique notamment, les coordonnées de u sont égales à ses composantes. Il faut cependant toujours garder à l'esprit qu'un vecteur et la matrice associée à celui-ci sont des objets différents.

Définition 1.1.3. Soit $\mathcal{F} = (u_1, \dots, u_p)$ une famille de vecteurs de E . La matrice de \mathcal{F} dans la base \mathcal{B} est la matrice notée $M_{\mathcal{B}}(\mathcal{F})$ définie par :

$$M_{\mathcal{B}}(\mathcal{F}) = (M_{\mathcal{B}}(u_1), \dots, M_{\mathcal{B}}(u_p)) \in M_{n,p}(\mathbb{K}).$$

Exemple 1.1.4. Notamment, $M_{\mathcal{B}}(\mathcal{B}) = I_n$.

Proposition 1.1.5. (\mathcal{F} est une base de E) \Leftrightarrow ($M_{\mathcal{B}}(\mathcal{F})$ est inversible).

1.1.6.

Soit E un espace vectoriel de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E .

Soit F un espace vectoriel de dimension p et $\mathcal{B}' = (e'_1, \dots, e'_p)$ une base de F .

Soit $f : E \rightarrow F$ une application linéaire.

La matrice de f depuis la base \mathcal{B} vers la base \mathcal{B}' est la matrice notée $M_{\mathcal{B}'}^{\mathcal{B}}(f)$ définie par :

$$M_{\mathcal{B}'}^{\mathcal{B}}(f) = M_{\mathcal{B}'}(f(\mathcal{B})) = (M_{\mathcal{B}'}(f(e_1)), \dots, M_{\mathcal{B}'}(f(e_n))) \in M_{p,n}(\mathbb{K}).$$

Exemple 1.1.7. Notamment, $M_{\mathcal{B}}^{\mathcal{B}}(\text{id}_E) = I_n$.

Tout l'intérêt de cette notion réside dans la formule suivante :

Proposition 1.1.8. $\forall u \in E, M_{\mathcal{B}'}(f(u)) = M_{\mathcal{B}'}^{\mathcal{B}}(f)M_{\mathcal{B}}(u)$.

ainsi que dans la propriété suivante :

Proposition 1.1.9. *Les applications $M_{\mathcal{B}} : E \rightarrow M_{n,1}(\mathbb{K})$ et $M_{\mathcal{B}'}^{\mathcal{B}} : \mathcal{L}(E, F) \rightarrow M_{p,n}(\mathbb{K})$ sont des isomorphismes.*

Remarque 1.1.10. En particulier, $\dim \mathcal{L}(E, F) = np$.

1.2 Propriétés

Proposition 1.2.1. *Soient E, F, G des espaces vectoriels de dimension finie et $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ des bases de E, F, G respectivement. Soient $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. Alors :*

$$M_{\mathcal{B}''}^{\mathcal{B}''}(g \circ f) = M_{\mathcal{B}''}^{\mathcal{B}'}(g)M_{\mathcal{B}'}^{\mathcal{B}}(f).$$

Démonstration. Soit $u \in E$. On a :

$$M_{\mathcal{B}''}^{\mathcal{B}''}(g \circ f)M_{\mathcal{B}}(u) = M_{\mathcal{B}''}(g \circ f(u)) = M_{\mathcal{B}''}(g(f(u))) = M_{\mathcal{B}''}^{\mathcal{B}'}(g)M_{\mathcal{B}'}(f(u)) = M_{\mathcal{B}''}^{\mathcal{B}'}(g)M_{\mathcal{B}'}^{\mathcal{B}}(f)M_{\mathcal{B}}(u).$$

Par bijectivité de $M_{\mathcal{B}}$, on a donc : $\forall X \in M_{n,1}(\mathbb{K}), M_{\mathcal{B}''}^{\mathcal{B}'}(g \circ f)X = M_{\mathcal{B}''}^{\mathcal{B}'}(g)M_{\mathcal{B}'}^{\mathcal{B}}(f)X$, d'où la formule voulue. \square

Proposition 1.2.2. *Avec les notations précédentes, soit $f \in \mathcal{L}(E, F)$ bijective. Alors $M_{\mathcal{B}'}^{\mathcal{B}}(f)$ est inversible, et :*

$$M_{\mathcal{B}}^{\mathcal{B}'}(f^{-1}) = (M_{\mathcal{B}'}^{\mathcal{B}}(f))^{-1}.$$

Démonstration. On a $f^{-1} \circ f = \text{id}_E$, donc d'après la proposition précédente :

$$M_{\mathcal{B}}^{\mathcal{B}'}(f^{-1})M_{\mathcal{B}'}^{\mathcal{B}}(f) = M_{\mathcal{B}}^{\mathcal{B}}(\text{id}_E) = I_n,$$

donc $M_{\mathcal{B}'}^{\mathcal{B}}(f)$ est inversible, d'inverse $M_{\mathcal{B}}^{\mathcal{B}'}(f^{-1})$. \square

2 Changements de bases

2.1 Matrices de passage

Définition 2.1.1. Soient \mathcal{B}_1 et \mathcal{B}_2 deux bases de E . La *matrice de passage* de la base \mathcal{B}_1 à la base \mathcal{B}_2 est la matrice notée $P_{\mathcal{B}_1}^{\mathcal{B}_2}$ définie par :

$$P_{\mathcal{B}_1}^{\mathcal{B}_2} = M_{\mathcal{B}_1}(\mathcal{B}_2).$$

Exemple 2.1.2. Notamment, $P_{\mathcal{B}}^{\mathcal{B}} = I_n$.

Remarque 2.1.3. Cette dénomination est particulièrement malheureuse, comme on peut le constater ci-après.

2.2 Formules de changement de base

Comme leur nom l'indique, les matrices de passage permettent de faire passer la matrice d'un vecteur dans une base \mathcal{B}_1 à la matrice de ce vecteur dans une base \mathcal{B}_2 ... mais la formule correspondante est assez contre-intuitive :

Proposition 2.2.1. $\forall u \in E, P_{\mathcal{B}_1}^{\mathcal{B}_2} M_{\mathcal{B}_2}(u) = M_{\mathcal{B}_1}(u)$.

Démonstration. Notons $\mathcal{B}_2 = (e_1, \dots, e_n)$ et $M_{\mathcal{B}_2}(u) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$. Alors $u = \lambda_1 e_1 + \dots + \lambda_n e_n$, donc par isomorphisme :

$$M_{\mathcal{B}_1}(u) = \lambda_1 M_{\mathcal{B}_1}(e_1) + \dots + \lambda_n M_{\mathcal{B}_1}(e_n) = (M_{\mathcal{B}_1}(e_1) \quad \dots \quad M_{\mathcal{B}_1}(e_n)) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = P_{\mathcal{B}_1}^{\mathcal{B}_2} M_{\mathcal{B}_2}(u).$$

□

Proposition 2.2.2. La matrice $P_{\mathcal{B}_1}^{\mathcal{B}_2}$ est inversible, et $(P_{\mathcal{B}_1}^{\mathcal{B}_2})^{-1} = P_{\mathcal{B}_2}^{\mathcal{B}_1}$.

On peut alors déterminer la formule de changement de base pour la matrice d'une application linéaire :

Proposition 2.2.3. Soit $f : E \rightarrow F$. Alors :

$$M_{\mathcal{B}_2'}^{\mathcal{B}_2}(f) = P_{\mathcal{B}_2'}^{\mathcal{B}_1'} M_{\mathcal{B}_1'}^{\mathcal{B}_1}(f) P_{\mathcal{B}_1}^{\mathcal{B}_2}.$$

Démonstration. Soit u dans E . D'une part, $M_{\mathcal{B}_2'}(f(u)) = M_{\mathcal{B}_2'}^{\mathcal{B}_2}(f) M_{\mathcal{B}_2}(u)$, et d'autre part :

$$M_{\mathcal{B}_2'}(f(u)) = P_{\mathcal{B}_2'}^{\mathcal{B}_1'} M_{\mathcal{B}_1'}(f(u)) = P_{\mathcal{B}_2'}^{\mathcal{B}_1'} M_{\mathcal{B}_1'}^{\mathcal{B}_1}(f) M_{\mathcal{B}_1}(u) = P_{\mathcal{B}_2'}^{\mathcal{B}_1'} M_{\mathcal{B}_1'}^{\mathcal{B}_1}(f) P_{\mathcal{B}_1}^{\mathcal{B}_2} M_{\mathcal{B}_2}(u).$$

□

3 Rang d'une matrice

Vu la construction effectuée, les concepts utilisés dans l'étude des applications linéaires peuvent être appliqués aux matrices :

Définition 3.0.1. Soit $M \in M_{p,n}(\mathbb{K})$. Le *noyau* de M est le sous-espace vectoriel de $M_{n,1}(\mathbb{K})$ défini par :

$$\text{Ker } M = \{X \in M_{n,1}(\mathbb{K}) \mid MX = 0_{p,1}\}.$$

L'*image* de M est le sous-espace vectoriel de $M_{p,1}(\mathbb{K})$ défini par :

$$\text{Im } M = \{MX \mid X \in M_{n,1}(\mathbb{K})\}.$$

Remarque 3.0.2. Ces espaces peuvent être naturellement identifiés au noyau et à l'image de l'application $f_M : \mathbb{K}^n \rightarrow \mathbb{K}^p$ canoniquement associée à M .

Définition 3.0.3. Le *rang* de M est la dimension de $\text{Im } M$: $\text{rg } M = \dim(\text{Im } M)$.

Remarque 3.0.4. Cette notion coïncide avec le rang d'une matrice défini, dans le cadre de l'étude des systèmes linéaires, comme le nombre de ses pivots ! Rappelons alors que :

$$\forall P \in GL_p(\mathbb{K}), \text{rg}(PM) = \text{rg}(M) \quad \text{et} \quad \text{rg}(M^\top) = \text{rg}(M).$$

Théorème 3.0.5 (Théorème du rang matriciel). $\dim(\text{Ker } M) + \text{rg } M = n$.

4 Déterminant d'une matrice

4.1 Définition

Définition 4.1.1. Soit M une matrice dans $M_2(\mathbb{R})$ de colonnes C_1 et C_2 , et soient u et v les vecteurs de \mathbb{R}^2 dont les matrices dans la base canonique sont respectivement C_1 et C_2 . Le *déterminant* de M est l'aire algébrique du parallélogramme dirigé par u et v dans le plan \mathbb{R}^2 . On a ainsi :

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Proposition 4.1.2. L'application $\det : M_2(\mathbb{R}) \rightarrow \mathbb{R}$ ainsi définie est linéaire et antisymétrique par rapport aux colonnes de M . De plus, $\det(I_2) = 1$ et :

$$(\det M = 0) \Leftrightarrow (\text{Les vecteurs } u \text{ et } v \text{ sont colinéaires}).$$

Définition 4.1.3. Soit M une matrice dans $M_3(\mathbb{R})$ de colonnes C_1 , C_2 et C_3 , et soient u , v et w les vecteurs de \mathbb{R}^3 dont les matrices dans la base canonique sont respectivement C_1 , C_2 et C_3 . Le *déterminant* de M est le volume algébrique du parallélépipède dirigé par u , v et w dans l'espace \mathbb{R}^3 .

Proposition 4.1.4. L'application $\det : M_3(\mathbb{R}) \rightarrow \mathbb{R}$ ainsi définie est linéaire et antisymétrique par rapport aux colonnes de M . De plus, $\det(I_3) = 1$ et :

$$(\det M = 0) \Leftrightarrow (\text{La famille } (u, v, w) \text{ est liée}).$$

Plus généralement, on peut montrer que :

Proposition 4.1.5. Il existe une unique application $\det : M_n(\mathbb{K}) \rightarrow \mathbb{K}$, appelée *déterminant*, telle que :

1. \det est linéaire par rapport aux colonnes : $\forall k \in \llbracket 1, n \rrbracket$, $\forall (C_k, D_k) \in M_{n,1}(\mathbb{K})^2$, $\forall (\lambda, \mu) \in \mathbb{K}^2$,

$$\det(C_1 \cdots (\lambda C_k + \mu D_k) \cdots C_n) = \lambda \det(C_1 \cdots C_k \cdots C_n) + \mu \det(C_1 \cdots D_k \cdots C_n),$$

2. \det est antisymétrique par rapport aux colonnes : $\forall i \neq j \in \llbracket 1, n \rrbracket^2$,

$$\det(C_1 \cdots C_i \cdots C_j \cdots C_n) = -\det(C_1 \cdots C_j \cdots C_i \cdots C_n),$$

3. $\det(I_n) = 1$.

Remarque 4.1.6.

- Par extension des observations précédentes, $\det(C_1 \cdots C_n)$ est égal au volume algébrique du *pavé* dirigé par les vecteurs canoniquement associés à C_1, \dots, C_n dans \mathbb{K}^n .
- On notera donc que la notion de déterminant n'a de sens que pour une matrice *carrée*.
- La fonction \det ainsi définie est également linéaire par rapport aux *lignes*, ce qui s'avérera utile par la suite.

4.2 Propriétés

Soit M dans $M_n(\mathbb{K})$.

Proposition 4.2.1. Si deux colonnes de M sont égales, alors $\det M = 0$.

Démonstration. C'est immédiat par antisymétrie par rapport aux colonnes. \square

Proposition 4.2.2. $\forall \lambda \in \mathbb{K}$, $\det(\lambda M) = \lambda^n \det M$.

Démonstration. C'est immédiat par linéarité par rapport aux colonnes. \square

Remarque 4.2.3. Attention, le déterminant **n'est pas** une application linéaire. Par exemple, pour $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, on a $\det A = \det B = 0$, mais $\det(A + B) = \det I_2 = 1$.

De même :

Proposition 4.2.4. $\forall A \in D_n(\mathbb{K})$, $\det(A) = \prod_{i=1}^n a_{ii}$.

Pour aller au-delà de ces premiers calculs, on aura besoin des deux propriétés suivantes, fondamentales :

Proposition 4.2.5 (Comportement du déterminant sous transformations élémentaires).

Sous transvection, dilatation et permutation, le déterminant se comporte comme suit :

$$\forall i \neq j \in \llbracket 1, n \rrbracket^2, \forall \lambda \in \mathbb{K},$$

$$\det(MT_{ij}(\lambda)) = \det M, \quad \det(MD_i(\lambda)) = \lambda \det M, \quad \det(MP_{ij}) = -\det M,$$

c'est-à-dire que pour toute matrice d'opération élémentaire E : $\det(ME) = \det(M) \det(E)$.

Par conséquent :

Proposition 4.2.6. Soit $A \in T_n^+(\mathbb{K})$ ou $T_n^-(\mathbb{K})$, alors $\det(A) = \prod_{i=1}^n a_{ii}$.

d'où, d'après l'algorithme du pivot de Gauss-Jordan :

Proposition 4.2.7. $M \in GL_n(\mathbb{K}) \Leftrightarrow \det M \neq 0$.

Remarque 4.2.8. On étend ainsi la propriété observée en dimension 2 et 3 : $\det M \neq 0$ si et seulement si les vecteurs canoniquement associés aux colonnes de M forment une base de \mathbb{K}^n .

On déduit également des propriétés ci-dessus que :

Proposition 4.2.9. $\forall A, B \in M_n(\mathbb{K})$, $\det(AB) = \det A \cdot \det B$.

Démonstration. Cette propriété est immédiatement vraie pour des matrices non inversibles. Sinon, d'après le théorème de Gauss-Jordan matriciel, A et B sont des produits de matrices de transformations élémentaires, sur les-quelles on montre par calcul que cette propriété est vraie. \square

Par suite :

Proposition 4.2.10. $\forall A \in GL_n(\mathbb{K})$, $\det(A^{-1}) = \frac{1}{\det A}$.

Proposition 4.2.11. $\forall A \in M_n(\mathbb{K})$, $\det(A^\top) = \det A$.

Démonstration. C'est encore une fois vrai pour les matrices de transformations élémentaires, donc pour toutes les matrices d'après le théorème de Gauss-Jordan matriciel. \square

Enfin, la propriété suivante permet de relier entre eux les déterminants de taille n et $n - 1$, ce qui nous sera utile en pratique :

Proposition 4.2.12. Si $M = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{pmatrix}$, alors $\det M = \det N$.

Démonstration. Par opérations élémentaires sur la matrice N , et toujours d'après l'algorithme du pivot, on se ramène à une matrice triangulaire. \square

5 Calcul de déterminants

Deux méthodes de calcul du déterminant peuvent être utilisées, parfois au cours d'un même calcul, de façon complémentaire. Elles sont décrites ci-dessous ; pour simplifier les notations, on notera de façon usuelle :

$$\forall M \in M_n(\mathbb{K}), \quad \det(M) = |M|.$$

5.1 Par l'algorithme du pivot

Comme décrit dans la partie précédente, le comportement du déterminant sous transformations élémentaires est connu ; il est en particulier invariant sous transvection. L'échelonnement de l'algorithme du pivot de Gauss-Jordan permet alors de se ramener au calcul, immédiat, du déterminant d'une matrice triangulaire.

5.2 Par développement par rapport à une ligne ou une colonne

Prenons l'exemple d'une matrice de taille 3 : $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Par linéarité par rapport à la première colonne :

$$\det A = a \begin{vmatrix} 1 & b & c \\ 0 & e & f \\ 0 & h & i \end{vmatrix} + d \begin{vmatrix} 0 & b & c \\ 1 & e & f \\ 0 & h & i \end{vmatrix} + g \begin{vmatrix} 0 & b & c \\ 0 & e & f \\ 1 & h & i \end{vmatrix},$$

d'où par permutations :

$$\det A = a \begin{vmatrix} 1 & b & c \\ 0 & e & f \\ 0 & h & i \end{vmatrix} - d \begin{vmatrix} 1 & e & f \\ 0 & b & c \\ 0 & h & i \end{vmatrix} + g \begin{vmatrix} 1 & h & i \\ 0 & b & c \\ 0 & e & f \end{vmatrix},$$

et donc, par transvections :

$$\det A = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \begin{vmatrix} b & c \\ e & f \end{vmatrix}.$$

Il s'agit du *développement de $\det A$ par rapport à la première colonne*. L'intérêt de ce calcul réside dans le fait qu'il permet de se ramener au calcul d'un déterminant de taille moindre (précisément, un déterminant de taille n se décompose en combinaison linéaire de n déterminants de taille $n - 1$), qui sont individuellement plus simples à calculer.

La même démarche peut être appliquée au calcul d'un déterminant de toute taille, par rapport à toute ligne ou à toute colonne :

Définition 5.2.1. Soit $A \in M_n(\mathbb{K})$. Pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$, on note $A_{ij} \in M_{n-1}(\mathbb{K})$ la sous-matrice de A obtenue en ôtant la ligne i et la colonne j de A .

Proposition 5.2.2 (Formules de développement du déterminant).

Soit $A \in M_n(\mathbb{K})$. Soient k et l dans $\llbracket 1, n \rrbracket$.

1. *Le développement du déterminant de A par rapport à la ligne k s'écrit :*

$$\det A = \sum_{j=1}^n (-1)^{k+j} a_{kj} \det(A_{kj}).$$

2. *Le développement du déterminant de A par rapport à la colonne l s'écrit :*

$$\det A = \sum_{i=1}^n (-1)^{i+l} a_{il} \det(A_{il}).$$

Remarque 5.2.3. Une bonne pratique consiste à identifier le signe en facteur de chaque coefficient. Celui-ci est égal à $(-1)^{i+j}$, c'est-à-dire que les signes se présentent en damier à partir d'un + sur le premier coefficient.

6 Utilisations du déterminant

6.1 Déterminant d'un endomorphisme

La notion de déterminant d'un endomorphisme est bien définie, d'après la propriété suivante :

Proposition 6.1.1. *Soit $f \in \mathcal{L}(E)$ et soient \mathcal{B} et \mathcal{B}' deux bases de E . Alors :*

$$\det(M_{\mathcal{B}}^{\mathcal{B}}(f)) = \det(M_{\mathcal{B}'}^{\mathcal{B}'}(f)).$$

Démonstration. C'est une conséquence immédiate des propriétés du déterminant et de la formule de changement de base :

$$M_{\mathcal{B}'}^{\mathcal{B}'}(f) = P_{\mathcal{B}'}^{\mathcal{B}} M_{\mathcal{B}}^{\mathcal{B}}(f) P_{\mathcal{B}}^{\mathcal{B}'}.$$

□

Définition 6.1.2. On note $\det f$, et on appelle *déterminant de l'endomorphisme* $f \in \mathcal{L}(E)$, la valeur commune de ces déterminants.

En toute logique, les propriétés du déterminant s'écrivent dans ce cadre :

Proposition 6.1.3. $f \in GL(E) \Leftrightarrow \det f \neq 0$.

Proposition 6.1.4. $\forall f, g \in \mathcal{L}(E), \det(f \circ g) = \det f \cdot \det g$.

Proposition 6.1.5. $\forall f \in GL(E), \det(f^{-1}) = \frac{1}{\det f}$.

6.2 Résolution de systèmes linéaires

Soit $AX = B$ un système linéaire de taille n , que l'on suppose de Cramer.

Notons $A = (C_1 \cdots C_n)$ les colonnes de A et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ les inconnues du système. On a alors :

$$x_1 C_1 + \cdots + x_n C_n = B.$$

Notons alors, pour tout j dans $\llbracket 1, n \rrbracket$, $A_j = (C_1 \cdots C_{j-1} \ B \ C_{j+1} \cdots C_n)$ la matrice obtenue à partir de A en remplaçant la colonne j par B . Alors :

$$\begin{aligned} \det A_j &= \det(C_1 \cdots C_{j-1} (x_1 C_1 + \cdots + x_n C_n) C_{j+1} \cdots C_n) \\ &= \sum_{k=1}^n x_k \det(C_1 \cdots C_{j-1} C_k C_{j+1} \cdots C_n) \quad \text{par linéarité du déterminant par rapport aux colonnes} \\ &= x_j \det A \quad \text{par annulation du déterminant lorsque deux colonnes sont égales.} \end{aligned}$$

On obtient ainsi les :

Proposition 6.2.1 (Formules de Cramer). *Avec les notations précédentes, les solutions du système sont données par :*

$$\forall j \in \llbracket 1, n \rrbracket, \quad x_j = \frac{\det A_j}{\det A}.$$

Chapitre 21

Intégration

Nous avons déjà consacré un chapitre à la primitivation, et mentionné son lien avec l'intégration. On revient ici sur la construction de l'intégrale, due à Riemann.

1 Définition

Dans tout le chapitre, sauf mention contraire, on considère un segment $I = [a, b]$ de \mathbb{R} , avec $a < b$, et $f : I \rightarrow \mathbb{R}$ une fonction.

1.1 Fonctions en escalier

On rappelle la construction des *fonctions en escalier*, déjà rencontrées :

Définition 1.1.1. Une *subdivision* de $[a, b]$ est une famille (x_0, x_1, \dots, x_n) telle que :

$$a = x_0 < x_1 < \dots < x_n = b.$$

Exemple 1.1.2. La n -*subdivision régulière* de $I = [a, b]$ est la subdivision (x_0, x_1, \dots, x_n) où :

$$\forall k \in \llbracket 0, n \rrbracket, x_k = a + k \frac{b - a}{n}.$$

Définition 1.1.3. On dit que $f : I \rightarrow \mathbb{R}$ est *en escalier* s'il existe une subdivision (x_0, x_1, \dots, x_n) de I telle que, pour tout k dans $\llbracket 0, n - 1 \rrbracket$, f est constante sur $]x_k, x_{k+1}[$. On dit que la subdivision (x_0, x_1, \dots, x_n) est *adaptée* à f .

Exemple 1.1.4. La fonction partie entière est en escalier sur \mathbb{R} .

Définition 1.1.5. Soit $f : I \rightarrow \mathbb{R}$ en escalier, de subdivision adaptée (x_0, x_1, \dots, x_n) . On note, pour tout k dans $\llbracket 0, n - 1 \rrbracket$, y_k la valeur prise par f sur $]x_k, x_{k+1}[$.

L'*intégrale sur I de f* est alors le nombre, noté $\int_I f$ ou $\int_a^b f$, défini par :

$$\int_I f = \sum_{k=0}^{n-1} y_k (x_{k+1} - x_k).$$

Remarque 1.1.6. Il s'agit de l'*aire algébrique* de la partie du plan située entre le graphe de f et l'axe des abscisses.

Proposition 1.1.7. L'*intégrale sur I de f* ne dépend pas de la subdivision choisie (tant que celle-ci est adaptée à f).

1.2 Fonctions continues

On considère à partir de maintenant que la fonction f est continue sur I .

Théorème 1.2.1 (Riemann). *On note :*

$$\mathcal{E}_1 = \{g : I \rightarrow \mathbb{R} \mid g \text{ est en escalier et } g \leq f\} \quad \text{et} \quad \mathcal{E}_2 = \{h : I \rightarrow \mathbb{R} \mid h \text{ est en escalier et } h \geq f\},$$

et :

$$J_1 = \left\{ \int_I g \mid g \in \mathcal{E}_1 \right\} \quad \text{et} \quad J_2 = \left\{ \int_I h \mid h \in \mathcal{E}_2 \right\}.$$

Alors l'ensemble J_1 admet une borne supérieure, l'ensemble J_2 admet une borne inférieure, et $\sup J_1 = \inf J_2$.

Démonstration. D'après le théorème de Weierstrass, f est majorée par un réel M et minorée par un réel m sur I . J_1 est alors une partie non-vide de \mathbb{R} majorée par $M(b-a)$, donc admet une borne supérieure α_1 . De même, J_2 est une partie non-vide de \mathbb{R} minorée par $m(b-a)$, donc admet une borne inférieure α_2 .

On a alors : $\forall g \in \mathcal{E}_1, \forall h \in \mathcal{E}_2, \int_I g \leq \int_I h$. Donc, par passage à la borne supérieure dans $\mathcal{E}_1 : \forall h \in \mathcal{E}_2, \alpha_1 \leq \int_I h$. Donc, par passage à la borne inférieure dans $\mathcal{E}_2 : \alpha_1 \leq \alpha_2$.

Il reste à montrer que $\alpha_2 \leq \alpha_1$. On se restreint au cas où f est C^1 . Soient C une borne de f' et $n \in \mathbb{N}^*$, considérons la subdivision régulière de I en n intervalles. Pour tout $k \in \llbracket 0, n-1 \rrbracket$, notons m_k, M_k un minorant et un majorant de f sur $[x_k, x_{k+1}]$. D'après l'inégalité des accroissements finis : $M_k - m_k \leq C \frac{b-a}{n}$. Considérons les fonctions h_n égale à M_k et g_n égale à m_k sur $[x_k, x_{k+1}]$ pour tout $k \in \llbracket 0, n-1 \rrbracket$. On a donc :

$$\alpha_2 \leq \int_I h_n \leq \int_I g_n + C \frac{(b-a)^2}{n} \leq \alpha_1 + C \frac{(b-a)^2}{n}.$$

Par passage à la limite $n \rightarrow +\infty$, on a donc : $\alpha_2 \leq \alpha_1$. □

Définition 1.2.2. L'intégrale sur I de f , notée $\int_I f$ ou $\int_a^b f$, est la borne supérieure de l'ensemble J_1 défini ci-dessus (ou de manière équivalente, la borne inférieure de J_2).

Première conséquence de cette construction : les propriétés suivantes de l'intégrale, immédiates pour les fonctions en escalier, restent valables pour les fonctions continues :

Proposition 1.2.3 (Propriétés de l'intégrale). *Soient $f, g : I \rightarrow \mathbb{R}$ continues.*

1. *Linéarité* : $\forall (\lambda, \mu) \in \mathbb{R}^2, \int_I (\lambda f + \mu g) = \lambda \int_I f + \mu \int_I g$,
2. *Croissance* : Si $f \leq g$, alors $\int_I f \leq \int_I g$,
3. *Positivité* : Si $f \geq 0$, alors $\int_I f \geq 0$, et $= 0$ si et seulement si $f = 0$,
4. *Relation de Chasles* : $\forall c \in [a, b], \int_a^b f = \int_a^c f + \int_c^b f$,
5. *Inégalité triangulaire* : $\left| \int_I f \right| \leq \int_I |f|$.

Remarque 1.2.4. Dans l'équivalence du 3. ci-dessus, la continuité et la positivité de f sont des hypothèses essentielles.

Deuxième conséquence de cette construction, avec les notations de la démonstration ci-dessus :

Corollaire 1.2.5. $\int_I f = \lim_{n \rightarrow +\infty} \int_I g_n = \lim_{n \rightarrow +\infty} \int_I h_n$.

Et plus généralement :

2 Sommes de Riemann

Soit $f : I \rightarrow \mathbb{R}$ continue, et soit $n \in \mathbb{N}^*$.

Définition 2.0.1. Soit (x_0, x_1, \dots, x_n) la n -subdivision régulière de I .

La n ème somme de Riemann à gauche associée à f est le nombre :

$$G_n(f) = \frac{b-a}{n} \sum_{k=0}^{n-1} f(x_k),$$

et la n ème somme de Riemann à droite associée à f est le nombre :

$$D_n(f) = \frac{b-a}{n} \sum_{k=1}^n f(x_k).$$

Théorème 2.0.2 (Riemann). *Les suites $(G_n(f))$ et $(D_n(f))$ sont convergentes vers $\int_I f$.*

Démonstration. C'est une conséquence directe du corollaire ci-dessus. En voici une preuve indépendante : On se restreint au cas où f est de classe C^1 . D'après le théorème de Weierstrass, il existe un minorant m et un majorant M de f' . D'après l'inégalité des accroissements finis, on a alors :

$$\forall t \in [x_k, x_{k+1}], \quad m(t - x_k) \leq f(t) - f(x_k) \leq M(t - x_k),$$

donc :

$$m \frac{(b-a)^2}{2n^2} \leq \int_{x_k}^{x_{k+1}} f - \frac{b-a}{n} f(x_k) \leq M \frac{(b-a)^2}{2n^2},$$

d'où :

$$m \frac{(b-a)^2}{2n} \leq \int_I f - G_n(f) \leq M \frac{(b-a)^2}{2n}.$$

□

Exemple 2.0.3. $\sum_{k=1}^n \frac{1}{n+k}$, $\sum_{k=1}^n \frac{n}{n^2+k^2}$, $\sqrt[n]{\prod_{k=1}^n \left(1 + \frac{k}{n}\right)}$. Inversement, $\int_0^1 e^t dt$.

3 Théorème fondamental de l'analyse

Définition 3.0.1. La valeur moyenne de f est le réel $\mu = \frac{1}{b-a} \int_a^b f$.

Proposition 3.0.2. $\exists c \in [a, b]$, $f(c) = \mu$.

Démonstration. D'après le théorème de Weierstrass, il existe m et M réels tels que : $\forall x \in [a, b]$, $m \leq f(x) \leq M$, donc : $m \leq \mu \leq M$. Donc $\mu \in [m, M] = f([a, b])$. □

On déduit de cette simple propriété le *théorème fondamental de l'analyse*, déjà rencontré :

Théorème 3.0.3 (Théorème fondamental de l'analyse). *La fonction $\Phi : x \mapsto \int_a^x f$ est une primitive de f .*

Démonstration. Soit $x_0 \in [a, b]$, alors : $\frac{\Phi(x) - \Phi(x_0)}{x - x_0} = \mu_{[x_0, x]} = f(c_x) \xrightarrow{x \rightarrow x_0} f(x_0)$. □

Remarque 3.0.4. On rappelle que les primitives de f sont exactement les fonctions $\Phi + c$, où $c \in \mathbb{R}$.

Les conséquences, nombreuses, de ce théorème sur le calcul d'intégrales ont été exposées dans le chapitre sur les primitives. On rajoutera simplement ici la formule ci-après, obtenue par intégrations par parties successives :

Théorème 3.0.5 (Formule de Taylor avec reste intégral). *Soit $f : I \rightarrow \mathbb{R}$ de classe C^{n+1} . Alors :*

$$\forall x \in I, \quad f(x) = f(a) + f'(a)(x-a) + \cdots + f^{(n)}(a) \frac{(x-a)^n}{n!} + \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt.$$

Démonstration. On raisonne par récurrence. Pour $n = 0$, soit f de classe C^1 . L'assertion s'écrit :

$$\forall x \in I, \quad f(x) = f(a) + \int_a^x f'(t) dt,$$

ce qui est vrai d'après le théorème fondamental de l'analyse.

Soit $n \in \mathbb{N}^*$, supposons l'assertion vraie au rang $n - 1$. Soit f de classe C^{n+1} . Alors, par intégration par parties :

$$\int_a^x f^{(n)}(t) \frac{(x-t)^{n-1}}{(n-1)!} dt = \left[-f^{(n)}(t) \frac{(x-t)^n}{n!} \right]_a^x + \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt = f^{(n)}(a) \frac{(x-a)^n}{n!} + \int_a^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt,$$

donc l'assertion est vraie au rang n . Par récurrence, l'assertion est vraie pour tout $n \in \mathbb{N}$. \square

4 Intégration des fonctions complexes

Comme d'habitude, on peut étendre les raisonnements précédents aux fonctions continues $f : I \rightarrow \mathbb{C}$, jusqu'à l'obstacle de l'absence de relation d'ordre sur \mathbb{C} .

Définition 4.0.1. L'intégrale sur I de $f : I \rightarrow \mathbb{C}$ est définie par :

$$\int_I f = \int_I \operatorname{Re}(f) + i \int_I \operatorname{Im}(f).$$

Exemple 4.0.2. $\int_0^\pi e^{it} dt$

La linéarité et la relation de Chasles restent vraies, tandis que la croissance et la positivité n'ont plus de sens. Notons que l'inégalité triangulaire est conservée !

Démonstration. Notons $f = re^{i\theta}$ et $\int_a^b f = \rho e^{i\alpha}$. Alors :

$$\left| \int_a^b f \right| = \rho = \int_a^b \underbrace{re^{i(\theta-\alpha)}}_{A+iB} = \int_a^b A + i \underbrace{\int_a^b B}_{=0 \text{ car } \rho \in \mathbb{R}} = \int_a^b r \cos(\theta - \alpha) \leq \int_a^b r = \int_a^b |f|.$$

\square

Appliquée à f' , on en déduit l'inégalité des accroissements finis pour les fonctions complexes.

Chapitre 22

Variables aléatoires

Dans notre premier chapitre sur les probabilités, la notion centrale était celle d'événement. On la remplace ici par la notion de *variable aléatoire*, plus simple à manipuler et qu'il est possible d'étudier quantitativement.

1 Variables aléatoires

Dans tout le chapitre, sauf mention contraire, on considère une expérience aléatoire d'univers Ω fini.

1.1 Définition

Définition 1.1.1. Une *variable aléatoire* X est une fonction $X : \Omega \rightarrow E$ où E est un ensemble (en général, $E = \mathbb{R}$).

On note $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$ l'ensemble image de X . C'est *l'ensemble des valeurs prises par X* .

Remarque 1.1.2. La notion de variable aléatoire a donc pour objectif de *quantifier* les résultats de l'expérience aléatoire considérée. Ainsi, dans le cas d'un lancer de dé, si X désigne le numéro obtenu (c'est-à-dire, formellement, que X est la fonction qui à chaque issue associe le nombre correspondant), on a $X(\Omega) = \llbracket 1, 6 \rrbracket$. Pour un lancer de pile ou face, on pose généralement $X(\Omega) = \llbracket 0, 1 \rrbracket$.

La notion de variable aléatoire remplace en partie celle d'événement, avec le lien suivant :

Définition 1.1.3. Soit X une variable aléatoire sur Ω , et soit x dans $X(\Omega)$. On note « $X = x$ » l'événement :

$$\text{« } X = x \text{ »} = X^{-1}(x) = \{\omega \in \Omega \mid X(\omega) = x\}.$$

On notera qu'il s'agit bien d'une partie de l'univers Ω .

On note de même « $X \leq x$ » l'événement :

$$\text{« } X \leq x \text{ »} = X^{-1}(\llbracket -\infty, x \rrbracket) = \{\omega \in \Omega \mid X(\omega) \leq x\},$$

et plus généralement, pour toute partie I de \mathbb{R} , on note « $X \in I$ » l'événement :

$$\text{« } X \in I \text{ »} = X^{-1}(I) = \{\omega \in \Omega \mid X(\omega) \in I\}.$$

Définition 1.1.4. Soient (Ω, \mathbb{P}) un univers probabilisé et X une variable aléatoire sur Ω . Pour tout x dans $X(\Omega)$, on note $\mathbb{P}(X = x)$ la probabilité de l'événement « $X = x$ », et de même pour les autres notations introduites ci-dessus.

Proposition 1.1.5. *La famille $(\text{« } X = x \text{ »})_{x \in X(\Omega)}$ est un système complet d'événements.*

Définition 1.1.6. La *loi de probabilité*, ou plus simplement la *loi*, de la variable aléatoire X , est la fonction

$$\mathbb{P}_X : \begin{cases} X(\Omega) & \rightarrow [0, 1] \\ x & \mapsto \mathbb{P}(X = x) \end{cases} .$$

Exemple 1.1.7. On tire deux boules d'une urne contenant 3 boules bleues, 1 boule rouge et 6 boules jaunes, et on considère la variable X donnant le nombre de boules bleues obtenues. On a alors $X(\Omega) = \{0, 1, 2\}$ et la loi de X est donnée par :

$$\mathbb{P}(X = 0) = \frac{\binom{7}{2}}{\binom{10}{2}} = \frac{7}{15}, \quad \mathbb{P}(X = 1) = \frac{\binom{3}{1} \binom{7}{1}}{\binom{10}{2}} = \frac{7}{15}, \quad \mathbb{P}(X = 2) = \frac{\binom{3}{2}}{\binom{10}{2}} = \frac{1}{15}.$$

Proposition 1.1.8. *Par construction, on a :*

$$\sum_{x \in X(\Omega)} \mathbb{P}_X(x) = 1.$$

Définition 1.1.9. La *fonction de répartition* de la variable aléatoire X est la fonction

$$F_X : \begin{cases} X(\Omega) & \rightarrow [0, 1] \\ x & \mapsto \mathbb{P}(X \leq x) \end{cases} .$$

Exemple 1.1.10. Lors d'un jet de deux dés, on note X le plus grand numéro obtenu. Il est dans ce cas plus simple de déterminer la fonction de répartition de X , puis d'en déduire la loi de X .

Proposition 1.1.11. *La fonction de répartition F_X est une fonction croissante.*

1.2 Espérance, variance, écart-type

On a vu qu'une variable aléatoire X se caractérise par l'ensemble $X(\Omega)$ des valeurs prises et par sa loi de probabilité \mathbb{P}_X . Des *indicateurs* peuvent être associés à cet ensemble de données, à commencer par la *valeur moyenne* prise par X :

Définition 1.2.1. L'*espérance* de X , ou *moyenne pondérée* de X , est le nombre réel $E(X)$ défini par :

$$E(X) = \sum_{x \in X(\Omega)} x \cdot \mathbb{P}(X = x).$$

Exemple 1.2.2. L'espérance du numéro X obtenu lors d'un lancer de dé à 6 faces équilibré est $E(X) = 3,5$.

Remarque 1.2.3. L'espérance de X donne une indication sur l'endroit, sur la droite réelle, autour duquel se situent les valeurs prises par X : c'est un *paramètre de position* de la variable X . Remarquablement, une première estimation de la *dispersion* des valeurs de X est donnée par :

Théorème 1.2.4 (Inégalité de Markov). *Soit X une variable aléatoire à valeurs positives, et soit $a \in \mathbb{R}_+^*$. Alors :*

$$\mathbb{P}(X \geq a) \leq \frac{E(X)}{a}.$$

Démonstration. On a : $E(X) \geq \sum_{x \geq a} x \cdot \mathbb{P}(X = x) \geq a \cdot \mathbb{P}(X \geq a)$. □

Remarque 1.2.5. Cette formule n'a bien sûr d'intérêt que pour $a > E(X)$.

On peut affiner la connaissance de X en se demandant à quel point ces valeurs sont *étagées* autour de l'espérance. C'est l'objectif de la *variance* :

Définition 1.2.6. La *variance* de X est le nombre réel positif $V(X)$ défini par :

$$V(X) = E((X - E(X))^2) = \sum_{x \in X(\Omega)} (x - E(X))^2 \cdot \mathbb{P}(X = x).$$

Remarque 1.2.7. La variance est donc la *moyenne des carrés des écarts à la moyenne* de la variable X : c'est un *paramètre de dispersion*. Par construction, elle est toujours positive.

L'espérance et la variance sont appelés des *moments* de la variable X .

Le paramètre homogène à la variable correspondant à la variance est alors sa racine, l'*écart-type* :

Définition 1.2.8. L'*écart-type* de X est le nombre réel positif $\sigma(X)$ défini par :

$$\sigma(X) = \sqrt{V(X)}.$$

Remarque 1.2.9. Comme on l'a dit, cette notion détermine la dispersion des valeurs prises par X : la plupart des valeurs (pondérées) prises par X est ainsi à une distance inférieure à l'écart-type $\sigma(X)$ de la valeur moyenne $E(X)$. Pour une loi normale, on a typiquement :

$$\mathbb{P}(|X - E(X)| \leq \sigma(X)) \simeq 0,6 \quad \text{et} \quad \mathbb{P}(|X - E(X)| \leq 2\sigma(X)) \simeq 0,95.$$

Plus généralement, on a :

Théorème 1.2.10 (Inégalité de Bienaymé-Tchebychev). *Soit X une variable aléatoire, et soit $a \in \mathbb{R}_+^*$. Alors :*

$$\mathbb{P}(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}.$$

Démonstration. On applique l'inégalité de Markov à $Y = (X - E(X))^2$ et a^2 . □

Remarque 1.2.11. Cette formule n'a bien sûr d'intérêt que pour $a > \sigma(X)$.

1.3 Opérations sur les variables aléatoires

L'intérêt de la notion de variable aléatoire réside également dans sa maniabilité, en particulier au niveau du *changement de variable* :

Proposition 1.3.1. *Soient $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire et $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction. Alors $Y = f \circ X$ (on note en général $Y = f(X)$) est une variable aléatoire sur Ω , et sa loi de probabilité est donnée par :*

$$\forall y \in Y(\Omega), \quad \mathbb{P}_Y(y) = \mathbb{P}_X(f^{-1}(\{y\})) = \sum_{x \in X(\Omega) \mid f(x)=y} \mathbb{P}_X(x).$$

L'espérance de la nouvelle variable Y est alors donnée par le :

Théorème 1.3.2 (Théorème de transfert). *Avec les notations de la propriété ci-dessus :*

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x) \cdot \mathbb{P}(X = x).$$

Dans le cas particulier où $Y = (X - E(X))^2$, on obtient ainsi la formule :

Proposition 1.3.3. $V(X) = E(X^2) - E(X)^2$.

De même, on déduit du théorème de transfert les règles de transformation suivantes :

Proposition 1.3.4. *Soient $a, b \in \mathbb{R}$. Alors :*

$$E(aX + b) = aE(X) + b, \quad V(aX + b) = a^2V(X), \quad \sigma(aX + b) = |a|\sigma(X).$$

Remarque 1.3.5. On voit en particulier que l'espérance est une application *linéaire* en la variable aléatoire. Étant donné une variable X , la variable $Y = X - E(X)$ est donc d'espérance nulle : c'est la *variable centrée* associée à X . De même, la variable $X^* = \frac{X - E(X)}{\sigma(X)}$ a une espérance nulle et une variance de 1 : c'est la *variable centrée réduite* associée à X .

2 Lois usuelles

Les trois lois suivantes, souvent rencontrées, sont à connaître ; les formules donnant leur espérance et leur variance peuvent être directement utilisées.

2.1 Loi uniforme

Soient $a, b \in \mathbb{Z}$ avec $a \leq b$. La *loi uniforme sur $\llbracket a, b \rrbracket$* est celle du résultat d'un lancer de dé équilibré dont les faces sont numérotées de a à b :

Définition 2.1.1. On dit qu'une variable aléatoire X suit la *loi uniforme* sur $\llbracket a, b \rrbracket$, et on note $X \hookrightarrow \mathcal{U}(a, b)$, si $X(\Omega) = \llbracket a, b \rrbracket$ et si :

$$\forall (k, l) \in X(\Omega)^2, \quad \mathbb{P}(X = k) = \mathbb{P}(X = l),$$

c'est-à-dire :

$$\forall k \in X(\Omega), \quad \mathbb{P}(X = k) = \frac{1}{b - a + 1}.$$

Proposition 2.1.2. Soit X telle que $X \hookrightarrow \mathcal{U}(a, b)$. Alors :

$$E(X) = \frac{a + b}{2} \quad \text{et} \quad V(X) = \frac{n^2 - 1}{12},$$

où $n = b - a + 1$.

2.2 Loi de Bernoulli

Soit $p \in [0, 1]$. La *loi de Bernoulli de paramètre p* est la loi indicatrice d'un événement de probabilité p :

Définition 2.2.1. On dit qu'une variable aléatoire X suit la *loi de Bernoulli de paramètre p* , et on note $X \hookrightarrow \mathcal{B}(p)$, si $X(\Omega) = \{0, 1\}$ et si :

$$\mathbb{P}(X = 1) = p \quad \text{et} \quad \mathbb{P}(X = 0) = 1 - p.$$

On note en général $q = 1 - p$.

Proposition 2.2.2. Soit X telle que $X \hookrightarrow \mathcal{B}(p)$. Alors :

$$E(X) = p \quad \text{et} \quad V(X) = pq.$$

2.3 Loi binomiale

Soient $p \in [0, 1]$ et $n \in \mathbb{N}^*$. La *loi binomiale de paramètres n et p* est la loi du nombre de succès parmi n expériences de Bernoulli de paramètre p , indépendantes les unes des autres :

Définition 2.3.1. On dit qu'une variable aléatoire X suit la *loi binomiale de paramètres n et p* , et on note $X \hookrightarrow \mathcal{B}(n, p)$, si $X(\Omega) = \llbracket 0, n \rrbracket$ et si :

$$\forall k \in \llbracket 0, n \rrbracket, \quad \mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k},$$

où $q = 1 - p$.

Proposition 2.3.2. Soit X telle que $X \hookrightarrow \mathcal{B}(n, p)$. Alors :

$$E(X) = np \quad \text{et} \quad V(X) = npq.$$

3 Couples de variables aléatoires

3.1 Propriétés

Définition 3.1.1. Soient X et Y deux variables aléatoires sur le même univers Ω . La *loi conjointe* du couple (X, Y) est l'application

$$\mathbb{P}_{X,Y} : \begin{cases} X(\Omega) \times Y(\Omega) & \rightarrow [0, 1] \\ (x, y) & \mapsto p_{xy} = \mathbb{P}(X = x \cap Y = y) \end{cases}.$$

Exemple 3.1.2. Lors d'un lancer de deux dés, on note X le plus petit numéro obtenu et Y le plus grand numéro obtenu.

Proposition 3.1.3. *Les lois des variables X et Y peuvent être déduites de la loi conjointe du couple par :*

$$\forall x \in X(\Omega), \quad \mathbb{P}_X(x) = \sum_{y \in Y(\Omega)} \mathbb{P}_{X,Y}(x, y),$$

et de même :

$$\forall y \in Y(\Omega), \quad \mathbb{P}_Y(y) = \sum_{x \in X(\Omega)} \mathbb{P}_{X,Y}(x, y).$$

Ce sont les lois marginales du couple (X, Y) .

De même que précédemment, on peut pratiquer des opérations sur un couple de variables aléatoires :

Théorème 3.1.4 (Théorème de transfert généralisé). *Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Alors :*

$$E(f(X, Y)) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} f(x, y) \cdot \mathbb{P}(X = x \cap Y = y).$$

Notamment, on retrouve la linéarité de l'espérance déjà mentionnée :

Proposition 3.1.5. *Soient $a, b \in \mathbb{R}$. Alors $E(aX + bY) = aE(X) + bE(Y)$.*

Remarque 3.1.6. La variance obéit à une règle de transformation plus compliquée, dont on verra qu'elle dépend de l'*indépendance* ou non des variables considérées.

Enfin, comme pour les événements, on peut définir une notion de probabilité *conditionnelle* :

Définition 3.1.7. La *loi conditionnelle* de X sachant Y est donnée par :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}_{Y=y}(X = x) = \frac{\mathbb{P}(X = x \cap Y = y)}{\mathbb{P}(Y = y)}.$$

3.2 Indépendance

Définition 3.2.1. Deux variables aléatoires X et Y sont dites indépendantes si :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}_{Y=y}(X = x) = \mathbb{P}(X = x),$$

c'est-à-dire :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y).$$

On note alors : $X \perp\!\!\!\perp Y$.

Cette notion d'*indépendance* se rattache à celle vue pour les événements grâce à la propriété suivante :

Proposition 3.2.2. *Deux variables aléatoires X et Y sont indépendantes si et seulement si, pour tous $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements « $X \in A$ » et « $Y \in B$ » sont indépendants.*

Proposition 3.2.3. Soient X et Y deux variables aléatoires et $f, g : \mathbb{R} \rightarrow \mathbb{R}$ deux fonctions. Si les variables X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ le sont également.

Proposition 3.2.4. Si X et Y sont indépendantes, alors :

$$E(XY) = E(X)E(Y) \quad \text{et} \quad V(X + Y) = V(X) + V(Y).$$

Remarque 3.2.5. En général, la variance n'est donc **pas** linéaire, et on a :

$$V(X + Y) = V(X) + 2\text{Cov}(X, Y) + V(Y),$$

où $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$ est la *covariance* des variables X et Y .

Comme pour les événements, la notion d'indépendance se généralise à une famille de variables aléatoires :

Définition 3.2.6. Soient $n \in \mathbb{N}^*$ et X_1, \dots, X_n des variables aléatoires sur un même univers Ω . Les variables X_1, \dots, X_n sont *mutuellement indépendantes* si :

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega), \quad \mathbb{P}\left(\bigcap_{i=1}^n X_i = x_i\right) = \prod_{i=1}^n \mathbb{P}(X_i = x_i).$$

Remarque 3.2.7. Si X_1, \dots, X_n sont mutuellement indépendantes, alors, pour tout $(A_1, \dots, A_n) \subset X_1(\Omega) \times \dots \times X_n(\Omega)$, les événements $(\langle X_i \in A_i \rangle)_{1 \leq i \leq n}$ sont mutuellement indépendants.

En particulier, on retrouve le lien entre loi de Bernoulli et loi binomiale déjà mentionné :

Proposition 3.2.8. Soient $p \in [0, 1]$, $n \in \mathbb{N}^*$ et $X_1, \dots, X_n \hookrightarrow \mathcal{B}(p)$ mutuellement indépendantes. Alors :

$$X_1 + \dots + X_n \hookrightarrow \mathcal{B}(n, p).$$

Chapitre 23

Séries numériques

Parmi les différents types de suites rencontrés cette année, on a vu des suites dont le terme général est une somme. Ce sont les *séries*, dont nous développons ici le formalisme et quelques méthodes d'étude spécifiques.

1 Généralités

1.1 Définition

Soit $(u_n)_{n \geq 0}$ une suite réelle.

Définition 1.1.1. La *série* de terme général u_n est la suite de terme général :

$$S_n = \sum_{k=0}^n u_k = u_0 + u_1 + \cdots + u_n.$$

On note alors $\sum_{n \geq 0} u_n = (S_n)_{n \geq 0}$. Pour $n \in \mathbb{N}$, le nombre S_n est appelé *somme partielle* au rang n de la série.

Remarque 1.1.2. Cette notation se généralise naturellement au cas d'une suite $(u_n)_{n \geq n_0}$, quel que soit $n_0 \in \mathbb{N}$.

Définition 1.1.3. On dit que la série $\sum_{n \geq 0} u_n$ converge vers un réel S si la suite $(S_n)_{n \geq 0}$ converge vers S . On note alors :

$$\sum_{n=0}^{+\infty} u_n = S.$$

Remarque 1.1.4. On remarquera que la représentation ci-dessus de la limite comme une somme infinie n'est qu'une **notation** ; on ne sait a priori pas manipuler une telle somme.

Définition 1.1.5. Soit $\sum_{n \geq 0} u_n$ une série convergente vers $S \in \mathbb{R}$, et soit $(S_n)_{n \geq 0}$ la suite de ses sommes partielles.

Le *reste* au rang $n \in \mathbb{N}$ de la série est le nombre R_n défini par :

$$R_n = S - S_n = \sum_{k=n+1}^{+\infty} u_k.$$

Proposition 1.1.6. Soit $\sum_{n \geq 0} u_n$ une série convergente. Alors la suite $(R_n)_{n \geq 0}$ de ses restes converge vers 0.

Exemple 1.1.7. $\sum_{n \geq 1} \frac{1}{n(n+1)}$.

Remarque 1.1.8. On prendra garde à la terminologie : les trois notations $\sum_{n \geq 0} u_n$, $\sum_{k=0}^n u_k$ et $\sum_{n=0}^{+\infty} u_n$ désignent des objets complètement différents ! Ce sont respectivement une série (c'est-à-dire une suite), une somme partielle (c'est-à-dire une somme finie de termes) et une somme de série (c'est-à-dire une limite).

1.2 Propriétés

Proposition 1.2.1 (Linéarité de la limite). *Soient deux séries convergentes de termes généraux respectifs u_n et v_n , et soient λ et μ des réels. Alors la série de terme général $\lambda u_n + \mu v_n$ est convergente, et on a :*

$$\sum_{n=0}^{+\infty} \lambda u_n + \mu v_n = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n.$$

Un premier résultat fondamental d'étude des séries est le suivant :

Proposition 1.2.2. *Si la série $\sum_{n \geq 0} u_n$ converge, alors la suite $(u_n)_{n \geq 0}$ converge vers 0.*

Remarque 1.2.3. Attention, la réciproque de cette propriété est **fausse** : ainsi, la suite de terme général $\frac{1}{n}$ converge vers 0, mais comme on va le voir, la série correspondante (appelée *série harmonique*) est divergente.

Définition 1.2.4. Si la suite de terme général u_n ne converge pas vers 0, on dit que la série $\sum_{n \geq 0} u_n$ *diverge grossièrement*.

Exemple 1.2.5. $\sum_{n \geq 0} \cos n.$

1.3 Cas particuliers

Définition 1.3.1. Une *série télescopique* est une série dont le terme général est de la forme $u_n = a_{n+1} - a_n$.

Proposition 1.3.2. *La série $\sum_{n \geq 0} a_{n+1} - a_n$ converge si et seulement si la suite $(a_n)_{n \geq 0}$ converge.*

Exemple 1.3.3. $\sum_{n \geq 1} \ln \left(1 + \frac{1}{n} \right).$

Définition 1.3.4. Une *série géométrique* est une série dont le terme général est de la forme $u_n = \lambda \cdot q^n$, où $\lambda, q \in \mathbb{R}$.

Théorème 1.3.5. *Avec les notations ci-dessus, une série géométrique est convergente si et seulement si $|q| < 1$. On a alors :*

$$\sum_{n=0}^{+\infty} u_n = \frac{\lambda}{1 - q}.$$

2 Séries à termes positifs

Dans le cas particulier où tous les termes de la série sont positifs, on dispose de résultats spécifiques. Dans cette partie, on se place donc dans le cas où :

$$\forall n \geq 0, \quad u_n \geq 0.$$

2.1 Comparaison entre séries

La première observation est immédiate, mais va guider les résultats suivants :

Proposition 2.1.1. Soit $\sum_{n \geq 0} u_n$ une série à termes positifs. Alors la suite $(S_n)_{n \geq 0}$ de ses sommes partielles est croissante ; de plus, si la série converge, alors la suite $(R_n)_{n \geq 0}$ de ses restes est décroissante (vers 0).

Par conséquent :

Théorème 2.1.2. Une série $\sum_{n \geq 0} u_n$ à termes positifs est convergente si et seulement si la suite $(S_n)_{n \geq 0}$ de ses sommes partielles est majorée.

Exemple 2.1.3. La série harmonique $\sum_{n \geq 1} \frac{1}{n}$ diverge.

On déduit de ce résultat le critère de convergence suivant :

Théorème 2.1.4 (Théorème de comparaison). Soient $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ deux séries à termes positifs telles que :

$$\forall n \geq 0, 0 \leq u_n \leq v_n.$$

Alors :

1. Si la série $\sum_{n \geq 0} v_n$ converge, alors la série $\sum_{n \geq 0} u_n$ converge, et on a :

$$\sum_{n=0}^{+\infty} u_n \leq \sum_{n=0}^{+\infty} v_n.$$

2. Si la série $\sum_{n \geq 0} u_n$ diverge, alors la série $\sum_{n \geq 0} v_n$ diverge.

Ce critère, spécifique aux séries à termes positifs, peut être affiné comme suit :

Théorème 2.1.5 (Théorème d'équivalence). Soient $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ deux séries à termes positifs telles que :

$$u_n \underset{n \rightarrow +\infty}{\sim} v_n.$$

Alors la série $\sum_{n \geq 0} u_n$ converge si et seulement si la série $\sum_{n \geq 0} v_n$ converge.

Exemple 2.1.6.

- La série $\sum_{n \geq 1} \frac{1}{n^2}$ converge.
- La série $\sum_{n \geq 0} \frac{n}{n^2 + 1}$ diverge.

Remarque 2.1.7. Comme le théorème de comparaison, le théorème d'équivalence n'est valable que pour les séries à termes positifs. Ainsi, les termes généraux des séries $\sum_{n \geq 1} \frac{(-1)^n}{\sqrt{n}}$ et $\sum_{n \geq 1} \frac{(-1)^n}{\sqrt{n}} + \frac{1}{n}$ sont équivalents ; néanmoins, la première série est convergente, tandis que la seconde est divergente.

2.2 Comparaison série-intégrale

Dans l'esprit de notre étude des sommes de Riemann, on peut établir le résultat suivant :

Théorème 2.2.1 (Théorème de comparaison série-intégrale). *Soit $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$ une fonction décroissante positive. Notons : $\forall n \in \mathbb{N}^*, u_n = f(n)$. Alors, pour tout $n \in \mathbb{N}^*$:*

$$\int_1^{n+1} f \leq \sum_{k=1}^n u_k \leq u_1 + \int_1^n f.$$

Exemple 2.2.2.

- $\sum_{n \geq 1} \frac{1}{n\sqrt{n}}$ converge,
- $\sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln n$.

2.3 Séries de Riemann

Le théorème de comparaison série-intégrale est spécifiquement taillé pour établir le critère de convergence ci-dessous, relatif aux *séries de Riemann*. Utilisé avec le théorème d'équivalence par exemple, ce critère permet d'établir la convergence ou la divergence de nombreuses suites :

Définition 2.3.1. Soit $s \in \mathbb{R}$. La *série de Riemann* de paramètre s est la série $\sum_{n \geq 1} \frac{1}{n^s}$.

Théorème 2.3.2. *La série de Riemann de paramètre s converge si et seulement si $s > 1$.*

Exemple 2.3.3. La série $\sum_{n \geq 1} \frac{1}{n^2}$ converge, et $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Remarque 2.3.4. On peut ainsi considérer la fonction $\zeta : s \mapsto \sum_{n=1}^{+\infty} \frac{1}{n^s}$, définie sur $]1, +\infty[$. Cette fonction, appelée *fonction zêta de Riemann*, a des propriétés remarquables et est notamment liée à la répartition des nombres premiers. Les valeurs $\zeta(2p)$, pour $p \in \mathbb{N}^*$, sont connues explicitement depuis Euler. En revanche, les valeurs $\zeta(2p+1)$, pour $p \in \mathbb{N}^*$, restent très mal connues ; tout juste sait-on (Apéry, 1978) que $\zeta(3)$ est irrationnel.

2.4 Développement décimal d'un réel

Pour clore ce panorama des séries à termes positifs, mentionnons leur lien avec l'écriture décimale des nombres réels :

Proposition 2.4.1. *Soit $(a_n)_{n \geq 0}$ une suite telle que : $\forall n \geq 1, a_n \in \llbracket 0, 9 \rrbracket$. Alors la série $\sum_{n \geq 0} \frac{a_n}{10^n}$ converge.*

Définition 2.4.2. Avec les notations précédentes, notons x_a la somme de la série $\sum_{n \geq 0} \frac{a_n}{10^n}$. L'écriture $a_0, a_1 a_2 a_3 \dots$ est appelée *développement décimal* du réel x_a .

Exemple 2.4.3. $\frac{1}{3} = 0, 3333 \dots$

Le développement décimal d'un réel est unique, au cas suivant près :

Proposition 2.4.4. Si la suite $(a_n)_{n \geq 0}$ est stationnaire égale à 9 à partir d'un certain rang $n_0 \geq 1$, alors $x_a = x_b$, où $(b_n)_{n \geq 0}$ est la suite de terme général :

$$b_n = \begin{cases} a_n & \text{si } n < n_0 - 1 \\ a_{n_0-1} + 1 & \text{si } n = n_0 - 1 \\ 0 & \text{si } n \geq n_0 \end{cases}.$$

On dit alors que le développement stationnaire à 9 est *impropre*, et on lui préfère le second développement. Ainsi, on privilégie l'écriture 1 à l'écriture 0,99999....

Mentionnons enfin le résultat suivant :

Théorème 2.4.5. Un nombre est rationnel si et seulement si son développement décimal est périodique.

3 Convergence absolue

On revient au cas des séries de terme général de signe quelconque. La première tentation est de se ramener au cas, désormais bien connu, des séries à termes positifs.

Définition 3.0.1. Soit $\sum_{n \geq 0} u_n$ une série de terme général de signe quelconque. Cette série est dite *absolument convergente* si la série $\sum_{n \geq 0} |u_n|$ est convergente.

Stratégie semi-payante, puisque nous avons le résultat :

Théorème 3.0.2. Si la série $\sum_{n \geq 0} u_n$ converge absolument, alors cette série converge. On a alors, de plus :

$$\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n|.$$

Remarque 3.0.3. La réciproque de ce résultat est (absolument) **fausse** : comme la formule de Taylor avec reste intégral nous le montre, la série $\sum_{n \geq 1} \frac{(-1)^{n+1}}{n}$ converge, de somme $\ln 2$, mais ne converge pas absolument.

Ce théorème nous permet d'obtenir un critère de convergence de puissance comparable à celui du théorème d'équivalence, mais cette fois sans condition sur le signe du terme général :

Théorème 3.0.4. Soient $\sum_{n \geq 0} u_n$ et $\sum_{n \geq 0} v_n$ deux séries telles que $\sum_{n \geq 0} v_n$ est à termes positifs et :

$$u_n = \underset{n \rightarrow +\infty}{O}(v_n).$$

Si $\sum_{n \geq 0} v_n$ converge, alors $\sum_{n \geq 0} u_n$ converge.

Exemple 3.0.5. La suite de terme général $u_n = \frac{n^n e^{-n} \sqrt{n}}{n!}$ converge vers un réel non nul k . La formule de Wallis nous donne $k = \frac{1}{\sqrt{2\pi}}$, d'où la mirifique *formule de Stirling* :

$$n! \underset{n \rightarrow +\infty}{\sim} \sqrt{2\pi n} \left(\frac{n}{e} \right)^n.$$

Chapitre 24

Espaces préhilbertiens

On enrichit la structure d'espace vectoriel d'un nouvel outil : le *produit scalaire*, qui permet de quantifier la position relative de deux vecteurs. Dans tout le chapitre, sauf mention contraire, on considère un \mathbb{R} -espace vectoriel E .

1 Produit scalaire

1.1 Définition

Le produit scalaire couramment utilisé dans \mathbb{R}^2 est l'opération qui à deux vecteurs $u = (x_1, y_1)$ et $v = (x_2, y_2)$ associe le nombre :

$$\langle u, v \rangle = x_1 x_2 + y_1 y_2.$$

Cette opération sert notamment à vérifier l'orthogonalité de deux vecteurs, permet de calculer la norme d'un vecteur ou encore le cosinus de l'angle entre deux vecteurs. Il en est de même dans \mathbb{R}^3 en rajoutant une composante selon z au produit ci-dessus.

Plus généralement, nous appellerons produit scalaire toute opération qui vérifie les hypothèses suivantes :

Définition 1.1.1. Un *produit scalaire* φ sur E est une **forme bilinéaire symétrique définie positive** :

1. **forme** : φ est une application de $E \times E$ dans \mathbb{R} :

$$\varphi : \begin{cases} E \times E & \rightarrow \mathbb{R} \\ (u, v) & \mapsto \varphi(u, v) \end{cases},$$

2. **bilinéaire** : l'application φ est linéaire à gauche :

$$\forall (u_1, u_2, v) \in E^3, \forall (\lambda, \mu) \in \mathbb{R}^2, \quad \varphi(\lambda u_1 + \mu u_2, v) = \lambda \varphi(u_1, v) + \mu \varphi(u_2, v),$$

et linéaire à droite :

$$\forall (u, v_1, v_2) \in E^3, \forall (\lambda, \mu) \in \mathbb{R}^2, \quad \varphi(u, \lambda v_1 + \mu v_2) = \lambda \varphi(u, v_1) + \mu \varphi(u, v_2),$$

3. **symétrique** : $\forall (u, v) \in E^2, \varphi(u, v) = \varphi(v, u)$,

4. **définie positive** : $\forall u \in E, \varphi(u, u) \geq 0$, et $(\varphi(u, u) = 0 \Rightarrow u = 0_E)$.

Exemple 1.1.2.

- Dans $E = \mathbb{R}^n$, l'application φ définie pour $u = (a_1, \dots, a_n)$ et $v = (b_1, \dots, b_n) \in E$ par :

$$\varphi(u, v) = a_1 b_1 + \dots + a_n b_n,$$

- Dans $E = \mathbb{R}_n[X]$, l'application φ définie pour P et $Q \in E$ par :

$$\varphi(P, Q) = P(0)Q(0) + P(1)Q(1) + \cdots + P(n)Q(n),$$

- Dans $E = C^0([0, 1], \mathbb{R})$, l'application φ définie pour f et $g \in E$ par :

$$\varphi(f, g) = \int_0^1 f(t)g(t)dt,$$

- Dans $E = M_n(\mathbb{R})$, l'application φ définie pour A et $B \in E$ par :

$$\varphi(A, B) = \sum_{1 \leq i, j \leq n} a_{ij}b_{ij}.$$

Remarque 1.1.3. On notera en général $\varphi = \langle \cdot, \cdot \rangle$, c'est-à-dire : $\forall (u, v) \in E^2$, $\varphi(u, v) = \langle u, v \rangle$. On rencontre souvent d'autres notations plus ou moins similaires : $\langle \cdot | \cdot \rangle$ (très utilisée en physique quantique), (\cdot, \cdot) , $(\cdot | \cdot)$, etc.

Définition 1.1.4. Un espace vectoriel muni d'un produit scalaire est un *espace préhilbertien*.
Un espace vectoriel de dimension finie muni d'un produit scalaire est un *espace euclidien*.

1.2 Norme

L'hypothèse de définition de positivité du produit scalaire permet de l'utiliser pour définir une *norme* sur E , c'est-à-dire un outil de mesure de la *longueur* des vecteurs, et donc de la *distance* entre deux vecteurs :

Définition 1.2.1. Soit (E, φ) un espace préhilbertien. La *norme* associée à φ est l'application :

$$N : \begin{cases} E & \rightarrow \mathbb{R}_+ \\ u & \mapsto \sqrt{\varphi(u, u)} \end{cases} .$$

Remarque 1.2.2. De même que le produit scalaire, on notera en général $N = \|\cdot\|$, c'est-à-dire : $\forall u \in E$, $N(u) = \|u\|$, et donc : $\|u\| = \sqrt{\langle u, u \rangle}$.

Proposition 1.2.3. $\forall u, v \in E$, $\|u + v\|^2 = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2$.

Proposition 1.2.4. La norme N est :

1. absolument homogène : $\forall u \in E$, $\forall \lambda \in \mathbb{R}$, $\|\lambda u\| = |\lambda| \|u\|$,
2. séparante : $\forall u \in E$, $\|u\| = 0 \Rightarrow u = 0_E$,
3. sous-additive, c'est-à-dire que $\|\cdot\|$ satisfait l'inégalité triangulaire :

$$\forall (u, v) \in E^2, \|u + v\| \leq \|u\| + \|v\|.$$

Cette dernière propriété est une conséquence de la célèbre :

Proposition 1.2.5 (Inégalité de Cauchy-Schwarz). Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien. Alors :

$$\forall (u, v) \in E^2, |\langle u, v \rangle| \leq \|u\| \cdot \|v\|,$$

avec égalité si et seulement si u et v sont colinéaires.

Remarque 1.2.6. Toute application vérifiant ces trois propriétés est une *norme* sur E , qu'elle provienne d'un produit scalaire (on dit alors qu'elle est *euclidienne*) ou non. Ainsi, les trois applications suivantes sont des normes sur \mathbb{R}^2 , mais seule $\|\cdot\|_2$ provient d'un produit scalaire :

$$\forall u = (x, y) \in \mathbb{R}^2, \|u\|_1 = |x| + |y|, \|u\|_2 = \sqrt{x^2 + y^2}, \|u\|_\infty = \max(|x|, |y|).$$

Exemple 1.2.7. L'ensemble E des variables aléatoires centrées sur un univers Ω forme un espace vectoriel, sur lequel la covariance est un produit scalaire. La norme associée est l'écart-type.

On a vu que la norme provient du produit scalaire :

$$\forall u \in E, \|u\| = \sqrt{\langle u, u \rangle}.$$

Inversement, le produit scalaire peut être obtenu à partir de la norme grâce aux formules suivantes :

Proposition 1.2.8 (Identités de polarisation). *Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien. Alors :*

$$\forall (u, v) \in E^2, \quad \langle u, v \rangle = \frac{1}{2} (\|u + v\|^2 - \|u\|^2 - \|v\|^2) = \frac{1}{2} (\|u\|^2 + \|v\|^2 - \|u - v\|^2).$$

Mentionnons au passage que la réorganisation de ces identités fournit la très géométrique :

Proposition 1.2.9 (Identité du parallélogramme). $\forall (u, v) \in E^2, \|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2)$.

Remarque 1.2.10. D'après Cauchy-Schwarz : $\forall (u, v) \in E^2, -1 \leq \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|} \leq 1$. L'angle entre u et v est alors défini par :

$$\theta(u, v) = \arccos \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}.$$

Exemple 1.2.11. Avec le produit scalaire défini ci-dessus sur $E = C^0([0, 1], \mathbb{R})$: $\theta(t \mapsto t, t \mapsto t^4) = \frac{\pi}{6}$.

2 Orthogonalité

On se place dans cette partie dans un espace préhilbertien $(E, \langle \cdot, \cdot \rangle)$.

2.1 Vecteurs et sous-espace orthogonaux

Définition 2.1.1. Deux vecteurs u et v sont *orthogonaux* si leur produit scalaire est nul.

Exemple 2.1.2. Dans $E = C^0([0, 1], \mathbb{R})$, les fonctions $f : t \mapsto 1$ et $g : t \mapsto t - \frac{1}{2}$ sont orthogonales pour le produit scalaire défini plus haut.

Les formules précédemment établies donnent directement le célèbre :

Théorème 2.1.3 (Théorème de Pythagore). *Soient u et v deux vecteurs orthogonaux. Alors :*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

L'orthogonalité donne une nouvelle structure à notre espace vectoriel :

Définition 2.1.4. Soit F un sous-espace vectoriel de E . L'*orthogonal de F* , noté F^\perp , est l'ensemble des vecteurs orthogonaux à tous les vecteurs de F :

$$F^\perp = \{u \in E \mid \forall v \in F, \langle u, v \rangle = 0\}.$$

Proposition 2.1.5. *L'ensemble F^\perp est un sous-espace vectoriel de E .*

Proposition 2.1.6. *Les sous-espaces vectoriels F et F^\perp sont en somme directe.*

2.2 Familles orthogonales

On peut généraliser la notion d'orthogonalité à une famille de vecteurs :

Définition 2.2.1. Une famille (e_1, \dots, e_p) de vecteurs de E est *orthogonale* si :

$$\forall i \neq j \in \llbracket 1, p \rrbracket^2, \quad \langle e_i, e_j \rangle = 0.$$

Théorème 2.2.2 (Théorème de Pythagore généralisé). *Soit (e_1, \dots, e_p) une famille orthogonale de vecteurs de E . Alors :*

$$\|e_1 + \dots + e_p\|^2 = \|e_1\|^2 + \dots + \|e_p\|^2.$$

Proposition 2.2.3. *Une famille orthogonale ne contenant pas le vecteur nul est libre.*

Les familles orthogonales (ne contenant pas le vecteur nul) sont donc de bonnes candidates pour constituer des bases. On va de plus les *normaliser* :

Définition 2.2.4. Soit $u \in E$ non nul. La *normalisation* de u est le vecteur $\frac{u}{\|u\|}$.

Remarque 2.2.5. La normalisation de u est donc un vecteur de norme 1. On dit également que ce vecteur est *normé*.

Définition 2.2.6. Une *famille orthonormée* (e_1, \dots, e_p) de vecteurs de E est une famille orthogonale de vecteurs normés. On a ainsi :

$$\forall (i, j) \in \llbracket 1, p \rrbracket^2, \quad \langle e_i, e_j \rangle = \delta_{ij},$$

où $\delta_{ij} = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j \end{cases}$ est le *symbole de Kronecker*.

3 Espaces euclidiens

On se place dans cette partie dans un espace euclidien $(E, \langle \cdot, \cdot \rangle)$ de dimension n .

3.1 Bases orthonormées

D'après ce qui a été dit précédemment, on peut directement prendre comme définition :

Définition 3.1.1. Une *base orthonormée* de E est une famille orthonormée de n vecteurs de E .

C'est bien sûr une base de E . Ce type de base a de multiples avantages, à commencer par la simplicité de la décomposition des vecteurs :

Théorème 3.1.2. *Soit (e_1, \dots, e_n) une base orthonormée de E , et soit u dans E . Alors :*

$$u = \sum_{i=1}^n \langle u, e_i \rangle e_i.$$

Proposition 3.1.3. $\forall u \in E, \|u\|^2 = \sum_{i=1}^n \langle u, e_i \rangle^2$.

Remarque 3.1.4. Attention, ces formules sont spécifiques aux bases orthonormées !

Exemple 3.1.5. À la marge de notre cadre d'étude, considérons $E = C^0([0, 2\pi], \mathbb{C})$ muni du produit scalaire (en fait *hermitien*) :

$$\forall f, g \in E, \quad \langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} dt.$$

La famille $(t \mapsto e^{int})_{n \in \mathbb{Z}}$ est une base orthonormée de cet espace. Le théorème ci-dessus y porte le nom de *décomposition en série de Fourier*, et la proposition celui de *formule de Parseval*.

3.2 Orthonormalisation

L'intérêt de disposer d'une base orthonormée étant établi, se pose la question de la *construction* d'une telle base. On sait, depuis notre construction de la dimension, que E admet des bases, toutes de même cardinal n . Considérons donc une base quelconque (e_1, \dots, e_n) de E ; nous allons l'*orthonormaliser*.

Proposition 3.2.1 (Orthogonalisation). *On définit, par récurrence, la famille (e'_1, \dots, e'_n) par :*

$$\begin{aligned} e'_1 &= e_1, \\ \forall k \in \llbracket 2, n \rrbracket, \quad e'_k &= e_k - \sum_{i=1}^{k-1} \frac{\langle e_k, e'_i \rangle}{\|e'_i\|^2} e'_i. \end{aligned}$$

La famille (e'_1, \dots, e'_n) est alors une base orthogonale de E . De plus :

$$\forall k \in \llbracket 1, n \rrbracket, \quad \text{Vect}(e'_1, \dots, e'_k) = \text{Vect}(e_1, \dots, e_k).$$

Remarque 3.2.2. Après avoir calculé e'_k , on peut, pour simplifier les calculs suivants, multiplier celui-ci par un scalaire (non nul) arbitraire; seule la direction des vecteurs nous intéresse à ce stade.

Proposition 3.2.3 (Normalisation). *On définit la famille (e''_1, \dots, e''_n) par :*

$$\forall k \in \llbracket 1, n \rrbracket, \quad e''_k = \frac{e'_k}{\|e'_k\|}.$$

La famille (e''_1, \dots, e''_n) est alors une base orthonormée de E . De plus :

$$\forall k \in \llbracket 1, n \rrbracket, \quad \text{Vect}(e''_1, \dots, e''_k) = \text{Vect}(e_1, \dots, e_k).$$

Ce procédé porte le nom d'*orthonormalisation de Gram-Schmidt*.

Exemple 3.2.4. Orthonormalisation de la base $((1, 1, 1), (1, 1, 0), (1, 0, 0))$ de \mathbb{R}^3 .

3.3 Projecteurs orthogonaux

Soit F un sous-espace vectoriel de E . En considérant une base orthonormée de E adaptée à F , on a :

Proposition 3.3.1. *Les sous-espaces F et F^\perp sont supplémentaires dans E .*

On peut donc définir le *projecteur orthogonal* sur F comme suit :

Définition 3.3.2. *Le projecteur orthogonal sur F , noté $p_{F,\perp}$, est le projecteur sur F parallèlement à F^\perp .*

Proposition 3.3.3. *Soit (e_1, \dots, e_p) une base orthonormée de F . Soit $u \in E$. Alors :*

$$p_{F,\perp}(u) = \sum_{i=1}^p \langle u, e_i \rangle e_i.$$

Démonstration. C'est une application directe du théorème 3.1.2. □

Le vecteur $p_{F,\perp}(u)$, appelé *projeté* ou *projection* orthogonale de u sur F , a la propriété remarquable de permettre de déterminer la distance de u à F :

Définition 3.3.4. Soit u dans E . La *distance de u à F* est définie par :

$$d(u, F) = \inf_{v \in F} \|u - v\|.$$

D'après le théorème de Pythagore, on a alors :

Théorème 3.3.5. *Le projeté orthogonal de u sur F réalise la distance de u à F :*

$$d(u, F) = \|u - p_{F,\perp}(u)\| = \|p_{F^\perp,\perp}(u)\|.$$

Démonstration. Soit $v \in F$. Alors : $u - v = \underbrace{(u - p_{F,\perp}(u))}_{\in F^\perp} + \underbrace{(p_{F,\perp}(u) - v)}_{\in F}$, donc, d'après Pythagore :

$$\|u - v\|^2 = \|u - p_{F,\perp}(u)\|^2 + \|p_{F,\perp}(u) - v\|^2 \geq \|u - p_{F,\perp}(u)\|^2,$$

donc $\|u - v\| \geq \|u - p_{F,\perp}(u)\|$, avec égalité si et seulement si $v = p_{F,\perp}(u)$. \square

4 Représentation matricielle du produit scalaire

On a vu dans la partie précédente comment orthonormaliser une base pour un produit scalaire donné. Réciproquement, on peut se poser la question suivante : étant donné une base de E , existe-t-il un produit scalaire sur E pour lequel cette base est orthonormée ? La réponse est positive, et s'appuie sur le résultat suivant :

Proposition 4.0.1. *Considérons $E = \mathbb{R}^n$ muni du produit scalaire usuel $\langle \cdot, \cdot \rangle$. Notons β_0 la base canonique de \mathbb{R}^n , puis, pour $u \in E$, $U = M_{\beta_0}(u)$. Alors :*

$$\forall (u, v) \in E^2, \quad \langle u, v \rangle = U^\top V.$$

Cette formule se généralise :

Proposition 4.0.2. *Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien. Soit $\beta = (e_1, \dots, e_n)$ une base de E . Notons, pour $u \in E$, $U = M_\beta(u)$. Alors :*

$$\forall (u, v) \in E^2, \quad \langle u, v \rangle = U^\top A V,$$

où $A = (\langle e_i, e_j \rangle)_{1 \leq i, j \leq n} \in M_n(\mathbb{R})$ est la matrice de Gram de $\langle \cdot, \cdot \rangle$ dans β .

Démonstration. Notons $u = \sum_{i=1}^n u_i e_i$ et $v = \sum_{j=1}^n v_j e_j$, alors :

$$\langle u, v \rangle = \sum_{i=1}^n \sum_{j=1}^n u_i v_j \langle e_i, e_j \rangle = (u_1 \cdots u_n) A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

\square

Remarque 4.0.3. Avec les notations choisies pour la représentation matricielle, en notant $\varphi = \langle \cdot, \cdot \rangle$, $G_\beta^\beta(\varphi)$ sa matrice de Gram dans la base β , et $M^\beta(u) = (M_\beta(u))^\top$, on a :

$$\forall (u, v) \in E^2, \quad \langle u, v \rangle = M^\beta(u) G_\beta^\beta(\varphi) M_\beta(v).$$

La formule de changement de base pour la matrice de Gram est alors directement :

Proposition 4.0.4. $G_{\beta_2}^{\beta_2}(\varphi) = (P_{\beta_1}^{\beta_2})^\top G_{\beta_1}^{\beta_1}(\varphi) P_{\beta_1}^{\beta_2}$

Démonstration. On a, d'une part, $\langle u, v \rangle = M^{\beta_2}(u)G_{\beta_2}^{\beta_2}(\varphi)M_{\beta_2}(v)$, et d'autre part :

$$\begin{aligned}\langle u, v \rangle &= M^{\beta_1}(u)G_{\beta_1}^{\beta_1}(\varphi)M_{\beta_1}(v) \\ &= \left(P_{\beta_1}^{\beta_2}M_{\beta_2}(u) \right)^{\top} G_{\beta_1}^{\beta_1}(\varphi)P_{\beta_1}^{\beta_2}M_{\beta_2}(v) \\ &= M^{\beta_2}(u) \left(P_{\beta_1}^{\beta_2} \right)^{\top} G_{\beta_1}^{\beta_1}(\varphi)P_{\beta_1}^{\beta_2}M_{\beta_2}(v).\end{aligned}$$

□

On a alors tous les outils pour répondre à la question initiale. En effet, on a évidemment :

Proposition 4.0.5. *Une base β de E est orthonormée pour φ si et seulement si $G_{\beta}^{\beta}(\varphi) = I_n$.*

Donc, étant donné une base de référence β_0 de E (la base canonique lorsqu'elle est définie par exemple), on a donc :

Proposition 4.0.6. *Soit β une base de E . Cette base est orthonormée pour le produit scalaire φ de matrice de Gram $G_{\beta_0}^{\beta_0}(\varphi) = \left(P_{\beta}^{\beta_0} \right)^{\top} P_{\beta}^{\beta_0}$.*

Exemple 4.0.7. Considérons la base $\beta = ((1, 1), (0, 1))$ de \mathbb{R}^2 .

On a : $P_{\beta_0}^{\beta} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, donc $P_{\beta}^{\beta_0} = \left(P_{\beta_0}^{\beta} \right)^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, et donc : $A = \left(P_{\beta}^{\beta_0} \right)^{\top} P_{\beta}^{\beta_0} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$.

La base β est donc orthonormée pour le produit scalaire $\langle \cdot, \cdot \rangle$ défini par :

$$\forall u = (x_1, y_1), v = (x_2, y_2) \in \mathbb{R}^2, \quad \langle u, v \rangle = (x_1 \ y_1) \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = 2x_1x_2 - x_1y_2 - x_2y_1 + y_1y_2.$$

Chapitre 25

Fonctions de deux variables

Ce chapitre est une introduction à l'étude des fonctions de deux variables, qui seront étudiées plus systématiquement en deuxième année. Nous allons notamment étendre à ces fonctions les notions de continuité et de dérivabilité, définies précédemment pour les fonctions à une variables.

Dans tout le chapitre, $\|\cdot\|$ désigne la norme euclidienne sur \mathbb{R}^2 et $\langle \cdot, \cdot \rangle$ le produit scalaire associé.

1 Généralités

1.1 Exemples

Définition 1.1.1. On appelle *fonction de deux variables* une fonction du type :

$$f : \begin{array}{ccc} \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ (x, y) & \mapsto & f(x, y) \end{array} .$$

Exemples 1.1.2.

- $f : (x, y) \mapsto x^2 + y^2$,
- $f : (x, y) \mapsto x$,
- $f : (x, y) \mapsto \frac{1}{\sqrt{x - y}}$.

Remarque 1.1.3. Il est crucial de comprendre que les variables x et y permettant d'écrire la fonction sont, comme pour les fonctions d'une variable, des *variables muettes*. On parle généralement de *première variable* et de *seconde variable* de la fonction.

Fixer l'une des variables permet de se ramener à l'étude d'une fonction réelle :

Définition 1.1.4. Soient (x_0, y_0) dans \mathbb{R}^2 et $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de deux variables. Les *applications partielles* associées à f en $a = (x_0, y_0)$ sont les applications :

$$f_{1,a} : \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & f(x, y_0) \end{array} \quad \text{et} \quad f_{2,a} : \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ y & \mapsto & f(x_0, y) \end{array} .$$

Représentation des fonctions de deux variables : La représentation complète d'une fonction f de deux variables se fait, par extension du principe de représentation des fonctions réelles, dans \mathbb{R}^3 . Le graphe de f est alors l'ensemble :

$$\Gamma_f = \{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y)\}.$$

Il est parfois plus simple et plus parlant de représenter des *coupes* de cette représentation, en fixant l'une des variables. Lorsque la variable z est fixée, on parle de *lignes de niveau*.

1.2 Ouverts de \mathbb{R}^2

L'étude de la régularité des fonctions de deux variables (continuité, dérivabilité...) nécessite de définir un analogue dans \mathbb{R}^2 de la notion d'intervalle, bien connue dans \mathbb{R} . On définit ainsi une *topologie* de \mathbb{R}^2 .

Définition 1.2.1. Soit $a \in \mathbb{R}^2$ et $r \geq 0$. La *boule ouverte* et la *boule fermée* de centre a et de rayon r sont respectivement les ensembles :

$$B(a, r) = \{p \in \mathbb{R}^2 \mid \|p - a\| < r\} \quad \text{et} \quad \overline{B}(a, r) = \{p \in \mathbb{R}^2 \mid \|p - a\| \leq r\}.$$

Définition 1.2.2. Soit U une partie de \mathbb{R}^2 . On dit que U est :

- un *ensemble borné de \mathbb{R}^2* , ou plus simplement un *borné de \mathbb{R}^2* , s'il existe $M \geq 0$ tel que $U \subset B(0_{\mathbb{R}^2}, M)$,
- un *ensemble ouvert de \mathbb{R}^2* , ou plus simplement un *ouvert de \mathbb{R}^2* , si :

$$\forall a \in U, \exists r > 0, B(a, r) \subset U.$$

Remarque 1.2.3. Dans l'assertion ci-dessus, l'inégalité stricte $r > 0$ est fondamentale : elle indique que tout point de U est à l'*intérieur* de U , comme dans un intervalle ouvert.

Le *bord* d'un ouvert de \mathbb{R}^2 peut alors être défini via la notion d'*adhérence* :

Définition 1.2.4. Soit U un ouvert de \mathbb{R}^2 et soit $a \in \mathbb{R}^2$. On dit que a est *adhérent à U* si : $\forall \varepsilon > 0, B(a, \varepsilon) \cap U \neq \emptyset$.

2 Continuité

La notion d'ouvert permet de définir la continuité d'une fonction de deux variables, directement à partir de la définition établie pour les fonctions réelles.

Soient U un ouvert de \mathbb{R}^2 , $f : U \rightarrow \mathbb{R}$, et $a \in \mathbb{R}^2$ adhérent à U .

Définition 2.0.1. Soit $l \in \mathbb{R}$. On dit que f a pour limite l en a , et on note $f(u) \xrightarrow{u \rightarrow a} l$, si :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall u \in U, (\|u - a\| \leq \delta) \Rightarrow (|f(u) - l| \leq \varepsilon).$$

Remarque 2.0.2. On définit de la même façon, en adaptant les définitions vues pour les fonctions réelles, le cas de $l = \pm\infty$.

Par suite :

Définition 2.0.3.

Soit $a \in U$. On dit que f est *continue en a* si $f(u) \xrightarrow{u \rightarrow a} f(a)$.

On dit que f est *continue sur U* si f est continue en a pour tout $a \in U$.

On retrouve certaines propriétés des fonctions réelles continues :

Proposition 2.0.4. Si f est continue en a , alors f est bornée au voisinage de a .

Proposition 2.0.5. Si f et g sont continues, alors $f + g$, λf , $f \cdot g$, $\frac{f}{g}$ sont continues lorsqu'elles sont définies.

On a vu que la représentation d'une fonction f de deux variables peut passer par celle de ses applications partielles f_1 et f_2 . Cependant, l'étude de f_1 et f_2 , et en particulier leur continuité, **ne suffit pas** à établir la continuité de f , comme le montre la propriété et la remarque suivantes :

Proposition 2.0.6. Si f est continue en $a = (x_0, y_0) \in U$, alors $f_{1,a}$ et $f_{2,a}$ sont continues en x_0 et y_0 respectivement.

Remarque 2.0.7. La réciproque de cette propriété est **fausse**, comme le montre le cas de la fonction $f : (x, y) \mapsto \frac{xy}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$, 0 si $(x, y) = (0, 0)$, en $(0, 0)$.

3 Dérivabilité

On a vu que la définition de la continuité s'étend naturellement des fonctions d'une variable aux fonctions de deux variables, et que la continuité d'une fonction de deux variables n'est pas équivalente à celle de ses applications partielles. Pour la dérivabilité, c'est tout l'inverse : la notion de taux d'accroissement est mal définie pour les fonctions de deux variables, et tout repose donc sur l'étude des applications partielles.

Soient U un ouvert de \mathbb{R}^2 , $f : U \rightarrow \mathbb{R}$, et $a = (x_0, y_0) \in U$.

Définition 3.0.1. Soit $v \in \mathbb{R}^2$. La *dérivée de f en a selon le vecteur v* , notée $\partial_v f(a)$ est la dérivée en 0 de la fonction $t \mapsto f(a + tv)$ (si elle existe).

Proposition 3.0.2. $\forall v \in \mathbb{R}^2, \forall \lambda \in \mathbb{R}, \partial_{\lambda v} f = \lambda \partial_v f$.

Les *dérivées partielles* de f sont alors les dérivées de f selon $e_1 = (1, 0)$ et $e_2 = (0, 1)$ respectivement, ou, de façon équivalente :

Définition 3.0.3. Si $f_{1,a}$ est dérivable en x_0 , on appelle *première dérivée partielle de f en a* ou *dérivée partielle de f par rapport à x en a* , et on note $\partial_1 f(a)$ ou $\frac{\partial f}{\partial x}(a)$, le nombre dérivé $f'_{1,a}(x_0)$ (s'il existe) :

$$\partial_1 f(a) = \frac{\partial f}{\partial x}(a) = f'_{1,a}(x_0).$$

De façon similaire, on appelle *seconde dérivée partielle de f en a* ou *dérivée partielle de f par rapport à y en a* , et on note $\partial_2 f(a)$ ou $\frac{\partial f}{\partial y}(a)$, le nombre dérivé $f'_{2,a}(y_0)$ (s'il existe) :

$$\partial_2 f(a) = \frac{\partial f}{\partial y}(a) = f'_{2,a}(y_0).$$

Lorsqu'elles sont définies sur U , $\partial_1 f$ et $\partial_2 f$ sont alors les *fonctions dérivées partielles* de f .

Exemples 3.0.4.

- $f : (x, y) \mapsto \sqrt{x^2 + y^2}$, non dérivable en $(0, 0)$,
- $f : (x, y) \mapsto \arctan\left(\frac{y}{x}\right)$.

On retrouve alors le lien entre nombre dérivé et extremum :

Proposition 3.0.5. Si f admet un extremum local en a , alors $\partial_1 f(a) = \partial_2 f(a) = 0$.

Remarque 3.0.6. Si $\partial_1 f(a) = \partial_2 f(a) = 0$, on dit que a est un *point critique* de f .

Comme pour les fonctions d'une variable, la réciproque de la proposition précédente est fausse, comme le montre le cas de la fonction $f : (x, y) \mapsto x^2 - y^2$.

Définition 3.0.7. On dit que f est de classe C^1 sur U , et on note $f \in C^1(U, \mathbb{R})$, si $\partial_1 f$ et $\partial_2 f$ existent et sont continues sur U .

On retrouve dans ce cas une notion de *développement limité* :

Théorème 3.0.8. Soit $f \in C^1(U, \mathbb{R})$ et $a = (x_0, y_0) \in U$. On a :

$$\begin{aligned} \forall u = (x, y) \in U, \quad f(u) &= f(a) + \partial_1 f(a)(x - x_0) + \partial_2 f(a)(y - y_0) + \underset{u \rightarrow a}{o} (\|u - a\|) \\ &= f(a) + \langle \nabla f(a), u - a \rangle + \underset{u \rightarrow a}{o} (\|u - a\|), \end{aligned}$$

où $\nabla f(a) := (\partial_1 f(a), \partial_2 f(a)) \in \mathbb{R}^2$ est le *gradient* de f en a .

Remarques 3.0.9.

- La notion de gradient est très pratique dans l'étude des fonctions de plusieurs variables. Par exemple, la proposition 3.0.5 se reformule comme suit : Si f admet un extremum local en a , alors $\nabla f(a) = 0_{\mathbb{R}^2}$.
- Le développement limité nous donne directement l'équation cartésienne du *plan tangent* à la surface représentative de f au point a :

$$z = f(a) + \partial_1 f(a)(x - x_0) + \partial_2 f(a)(y - y_0).$$

- D'après le théorème de représentation de Riesz, il existe une et une seule forme linéaire de \mathbb{R}^2 égale à $\langle \nabla f(a), \cdot \rangle$. Cette forme est appelée *differential de f en a* , et généralement notée df_a . On a alors :

$$\forall u \in U, f(u) = f(a) + df_a(u - a) + \underset{u \rightarrow a}{o}(\|u - a\|).$$

Avec la notion de gradient, et d'après le développement ci-dessus, on obtient les règles de composition suivantes :

Proposition 3.0.10. Soit $f \in C^1(U, \mathbb{R})$ et $x, y \in C^1(I, \mathbb{R})$ telles que $\gamma : t \mapsto (x(t), y(t))$ vérifie $\gamma(I) \subset U$. Alors $f \circ \gamma : I \rightarrow \mathbb{R}$ est de classe C^1 et :

$$\begin{aligned} \forall t \in I, \quad (f \circ \gamma)'(t) &= \partial_1 f(\gamma(t)) \cdot x'(t) + \partial_2 f(\gamma(t)) \cdot y'(t) \\ &= \langle \nabla f(\gamma(t)), \gamma'(t) \rangle. \end{aligned}$$

Proposition 3.0.11. Soient $f, \varphi, \psi \in C^1(U, \mathbb{R})$ telles que $(\varphi, \psi)(U) \subset U$. Alors $F = f \circ (\varphi, \psi) : U \rightarrow \mathbb{R}$ est de classe C^1 et :

$$\forall a \in U, \quad \begin{cases} \partial_1 F(a) &= \partial_1 f(\varphi(a), \psi(a)) \cdot \partial_1 \varphi(a) + \partial_2 f(\varphi(a), \psi(a)) \cdot \partial_1 \psi(a) \\ \partial_2 F(a) &= \partial_1 f(\varphi(a), \psi(a)) \cdot \partial_2 \varphi(a) + \partial_2 f(\varphi(a), \psi(a)) \cdot \partial_2 \psi(a) \end{cases},$$

c'est-à-dire :

$$\forall a \in U, \quad \begin{cases} \partial_1 F(a) &= \langle \nabla f(\varphi(a), \psi(a)), \partial_1(\varphi, \psi)(a) \rangle \\ \partial_2 F(a) &= \langle \nabla f(\varphi(a), \psi(a)), \partial_2(\varphi, \psi)(a) \rangle \end{cases}.$$