

Licence

Prépas

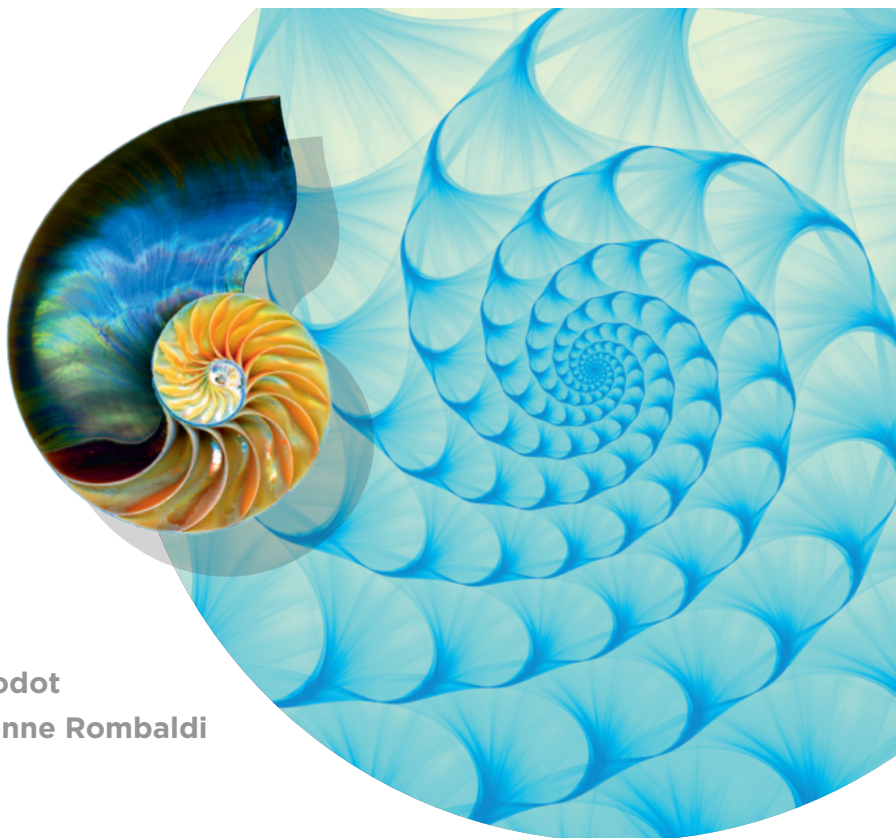
Capes

FORMULAIRE

DE

MATHS

AVEC RÉSUMÉS DE COURS



Olivier Rodot

Jean-Étienne Rombaldi

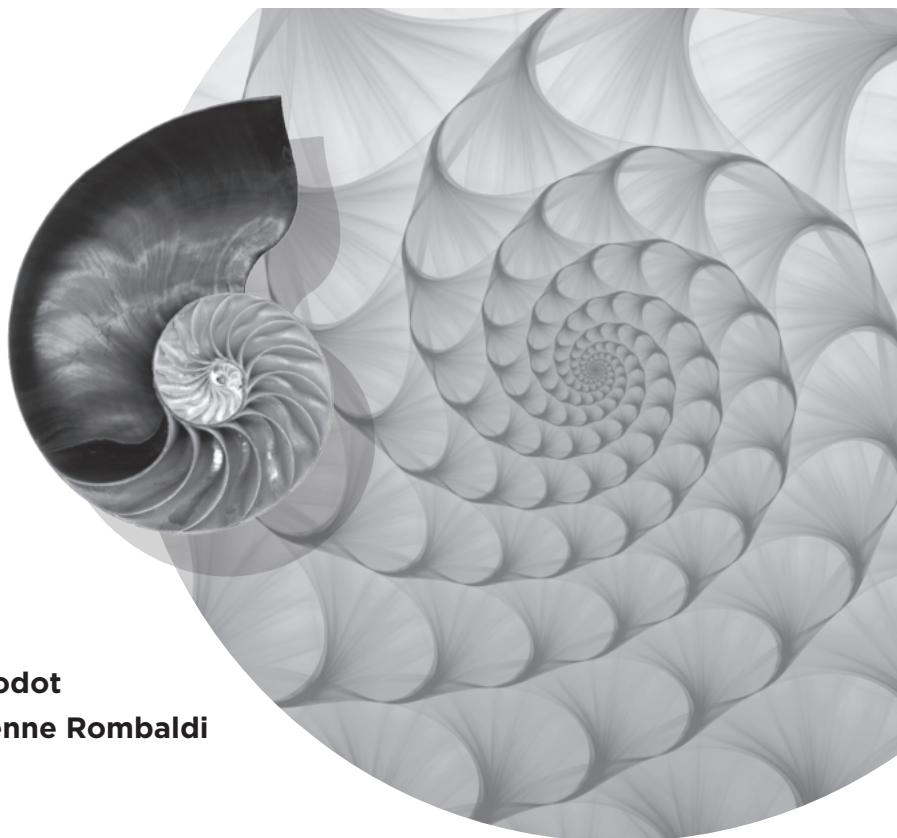
Licence

Prépas

Capes

FORMULAIRE DE MATHS

AVEC RÉSUMÉS DE COURS



Olivier Rodot
Jean-Étienne Rombaldi

Chez le même éditeur (extrait du catalogue)

DARRACQ M.-C. & ROMBALDI J.-É., *Analyse pour la Licence*

DARRACQ M.-C. & ROMBALDI J.-É., *Algèbre et géométrie pour la Licence*

DARRACQ M.-C. & ROMBALDI J.-É., *Probabilités pour la Licence*

DARRACQ M.-C. & ROMBALDI J.-É., *Mathématiques pour le Capes. Analyse*

DARRACQ M.-C. & ROMBALDI J.-É., *Mathématiques pour le Capes. Algèbre et géométrie*

DARRACQ M.-C. & ROMBALDI J.-É., *Mathématiques pour le Capes. Probabilités*

Pour toute information sur notre fonds et les nouveautés dans votre domaine
de spécialisation, consultez notre site web :

www.deboecksuperieur.com

En couverture : Coupe d'un nautille © Getty Images

Relectures : François Pantigny et Jérôme Cartailier

Maquette intérieure : Hervé Soulard/Nexeme, François Pantigny et Olivier Rodot

Mise en pages des auteurs

Maquette de couverture : Primo&Primo

Dépôt légal :

Bibliothèque royale de Belgique : 2022/13647/098

Bibliothèque nationale, Paris : août 2022

ISBN : 978-2-8073-3988-0

Tous droits réservés pour tous pays.

Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme ou de quelque manière que ce soit.

© De Boeck Supérieur SA, 2022 - Rue du Bosquet 7, B1348 Louvain-la-Neuve

De Boeck Supérieur - 5 allée de la 2^e DB, 75015 Paris

Table des matières

ANALYSE

par Olivier RODOT

1	Suites numériques	15
1.1	Borne supérieure, borne inférieure	15
1.2	Généralités sur les suites réelles	16
1.3	Limite d'une suite réelle	18
1.4	Suites monotones et suites adjacentes	20
1.5	Traductions séquentielles	21
1.6	Extension aux suites complexes	22
1.7	Suites extraites	24
1.8	Suites arithmético-géométriques	25
1.9	Suites récurrentes linéaires d'ordre 2	27
1.10	Suites récurrentes $u_{n+1} = f(u_n)$	28
2	Fonctions numériques d'une variable réelle	29
2.1	Rappels du vocabulaire usuel sur les fonctions	29
2.2	Compléments sur les fonctions usuelles	32
2.3	Limite d'une fonction en un point	35
2.4	Extension aux fonctions à valeurs complexes	41
2.5	Continuité	42
2.6	Théorèmes relatifs à la continuité	44
2.7	Continuité par morceaux	47
2.8	Continuité uniforme et fonction lipschitzienne	48
2.9	Dérivabilité	49
2.10	Fonctions de classe $\mathcal{C}^k$	52
2.11	Extremum local	53
2.12	Théorèmes relatifs à la dérivabilité	55
2.13	Convexité	57

2.14	Comparaison asymptotique	60
2.15	Comparaison asymptotique (cas des suites)	65
2.16	Développements limités	68
3	Intégration	74
3.1	Primitives	74
3.2	Intégration d'une fonction continue	75
3.3	Intégration d'une fonction continue par morceaux	76
3.4	Sommes de Riemann	79
3.5	Intégration par parties et par changement de variable	79
3.6	Formules de Taylor	80
3.7	Intégrabilité au sens de Riemann	80
4	Espaces vectoriels normés	83
4.1	Normes	83
4.2	Distance	84
4.3	Produit d'espaces vectoriels normés	85
4.4	Suite d'éléments d'un espace vectoriel normé	85
4.5	Boules et sphères	87
4.6	Normes équivalentes	88
4.7	Voisinage, ouvert, fermé	89
4.8	Caractérisations séquentielles	91
4.9	Partie dense	92
4.10	Voisinages relatifs, ouverts et fermés relatifs	92
4.11	Limites dans les espaces vectoriels normés	94
4.12	Continuité	96
4.13	Uniforme continuité	98
4.14	Applications lipschitziennes	98
4.15	Applications linéaires continues	99
4.16	Parties compactes	102
4.17	Continuité sur une partie compacte	103
4.18	Connexité par arcs	103
4.19	Espaces vectoriels normés de dimension finie	105
4.20	Complétude	107
5	Séries numériques	110
5.1	Convergence et divergence	110
5.2	Série géométrique	111
5.3	Somme et reste d'une série convergente	111
5.4	Somme de deux séries numériques et produit par un scalaire	112
5.5	Condition nécessaire de convergence et série harmonique	112
5.6	Séries à termes positifs	113
5.7	Séries à termes quelconques	116

5.8	Somation des relations de comparaison	119
5.9	Produit de deux séries numériques	120
5.10	Comparaison série-intégrale	120
5.11	Sommes de séries convergentes classiques	120
5.12	Familles sommables	121
6	Intégrales généralisées	126
6.1	Convergence et divergence	126
6.2	Linéarité, conjugaison, croissance et stricte positivité	128
6.3	Intégrales de Riemann	129
6.4	Cas des fonctions positives	129
6.5	Critères de comparaison	130
6.6	Comparaison série-intégrale	130
6.7	Intégration par parties et par changement de variable	131
6.8	Comportement asymptotique	132
6.9	Fonctions intégrables	132
6.10	Intégrabilité par comparaison et domination	133
6.11	Intégration des relations de comparaison	134
6.12	Quelques calculs d'intégrales généralisées	134
6.13	Suites et séries d'intégrales généralisées	135
6.14	Intégrales généralisées à paramètre	137
7	Équations différentielles linéaires	139
7.1	Équations différentielles linéaires d'ordre 1	139
7.2	Équations différentielles linéaires d'ordre 2 à coefficients constants	141
7.3	Exponentielle d'une matrice	143
7.4	Système différentiel linéaire d'ordre 1 à coefficients constants . .	144
7.5	Équations différentielles linéaires d'ordre 2	146
7.6	Équations différentielles linéaires vectorielles d'ordre 1	147
7.7	Équations différentielles linéaires d'ordre n	148
8	Suites de fonctions	150
8.1	Convergence simple et convergence uniforme	150
8.2	Convergence uniforme et continuité	152
8.3	Convergence uniforme et double limite	153
8.4	Intégration d'une suite de fonctions	153
8.5	Dérivation d'une suite de fonctions	154
8.6	Approximation uniforme	155
9	Séries de fonctions	156
9.1	Convergence simple et convergence uniforme	156
9.2	Convergence normale et convergence absolue	158
9.3	Convergence uniforme et continuité	160

9.4	Convergence uniforme et double limite	160
9.5	Intégration terme à terme sur un segment	161
9.6	Dérivation terme à terme	161
10	Séries entières	162
10.1	Rayon de convergence	162
10.2	Règles de d'Alembert et de Cauchy	164
10.3	Comparaisons des rayons de convergence	165
10.4	Opérations sur les séries entières	166
10.5	Dérivée et primitive d'une série entière	167
10.6	Convergence normale	168
10.7	Séries entières d'une variable réelle	168
10.8	Fonctions développables en série entière	171
10.9	Fonctions d'une variable réelle développables en série entière . . .	172
11	Séries de Fourier	176
11.1	Séries trigonométriques	176
11.2	Coefficients et séries de Fourier	177
11.3	Espace de Dirichlet	179
11.4	Inégalité de Bessel et lemme de Riemann-Lebesgue	180
11.5	Théorèmes de convergence	181
12	Fonctions vectorielles	184
12.1	Limite, continuité et dérivabilité	184
12.2	Opérations sur les fonctions dérivables	187
12.3	Intégration sur un segment	188
12.4	Sommes de Riemann	190
12.5	Intégrale fonction de sa borne supérieure	190
12.6	Formules de Taylor	191
12.7	Arcs paramétrés du plan	192
13	Calcul différentiel	197
13.1	Application différentiable et différentielle	197
13.2	Dérivées partielles	201
13.3	Matrice jacobienne et jacobien	204
13.4	Gradient	206
13.5	Applications de classe $\mathcal{C}^1$	207
13.6	Vecteurs tangents	208
13.7	Point critique et extremum	210
13.8	Applications de classe $\mathcal{C}^k$	211

13.9	Formule de Taylor-Young	212

ALGÈBRE

par Jean-Étienne ROMBALDI

14	Théorie des ensembles	217
14.1	Quelques notions de logique	217
14.2	Notions de base sur les ensembles	220
14.3	Les symboles $\sum$ et $\prod$	221
14.4	L'algèbre des parties d'un ensemble	222
14.5	Fonctions, injectivité, surjectivité, bijectivité	223
14.6	Relations d'ordre et d'équivalence	226
15	Groupes, anneaux et corps	227
15.1	Groupes monogènes	231
15.2	Groupes finis, théorème de Lagrange	232
15.3	Morphismes de groupes	232
15.4	Ordre d'un élément dans un groupe	234
15.5	Groupe des permutations d'un ensemble fini	235
15.6	Anneaux	239
15.7	Éléments inversibles dans un anneau	240
15.8	Sous-anneaux	241
15.9	Morphismes d'anneaux	241
15.10	Corps	243
15.11	Morphismes de corps	244
16	Arithmétique dans $\mathbb{Z}$	245
16.1	Théorème de division euclidienne dans $\mathbb{Z}$	245
16.2	Plus grand commun diviseur et plus petit commun multiple	246
16.3	L'algorithme d'Euclide	249
16.4	L'ensemble $\mathcal{P}$ des nombres premiers	250
16.5	Décomposition en facteurs premiers	251
16.6	Les théorèmes de Fermat et de Wilson	252
16.7	Les anneaux $\mathbb{Z}/n\mathbb{Z}$ et l'indicatrice d'Euler	252
17	Nombres complexes	255
17.1	Construction de $\mathbb{C}$	255
17.2	Conjugué et module d'un nombre complexe	256
17.3	Les équations de degré 2	257
17.4	Arguments d'un nombre complexe	258
17.5	Racines n-ièmes d'un nombre complexe	259

18	Espaces vectoriels	261
18.1	Sous-espaces vectoriels	262
18.2	Applications linéaires	264
18.3	Dualité	266
18.4	Sommes et sommes directes de sous-espaces vectoriels	268
18.5	Familles libres, familles génératrices et bases	270
18.6	Espaces vectoriels de dimension finie	271
18.7	Rang d'une famille de vecteurs ou d'une application linéaire . . .	275
18.8	Applications linéaires et calcul matriciel	277
18.9	Formules de changement de base	283
19	Déterminants	284
19.1	Opérations élémentaires. Matrices de dilatation et de transvection	284
19.2	Déterminants des matrices carrées	285
19.3	Déterminant d'une famille de vecteurs	288
19.4	Déterminant d'un endomorphisme	289
19.5	Systèmes linéaires	290
20	Polynômes à coefficients dans un corps	293
20.1	L'algèbre $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$	293
20.2	Division euclidienne dans $\mathbb{K}[X]$	296
20.3	Polynômes premiers entre eux	298
20.4	Racines des polynômes	299
20.5	Théorème de Leibniz et formule de Taylor	301
20.6	Polynômes irréductibles	303
20.7	Décomposition des fractions rationnelles en éléments simples . . .	304
20.8	Polynômes d'interpolation de Lagrange	306
21	Réduction des endomorphismes	307
21.1	Valeurs et vecteurs propres	307
21.2	Diagonalisation	308
21.3	Endomorphismes trigonalisables	309
21.4	Polynômes d'endomorphismes, polynôme minimal	310
21.5	Sous-espaces caractéristiques	314
22	Formes bilinéaires et quadratiques	317
22.1	Formes bilinéaires	317
22.2	Expression matricielle des formes bilinéaires.	318
22.3	Formes quadratiques	319
22.4	Théorème de réduction de Gauss	320
22.5	Orthogonalité, noyau et rang	321
22.6	Signature d'une forme quadratique réelle en dimension finie . . .	324
22.7	Quadriques dans $\mathbb{R}^n$ ou $\mathbb{C}^n$	325

23	Espaces préhilbertiens	327
23.1	Produit scalaire	327
23.2	Orthogonalité	329
23.3	Orthonormalisation de Gram-Schmidt	331
23.4	Adjoint d'un endomorphisme	332
23.5	Projection orthogonale sur un sous-espace de dimension finie	333
23.6	Symétries orthogonales dans les espaces euclidiens	335
23.7	Les endomorphismes symétriques réels	337
23.8	Isométries vectorielles	339
23.9	Orientation d'un espace euclidien	341
23.10	Produit vectoriel dans un espace euclidien	341
23.11	Isométries vectorielles en dimension 2	343
23.12	Isométries vectorielles en dimension 3	344
23.13	Espaces vectoriels hermitiens	345
23.14	Réduction des matrices normales	347
24	Géométrie affine et euclidienne	349
24.1	Espace affine associé à un espace vectoriel	349
24.2	Le groupe affine $GA(\mathcal{E})$ en dimension finie	353
24.3	Espaces affines euclidiens	355
24.4	Hyperplans dans un espace euclidien	357
24.5	Isométries affines d'un espace euclidien	358
24.6	Barycentres	359

PROBABILITÉS

par Jean-Étienne ROMBALDI

25	Dénombrement et probabilités	365
25.1	Dénombrement	365
25.2	Fonctions indicatrices d'ensembles	367
25.3	Expérience aléatoire	368
25.4	Tribus d'événements, espaces probabilisables	368
25.5	Espaces probabilisés	371
25.6	Espaces probabilisés discrets	372
25.7	Probabilités conditionnelles	373
25.8	Événements indépendants	374
26	Variables aléatoires réelles discrètes	376
26.1	Opérations sur les variables aléatoires réelles discrètes	379
26.2	Fonction de répartition d'une variable aléatoire discrète	380
26.3	Loi conjointe et lois marginales	381

26.4	Variables aléatoires discrètes indépendantes	381
26.5	Espérance, variance, moments d'une variable aléatoire discrète . .	383
26.6	Fonction génératrice d'une variable aléatoire discrète	387
27	Variable aléatoires réelles	390
27.1	Fonction de répartition d'une probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$	390
27.2	Variables aléatoires réelles	391
27.3	Variables aléatoires réelles indépendantes	393
27.4	Espérance d'une variable aléatoire réelle	394
27.5	Variance, covariance, moments d'ordre r	397
27.6	Loi faible des grands nombres	400
28	Variables aléatoires à densité	402
28.1	Densité de probabilité	402
28.2	Variables aléatoires à densité classiques	404
28.3	Espérance des variables aléatoires à densité	406
28.4	Formule de transfert, moments d'ordre r	407
28.5	Convergence en loi, théorème de la limite centrale	408
Annexe		
	Fonctions trigonométriques	413
	Formules de trigonométrie	414
	Dérivées usuelles	415
	Primitives usuelles	416
	Développements limités usuels	417
	Développements en séries entières usuels	417
Index		419

Avant-propos

Cet ouvrage dépasse le cadre classique d'un formulaire : il contient un résumé complet de cours comprenant souvent des remarques, parfois des exemples ou des contre-exemples. Il couvre le programme des classes préparatoires scientifiques, en particulier les programmes de MPSI et MP, mais aussi certaines notions et certains chapitres étudiés dans le premier cycle universitaire (séries de Fourier, variables aléatoires à densité). Il intéressera également les candidats au Capes.

L'ouvrage est divisé en trois parties : analyse, rédigée par Olivier RODOT, algèbre et probabilités, rédigées par Jean-Étienne ROMBALDI.

L'index est particulièrement soigné afin de permettre à l'étudiant de trouver immédiatement l'information cherchée.

En fin d'ouvrage sont rassemblés un formulaire de trigonométrie, les dérivées et primitives usuelles ainsi que les développements limités et les développements en séries entières classiques.

Les auteurs espèrent que ce formulaire constituera une aide précieuse pour l'étudiant souhaitant s'assurer de l'assimilation du très vaste programme de mathématiques des deux premières années post-baccalauréat.

Nous tenons à remercier très chaleureusement le travail très ingrat et méticuleux de relecture assuré par François PANTIGNY et Jérôme CARTAILLER. Leurs remarques pertinentes se sont révélées très utiles.

Nous remercions également François PANTIGNY pour sa proposition de mise en page, notamment au niveau des en-têtes de chapitres.

Enfin, nous remercions les éditions De Boeck et, en particulier, Alain LUGUET, pour la confiance témoignée lors de la rédaction de cet ouvrage.

N'hésitez pas à nous faire part de vos remarques ou suggestions en nous envoyant un courrier électronique à l'adresse suivante : `rodot.livre.maths@gmail.com`

Olivier RODOT et Jean-Étienne ROMBALDI

ANALYSE

par Olivier RODOT

Suites numériques

1

Les premières notions se concentrent sur l'étude des suites réelles puis s'étendront ensuite aux suites complexes.

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

1.1 Borne supérieure, borne inférieure

Définition

Soit A une partie de $\mathbb{R}$.

On dit qu'un réel M est un *majorant* de A si, pour tout $x \in A$, $x \leq M$.

Le plus petit (s'il existe) des majorants de A est appelé la *borne supérieure*¹ de A notée $\sup A$.

On dit qu'un réel m est un *minorant* de A si, pour tout $x \in A$, $x \geq m$.

Le plus grand (s'il existe) des minorants de A est appelé la *borne inférieure*² de A notée $\inf A$.

Par exemple $]0, +\infty[$ n'admet pas de borne supérieure mais admet 0 comme borne inférieure.

Ne pas confondre plus grand élément et borne supérieure : par exemple, 1 est la borne supérieure de $[0, 1[$ mais $[0, 1[$ n'admet pas de plus grand élément. On remarquera que la borne supérieure d'une partie A de $\mathbb{R}$ n'appartient pas nécessairement à A .

En revanche, on a le théorème qui suit.

1. ou *supremum*.

2. ou *infimum*.

Théorème

Soit A une partie de $\mathbb{R}$.

Si A admet un plus grand élément, alors A admet une borne supérieure vérifiant $\sup A = \max A$.

Si A admet un plus petit élément, alors A admet une borne inférieure vérifiant $\inf A = \min A$.

Théorème

Toute partie non vide et majorée de $\mathbb{R}$ admet une borne supérieure³.

Toute partie non vide et minorée de $\mathbb{R}$ admet une borne inférieure.

Théorème

Soient A une partie non vide et majorée de $\mathbb{R}$ et $M \in \mathbb{R}$. Alors les assertions suivantes sont équivalentes⁴ :

- (i) $M = \sup A$;
- (ii) $\begin{cases} M \text{ est un majorant de } A; \\ \forall \varepsilon > 0, \exists a \in A, a > M - \varepsilon. \end{cases}$

Soit A une partie non vide et minorée de $\mathbb{R}$ et $m \in \mathbb{R}$. Alors les assertions suivantes sont équivalentes⁵ :

- (i) $m = \inf A$;
- (ii) $\begin{cases} m \text{ est un minorant de } A; \\ \forall \varepsilon > 0, \exists a \in A, a < m + \varepsilon. \end{cases}$

1.2 Généralités sur les suites réelles

Définition

Soit $N \in \mathbb{N}$. On appelle *suite réelle*⁶ toute application

$$u : \{n \in \mathbb{N}, n \geq N\} \longrightarrow \mathbb{R}.$$

L'image d'un entier naturel $n \geq N$ par u est alors notée u_n et la suite u est également le plus souvent notée $(u_n)_{n \geq N}$.

3. Ce résultat est faux pour $\mathbb{Q}$. Par exemple, $A = \{x \in \mathbb{Q}, x^2 \leq 2\}$ est une partie non vide et majorée de $\mathbb{Q}$ n'admettant pas de borne supérieure (dans $\mathbb{Q}$).

4. L'assertion (ii) exprime le fait que M est un majorant de A et que pour tout $\varepsilon > 0$, $M - \varepsilon$ n'est plus un majorant de A ce qui correspond bien à la définition du *plus petit des majorants* de A .

5. L'assertion (ii) exprime le fait que m est un minorant de A et que pour tout $\varepsilon > 0$, $m + \varepsilon$ n'est plus un minorant de A ce qui correspond bien à la définition du *plus grand des minorants* de A .

Pour faciliter la lecture de ce chapitre, on prendra $N = 0$ sachant que les définitions et théorèmes s'adaptent aisément au cas général⁷.

L'ensemble des suites réelles est noté $\mathbb{R}^{\mathbb{N}}$.

On dit qu'une suite réelle (u_n) vérifie une propriété P_n à partir d'un certain rang si⁸

$$\exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies u_n \text{ vérifie la propriété } P_n.$$

Définition

Soit (u_n) une suite réelle. On dit que (u_n) est *majorée* si

$$\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \leq M.$$

Le réel M est alors appelé un *majorant* de (u_n) .

On dit que (u_n) est *minorée* si

$$\exists m \in \mathbb{R}, \forall n \in \mathbb{N}, u_n \geq m.$$

Le réel m est alors appelé un *minorant* de (u_n) .

On dit que (u_n) est *bornée* si (u_n) est majorée et minorée autrement dit si⁹

$$\exists (m, M) \in \mathbb{R}^2, \forall n \in \mathbb{N}, m \leq u_n \leq M.$$

Théorème

Soit (u_n) une suite réelle.

Alors (u_n) est bornée si, et seulement si, $(|u_n|)$ est majorée.

Une combinaison linéaire de suites réelles bornées est une suite réelle bornée de même que le produit de deux suites réelles bornées.

Définition

Soit (u_n) une suite réelle. On dit que (u_n) est *constante* si

$$\exists c \in \mathbb{R}, \forall n \in \mathbb{N}, u_n = c.$$

On dit que (u_n) est *stationnaire* si (u_n) est constante à partir d'un certain rang c'est-à-dire si¹⁰

$$\exists N \in \mathbb{N}, \exists c \in \mathbb{R}, \forall n \in \mathbb{N}, n \geq N \implies u_n = c.$$

6. On emploie aussi l'expression *suite de réels*.

7. Afin d'alléger l'écriture, (u_n) désignera dans le reste du chapitre et de l'ouvrage $(u_n)_{n \in \mathbb{N}}$.

8. Ce qui s'écrit souvent en abrégé : $\exists N \in \mathbb{N}, \forall n \geq N, u_n$ vérifie la propriété P_n .

9. Lorsque (u_n) est bornée, m et M sont des réels, ne dépendant donc pas de $n \in \mathbb{N}$, et ne sont pas uniques.

10. Ce qui s'écrit également en abrégé : $\exists N \in \mathbb{N}, \exists c \in \mathbb{R}, \forall n \geq N, u_n = c$.

Par exemple, la suite (u_n) définie pour tout $n \in \mathbb{N}$ par¹¹ $u_n = \left\lfloor \frac{2}{n+1} \right\rfloor + 1$ est stationnaire car pour tout $n \geq 2$, $u_n = 1$.

Définition

Soit (u_n) une suite réelle.

On dit que (u_n) est *croissante* si, pour tout $n \in \mathbb{N}$, $u_n \leq u_{n+1}$.

On dit que (u_n) est *décroissante* si, pour tout $n \in \mathbb{N}$, $u_n \geq u_{n+1}$.

On dit que (u_n) est *monotone* si (u_n) est croissante ou décroissante.

1.3 Limite d'une suite réelle

Définition

Soit (u_n) une suite réelle.

On dit que (u_n) admet^{12, 13} une *limite* finie $\ell \in \mathbb{R}$ si¹⁴

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies |u_n - \ell| < \varepsilon.$$

(u_n) admet une limite $\ell \in \mathbb{R}$ signifie donc que, pour tout $\varepsilon > 0$, tous les termes de la suite sont dans l'intervalle $]\ell - \varepsilon, \ell + \varepsilon[$ à partir d'un certain rang.

Définition

Soit (u_n) une suite réelle.

On dit (u_n) *tend vers* $+\infty$ si

$$\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies u_n > A.$$

On dit (u_n) *tend vers* $-\infty$ si

$$\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies u_n < A.$$

11. On rappelle que la fonction *partie entière*, notée $\lfloor \cdot \rfloor$, est définie pour tout $x \in \mathbb{R}$ par le plus grand entier relatif inférieur ou égal à x .

12. ou *converge vers* ou encore *tend vers*. Ces trois expressions synonymes sont à connaître.

13. Sous-entendu en $+\infty$ ou quand n tend vers $+\infty$.

14. En toute rigueur, on devrait écrire $\forall \varepsilon \in \mathbb{R}_+^*$ mais l'abus habituel d'écriture $\forall \varepsilon > 0$ sera pratiqué dans tout l'ouvrage pour ce réel strictement positif sous-entendu aussi petit que l'on veut.

Théorème

Soit (u_n) une suite réelle.

Si (u_n) tend vers $\ell \in \mathbb{R} \cup \{-\infty, +\infty\}$ alors ℓ est unique et sera notée¹⁵ $\lim_{n \rightarrow +\infty} u_n$ ou $\lim u_n$.

Définition

Soit (u_n) une suite réelle.

On dit que (u_n) *converge*¹⁶ si (u_n) admet une limite finie $\ell \in \mathbb{R}$.

On dit qu'elle *diverge*¹⁷ sinon.

Dire qu'une suite (u_n) diverge signifie : $\lim_{n \rightarrow +\infty} u_n = +\infty$ ou $\lim_{n \rightarrow +\infty} u_n = -\infty$ ou (u_n) n'admet pas de limite (finie).

Par exemple, la suite $(u_n) = ((-1)^n)$ diverge car (u_n) n'admet pas de limite.

Théorème

Toute suite réelle convergente est bornée.

La réciproque de ce théorème est fausse : il suffit pour s'en convaincre de reprendre l'exemple précédent $(u_n) = ((-1)^n)$.

Théorème

Soient $\lambda \in \mathbb{R}$, (u_n) et (v_n) deux suites réelles convergeant respectivement vers $\ell \in \mathbb{R}$ et $\ell' \in \mathbb{R}$. Alors

- $(\lambda u_n + v_n)$ converge et $\lim_{n \rightarrow +\infty} (\lambda u_n + v_n) = \lambda \ell + \ell'$;
- $(u_n v_n)$ converge et $\lim_{n \rightarrow +\infty} u_n v_n = \ell \ell'$;
- si $\ell' \neq 0$, $\frac{u_n}{v_n}$ converge¹⁸ et $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = \frac{\ell}{\ell'}$;
- si (u_n) est bornée et $\lim_{n \rightarrow +\infty} v_n = 0$ alors $\lim_{n \rightarrow +\infty} u_n v_n = 0$.

15. On écrira aussi $u_n \xrightarrow{n \rightarrow +\infty} \ell$ ou $u_n \longrightarrow \ell$.

16. ou est *convergente*.

17. ou est *divergente*.

18. On montre facilement que comme $\ell' \neq 0$, v_n ne s'annule pas à partir d'un certain rang.

Théorème

Soient (u_n) et (v_n) deux suites réelles convergeant respectivement vers deux réels ℓ et ℓ' . Si, à partir d'un certain rang, $u_n \leq v_n$ alors $\ell \leq \ell'$.

Ce théorème est faux pour les inégalités strictes : si $(u_n) = (1 - \frac{1}{n+1})$ et (v_n) est la suite constante égale à 1, pour tout $n \in \mathbb{N}$, on a $u_n < v_n$ mais $\lim_{n \rightarrow +\infty} u_n = 1$.

Théorème

Soit (u_n) une suite réelle convergeant vers un réel strictement positif. Alors (u_n) est strictement positive à partir d'un certain rang.

Théorème (limite finie par encadrement)

Soient (u_n) , (v_n) et (w_n) trois suites réelles.

Si, à partir d'un certain rang, $u_n \leq w_n \leq v_n$ avec (u_n) et (v_n) convergeant vers la même limite $\ell \in \mathbb{R}$, alors (w_n) converge vers ℓ .

Théorème (limite infinie par minoration ou majoration)

Soient (u_n) et (v_n) deux suites réelles vérifiant $v_n \geq u_n$ à partir d'un certain rang.

- Si $\lim_{n \rightarrow +\infty} u_n = +\infty$ alors $\lim_{n \rightarrow +\infty} v_n = +\infty$.
- Si $\lim_{n \rightarrow +\infty} v_n = -\infty$ alors $\lim_{n \rightarrow +\infty} u_n = -\infty$.

1.4 Suites monotones et suites adjacentes

Théorème de la limite monotone

Soit (u_n) une suite réelle.

Si (u_n) est monotone alors (u_n) admet une limite dans $\mathbb{R} \cup \{-\infty, +\infty\}$.

Plus précisément ¹⁹

- si (u_n) est croissante et majorée (respectivement croissante et non majorée), (u_n) est convergente (respectivement $\lim_{n \rightarrow +\infty} u_n = +\infty$);
- si (u_n) est décroissante et minorée (respectivement décroissante et non minorée), (u_n) est convergente (respectivement $\lim_{n \rightarrow +\infty} u_n = -\infty$).

Définition

On dit que deux suites réelles (u_n) et (v_n) sont *adjacentes* si (u_n) est croissante, (v_n) décroissante et $\lim_{n \rightarrow +\infty} (v_n - u_n) = 0$.

Théorème

Soient (u_n) et (v_n) deux suites adjacentes. Alors

- pour tout $n \in \mathbb{N}$, $u_n \leq v_n$;
- (u_n) et (v_n) sont convergentes et convergent vers une même limite $\ell \in \mathbb{R}$;
- pour tout $n \in \mathbb{N}$, $u_n \leq \ell \leq v_n$.

1.5 Traductions séquentielles

Théorème

Si A est une partie non vide et majorée (respectivement non vide et minorée) de $\mathbb{R}$ alors il existe une suite d'éléments de A convergeant vers la borne supérieure de A (respectivement la borne inférieure de A).

Si A est une partie non vide et non majorée (respectivement non vide et non minorée) de $\mathbb{R}$, alors il existe une suite d'éléments de A de limite $+\infty$ (respectivement $-\infty$).

Définition

Soit A une partie de $\mathbb{R}$.

On dit que A est *dense* dans $\mathbb{R}$ si A rencontre tout intervalle ouvert non vide de $\mathbb{R}$ c'est-à-dire si tout intervalle ouvert non vide de $\mathbb{R}$ contient au moins un point de A ²⁰.

19. L'hypothèse d'une suite *réelle* est capitale car ce théorème devient faux par exemple avec une suite de rationnels, qui même croissante et majorée, ne converge pas nécessairement vers un rationnel. Il suffit pour s'en convaincre de considérer la suite (u_n) définie par $u_0 = 2$ et pour tout $n \in \mathbb{N}^*$, $u_n = \frac{1}{2}(u_{n-1} + \frac{2}{u_{n-1}})$. Cette suite de rationnels est croissante et majorée mais ne converge pas vers un rationnel mais vers l'irrationnel $\sqrt{2}$.

20. On peut aussi dire plus simplement : A est *dense* dans $\mathbb{R}$ si entre deux réels distincts, il existe au moins un point de A .

Théorème

Soit A une partie de $\mathbb{R}$.

Alors les assertions suivantes sont équivalentes :

- (i) A est dense dans $\mathbb{R}$.
- (ii) $\forall x \in \mathbb{R}, \forall \varepsilon > 0, \exists a \in A, |x - a| < \varepsilon$.
- (iii) Pour tout réel x , il existe une suite d'éléments de A convergeant vers x .

Théorème

$\mathbb{Q}, \mathbb{R} \setminus \mathbb{Q}$ et $\mathbb{D}^{21}$ sont denses dans $\mathbb{R}$.

Ainsi, on peut toujours trouver un réel entre deux rationnels (ou entre deux irrationnels) distincts.

1.6 Extension aux suites complexes

Définition

Soit $N \in \mathbb{N}$. On appelle *suite complexe*²² toute application

$$u : \{n \in \mathbb{N}, n \geq N\} \longrightarrow \mathbb{C}.$$

Comme pour les suites réelles, on prendra $N = 0$ sachant que tous les résultats s'adaptent au cas général.

L'ensemble des suites complexes est noté $\mathbb{C}^{\mathbb{N}}$.

Définition

Soit (u_n) une suite complexe.

On dit que (u_n) admet^{23, 24} une *limite* $\ell \in \mathbb{C}$ si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies |u_n - \ell| < \varepsilon.$$

Comme pour les suites réelles, l'unicité de la limite éventuelle est assurée.

Une suite complexe (u_n) converge donc vers $\ell \in \mathbb{C}$ si la suite réelle $(|u_n - \ell|)$ converge vers 0.

21. $\mathbb{D}$ est l'ensemble des décimaux c'est-à-dire l'ensemble des rationnels de la forme $\frac{p}{10^n}$ où $p \in \mathbb{Z}$ et $n \in \mathbb{N}$.

22. On emploie aussi l'expression *suite de complexes*.

23. ou *converge vers* ou encore *tend vers*. Ces trois expressions synonymes sont à connaître.

24. Sous-entendu en $+\infty$ ou quand n tend vers $+\infty$.

Définition

Soit (u_n) une suite complexe.

On dit (u_n) *converge*²⁵ si (u_n) admet une limite $\ell \in \mathbb{C}$.

On dit qu'elle *diverge*²⁶ sinon.

Théorème

Soit (u_n) une suite complexe.

Alors (u_n) converge vers $\ell \in \mathbb{C}$ si, et seulement si, les suites réelles $(\operatorname{Re}(u_n))$ et $(\operatorname{Im}(u_n))$ convergent respectivement vers $\operatorname{Re}(\ell)$ et $\operatorname{Im}(\ell)$.

De plus dans ce cas, on a $\lim_{n \rightarrow +\infty} u_n = \lim_{n \rightarrow +\infty} (\operatorname{Re}(u_n)) + i \lim_{n \rightarrow +\infty} (\operatorname{Im}(u_n))$.

Définition

Soit (u_n) une suite complexe. On dit que (u_n) est *bornée* si

$$\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, |u_n| \leq M.$$

Une combinaison linéaire de suites complexes bornées est une suite complexe bornée de même que le produit de deux suites complexes bornées.

Théorème

Toute suite complexe convergente est bornée.

Théorème

Soient $\lambda \in \mathbb{C}$, (u_n) et (v_n) deux suites complexes convergeant respectivement vers $\ell \in \mathbb{C}$ et $\ell' \in \mathbb{C}$. Alors

- $(\lambda u_n + v_n)$ converge et $\lim_{n \rightarrow +\infty} (\lambda u_n + v_n) = \lambda \ell + \ell'$;
- $(u_n v_n)$ converge et $\lim_{n \rightarrow +\infty} u_n v_n = \ell \ell'$;
- si $\ell' \neq 0$, $\frac{u_n}{v_n}$ converge²⁷ et $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = \frac{\ell}{\ell'}$;
- si (u_n) est bornée et $\lim_{n \rightarrow +\infty} v_n = 0$ alors $\lim_{n \rightarrow +\infty} u_n v_n = 0$.

25. ou est *convergente*.

26. ou est *divergente*.

27. On montre facilement que comme $\ell' \neq 0$, v_n ne s'annule pas à partir d'un certain rang.

1.7 Suites extraites

Définition

Soit (u_n) une suite réelle ou complexe. On appelle *suite extraite*²⁸ de (u_n) toute suite $(u_{\varphi(n)})$ où $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est strictement croissante.

Une telle application φ est appelée *extractrice*.

(u_{2n}) , (u_{3n}) , (u_{n^2}) sont par exemple des suites extraites de (u_n) .

Une composée de deux extractrices φ et ψ est une extractrice de sorte que $(u_{(\psi \circ \varphi)(n)})$ est une suite extraite de (u_n) .

Théorème

Soit φ une extractrice.

Alors, pour tout $n \in \mathbb{N}$, $\varphi(n) \geq n$.

Théorème

Soit (u_n) une suite complexe.

Si (u_n) converge vers $\ell \in \mathbb{C}$, toute suite extraite de (u_n) converge vers ℓ .

Soit (u_n) une suite réelle.

Si (u_n) tend vers $\ell \in \mathbb{R} \cup \{-\infty, +\infty\}$, toute suite extraite de (u_n) tend vers ℓ .

Corollaire

Soit (u_n) une suite complexe.

Si (u_n) admet deux suites extraites convergeant vers des limites distinctes, alors (u_n) diverge.

Soit (u_n) une suite réelle.

Si (u_n) admet deux suites extraites tendant vers des limites (finies ou infinies) distinctes, alors (u_n) diverge.

Ce corollaire est très utile : on montre ainsi aisément la divergence de la suite complexe²⁹ $(u_n) = (j^n)$ en considérant les deux suites extraites (u_{3n}) et (u_{3n+1}) convergeant respectivement vers 1 et j .

De même la suite réelle $(u_n) = ((-1)^n)$ diverge car les deux suites extraites (u_{2n}) et (u_{2n+1}) convergent respectivement vers 1 et -1 .

28. ou *sous-suite*.

29. j est le nombre complexe $e^{2i\pi/3}$.

Théorème

Soit (u_n) une suite réelle ou complexe. Si (u_{2n}) et (u_{2n+1}) sont convergentes et convergent vers la même limite alors (u_n) converge.

On a déjà vu que toute suite convergente est bornée et que la réciproque est fausse. On a néanmoins le théorème suivant.

Théorème de Bolzano-Weierstrass

Soit (u_n) une suite réelle ou complexe.

Si (u_n) est bornée, on peut extraire de (u_n) une sous-suite convergente.

1.8 Suites arithmético-géométriques

Définition

Soit (u_n) une suite réelle ou complexe.

On dit que (u_n) est *arithmétique* si³⁰ : $\exists r \in \mathbb{K}, \forall n \in \mathbb{N}, u_{n+1} = u_n + r$.
 r est alors appelé la *raison* de la suite (u_n) .

Théorème

Soit (u_n) une suite arithmétique de raison $r \in \mathbb{K}$. Alors³¹

- $\forall n \in \mathbb{N}, u_n = u_0 + nr$;
- Si (u_n) est réelle alors $\begin{cases} \text{si } r = 0, (u_n) \text{ est constante ;} \\ \text{si } r > 0, (u_n) \text{ est strictement croissante ;} \\ \text{si } r < 0, (u_n) \text{ est strictement décroissante.} \end{cases}$

Définition

Soit (u_n) une suite réelle ou complexe.

On dit que (u_n) est *géométrique* si³² : $\exists q \in \mathbb{K}, \forall n \in \mathbb{N}, u_{n+1} = qu_n$.
 q est alors appelé³³ la *raison* de la suite (u_n) .

30. Il est sous-entendu que $r \in \mathbb{R}$ si (u_n) est une suite réelle et $r \in \mathbb{C}$ si (u_n) est une suite complexe.

31. Plus généralement, pour tout $(n, p) \in \mathbb{N}^2$, $u_n = u_p + (n - p)r$.

32. Il est sous-entendu que $q \in \mathbb{R}$ si (u_n) est une suite réelle et $q \in \mathbb{C}$ si (u_n) est une suite complexe.

33. Si $u_0 = 0$, on ne peut pas parler de *raison* de la suite (u_n) car alors, pour tout $q \in \mathbb{K}$, (u_n) est la suite nulle.

Théorème

Soit (u_n) une suite géométrique de raison $q \in \mathbb{K}$. Alors³⁴

- $\forall n \in \mathbb{N}, u_n = q^n u_0$.
- Si (u_n) est réelle alors $\begin{cases} \text{si } q = 1, (u_n) \text{ est constante;} \\ \text{si } q > 1, (u_n) \text{ est strictement croissante;} \\ \text{si } 0 < q < 1, (u_n) \text{ est strictement décroissante.} \end{cases}$

Théorème

Soit $(u_n) = (q^n)$ où $q \in \mathbb{R}$. Alors

- si $q = 1$, (u_n) est constante égale à 1 donc converge vers 1 ;
- si $q > 1$, $\lim_{n \rightarrow +\infty} u_n = +\infty$;
- si $-1 < q < 1$, $\lim_{n \rightarrow +\infty} u_n = 0$;
- si $q \leq -1$, (u_n) n'a pas de limite donc diverge.

Théorème

Soit $(u_n) = (q^n)$ où $q \in \mathbb{C}$. Alors

- si $q = 1$, (u_n) est constante égale à 1 donc converge vers 1 ;
- si $|q| < 1$, $\lim_{n \rightarrow +\infty} u_n = 0$;
- dans tous les autres cas, (u_n) diverge.

Définition

Soit (u_n) une suite réelle ou complexe.

On dit que (u_n) est *arithmético-géométrique* si³⁵

$$\exists(q, r) \in \mathbb{K}^2, \forall n \in \mathbb{N}, u_{n+1} = qu_n + r.$$

Comme pour les suites géométriques et arithmétiques, il existe une formule générale donnant l'expression de (u_n) en fonction de n mais elle n'est pas à retenir.

La méthode pour déterminer u_n en fonction de n dans le cas où (u_n) n'est pas géométrique ($q \neq 1$) est la suivante : on résout l'équation $a = qa + r$ d'inconnue $a \in \mathbb{K}$ puis on vérifie que la suite $(v_n) = (u_n - a)$ est géométrique de raison q .

34. Plus généralement, si $q \neq 0$, pour tout $(n, p) \in \mathbb{N}^2$, $u_n = q^{n-p} u_p$.

35. Si $r = 0$, on retrouve la notion de suite géométrique et lorsque $q = 0$, on retrouve celle de suite arithmétique.

1.9 Suites récurrentes linéaires d'ordre 2

Définition

Soit (u_n) une suite réelle ou complexe

On dit que (u_n) est *récurrente linéaire d'ordre 2* si³⁶

$$\exists(a, b) \in \mathbb{K}^2 \setminus \{(0, 0)\}, \forall n \in \mathbb{N}, u_{n+2} + au_{n+1} + bu_n = 0.$$

L'équation $r^2 + ar + b = 0$ est appelée *équation caractéristique* de (u_n) .

Distinguons à présent le cas réel du cas complexe dans les théorèmes qui suivent.

Théorème

Soit (u_n) une suite *réelle* récurrente linéaire d'ordre 2 et notons Δ le discriminant de son équation caractéristique.

- Si $\Delta > 0$, l'équation caractéristique admet deux racines réelles r_1 et r_2 distinctes et

$$\exists(\lambda, \mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = \lambda r_1^n + \mu r_2^n.$$

- Si $\Delta = 0$, l'équation caractéristique admet une racine réelle r et

$$\exists(\lambda, \mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = (\lambda n + \mu)r^n.$$

- Si $\Delta < 0$, l'équation caractéristique admet deux racines complexes conjuguées $re^{i\theta}$ et $re^{-i\theta}$ et

$$\exists(\lambda, \mu) \in \mathbb{R}^2, \forall n \in \mathbb{N}, u_n = ((\lambda \cos(n\theta) + \mu \sin(n\theta))r^n).$$

Théorème

Soit (u_n) une suite *complexe* récurrente linéaire d'ordre 2 et notons Δ le discriminant de son équation caractéristique.

- Si $\Delta \neq 0$, l'équation caractéristique admet deux racines complexes r_1 et r_2 distinctes et

$$\exists(\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, u_n = \lambda r_1^n + \mu r_2^n.$$

- Si $\Delta = 0$, l'équation caractéristique admet une racine complexe r et

$$\exists(\lambda, \mu) \in \mathbb{C}^2, \forall n \in \mathbb{N}, u_n = (\lambda n + \mu)r^n.$$

36. On suppose dans cette définition que $(a, b) \neq (0, 0)$ car sinon, (u_n) est la suite nulle dès que $n \geq 2$.

1.10 Suites récurrentes $u_{n+1} = f(u_n)$

Définition

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

On dit que I est *stable* par f si $f(I) \subset I$ c'est-à-dire si

$$\forall x \in I, f(x) \in I.$$

On dit que $x \in I$ est un *point fixe* de f si $f(x) = x$.

Théorème

Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{R}$.

Si I est stable par f , il existe une unique suite (u_n) telle que

$$u_0 = a \text{ et, pour tout } n \in \mathbb{N}, u_{n+1} = f(u_n).$$

De plus, pour tout $n \in \mathbb{N}$, $u_n \in I$.

Le théorème qui suit rassemble les propriétés à connaître sur la monotonie et la limite éventuelle de ce type de suites.

Théorème

Soient I un intervalle de $\mathbb{R}$, $a \in I$, f une fonction de I vers $\mathbb{R}$ tels que I est stable par f et (u_n) définie par $u_0 = a$ et, pour tout $n \in \mathbb{N}$, $u_{n+1} = f(u_n)$.

- Si, pour tout $x \in I$, $f(x) \geq x$ (respectivement $f(x) \leq x$), (u_n) est croissante (respectivement décroissante).
- Si f est croissante sur I , (u_n) est monotone³⁷.
- Si f est décroissante sur I , (u_{2n}) et (u_{2n+1}) sont monotones³⁸.
- Si (u_n) converge vers $\ell \in I$ et si f est continue en ℓ alors ℓ est un point fixe de f .

37. Plus précisément, si $u_0 \leq u_1$, (u_n) est croissante et si $u_0 \geq u_1$, (u_n) est décroissante.

38. et de sens de variations contraires. Plus précisément, si $u_0 \leq u_2$, (u_{2n}) croissante et (u_{2n+1}) décroissante, et si $u_0 \geq u_2$, (u_{2n}) décroissante et (u_{2n+1}) croissante.

Fonctions numériques d'une variable réelle

2

Ce chapitre aborde l'étude des fonctions numériques¹ d'une variable réelle, plus précisément des fonctions définies sur une partie de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

À partir de la section 3, les fonctions sont définies sur un intervalle² de $\mathbb{R}$ non vide et non réduit à un point et sont d'abord, dans la première section, à valeurs réelles. La seconde section propose d'étendre certaines des notions à des fonctions à valeurs complexes. Par la suite, sauf cas particuliers³, les fonctions seront à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

2.1 Rappels du vocabulaire usuel sur les fonctions

Soit X une partie non vide⁴ de $\mathbb{R}$.

On dira par la suite que f est une fonction *de* X vers $\mathbb{K}$ si f est *définie* sur X à valeurs dans $\mathbb{K}$ c'est-à-dire, pour tout $x \in X$, $f(x)$ est définie⁵ et $f(x) \in \mathbb{K}$.

Lorsque $\mathbb{K} = \mathbb{R}$ (respectivement $\mathbb{K} = \mathbb{C}$), on parle de *fonction réelle* (respectivement *fonction complexe*).

Lorsque f est une fonction de X vers $\mathbb{K}$, on emploiera parfois la notation

$$f : X \longrightarrow \mathbb{K} \quad \text{ou} \quad f : \begin{cases} X & \longrightarrow \mathbb{K} \\ x & \longmapsto f(x) \end{cases} .$$

1. Le terme *numérique* fait référence à l'espace d'arrivée en l'occurrence $\mathbb{R}$ ou $\mathbb{C}$.

2. Le programme officiel de MPSI restreint l'étude des notions de ce chapitre à des intervalles, sachant que ces notions s'étendent à des ensembles qui ne sont pas nécessairement des intervalles mais des parties quelconques de $\mathbb{R}$ non vides et non réduit à un point.

3. comme par exemple les paragraphes énonçant les principaux théorèmes sur la continuité ou la dérivabilité.

4. Toutes les parties de $\mathbb{R}$ de cette section seront considérées comme non vide.

5. Autrement dit, le *domaine de définition* de f est X .

Définition

Soient $(O, \vec{i}, \vec{j})$ un repère du plan, X une partie de $\mathbb{R}$ et f une fonction de X vers $\mathbb{R}$. On appelle *graphe*⁶ de f la courbe du plan d'équation $y = f(x)$ c'est-à-dire l'ensemble noté Γ_f ou C_f défini par⁷

$$\Gamma_f = \{(x, y) \in \mathbb{R}^2, x \in X \text{ et } y = f(x)\}.$$

Définition

Soient X une partie de $\mathbb{R}$, A une partie de X et f une fonction de X vers $\mathbb{K}$. On appelle *restriction* de f à A la fonction de A vers $\mathbb{K}$, notée $f|_A$, définie par

$$\forall x \in A, f|_A(x) = f(x).$$

On appelle *prolongement* (ou *extension*) de f à $Y \supset X$ toute fonction g de Y vers $\mathbb{K}$ vérifiant

$$\forall x \in X, g(x) = f(x).$$

Définition

Soient $\lambda \in \mathbb{K}$, X une partie de $\mathbb{R}$ et f et g deux fonctions de X vers $\mathbb{K}$. Alors $\lambda f + g$ et fg sont deux fonctions de X vers $\mathbb{K}$ définies pour tout $x \in X$ par $(\lambda f + g)(x) = \lambda f(x) + g(x)$ et $(fg)(x) = f(x)g(x)$.

Définition

Soient X et Y deux parties de $\mathbb{R}$, f une fonction de X vers $\mathbb{R}$ telle que⁸ $f(X) \subset Y$ et g une fonction de Y vers $\mathbb{K}$.

Alors $g \circ f$ est une fonction de X vers $\mathbb{K}$ définie pour tout $x \in X$ par

$$(g \circ f)(x) = g(f(x)).$$

Définition

Soient $T \in \mathbb{R}_+^*$, X une partie de $\mathbb{R}$ et f une fonction de X vers $\mathbb{K}$.

On dit que f est *paire* si, pour tout $x \in X$, $-x \in X$ et $f(-x) = f(x)$.

On dit que f est *impaire* si, pour tout $x \in X$, $-x \in X$ et $f(-x) = -f(x)$.

On dit que f est *T-périodique* si, pour tout $x \in X$, $x + T \in X$, $x - T \in X$ et $f(x + T) = f(x)$. On dit alors que T est une période⁹ de f .

6. ou *courbe représentative* de f .

7. noté aussi $\Gamma_f = \{(x, f(x)); x \in X\}$.

8. Lorsque $f(X) \subset Y$, on dit que f est à *valeurs dans* Y .

9. En général T n'est pas unique sauf si par exemple f est continue et non constante sur X .

Théorème

Soient $(O, \vec{i}, \vec{j})$ un repère du plan, X une partie de $\mathbb{R}$ et f une fonction de X vers $\mathbb{R}$.

- Si f est paire, son graphe est symétrique par rapport à l'axe des ordonnées.
- Si f est impaire, son graphe est symétrique par rapport à O .

Définition

Soient X une partie de $\mathbb{R}$ et f une fonction de X vers $\mathbb{R}$.

On dit que f est *croissante* si : $\forall (x, y) \in X^2, x \leq y \implies f(x) \leq f(y)$.

On dit que f est *décroissante* si : $\forall (x, y) \in X^2, x \leq y \implies f(x) \geq f(y)$.

On dit que f est *monotone* si f est croissante ou décroissante.

Les définitions de f strictement croissante (respectivement décroissante) s'obtiennent en remplaçant dans chacune des définitions ci-dessus les inégalités larges par des inégalités strictes.

Définition

Soient X une partie de $\mathbb{R}$ et f une fonction de X vers $\mathbb{R}$.

On dit que f est *majorée* si : $\exists M \in \mathbb{R}, \forall x \in X, f(x) \leq M$.

Le réel M est alors appelé un *majorant* de f .

On dit que f est *minorée* si : $\exists m \in \mathbb{R}, \forall x \in X, f(x) \geq m$.

Le réel m est alors appelé un *minorant* de f .

On dit que f est *bornée* si f est majorée et minorée¹⁰.

Définition

Soient X une partie de $\mathbb{R}$, $a \in X$ et f une fonction de X vers $\mathbb{R}$.

On dit que f admet un *maximum en a* si, pour tout $x \in X$, $f(x) \leq f(a)$.

Le réel $f(a)$ est alors appelé le *maximum de f* , noté $\max_{x \in X} f(x)$ ou $\max_X f$.

On dit que f admet un *minimum en a* si, pour tout $x \in X$, $f(x) \geq f(a)$.

Le réel $f(a)$ est alors appelé le *minimum de f* , noté $\min_{x \in X} f(x)$ ou $\min_X f$.

10. On montre aisément que f bornée équivaut à $|f|$ majorée.

2.2 Compléments sur les fonctions usuelles

Les fonctions puissances $x \mapsto x^\alpha$ (où $x \in \mathbb{R}_+^*$), exponentielle, logarithme népérien et trigonométriques $\sin$, $\cos$ et $\tan$ sont supposées connues¹¹.

Définition (fonction partie entière)

On appelle fonction *partie entière* la fonction de $\mathbb{R}$ vers $\mathbb{R}$ définie pour tout $x \in \mathbb{R}$ par le plus grand entier relatif inférieur ou égal à x .

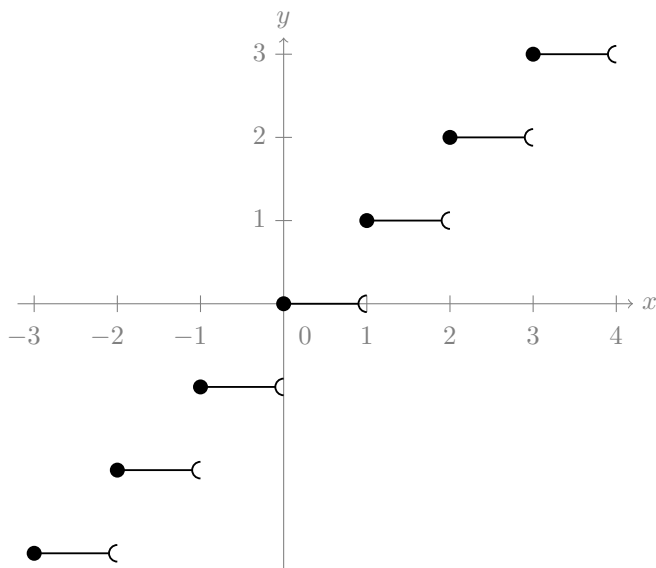
La partie entière d'un réel x sera notée $\lfloor x \rfloor$.

Par exemple, $\lfloor \pi \rfloor = 3$ et $\lfloor -\pi \rfloor = -4$.

Théorème (caractérisation de la partie entière)

Soit $x \in \mathbb{R}$. Alors

- $\lfloor x \rfloor$ est l'unique entier relatif tel que $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$;
- $\lfloor x \rfloor$ est l'unique entier relatif tel que $x - 1 < \lfloor x \rfloor \leq x$.



11. On trouvera les formules trigonométriques usuelles page 414.

Définition (fonctions hyperboliques)

On appelle *sinus hyperbolique*, notée sh, la fonction de $\mathbb{R}$ vers $\mathbb{R}$ définie pour tout $x \in \mathbb{R}$ par

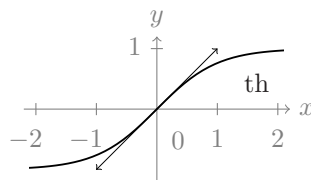
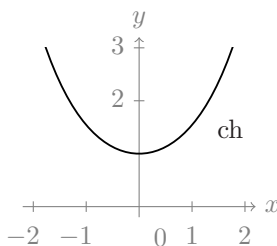
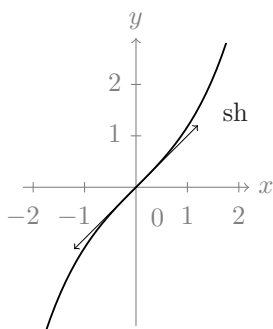
$$\text{sh}(x) = \frac{e^x - e^{-x}}{2}.$$

On appelle *cosinus hyperbolique*, notée ch, la fonction de $\mathbb{R}$ vers $\mathbb{R}$ définie pour tout $x \in \mathbb{R}$ par

$$\text{ch}(x) = \frac{e^x + e^{-x}}{2}.$$

On appelle *tangente hyperbolique*, notée th, la fonction de $\mathbb{R}$ vers $\mathbb{R}$ définie pour tout $x \in \mathbb{R}$ par

$$\text{th}(x) = \frac{\text{sh}(x)}{\text{ch}(x)}.$$



Théorème (propriétés des fonctions hyperboliques)

- La fonction sh est impaire, croissante sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$,

$$(\text{sh})'(x) = \text{ch}(x).$$

- La fonction ch est paire, croissante sur $\mathbb{R}_+$, décroissante sur $\mathbb{R}_-$ et pour tout $x \in \mathbb{R}$,

$$(\text{ch})'(x) = \text{sh}(x).$$

- La fonction th est impaire, croissante sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$,

$$(\text{th})'(x) = \frac{1}{\text{ch}^2(x)} = 1 - \text{th}^2(x).$$

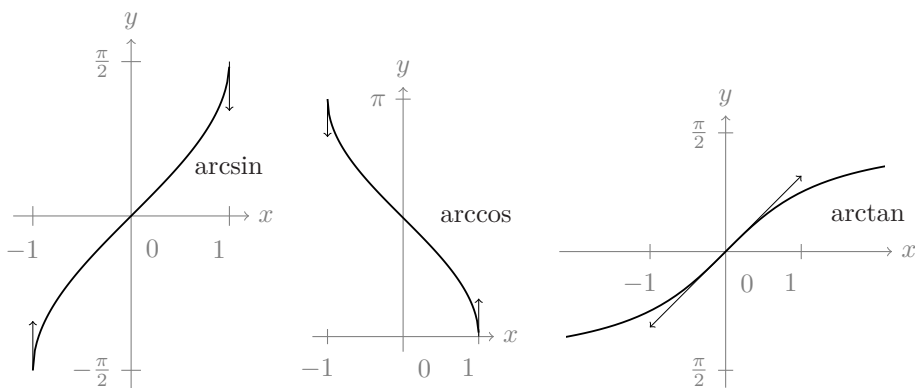
- Pour tout $x \in \mathbb{R}$, $\text{ch}^2(x) - \text{sh}^2(x) = 1$.

Définition (fonctions réciproques des fonctions trigonométriques)

La fonction sinus est bijective de $[-\frac{\pi}{2}, \frac{\pi}{2}]$ vers $[-1, 1]$. Sa bijection réciproque est appelée ¹² *arcsinus* et notée $\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$.

La fonction cosinus est bijective de $[0, \pi]$ vers $[-1, 1]$. Sa bijection réciproque est appelée *arccosinus* et notée $\arccos : [-1, 1] \rightarrow [0, \pi]$.

La fonction tangente est bijective de $]-\frac{\pi}{2}, \frac{\pi}{2}[$ vers $\mathbb{R}$. Sa bijection réciproque est appelée *arctangente* et notée $\arctan : \mathbb{R} \rightarrow]-\frac{\pi}{2}, \frac{\pi}{2}[$.

**Théorème (propriétés de la fonction arcsinus)**

• $\arcsin(0) = 0$, $\arcsin(\frac{1}{2}) = \frac{\pi}{6}$, $\arcsin(\frac{\sqrt{2}}{2}) = \frac{\pi}{4}$, $\arcsin(\frac{\sqrt{3}}{2}) = \frac{\pi}{3}$ et $\arcsin(1) = \frac{\pi}{2}$.

• La fonction arcsinus est impaire, croissante sur $[-1, 1]$ et, pour tout $x \in]-1, 1[$,

$$(\arcsin)'(x) = \frac{1}{\sqrt{1-x^2}}.$$

• $\begin{cases} \text{Pour tout } x \in [-1, 1], \sin(\arcsin(x)) = x. \\ \text{Pour tout } x \in [-\frac{\pi}{2}, \frac{\pi}{2}], \arcsin(\sin(x)) = x. \end{cases}$

12. Le nom *arcsinus* vient du fait que $\arcsin(x)$ est l'arc de $[-\frac{\pi}{2}, \frac{\pi}{2}]$ dont le sinus est x .

Théorème (propriétés de la fonction arccosinus)

• $\arccos(0) = \frac{\pi}{2}$, $\arccos(\frac{1}{2}) = \frac{\pi}{3}$, $\arccos(\frac{\sqrt{2}}{2}) = \frac{\pi}{4}$, $\arccos(\frac{\sqrt{3}}{2}) = \frac{\pi}{6}$ et $\arccos(1) = 0$.

• La fonction arccosinus est décroissante sur $[-1, 1]$ et, pour tout $x \in]-1, 1[$,

$$(\arccos)'(x) = -\frac{1}{\sqrt{1-x^2}}.$$

- $\begin{cases} \text{Pour tout } x \in [-1, 1], \cos(\arccos(x)) = x. \\ \text{Pour tout } x \in [0, \pi], \arccos(\cos(x)) = x. \end{cases}$
- Pour tout $x \in [-1, 1]$, $\arccos(x) + \arcsin(x) = \frac{\pi}{2}$.

Théorème (propriétés de la fonction arctangente)

• $\arctan(0) = 0$, $\arctan(\frac{1}{\sqrt{3}}) = \frac{\pi}{6}$, $\arctan(\sqrt{3}) = \frac{\pi}{3}$ et $\arctan(1) = \frac{\pi}{4}$.

• La fonction arctangente est impaire, croissante sur $\mathbb{R}$ et, pour tout $x \in \mathbb{R}$,

$$(\arctan)'(x) = \frac{1}{1+x^2}.$$

- $\begin{cases} \text{Pour tout } x \in \mathbb{R}, \tan(\arctan(x)) = x. \\ \text{Pour tout } x \in]-\frac{\pi}{2}, \frac{\pi}{2}[, \arctan(\tan(x)) = x. \end{cases}$

2.3 Limite d'une fonction en un point

Définition

On appelle^{13, 14, 15} *segment* de $\mathbb{R}$ toute partie de $\mathbb{R}$ de la forme $[a, b]$ où $(a, b) \in \mathbb{R}^2$.

On appelle *intervalle* de $\mathbb{R}$, toute partie I de $\mathbb{R}$ vérifiant

$$\forall (a, b) \in I^2 \text{ tel que } a \leq b, [a, b] \subset I.$$

13. Pour tout $a \in \mathbb{R}$, le singleton $\{a\}$ est considéré comme un segment réduit à un point.

14. $[a, b] = \{x \in \mathbb{R}, a \leq x \leq b\}$.

15. Dans toute la suite, dans l'écriture $[a, b]$, il est sous-entendu que $a \leq b$ de sorte que par exemple l'écriture $[1, 0]$ sera proscrite.

Mis à part $\mathbb{R}$ et $\emptyset$, si a et b sont deux réels, tout intervalle de $\mathbb{R}$ est de l'une des formes suivantes :

$$\begin{aligned}]a, b[&= \{x \in \mathbb{R}, a < x < b\}; [a, +\infty[= \{x \in \mathbb{R}, x \geq a\}; \\]a, b] &= \{x \in \mathbb{R}, a < x \leq b\};]-\infty, b] = \{x \in \mathbb{R}, x \leq b\}; \\]a, +\infty[&= \{x \in \mathbb{R}, x > a\};]-\infty, b[= \{x \in \mathbb{R}, x < b\}; \\ [a, b] &= \{x \in \mathbb{R}, a \leq x \leq b\}; [a, b[= \{x \in \mathbb{R}, a \leq x < b\}. \end{aligned}$$

Définition

Soit $(a, b) \in \mathbb{R}^2$.

On dit qu'un intervalle ¹⁶ de $\mathbb{R}$ est *ouvert* s'il est de la forme $]a, b[,]a, +\infty[,]-\infty, b[, \emptyset$ ou $\mathbb{R}$.

On dit qu'un intervalle de $\mathbb{R}$ est *fermé* s'il est de la forme $[a, b], [a, +\infty[,]-\infty, b], \emptyset$ ou $\mathbb{R}$.

On dit qu'un intervalle de $\mathbb{R}$ est *semi-ouvert* s'il est de la forme $[a, b[$ ou $]a, b]$.

Définition

On dit qu'une partie V de $\mathbb{R}$ est un *voisinage d'un réel a* s'il existe un réel $\alpha > 0$ tel que ¹⁷ $]a - \alpha, a + \alpha[\subset V$.

On dit qu'une partie V de $\mathbb{R}$ est un *voisinage de $+\infty$* s'il existe un réel A tel que $]A, +\infty[\subset V$.

On dit qu'une partie V de $\mathbb{R}$ est un *voisinage de $-\infty$* s'il existe un réel A tel que $]-\infty, A[\subset V$.

Définition

Soit I un intervalle de $\mathbb{R}$.

On dit que ¹⁸ $a \in \mathbb{R} \cup \{-\infty, +\infty\}$ est un *point adhérent* ¹⁹ à I si ²⁰ tout voisinage de a rencontre I c'est-à-dire si tout voisinage de a contient au moins un point de I .

On dit que $a \in \mathbb{R}$ est un *point intérieur* de I si ²¹ I est un voisinage de a c'est-à-dire s'il existe un réel $\alpha > 0$ tel que $]a - \alpha, a + \alpha[\subset I$.

16. $\emptyset$ et $\mathbb{R}$ sont à la fois ouverts et fermés.

17. En particulier, pour tout réel $\alpha > 0$, $]a - \alpha, a + \alpha[$ est un voisinage de a .

18. $\mathbb{R} \cup \{-\infty, +\infty\}$ se note aussi $\overline{\mathbb{R}}$ et est appelé *droite numérique achevée*.

19. La notion de *point adhérent* n'est pas indispensable à la compréhension de ce chapitre. Elle n'interviendra donc qu'en notes de bas de page.

20. $a \in \mathbb{R} \cup \{-\infty, +\infty\}$ *adhérent* à I signifie en termes simples que a est un point de I ou une extrémité finie ou infinie de I .

21. a *point intérieur* de I signifie en termes simples que a est un point de I mais pas une extrémité de I .

Par exemple $+\infty$ et 0 sont adhérents à $]0, +\infty[$. 1 n'est pas un point intérieur de $[0, 1]$, ni de $[0, 1[$ mais tout réel $x \in]0, 1[$ est un point intérieur de $[0, 1]$.

Définition

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

1. Soit ²² a un point de I ou une extrémité finie de I .

On dit que f admet une *limite finie* en a si

$$\exists \ell \in \mathbb{R}, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies |f(x) - \ell| < \varepsilon.$$

2. On dit que f admet une *limite finie* en $+\infty$ si

$$\exists \ell \in \mathbb{R}, \forall \varepsilon > 0, \exists A \in \mathbb{R}, \forall x \in I, x > A \implies |f(x) - \ell| < \varepsilon.$$

On dit que f admet une *limite finie* en $-\infty$ si

$$\exists \ell \in \mathbb{R}, \forall \varepsilon > 0, \exists A \in \mathbb{R}, \forall x \in I, x < A \implies |f(x) - \ell| < \varepsilon.$$

Ainsi, dire que f admet une limite en a signifie : il existe $\ell \in \mathbb{R}$ tel que l'écart entre $f(x)$ et ℓ peut être rendu aussi petit que l'on veut pourvu que x soit suffisamment proche de a . La première définition ci-dessus peut également s'écrire de manière équivalente de la façon suivante :

$$\exists \ell \in \mathbb{R}, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, x \in]a - \alpha, a + \alpha[\implies f(x) \in]\ell - \varepsilon, \ell + \varepsilon[.$$

Définition

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

1. Soit a un point de I ou une extrémité finie de I .

On dit que f *tend vers* $+\infty$ en a si ²³

$$\forall A \in \mathbb{R}, \exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies f(x) > A.$$

2. On dit que f *tend vers* $+\infty$ en $+\infty$ si ²⁴

$$\forall A \in \mathbb{R}, \exists B \in \mathbb{R}, \forall x \in I, x > B \implies f(x) > A.$$

On dit que f *tend vers* $+\infty$ en $-\infty$ si

$$\forall A \in \mathbb{R}, \exists B \in \mathbb{R}, \forall x \in I, x < B \implies f(x) > A.$$

Une fonction f peut donc admettre une limite finie ou infinie en a mais aussi ne pas avoir de limite (cf. exemple page 39).

²². On pourrait aussi dire : soit a un réel *adhérent* à I .

²³. La définition de f *tend vers* $-\infty$ en a est analogue en remplaçant simplement $f(x) > A$ par $f(x) < A$.

²⁴. La définition de f *tend vers* $-\infty$ en $+\infty$ est analogue en remplaçant simplement $f(x) > A$ par $f(x) < A$.

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et a un point²⁵ de I ou une extrémité finie ou infinie de I . Si f admet une limite finie en a alors celle-ci est unique et sera notée²⁶ $\lim_a f$ ou $\lim_{x \rightarrow a} f(x)$.

Si f est définie en a et admet une limite finie en a alors $\lim_a f = f(a)$.

Définition

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

1. Soit a un point de I ou une extrémité finie de I .

On dit que f vérifie une certaine propriété *au voisinage* de a s'il existe $\alpha > 0$ tel que f vérifie cette propriété sur $I \cap]a - \alpha, a + \alpha[$.

2. On dit que f vérifie une certaine propriété *au voisinage* de $+\infty$ s'il existe $A \in \mathbb{R}$ tel que f vérifie cette propriété sur $I \cap]A, +\infty[$.

On dit que f vérifie une certaine propriété *au voisinage* de $-\infty$ s'il existe $A \in \mathbb{R}$ tel que f vérifie cette propriété sur $I \cap]-\infty, A[$.

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et a un point de I ou une extrémité finie ou infinie de I .

Si f admet une limite finie en a alors f est bornée au voisinage de a .

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et a un point de I ou une extrémité finie de I .

1. On dit que f admet une *limite finie à gauche* en a si²⁷

$$\exists \ell \in \mathbb{R}, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, a - \alpha < x < a \implies |f(x) - \ell| < \varepsilon.$$

Elle sera alors notée $\lim_{a^-} f$ ou $\lim_{x \rightarrow a^-} f(x)$ ou $\lim_{x < a} f(x)$ ou encore $f(a^-)$.

2. On dit que f admet une *limite finie à droite* en a si²⁸

$$\exists \ell \in \mathbb{R}, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, a < x < a + \alpha \implies |f(x) - \ell| < \varepsilon.$$

Elle sera alors notée $\lim_{a^+} f$ ou $\lim_{x \rightarrow a^+} f(x)$ ou $\lim_{x > a} f(x)$ ou encore $f(a^+)$.

25. On peut aussi dire : soit a un point de $\mathbb{R} \cup \{-\infty, +\infty\}$ adhérent à I .

26. Si ℓ désigne cette limite, on écrira également $f \rightarrow \ell$ ou $f(x) \xrightarrow{x \rightarrow a} \ell$.

27. Autrement dit : la restriction de f à $I \cap]-\infty, a[$ admet une limite finie en a .

28. Autrement dit : la restriction de f à $I \cap]a, +\infty[$ admet une limite finie en a .

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et a un point de I ou une extrémité finie de I .

1. On suppose que f est définie en a . Alors f admet une limite finie $\ell \in \mathbb{R}$ en a si, et seulement si, f admet une limite finie à gauche et à droite en a égales à $f(a)$.
2. On suppose que f n'est pas définie en a . Alors f admet une limite finie $\ell \in \mathbb{R}$ en a si, et seulement si, f admet une limite finie à gauche et à droite en a égales à ℓ .

Donnons deux exemples de ce théorème.

$$\text{Soient } f : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sinon} \end{cases} \end{cases} \quad \text{et } g : \begin{cases} \mathbb{R}^* & \longrightarrow \mathbb{R} \\ x & \longmapsto 1 \end{cases}.$$

On a $f(0) = 1$ et des limites à droite et à gauche de f en 0 égales à $0 \neq f(0)$.

Donc f n'a pas de limite en 0. D'autre part, g n'est pas définie en 0, admet des limites à gauche et à droite égales à 1 donc g admet 1 comme limite en 0.

Théorème (caractérisation séquentielle de la limite)

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$, a un point de I ou une extrémité finie ou infinie de I et $\ell \in \mathbb{R} \cup \{-\infty, +\infty\}$.

Alors les assertions suivantes sont équivalentes :

- (i) $\lim_{x \rightarrow a} f(x) = \ell$.
- (ii) Pour toute suite (x_n) de I convergeant vers a , la suite $(f(x_n))$ converge vers ℓ .

Théorème (combinaison linéaire, produit et quotient des limites)

Soient I un intervalle de $\mathbb{R}$, a un point de I ou une extrémité finie ou infinie de I , $\lambda \in \mathbb{R}$, f et g deux fonctions de I vers $\mathbb{R}$ admettant des limites finies ℓ et ℓ' en a . Alors

- $\lambda f + g$ admet une limite finie en a et $\lim_a (\lambda f + g) = \lambda \ell + \ell'$;
- fg admet une limite finie en a et $\lim_a fg = \ell \ell'$;
- si $\ell' \neq 0$, $\frac{f}{g}$ admet²⁹ une limite finie en a et $\lim_a \frac{f}{g} = \frac{\ell}{\ell'}$.

29. On montre facilement que comme $\ell' \neq 0$, g ne s'annule pas au voisinage de a .

Théorème (composition des limites)

Soient I et J deux intervalles de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ telle que $f(I) \subset J$, g une fonction de J vers $\mathbb{R}$, a un point de I ou une extrémité finie ou infinie de I . Si $\lim_a f = b$ alors

- b est un point de J ou une extrémité finie ou infinie de J ;
- si $\lim_b g = \ell$ alors $\lim_a (g \circ f) = \ell$.

Énonçons le théorème relatif au passage à la limite dans des inégalités larges.

Théorème

Soient I un intervalle de $\mathbb{R}$, a un point de I ou une extrémité finie ou infinie de I , f et g deux fonctions de I vers $\mathbb{R}$ admettant des limites finies ℓ et ℓ' en a .

Si au voisinage de a , $f \leq g$, alors $\ell \leq \ell'$.

Ce théorème devient faux pour les inégalités strictes : on a par exemple pour tout $x \in \mathbb{R}_+^*$, $1 - \frac{1}{x} < 1$ mais $\lim_{x \rightarrow +\infty} (1 - \frac{1}{x}) = 1$ n'est pas strictement inférieure à 1.

Théorème (limite finie par encadrement)

Soient I un intervalle de $\mathbb{R}$, a un point de I ou une extrémité finie ou infinie de I , f , g et h trois fonctions de I vers $\mathbb{R}$.

Si³⁰ au voisinage de a , $f \leq h \leq g$ avec $\lim_a f = \ell \in \mathbb{R}$ et $\lim_a g = \ell$ alors $\lim_a h = \ell$.

Théorème (limite infinie par minoration ou majoration)

Soient I un intervalle de $\mathbb{R}$, a un point de I ou une extrémité finie ou infinie de I , f et g deux fonctions de I vers $\mathbb{R}$ vérifiant $g \geq f$ au voisinage de a .

- Si $\lim_a f = +\infty$ alors $\lim_a g = +\infty$.
- Si $\lim_a g = -\infty$ alors $\lim_a f = -\infty$.

30. L'hypothèse sera le plus souvent : pour tout $x \in I$, $f(x) \leq h(x) \leq g(x)$ mais cette inégalité seulement au voisinage de a est suffisante pour aboutir à la conclusion.

Théorème de la limite monotone

Soient $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{+\infty\}$ tels que $a < b$ et f une fonction de $]a, b[$ vers $\mathbb{R}$ monotone. Alors

- $\lim_a f$ et $\lim_b f$ existent et sont finies ou infinies³¹ ;
- pour tout $c \in]a, b[$, $\lim_{c^-} f$ et $\lim_{c^+} f$ existent et

$$\lim_{c^-} f \leq f(c) \leq \lim_{c^+} f \quad \text{si } f \text{ est croissante}$$

$$\lim_{c^+} f \leq f(c) \leq \lim_{c^-} f \quad \text{si } f \text{ est décroissante.}$$

2.4 Extension aux fonctions à valeurs complexes

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{C}$ et a un point de I ou une extrémité finie de I .

On dit que f admet une *limite* finie $\ell \in \mathbb{C}$ en a

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies |f(x) - \ell| < \varepsilon.$$

L'unicité de cette éventuelle limite se démontre de la même manière que le cas d'une fonction à valeurs réelles.

De même, on peut étendre la notion de limite en $+\infty$ et $-\infty$ avec exactement la même définition que pour les fonctions à valeurs réelles à ceci près que $|f(x) - \ell|$ désigne désormais le module de $f(x) - \ell$.

En revanche, la notion de limite infinie n'a plus de sens dans $\mathbb{C}$.

Lorsque f est une fonction de I vers $\mathbb{C}$, on définit les fonctions $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$ de I vers $\mathbb{R}$ pour tout $x \in I$ par

$$(\operatorname{Re}(f))(x) = \operatorname{Re}(f(x)) \quad \text{et} \quad (\operatorname{Im}(f))(x) = \operatorname{Im}(f(x)).$$

Ces deux nouvelles fonctions à valeurs réelles vérifient : $f = \operatorname{Re}(f) + i \operatorname{Im}(f)$.

31. Plus précisément, dans le cas par exemple d'une fonction f croissante, $\lim_a f$ est finie (et égale à la borne inférieure de f sur $]a, b[$) si f est minorée et égale à $-\infty$ sinon ; $\lim_b f$ est finie (et égale à la borne supérieure de f sur $]a, b[$) si f est majorée et égale à $+\infty$ sinon. Ce théorème est en fait l'analogue pour les fonctions du *théorème de la limite monotone* pour les suites : toute suite réelle croissante (respectivement décroissante) et majorée (respectivement minorée) converge.

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{C}$ et a un point de I ou une extrémité finie ou infinie de I .

Alors $\lim_a f = \ell \in \mathbb{C}$ si, et seulement si, $\begin{cases} \lim_a \operatorname{Re}(f) = \operatorname{Re}(\ell) \\ \lim_a \operatorname{Im}(f) = \operatorname{Im}(\ell) \end{cases}$

On peut également étendre la notion de fonction bornée d'une fonction f de I vers $\mathbb{C}$ de la façon suivante : f est *bornée* sur I si

$$\exists K \in \mathbb{R}_+, \forall x \in I, |f(x)| \leq K.$$

avec $|f(x)|$ qui désigne désormais le module de $f(x)$.

Comme pour le cas d'une fonction à valeurs réelles, toute fonction d'un intervalle I à valeurs complexes admettant une limite finie en un point a de I (ou une extrémité finie ou infinie de I) est bornée au voisinage de a .

Les notions de limite à gauche et à droite sont également étendues de même que la caractérisation séquentielle de la limite.

2.5 Continuité

À partir de maintenant, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et $a \in I$. On dit que³² f est *continue en a* si $\lim_a f = f(a)$ c'est-à-dire si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies |f(x) - f(a)| < \varepsilon.$$

On dit que f est *continue sur I* si f est continue en tout point de I .

Contrairement à la définition de la limite d'une fonction f de I vers $\mathbb{K}$ en un point a où a n'est pas nécessairement un point de I , la notion de continuité en un point a n'a pas de sens si ce point n'est pas dans I .

En vertu du théorème page 38, on pourrait simplement donner comme définition de la continuité de f , l'existence d'une limite finie de f en a .

32. Il existe des fonctions qui ne sont continues en aucun point de $\mathbb{R}$. C'est le cas par exemple de la *fonction caractéristique* de $\mathbb{Q}$, appelée également *fonction de Dirichlet* et notée parfois $1_{\mathbb{Q}}$,

définie par :
$$\begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{sinon} \end{cases} \end{cases}.$$

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{C}$ et $a \in I$.

Alors f est *continue en a* (respectivement sur I) si, et seulement si, les fonctions à valeurs réelles $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$ sont continues en a (respectivement sur I).

Définition (prolongement par continuité)

Soient I un intervalle de $\mathbb{R}$, a un réel n'appartenant pas à I et f une fonction de I vers $\mathbb{K}$.

On dit que f est *prolongeable par continuité en a* si $\lim_a f$ existe et est finie.

On définit alors le *prolongement* $\tilde{f}$ de f , continu sur $I \cup \{a\}$, défini par

$$\tilde{f} : \begin{cases} I \cup \{a\} & \longrightarrow \mathbb{K} \\ x & \longmapsto \begin{cases} f(x) & \text{si } x \neq a \\ \lim_a f & \text{sinon} \end{cases} \end{cases}$$

Par exemple, la fonction $f : x \mapsto \frac{\sin(x)}{x}$ définie sur l'intervalle $\mathbb{R}_+^*$ se prolonge par continuité en 0 car $\lim_0 f = 1$.

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et $a \in I$.

On dit que f est *continue à gauche en a* si $\lim_{a^-} f = f(a)$.

On dit que f est *continue à droite en a* si $\lim_{a^+} f = f(a)$.

Par exemple la fonction *partie entière*³³ est continue à droite en 1 mais n'est pas continue à gauche en 1 car $\lfloor 1 \rfloor = 1$ et $\lim_{x \rightarrow 1^-} \lfloor x \rfloor = 0 \neq \lfloor 1 \rfloor$.

Théorème

Soient I un intervalle de $\mathbb{R}$, $a \in I$ qui n'est pas une extrémité de I et f une fonction de I vers $\mathbb{K}$. Alors f est continue en a si, et seulement si, f est continue à gauche et à droite en a .

Donnons à présent la caractérisation séquentielle de la continuité en un point.

33. On rappelle que la fonction *partie entière*, notée $\lfloor \cdot \rfloor$ est définie pour tout $x \in \mathbb{R}$ par le plus grand entier relatif inférieur ou égal à x . Elle vérifie en particulier pour tout $x \in \mathbb{R}$: $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

Théorème (caractérisation séquentielle de la continuité)

Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{K}$.

Alors les assertions suivantes sont équivalentes :

- (i) f est continue en a .
- (ii) Pour toute suite (x_n) de I convergeant vers a , la suite $(f(x_n))$ converge vers $f(a)$.

Théorème

Soient I un intervalle de $\mathbb{R}$, $a \in I$, $\lambda \in \mathbb{K}$, f et g deux fonctions de I vers $\mathbb{K}$ continues en a (respectivement sur I). Alors

- $\lambda f + g$ est continue en a (respectivement sur I) ;
- fg est continue en a (respectivement sur I) ;
- si $g(a) \neq 0$, g ne s'annule pas au voisinage de a et $\frac{f}{g}$ est continue en a .

Théorème

Soient I et J deux intervalles de $\mathbb{R}$, $a \in I$, f une fonction de I vers $\mathbb{R}$ telle que $f(I) \subset J$ et g une fonction de J vers $\mathbb{K}$.

Si f est continue en a (respectivement sur I) et g est continue sur $f(a)$ (respectivement sur J) alors $g \circ f$ est continue en a (respectivement sur I).

2.6 Théorèmes relatifs à la continuité

Dans cette section, toutes les fonctions sont à valeurs dans $\mathbb{R}$.

Théorème des valeurs intermédiaires

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ avec $a < b$ et f une fonction de I vers $\mathbb{R}$ continue sur I .

1. Tout réel λ strictement compris entre $f(a)$ et $f(b)$ possède au moins un antécédent par f sur $]a, b[$ autrement dit³⁴ il existe au moins un réel $c \in]a, b[$ tel que $f(c) = \lambda$.
2. Si f est de plus strictement monotone, tout réel λ strictement compris entre $f(a)$ et $f(b)$ possède un unique antécédent par f sur $]a, b[$ autrement dit³⁵ il existe un unique réel $c \in]a, b[$ tel que $f(c) = \lambda$.

34. ou encore l'équation $f(x) = \lambda$ admet au moins une solution dans $]a, b[$.

35. ou encore l'équation $f(x) = \lambda$ admet une unique solution dans $]a, b[$.

Énonçons son corollaire immédiat, également appelé parfois *théorème des valeurs intermédiaires*.

Corollaire

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ avec $a < b$ et f une fonction de I vers $\mathbb{R}$ continue sur I .

Si ³⁶ $f(a)f(b) < 0$, il existe au moins ³⁷ un réel $c \in]a, b[$ tel que $f(c) = 0$

Énonçons un autre corollaire du théorème des valeurs intermédiaires concernant l'image d'un intervalle par une fonction continue.

Corollaire

L'image d'un intervalle ³⁸ par une fonction continue est un intervalle.

L'hypothèse de continuité est indispensable pour assurer la validité du théorème. Il suffit par exemple de reprendre l'exemple de la fonction *partie entière* : en la notant E , on a $E([0, 1]) = \{0, 1\}$.

L'image d'un intervalle borné par une fonction continue n'est pas nécessairement un intervalle borné : prenons par exemple la fonction $f : \begin{cases}]0, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto \frac{1}{x} \end{cases}$ pour laquelle on a $f([0, 1]) = [1, +\infty[$.

En ajoutant la condition de stricte monotonie, on peut affiner le corollaire précédent qui donne le théorème suivant.

Théorème

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ avec $a < b$ et f une fonction de I vers $\mathbb{R}$ continue sur I .

Si f est strictement croissante, $f([a, b]) = [f(a), f(b)]$.

Si f est strictement décroissante, $f([a, b]) = [f(b), f(a)]$.

On peut encore prolonger le théorème précédent à des intervalles qui ne sont pas des segments : avec les mêmes hypothèses que le théorème, si f est strictement croissante, par exemple $f([a, b[) = [f(a), \lim_b f[$ ou encore $f(]a, +\infty[) =]\lim_a f, \lim_{+\infty} f[$.

36. ce qui signifie $f(a)$ et $f(b)$ de signes opposés.

37. Si f est de plus strictement monotone, ce réel est unique.

38. Ce corollaire est également vérifié dans le cas d'un intervalle réduit à un point : par exemple, pour la fonction constante égale à 1 sur $\mathbb{R}$, on a $f(\mathbb{R}) = \{1\}$.

Théorème des bornes atteintes

Soit f une fonction de $[a, b]$ vers $\mathbb{R}$ continue sur le segment $[a, b]$.
Alors f est bornée sur $[a, b]$ et atteint ses bornes³⁹.

Corollaire

L'image d'un segment par une fonction continue est un segment⁴⁰.

Le théorème n'est plus vérifié si la fonction n'est pas continue sur l'intégralité du segment.

Il suffit pour s'en convaincre de reprendre l'exemple de la *partie entière* : en la notant E , elle est continue sur $[0, 1[$ mais pas sur $[0, 1]$ et on a $E([0, 1]) = \{0, 1\}$.

D'autre part, pour une telle fonction f continue sur le segment $[a, b]$, on n'a pas nécessairement $f([a, b]) = [f(a), f(b)]$.

Par exemple en prenant la fonction $f : x \mapsto x^2$, on a $f([-1, 2]) = [0, 4]$ et $[0, 4] \neq [f(-1), f(2)]$.

Le théorème qui suit est une autre conséquence du théorème des valeurs intermédiaires.

Théorème

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

Si f est injective et continue sur I alors f est strictement monotone sur I .

Pour conclure cette section, énonçons le théorème relatif à la continuité de la fonction réciproque.

Théorème

Soient I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{R}$ continue et strictement monotone sur I . Alors

1. f est bijective de I sur $f(I)$.
2. La bijection réciproque f^{-1} de f est continue et strictement monotone sur $f(I)$, de même monotonie que f .

39. En particulier, f admet un minimum et un maximum.

40. Le segment peut être réduit à un point : par exemple, en prenant la fonction f constante égale à 1 sur $\mathbb{R}$, $f([0, 1]) = \{1\}$ où $\{1\}$ est le segment réduit à un point.

2.7 Continuité par morceaux

Définition

Soient $n \in \mathbb{N}^*$, a et b deux réels tel que $a < b$.

On appelle *subdivision* de $[a, b]$ toute famille $(x_0, \dots, x_n) \in [a, b]^{n+1}$ telle que : $a = x_0 < x_1 < \dots < x_n = b$.

On appelle *pas* de la subdivision le réel strictement positif $\max_{i \in [0, n-1]} (x_{i+1} - x_i)$.

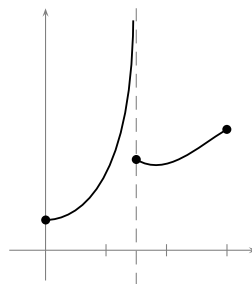
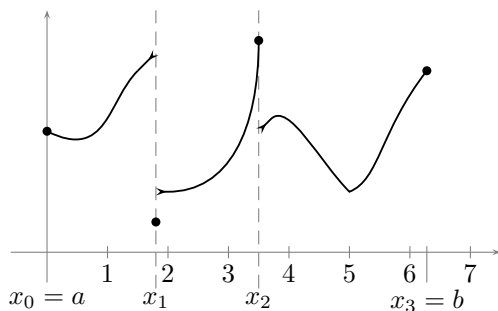
Lorsque toutes les différences $x_{i+1} - x_i$ sont égales, on dit que la subdivision est *régulière*.

Définition

Soient a et b deux réels tel que $a < b$ et f une fonction de $[a, b]$ vers $\mathbb{K}$.

On dit que f est *continue par morceaux* sur $[a, b]$ si f est continue sur $[a, b]$ sauf en un nombre fini (éventuellement nul⁴¹) de points en lesquels f admet une limite à droite et à gauche c'est-à-dire s'il existe $n \in \mathbb{N}^*$ et une subdivision $(x_0, \dots, x_n)$ de $[a, b]$ tels que f est continue sur chaque $]x_i, x_{i+1}[$ et admet⁴² des limites finies à droite en x_i et à gauche en x_{i+1} .

1



Une fonction continue par morceaux ...

... et une qui ne l'est pas.

41. Cette précision permet de voir immédiatement qu'une fonction *continue* est *continue par morceaux*.

42. On peut écrire plus simplement f admet des limites finies à droite et à gauche en chaque x_i mais il faut alors préciser qu'en $x_0 = a$ et $x_n = b$, f n'admet respectivement qu'une limite à droite et à gauche. Cette limite finie de f à droite et à gauche en chaque x_i n'est pas nécessairement égale à $f(x_i)$.

On peut aussi formuler la définition de la façon suivante : f est *continue par morceaux* sur $[a, b]$ s'il existe $n \in \mathbb{N}^*$ et $(x_0, \dots, x_n)$ une subdivision de $[a, b]$ tels que pour tout $i \in \llbracket 0, n-1 \rrbracket$, la restriction de f à $]x_i, x_{i+1}[$ est continue sur $]x_i, x_{i+1}[$ et prolongeable par continuité en x_i et x_{i+1} .

La fonction *partie entière* est par exemple continue par morceaux sur tout segment $[a, b]$. En revanche, la fonction $x \mapsto \frac{1}{x}$, certes prolongeable en 0 par un réel quelconque, n'est pas continue par morceaux sur $[0, 1]$ car ce prolongement est impossible par continuité.

Théorème

Soient a et b deux réels tels que $a < b$ et f une fonction de $[a, b]$ vers $\mathbb{K}$ continue par morceaux sur $[a, b]$. Alors f est bornée⁴³ sur $[a, b]$.

La notion de continuité par morceaux peut s'étendre à tout intervalle de $\mathbb{R}$ comme le précise la définition suivante.

Définition

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

On dit que f est *continue par morceaux* sur I si f est continue par morceaux sur tout segment inclus dans I .

2.8 Continuité uniforme et fonction lipschitzienne

Rappelons tout d'abord la définition de la continuité sur un intervalle : f est continue sur I si : $\forall x \in I, \forall \varepsilon > 0, \exists \alpha > 0, \forall y \in I, |x - y| < \alpha \implies |f(x) - f(y)| < \varepsilon$.

Définition

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

On dit que f est *uniformément continue* sur I si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall (x, y) \in I^2, |x - y| < \alpha \implies |f(x) - f(y)| < \varepsilon.$$

Contrairement à la définition de la continuité, le α ne dépend⁴⁴ pas du point x choisi dans I mais ne dépend que du ε .

43. mais n'atteint pas nécessairement ses bornes. Par exemple la fonction $f : \begin{cases} [0, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} x & \text{si } x \in [0, 1[\\ 0 & \text{si } x = 1 \end{cases} \end{cases}$ est continue par morceaux et bornée sur $[0, 1]$ mais ne peut atteindre son maximum en l'absence de ce dernier.

44. c'est-à-dire pour la continuité, la valeur de α est *a priori* modifiée à chaque nouvelle valeur de x dans I contrairement à la continuité uniforme où cet α est le même pour tous les x dans I .

Toute fonction uniformément continue sur I est continue sur I mais la réciproque est fausse : par exemple, la fonction $x \mapsto x^2$ est continue sur $\mathbb{R}_+$ mais n'est pas uniformément continue sur $\mathbb{R}_+$.

La somme de deux fonctions uniformément continues, le produit d'une fonction uniformément continue par un complexe et la composée de deux fonctions uniformément continues⁴⁵ sont uniformément continues.

En revanche, le produit de deux fonctions uniformément continues n'est pas nécessairement uniformément continue : prenons par exemple la fonction $f : x \mapsto x$. Alors f est uniformément continue sur $\mathbb{R}_+$ mais, comme on l'a déjà remarqué plus haut, $f^2 : x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}_+$.

Définition

Soient I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{K}$ et $K \in \mathbb{R}_+$. On dit que f est *K-lipschitzienne* sur I si

$$\forall (x, y) \in I^2, |f(x) - f(y)| \leq K|x - y|.$$

On dit que f est *lipschitzienne* sur I s'il existe $K \in \mathbb{R}_+$ tel que f est *K-lipschitzienne* sur I .

On dit que f est *contractante* sur I s'il existe $K \in [0, 1[$ tel que f est *K-lipschitzienne* sur I .

Théorème

Soient I un intervalle de $\mathbb{R}$ et f une fonction lipschitzienne de I vers $\mathbb{K}$. Alors f est uniformément continue sur I .

La réciproque du théorème ci-dessus est fausse.

Par exemple la fonction $f : x \mapsto \sqrt{x}$ est uniformément continue sur $\mathbb{R}_+$ mais n'est pas lipschitzienne.

Théorème de Heine

Toute fonction continue sur un segment $[a, b]$ est uniformément continue sur ce segment.

2.9 Dérivabilité

45. à condition que la définition de cette composée soit possible c'est-à-dire que l'intervalle d'arrivée de l'une soit incluse dans celui de départ de l'autre.

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et $a \in I$.

On dit que f est *dérivable en a* si $\frac{f(x)-f(a)}{x-a}$, appelé *taux d'accroissement* de f entre a et x , admet une limite finie lorsque x tend vers a .

Cette limite est alors notée $f'(a)$ et appelée *nombre dérivé* de f en a .

Lorsque f est dérivable en tout point de I , on appelle *dérivée* de f la fonction notée f' définie par $f' : \begin{cases} I & \longrightarrow \mathbb{K} \\ x & \longmapsto f'(x) \end{cases}$.

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et $a \in I$. Alors les assertions suivantes sont équivalentes :

- (i) f est dérivable en a .
- (ii) $\frac{f(a+h)-f(a)}{h}$ admet une limite finie lorsque h tend vers 0.
- (iii) Il existe $\ell \in \mathbb{R}$ et une fonction ε de I vers $\mathbb{K}$ tels que $\lim_0 \varepsilon = 0$ et pour tout réel h vérifiant $a+h \in I$,

$$f(a+h) = f(a) + \ell h + h\varepsilon(h).$$

Lorsque f est dérivable en a , le réel ℓ du (iii) est $f'(a)$ et l'égalité

$$f(a+h) = f(a) + f'(a)h + h\varepsilon(h)$$

est appelée *développement limité* de f à l'ordre 1 en a .

Théorème

Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{K}$ dérivable en a . Alors f est continue en a .

La réciproque de ce théorème est fautive⁴⁶ comme le montre l'exemple de la fonction valeur absolue qui n'est pas dérivable en 0.

46. Il existe même des fonctions continues en tout point de $\mathbb{R}$ et dérivables en aucun point de $\mathbb{R}$. C'est le cas par exemple de la fonction : $\begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \sum_{n=0}^{+\infty} \frac{1}{2^n} \cos(3^n x) \end{cases}$ qui utilise la notion de *série de fonctions*.

Théorème

Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{K}$ dérivable en a .

Alors f dérivable en a si, et seulement si, $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$ sont dérivables en a .

De plus, dans ce cas, on a $f'(a) = (\operatorname{Re}(f))'(a) + i (\operatorname{Im}(f))'(a)$.

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et a un point intérieur⁴⁷ de I .

On dit que f est *dérivable à gauche* en a si $\frac{f(x)-f(a)}{x-a}$ admet une limite finie lorsque x tend vers a^- . Cette limite est alors notée $f'_g(a)$.

On dit que f est *dérivable à droite* en a si $\frac{f(x)-f(a)}{x-a}$ admet une limite finie lorsque x tend vers a^+ . Cette limite est alors notée $f'_d(a)$.

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et a un point intérieur de I .

Alors f dérivable en a si, et seulement si, f est dérivable à gauche et à droite en a avec $f'_g(a) = f'_d(a)$.

Les trois théorèmes qui suivent établissent les opérations sur les fonctions dérivables.

Théorème

Soient I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$, dérivables sur I , et $\lambda \in \mathbb{K}$. Alors

- $\lambda f + g$ est dérivable sur I et $(\lambda f + g)' = \lambda f' + g'$.
- fg est dérivable sur I et $(fg)' = f'g + fg'$.
- Si g ne s'annule pas sur I , $\frac{f}{g}$ est dérivable sur I et $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.

47. On peut aussi choisir comme hypothèse I intervalle ouvert de $\mathbb{R}$ et prendre alors un point quelconque a de I .

Théorème

Soient I et J deux intervalles de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ telle que $f(I) \subset J$ et g une fonction de J vers $\mathbb{K}$.

Si f est dérivable sur I et g dérivable sur J , $g \circ f$ est dérivable sur I et

$$(g \circ f)' = (g' \circ f) \times f'.$$

Théorème

Soient I un intervalle de $\mathbb{R}$, $J = f(I)$ et f une fonction de I vers $\mathbb{R}$ bijective de I sur J et dérivable sur I .

Si f' ne s'annule pas sur I , alors f^{-1} est dérivable sur J et

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$

L'hypothèse « f' ne s'annule pas sur I » est essentielle.

Par exemple la fonction $f : x \mapsto x^2$ est dérivable sur $\mathbb{R}_+$ mais $f^{-1} : x \mapsto \sqrt{x}$ n'est pas dérivable sur $\mathbb{R}_+$ car non dérivable en 0.

2.10 Fonctions de classe $\mathcal{C}^k$

Définition

Soient $k \in \mathbb{N}$, I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{K}$.

On dit que f est *k fois dérivable*⁴⁸ sur I si on peut déterminer sur I les dérivées successives de f jusqu'à sa *dérivée k -ième*⁴⁹ notée $f^{(k)}$.

Définition

Soient $k \in \mathbb{N}$, I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{K}$.

On dit que f est de *classe $\mathcal{C}^k$* sur I si f est k fois dérivable sur I et $f^{(k)}$ est continue sur I .

On dit que f est ⁵⁰ de *classe $\mathcal{C}^\infty$* sur I si, pour tout $k \in \mathbb{N}$, f est de classe $\mathcal{C}^k$ sur I .

On note ⁵¹ $\mathcal{C}^k(I, \mathbb{K})$ (respectivement $\mathcal{C}^\infty(I, \mathbb{K})$) l'ensemble des fonctions de classe $\mathcal{C}^k$ (respectivement de classe $\mathcal{C}^\infty$) sur I .

48. Pour $k = 0$, $f^{(0)} = f$ est appelée *dérivée d'ordre 0* (ou *dérivée 0-ième*) de f .

49. ou *dérivée d'ordre k* .

50. On dit aussi f est *indéfiniment dérivable*.

51. En particulier $\mathcal{C}^0(I, \mathbb{K})$ désigne l'ensemble des fonctions continues sur I .

Comme la dérivabilité implique la continuité, si f est k fois dérivable sur I , f , f' , f'' , $\dots$, $f^{(k-1)}$ sont continues sur I mais la continuité de $f^{(k)}$ sur I n'est pas assurée.

Par exemple, on peut montrer que la fonction $f : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} x^3 \sin\left(\frac{1}{x^2}\right) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases} \end{cases}$

est dérivable sur $\mathbb{R}$ mais f' n'est pas dérivable sur $\mathbb{R}$ car non dérivable en 0.

Théorème

Soient $k \in \mathbb{N}$, $\lambda \in \mathbb{K}$, I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$, de classe $\mathcal{C}^k$ (respectivement $\mathcal{C}^\infty$) sur I .

Alors $\lambda f + g$ est de classe $\mathcal{C}^k$ (respectivement $\mathcal{C}^\infty$) sur I et

$$(\lambda f + g)^{(k)} = \lambda f^{(k)} + g^{(k)}.$$

Théorème (formule de Leibniz)

Soient $k \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$, de classe $\mathcal{C}^k$ (respectivement $\mathcal{C}^\infty$) sur I .

Alors fg est de classe $\mathcal{C}^k$ (respectivement $\mathcal{C}^\infty$) sur I et on a la *formule de Leibniz* suivante :

$$(fg)^{(k)} = \sum_{i=0}^k \binom{k}{i} f^{(i)} g^{(k-i)}.$$

Si g ne s'annule pas sur I , $\frac{f}{g}$ est de classe $\mathcal{C}^k$ (respectivement $\mathcal{C}^\infty$) sur I .

Théorème

Soient $k \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, $J = f(I)$ et f une fonction de I vers $\mathbb{R}$ bijective⁵² de I sur J et de classe $\mathcal{C}^k$ (respectivement de classe $\mathcal{C}^\infty$) sur I .

Si f' ne s'annule pas sur I , alors f^{-1} est de classe $\mathcal{C}^k$ (respectivement de classe $\mathcal{C}^\infty$) sur J .

2.11 Extremum local

Dans cette section, les fonctions sont à valeurs réelles.

⁵². C'est notamment le cas si f est continue et strictement monotone sur I .

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et $a \in I$.

On dit que f admet un *maximum local* en a si

$$\exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies f(x) \leq f(a).$$

On dit que f admet un *minimum local* en a si

$$\exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies f(x) \geq f(a).$$

On dit que f admet un *extremum local* en a si f admet un maximum ou un minimum local en a .

On dit que a est un *point critique* de f si f est dérivable en a et $f'(a) = 0$.

Théorème (condition nécessaire d'un extremum local)

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et a un point intérieur⁵³ de I . Si f est dérivable en a et admet un extremum local en a alors a est un point critique de f .

L'hypothèse « a point intérieur de I » (ou bien « I intervalle ouvert ») est fondamentale pour assurer la validité du théorème. Par exemple la fonction $f : x \mapsto x$ est dérivable en 0 et 1, admet un minimum en 0, un maximum en 1 et pourtant $f'(0) \neq 0$ et $f'(1) \neq 0$.

Une fonction f peut admettre un extremum en a sans être dérivable en a . Par exemple la fonction $x \mapsto \sqrt{x}$ admet un minimum en 0 mais n'est pas dérivable en 0.

La réciproque du théorème est fautive : la fonction $f : \begin{cases} [-1, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto x^3 \end{cases}$ n'admet pas d'extremum et pourtant $f'(0) = 0$ et 0 est bien un point intérieur de $[-1, 1]$.

Théorème (condition suffisante d'un extremum local)

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et a un point intérieur de I . Si a est un point critique de f et f' s'annule en a en changeant de signe⁵⁴, alors f admet un extremum local en a .

53. On peut aussi supposer I intervalle ouvert et prendre alors un point quelconque a de I .

54. Dire que f' s'annule en a en changeant de signe signifie $f' > 0$ à gauche de a et $f' < 0$ à droite de a ou inversement c'est-à-dire : ou bien $f' > 0$ sur un certain intervalle $]a - \alpha, a[$ (avec $\alpha > 0$) et $f' < 0$ sur $]a, a + \alpha[$ (auquel cas f admet un maximum local en a) ou bien $f' < 0$ sur un certain intervalle $]a - \alpha, a[$ et $f' > 0$ sur $]a, a + \alpha[$ (auquel cas f admet un minimum local en a). Notons qu'il existe $\alpha > 0$ tel que $]a - \alpha, a[$ et $]a, a + \alpha[$ soient bien inclus dans I car a est un point intérieur de I .

2.12 Théorèmes relatifs à la dérivabilité

Dans cette section, les fonctions sont également à valeurs réelles.

Théorème de Rolle

Soient a et b deux réels tels que $a < b$, f une fonction de $[a, b]$ vers $\mathbb{R}$, continue sur $[a, b]$, dérivable sur $]a, b[$ et tels que $f(a) = f(b)$.

Alors il existe (au moins) un réel c dans $]a, b[$ tel que $f'(c) = 0$.

Très souvent, f est dérivable aussi en a et b mais cette hypothèse n'est pas nécessaire pour la validité du théorème.

En revanche, la continuité de f en a et b est indispensable.

Par exemple, la fonction $f : \begin{cases} [0, 1] & \longrightarrow [0, 1] \\ x & \longmapsto \begin{cases} x & \text{si } x \in [0, 1[\\ 0 & \text{sinon} \end{cases} \end{cases}$ est dérivable (donc

continue) sur $]0, 1[$ avec $f(0) = f(1) = 0$ mais pour tout $c \in]0, 1[$, $f'(c) = 1 \neq 0$. Si on retire, dans les hypothèses du théorème de Rolle, l'hypothèse $f(a) = f(b)$, on obtient l'égalité du théorème qui suit.

Théorème (égalité des accroissements finis)

Soient a et b deux réels tels que $a < b$, f une fonction de $[a, b]$ vers $\mathbb{R}$, continue sur $[a, b]$ et dérivable sur $]a, b[$.

Alors il existe (au moins) un réel c dans $]a, b[$ tel que

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

Dans le théorème suivant, les fonctions sont éventuellement à valeurs complexes.

Théorème (inégalité des accroissements finis)

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ dérivable sur I et tels que $|f'|$ est majorée par un réel K sur I .

Alors f est K -lipschitzienne sur I c'est-à-dire

$$\forall (x, y) \in I^2, |f(x) - f(y)| \leq K|x - y|.$$

Dans le cas particulier d'une fonction f de $[a, b]$ vers $\mathbb{R}$ (avec $a < b$) continue sur $[a, b]$, dérivable sur $]a, b[$ telle qu'il existe deux réels m et M vérifiant $m \leq f' \leq M$, on a alors

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

Les deux théorèmes qui suivent caractérisent les fonctions constantes et monotones sur un intervalle à l'aide de leurs dérivées sachant que le premier est valable pour les fonctions à valeurs complexes.

Théorème

Soient I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{K}$ dérivable sur I . Alors f est constante sur I si, et seulement si, f' est nulle sur I .

Théorème

Soient I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{R}$ dérivable sur I . Alors f est croissante (respectivement décroissante) sur I si, et seulement si, f' est positive ou nulle (respectivement négative ou nulle) sur I .

Une condition nécessaire et suffisante de stricte monotonie à l'aide la dérivée est plus subtile.

Énonçons d'abord une condition suffisante.

Théorème

Soient I un intervalle⁵⁵ de $\mathbb{R}$ et f une fonction de I vers $\mathbb{R}$ dérivable sur I . Si f' est strictement positive (respectivement strictement négative) sur I , sauf éventuellement en un nombre fini de points⁵⁶, alors f est strictement croissante (respectivement décroissante) sur I .

Ce théorème est faux si I n'est pas intervalle.

Par exemple la fonction $f : \begin{cases} \mathbb{R}^* & \longrightarrow \mathbb{R} \\ x & \longmapsto \frac{1}{x} \end{cases}$ est de dérivée strictement négative sur tout $\mathbb{R}^*$ car pour tout $x \in \mathbb{R}^*$, $f'(x) = -\frac{1}{x^2}$ mais f n'est pas strictement décroissante sur $\mathbb{R}^*$ puisque $f(-1) = -1 < 1 = f(1)$.

Le théorème qui suit donne enfin une condition nécessaire et suffisante de stricte monotonie.

55. On rappelle que tous les intervalles de ce chapitre sont non vides et non réduits à un point.

56. En particulier, si f' est strictement positive (respectivement négative) sur tout I , f est strictement croissante (respectivement décroissante) sur I .

Théorème

Soient I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{R}$ dérivable sur I de dérivée positive (respectivement négative) sur I .

Alors f est strictement croissante (respectivement strictement décroissante) sur I si, et seulement si⁵⁷, $\{[a, b] \subset I, a < b, \text{ tel que } f' = 0 \text{ sur } [a, b]\} = \emptyset$.

Énonçons enfin le *théorème de la limite de la dérivée* également appelée *théorème de dérivabilité d'un prolongement*.

Théorème de la limite de la dérivée

Soient I un intervalle de $\mathbb{R}$, $a \in I$, f une fonction de I vers $\mathbb{R}$ continue sur I et dérivable sur $I \setminus \{a\}$.

Si $\lim_{\substack{x \rightarrow a \\ x \neq a}} f'(x) = \ell \in \mathbb{R}$ alors⁵⁸ f est dérivable en a et $f'(a) = \ell$.

Si $\lim_{\substack{x \rightarrow a \\ x \neq a}} f'(x) = \ell \in \{-\infty, +\infty\}$ alors $\frac{f(x) - f(a)}{x - a} \xrightarrow{x \rightarrow a} \ell$ et le graphe de f présente une tangente verticale au point de coordonnées $(a, f(a))$.

Ne pas confondre les deux éventuelles limites $\lim_{\substack{x \rightarrow a \\ x \neq a}} f'(x)$ et $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$.

La première est liée à la continuité de f' en a tandis que la seconde concerne la dérivabilité de f en a .

D'autre part, la limite $\lim_{\substack{x \rightarrow a \\ x \neq a}} f'(x)$ peut ne pas exister, auquel cas le théorème ne permet pas de conclure quant à la dérivabilité de f en a .

Enfin, on n'utilise parfois ce théorème pour des limites à droite ou à gauche. En reprenant les mêmes premières hypothèses du théorème précédent, si cette fois-ci

$\lim_{\substack{x \rightarrow a \\ x < a}} f'(x) = \ell \in \mathbb{R}$ alors f est dérivable à gauche en a et $f'_g(a) = \ell$.

2.13 Convexité

Dans cette section, toutes les fonctions sont à valeurs réelles.

Pour tout $(a, b) \in \mathbb{R}^2$, on rappelle⁵⁹ que le *segment* $[a, b] = \{x \in \mathbb{R}, a \leq x \leq b\}$, ensemble qui peut également s'écrire sous la forme

$$[a, b] = \{(1 - \lambda)a + \lambda b; \lambda \in [0, 1]\}.$$

57. c'est-à-dire f' n'est la fonction nulle sur aucun segment $[a, b] \subset I$.

58. c'est-à-dire f dérivable en a et f' continue en a .

59. Cf. en particulier la note 15 page 35.

On rappelle aussi que si f est une fonction d'un intervalle I de $\mathbb{R}$ vers $\mathbb{R}$, le *graphe* de f est la partie du plan noté Γ_f définie par

$$\Gamma_f = \{(x, y) \in \mathbb{R}^2, x \in I \text{ et } y = f(x)\}.$$

Rappelons enfin qu'une partie E du plan est *convexe* si pour tout $(A, B) \in E^2$, $[A, B] \subset E$.

Définition

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

On appelle *corde* de f tout segment $[A, B]$ du plan où $A(a, f(a))$ et $B(b, f(b))$ avec $(a, b) \in I^2$.

On appelle *sécante* de f toute droite (AB) du plan où $A(a, f(a))$ et $B(b, f(b))$ avec a et b deux points distincts de I .

On appelle *épigraphe* de f la partie du plan notée Γ_f^+ définie par

$$\Gamma_f^+ = \{(x, y) \in \mathbb{R}^2, x \in I \text{ et } y \geq f(x)\}.$$

Une corde de f est donc une partie du plan de la forme

$$\left\{((1 - \lambda)a + \lambda b, (1 - \lambda)f(a) + \lambda f(b)); \lambda \in [0, 1]\right\}$$

où $(a, b) \in I^2$.

Définition

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

On dit que f est *convexe* sur I si son graphe est situé en dessous de toutes ses cordes c'est-à-dire si

$$\forall (a, b) \in I^2, \forall \lambda \in [0, 1], f((1 - \lambda)a + \lambda b) \leq (1 - \lambda)f(a) + \lambda f(b).$$

On dit que f est *concave*⁶⁰ sur I si son graphe est situé au-dessus de toutes ses cordes c'est-à-dire si

$$\forall (a, b) \in I^2, \forall \lambda \in [0, 1], f((1 - \lambda)a + \lambda b) \geq (1 - \lambda)f(a) + \lambda f(b).$$

Théorème

Soient I un intervalle de $\mathbb{R}$, a et b deux points de I avec $a < b$ et f une fonction de I vers $\mathbb{R}$, convexe sur I .

Alors le graphe de f est situé en dessous de sa sécante sur $[a, b]$ et au-dessus à l'extérieur de $[a, b]$.

60. On remarquera que f concave équivaut à $-f$ convexe et que les seules fonctions à la fois convexes et concaves sont les fonctions affines.

Théorème (inégalité de Jensen)

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

Si f est convexe alors pour tous points $a_1, \dots, a_n$ de I et tous réels positifs $\lambda_1, \dots, \lambda_n$ de somme égale à 1, on a

$$f\left(\sum_{i=1}^n \lambda_i a_i\right) \leq \sum_{i=1}^n \lambda_i f(a_i).$$

Notons que l'on retrouve la définition de la convexité de f en prenant $n = 2$ dans l'inégalité de Jensen.

Explicitons à présent des assertions équivalentes à la convexité d'une fonction.

Théorème

Soient I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{R}$. Alors les assertions suivantes sont équivalentes :

- (i) f est convexe sur I .
- (ii) L'épigraphe de f est une partie convexe du plan.
- (iii) Pour tout $a \in I$, la fonction⁶¹ $x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur $I \setminus \{a\}$.
- (iv) Pour tout $(a, b, c) \in I^3$ tel que $a < b < c$, on a l'*inégalité des pentes*⁶² suivante :

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}.$$

Ajoutons à présent une hypothèse supplémentaire de dérivabilité à la convexité.

Théorème

Soient I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{R}$ dérivable sur I . Alors les assertions suivantes sont équivalentes :

- (i) f est convexe sur I .
- (ii) f' est croissante⁶³ sur I .
- (iii) Le graphe de f est situé au-dessus de toutes ses tangentes⁶⁴.

61. On pourrait l'appeler *fonction pente d'origine A* car elle modélise la pente de la sécante (AX) où $A(a, f(a))$ et $X(x, f(x))$.

62. appelée également *lemme des trois pentes*.

63. assertion équivalente à f'' positive sur I si f est deux fois dérivable sur I .

64. c'est-à-dire : pour tout $a \in I$, pour tout $x \in I$, $f(x) \geq f(a) + f'(a)(x - a)$.

Définition

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ et a un point intérieur⁶⁵ de I . On dit que f admet un *point d'inflexion* en a si le graphe de f traverse sa tangente⁶⁶ au point $(a, f(a))$.

Théorème (condition nécessaire d'un point d'inflexion)

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ deux fois dérivable sur I et a un point intérieur de I .

Si f admet un point d'inflexion en a alors $f''(a) = 0$.

La réciproque de ce théorème est fausse. Par exemple la fonction $f : x \mapsto x^4$ vérifie $f''(0) = 0$ mais f n'admet pas de point d'inflexion en 0.

Théorème (condition suffisante d'un point d'inflexion)

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{R}$ deux fois dérivable sur I et a un point intérieur de I .

Si f'' s'annule en a en changeant de signe⁶⁷ alors f admet un point d'inflexion en a .

2.14 Comparaison asymptotique

Définition

Soient⁶⁸ I un intervalle de $\mathbb{R}$, f et g deux fonctions⁶⁹ de I vers $\mathbb{K}$ et a un point⁷⁰ de I ou une extrémité finie ou infinie de I .

On dit que f est *dominée* par g au voisinage de a s'il existe un voisinage de a et une fonction h de $I \cap V$ vers $\mathbb{K}$, bornée sur $I \cap V$ tels que $f = hg$ sur $I \cap V$. On écrit^{71, 72} alors $f = O_a(g)$ ou $f(x) = O(g(x))$.

65. Cf. notes 47 et 53 pages 51 et 54.

66. c'est-à-dire f change de convexité en a : f est par exemple convexe à gauche de a et concave à droite de a ou inversement.

67. Cf. note 54 page 54 en remplaçant f' par f'' .

68. Les définitions de cette section sont définies, comme dans tout le chapitre, sur un intervalle sachant qu'elles s'étendent à des ensembles qui ne sont pas nécessairement des intervalles mais des parties quelconques de $\mathbb{R}$ non vides et non réduit à un point.

69. On peut étendre cette définition, ainsi que les deux suivantes, à des fonctions définies au moins au voisinage de a .

70. On peut aussi dire : soit a un point de $\mathbb{R} \cup \{-\infty, +\infty\}$ adhérent à I .

71. Se lit « f est un grand O de g au voisinage de a ».

Même si au premier abord, cette définition (et surtout la notation) peuvent paraître délicates, le théorème qui suit en facilite la compréhension.

Théorème

Soient I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie ou infinie de I . On suppose de plus que g ne s'annule pas au voisinage de a (sauf éventuellement en a). Alors les assertions suivantes sont équivalentes :

- (i) Au voisinage de a , $f = O(g)$.
- (ii) $\frac{f}{g}$ est bornée au voisinage de a .
- (iii) • Si a est un réel :

$$\exists \alpha > 0, \exists K \in \mathbb{R}_+, \forall x \in I, |x - a| < \alpha \implies |f(x)| \leq K |g(x)|.$$

- Si $a = +\infty$:

$$\exists A > 0, \exists K \in \mathbb{R}_+, \forall x \in I, x > A \implies |f(x)| \leq K |g(x)|.$$

Par exemple, au voisinage de a , $f = O(1)$ signifie f bornée au voisinage de a .

Comme expliqué dans la note 71, il faut bien comprendre que la notation n'est pas à prendre au sens d'une vraie égalité. Par exemple, au voisinage de tout réel, $\sin(x) = O(1)$ et $\cos(x) = O(1)$ mais les fonctions sinus et cosinus sont pourtant distinctes.

Toujours dans le même état d'esprit, au voisinage d'un a , $O(g) - O(g) \neq 0$ mais $O(g) - O(g) = O(g)$.

La relation de domination n'est pas non plus symétrique : par exemple, au voisinage de $+\infty$, $\frac{1}{x} = O(1)$ mais 1 n'est pas un $O\left(\frac{1}{x}\right)$.

Par la suite, au voisinage d'un point a , l'écriture $O(g)$ désignera dans toute égalité n'importe quelle fonction dominée par g au voisinage de a .

72. ou même plus simplement : au voisinage de a , $f = O(g)$. On emploie aussi parfois la notation $f \underset{a}{\ll} g$ ou $f(x) \underset{x \rightarrow a}{\ll} g(x)$. Bien qu'historiquement, la notation $\ll$ fut introduite comme symbole de substitution pour la relation de domination, elle reste encore parfois utilisée pour noter la relation de *négligeabilité*.

73. Dans le cas $a = -\infty$, remplacer $x > A$ par $x < A$.

74. Il faut penser en ces termes : une fonction dominée par g au voisinage de a moins une fonction dominée par g au voisinage de a reste une fonction dominée par g au voisinage de a .

Théorème

Soient I un intervalle de $\mathbb{R}$, f , g et h et k quatre fonctions de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie ou infinie de I . Alors, au voisinage de a ,

- si $f = O(g)$ et $g = O(h)$ alors $f = O(h)$;
- si $f = O(h)$ et $g = O(h)$ alors $f + g = O(h)$;
- si $f = O(h)$ et $g = O(k)$ alors $fg = O(hk)$.

Définition

Soient I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie ou infinie de I .

On dit que f est *négligeable* devant g au voisinage de a s'il existe un voisinage de a et une fonction ε de $I \cap V$ vers $\mathbb{K}$ tels que $f = \varepsilon g$ sur $I \cap V$ avec $\lim_a \varepsilon = 0$.

On écrit ⁷⁵ alors $f = o(g)$ ou $f(x) \underset{x \rightarrow a}{=} o(g(x))$.

Comme pour la relation de domination, explicitons immédiatement le théorème permettant de clarifier cette notion.

Théorème

Soient I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie ou infinie de I . On suppose de plus que g ne s'annule pas au voisinage de a (sauf éventuellement en a). Alors les assertions suivantes sont équivalentes :

(i) Au voisinage de a , $f = o(g)$.

(ii) $\lim_a \frac{f}{g} = 0$.

(iii) • Si a est un réel :

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, |x - a| < \alpha \implies |f(x)| \leq \varepsilon |g(x)|.$$

• Si ⁷⁶ $a = +\infty$:

$$\forall \varepsilon > 0, \exists A > 0, \forall x \in I, x > A \implies |f(x)| \leq \varepsilon |g(x)|.$$

Par exemple, au voisinage de a , $f = o(1)$ signifie $\lim_a f = 0$.

⁷⁵. ou plus simplement : au voisinage de a , $f = o(g)$. Cette fausse égalité se lit « f est un petit o de g au voisinage de a ».

⁷⁶. Dans le cas $a = -\infty$, remplacer $x > A$ par $x < A$.

Comme pour la relation de domination, il faut bien comprendre que cette notation n'est pas une vraie égalité.

Par exemple, au voisinage de $+\infty$, $x = o(e^x)$ et $\ln(x) = o(e^x)$ mais les fonctions $\ln$ et $x \mapsto x$ sont distinctes.

De même, au voisinage d'un a , $o(g) - o(g)$ n'est pas égal à 0 mais reste un $o(g)$.

Par la suite, au voisinage d'un point a , l'écriture $o(g)$ désignera dans toute égalité n'importe quelle fonction négligeable devant g au voisinage de a .

Théorème

Soient I un intervalle de $\mathbb{R}$, f , g et h et k quatre fonctions de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie ou infinie de I . Alors, au voisinage de a ,

- si $f = o(g)$ alors $f = O(g)$;
- si $f = o(g)$ et $g = o(h)$ alors $f = o(h)$;
- si $f = o(h)$ et $g = o(h)$ alors $f + g = o(h)$;
- si $f = o(h)$ alors $fg = o(gh)$;
- si $f = o(h)$ et $g = o(k)$ alors $fg = o(hk)$;
- $\lim_a f = \ell \in \mathbb{K}$ si, et seulement si, au voisinage de a , $f = \ell + o(1)$.

Le théorème qui suit permet en particulier de réécrire le *théorème des croissances comparées* étudié en classe de Terminale à l'aide de cette relation de négligeabilité.

Théorème

Au voisinage de $+\infty$, on a

- pour tout $\alpha \in \mathbb{R}$, $x^\alpha = o(e^x)$;
- pour tout $(\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{R}$, $\ln^\beta(x) = o(x^\alpha)$;
- pour tout $(\alpha, \beta) \in \mathbb{R}^2$ tel que $\alpha < \beta$, $x^\alpha = o(x^\beta)$.

Au voisinage de 0, on a

- pour tout $(\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{R}$, $\ln^\beta(x) = o\left(\frac{1}{x^\alpha}\right)$;
- pour tout $(\alpha, \beta) \in \mathbb{R}^2$ tel que $\alpha < \beta$, $x^\beta = o(x^\alpha)$.

Définition

On dit que f est *équivalente* à g au voisinage de a s'il existe un voisinage de a et une fonction k de $I \cap V$ vers $\mathbb{K}$ tels que $f = kg$ sur $I \cap V$ avec $\lim_a k = 1$. On écrit alors $f \underset{a}{\sim} g$ ou $f(x) \underset{x \rightarrow a}{\sim} g(x)$.

Théorème

Soient I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie ou infinie de I . On suppose de plus que g ne s'annule pas au voisinage de a (sauf éventuellement en a). Alors les assertions suivantes sont équivalentes :

- (i) $f \underset{a}{\sim} g$.
- (ii) $\lim_a \frac{f}{g} = 1$.
- (iii) Au voisinage de a , $f = g + o(g)$.

Théorème

Soient I un intervalle de $\mathbb{R}$, f , g et h et k quatre fonctions de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie ou infinie de I . Alors

- la relation $\underset{a}{\sim}$ est une relation d'équivalence ;
- si $f \underset{a}{\sim} g$ alors $f = O(g)$ et $g = O(f)$;
- si $f \underset{a}{\sim} h$ et $g \underset{a}{\sim} k$ alors $fg \underset{a}{\sim} hk$;
- si $f \underset{a}{\sim} g$ et f et g à valeurs réelles alors f et g ont même signe⁷⁷ au voisinage de a ;
- si $\lim_a f = \ell \in \mathbb{K}$ avec $\ell \neq 0$ alors $f \underset{a}{\sim} \ell$;
- si $f \underset{a}{\sim} g$ et f et g à valeurs réelles strictement positives⁷⁸ alors, pour tout $\alpha \in \mathbb{R}$, $f^\alpha \underset{a}{\sim} g^\alpha$;
- si $f \underset{a}{\sim} g$ et $\lim_a g = \ell$ alors $\lim_a f = \ell$;
- si f , g et h sont à valeurs réelles et vérifient $f \leq g \leq h$ avec $f \underset{a}{\sim} h$ alors $g \underset{a}{\sim} f$.

$f \underset{a}{\sim} g$ n'est pas équivalente à $\lim_a (f - g) = 0$. Par exemple $x^2 + x \underset{x \rightarrow +\infty}{\sim} x^2$ mais $x^2 + x - x^2 = x$ tend vers $+\infty$ (et non vers 0) quand x tend vers $+\infty$.

77. c'est-à-dire si $f \underset{a}{\sim} g$ et $g > 0$ (respectivement $g < 0$) au voisinage de a alors $f > 0$ (respectivement $f < 0$ au voisinage de a).

78. au moins au voisinage de a .

79. y compris, dans le cas de fonctions à valeurs réelles, $\ell \in \{-\infty, +\infty\}$.

Les équivalents ne sont pas compatibles avec la somme : si $f \underset{a}{\sim} h$ et $g \underset{a}{\sim} k$, on ne peut pas en déduire que $f + g$ est équivalente à $h + k$ en a .

Par exemple $x + 1 \underset{x \rightarrow +\infty}{\sim} x$ et $1 - x \underset{x \rightarrow +\infty}{\sim} 1 - x$ mais $x + 1 + 1 - x = 2$ n'est pas équivalent à $x + 1 - x = 1$ en $+\infty$.

Si f et g possèdent la même limite en a , on ne peut pas en déduire que f est équivalente à g au voisinage de a .

Par exemple $\lim_{x \rightarrow +\infty} \ln(x) = +\infty$ et $\lim_{x \rightarrow +\infty} x^2 = +\infty$ mais $\ln(x)$ n'est pas équivalente à x^2 au voisinage de $+\infty$.

2.15 Comparaison asymptotique (cas des suites)

Cette section est une rapide adaptation aux suites de la section précédente et ne contient donc aucune remarque.

Définition

Soient (u_n) et (v_n) deux suites réelles ou complexes.

On dit que (u_n) est *dominée* par (v_n) s'il existe une suite (b_n) bornée telle que $u_n = b_n v_n$ à partir d'un certain rang.

On écrit alors⁸⁰ $u_n = O(v_n)$.

Théorème

Soient (u_n) et (v_n) deux suites réelles ou complexes. On suppose de plus que (v_n) ne s'annule pas à partir d'un certain rang. Alors les assertions suivantes sont équivalentes :

- (i) $u_n = O(v_n)$.
- (ii) $\frac{u_n}{v_n}$ est bornée à partir d'un certain rang.
- (iii) $\exists N \in \mathbb{N}, \exists K \in \mathbb{R}_+, \forall n \in \mathbb{N}, n \geq N \implies |u_n| \leq K |v_n|$.

Théorème

Soient (u_n) , (v_n) , (x_n) et (y_n) quatre suites réelles ou complexes. Alors

- si $u_n = O(v_n)$ et $v_n = O(x_n)$ alors $u_n = O(x_n)$;
- si $u_n = O(x_n)$ et $v_n = O(x_n)$ alors $u_n + v_n = O(x_n)$;
- si $u_n = O(x_n)$ et $v_n = O(y_n)$ alors $u_n v_n = O(x_n y_n)$.

80. Se lit « u_n est un grand O de v_n ».

Adaptons aux suites la relation de *négligeabilité*.

Définition

Soient (u_n) et (v_n) deux suites réelles ou complexes.

On dit que (u_n) est *négligeable* devant (v_n) s'il existe une suite (ε_n) convergeant vers 0 telle que $u_n = \varepsilon_n v_n$ à partir d'un certain rang.

On écrit alors ⁸¹ $u_n = o(v_n)$.

Théorème

Soient (u_n) et (v_n) deux suites réelles ou complexes. On suppose de plus que (v_n) ne s'annule pas à partir d'un certain rang. Alors les assertions suivantes sont équivalentes :

- (i) $u_n = o(v_n)$.
- (ii) $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 0$.
- (iii) $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies |u_n| \leq \varepsilon |v_n|$.

Théorème

Soient (u_n) , (v_n) , (x_n) et (y_n) quatre suites réelles ou complexes. Alors

- si $u_n = o(v_n)$ alors $u_n = O(v_n)$;
- si $u_n = o(v_n)$ et $v_n = o(x_n)$ alors $u_n = o(x_n)$;
- si $u_n = o(x_n)$ et $v_n = o(x_n)$ alors $u_n + v_n = o(x_n)$;
- si $u_n = o(x_n)$ et $v_n = o(y_n)$ alors $u_n v_n = o(x_n y_n)$;
- $\lim_{n \rightarrow +\infty} u_n = \ell \in \mathbb{K}$ si, et seulement si, $u_n = \ell + o(1)$.

Théorème

- Pour tout réel α , $n^\alpha = o(e^n)$;
- pour tout $(\alpha, \beta) \in \mathbb{R}_+^* \times \mathbb{R}$, $\ln^\beta(n) = o(n^\alpha)$;
- pour tout $(\alpha, \beta) \in \mathbb{R}^2$ tel que $\alpha < \beta$, $n^\alpha = o(n^\beta)$;
- pour tout réel α , $n^\alpha = o(n!)$ et $\alpha^n = o(n!)$;
- $n! = o(n^n)$.

81. Se lit « u_n est un petit o de v_n ».

Définition

Soient (u_n) et (v_n) deux suites réelles ou complexes.

On dit que (u_n) est *équivalente* à (v_n) s'il existe une suite (k_n) convergent vers 1 telle que $u_n = k_n v_n$ à partir d'un certain rang.

On écrit alors $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ ou $u_n \sim v_n$.

Théorème

Soient (u_n) et (v_n) deux suites réelles ou complexes. On suppose de plus que (v_n) ne s'annule pas à partir d'un certain rang. Alors les assertions suivantes sont équivalentes :

- (i) $u_n \sim v_n$.
- (ii) $\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = 1$.
- (iii) $u_n = v_n + o(v_n)$.

Théorème

Soient (u_n) , (v_n) , (x_n) et (y_n) quatre suites réelles ou complexes. Alors

- la relation $\sim$ est une relation d'équivalence ;
- si $u_n \sim v_n$ alors $u_n = O(v_n)$ et $v_n = O(u_n)$;
- si $u_n \sim x_n$ et $v_n \sim y_n$ alors $u_n v_n \sim x_n y_n$;
- si $u_n \sim v_n$ et (u_n) et (v_n) sont des suites réelles alors u_n et v_n ont même signe ⁸² à partir d'un certain rang ;
- si $\lim_{n \rightarrow +\infty} u_n = \ell \in \mathbb{K}$ avec $\ell \neq 0$ alors $u_n \sim \ell$;
- si $u_n \sim v_n$ et (u_n) et (v_n) sont des suites réelles strictement positives ⁸³ alors, pour tout $\alpha \in \mathbb{R}$, $u_n^\alpha \sim v_n^\alpha$;
- si $u_n \sim v_n$ et ⁸⁴ $\lim_{n \rightarrow +\infty} v_n = \ell$ alors $\lim_{n \rightarrow +\infty} u_n = \ell$;
- si (u_n) , (v_n) et (x_n) sont des suites réelles et vérifient $u_n \leq v_n \leq x_n$ avec $u_n \sim x_n$ alors $v_n \sim u_n$.

^{82.} c'est-à-dire si $u_n \sim v_n$ et $v_n > 0$ (respectivement $v_n < 0$) alors $u_n > 0$ (respectivement $u_n < 0$) à partir d'un certain rang.

^{83.} au moins à partir d'un certain rang.

^{84.} y compris, dans le cas d'une suite réelle, $\ell \in \{-\infty, +\infty\}$.

Théorème (formule de Stirling)

$$n! \sim n^n e^{-n} \sqrt{2\pi n}.$$

2.16 Développements limités

Définition

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, f une fonction⁸⁵ de I vers $\mathbb{K}$ et a un point⁸⁶ de I ou une extrémité finie de I .

On dit que f admet^{87, 88} un *développement limité* à l'ordre n en a s'il existe $(a_0, \dots, a_n) \in \mathbb{K}^{n+1}$ tel que, au voisinage de a ,

$$f(x) = a_0 + a_1(x-a) + \dots + a_n(x-a)^n + o((x-a)^n).$$

Le terme⁸⁹ $a_0 + a_1(x-a) + \dots + a_n(x-a)^n$ est alors appelé *partie régulière* du développement limité à l'ordre n de f en a .

Un développement limité est donc une approximation locale d'une fonction au voisinage d'un point par un polynôme.

Dans le cas particulier $a = 0$, on dit que f admet un *développement limité* à l'ordre n en 0 s'il existe $(a_0, \dots, a_n) \in \mathbb{K}^{n+1}$ tel que, au voisinage de 0 ,

$$f(x) = a_0 + a_1x + \dots + a_nx^n + o(x^n).$$

c'est-à-dire, au voisinage de a ,

$$f(x) - (a_0 + a_1x + \dots + a_nx^n) = o(x^n).$$

Comme l'explique le théorème qui suit, la partie régulière d'un développement limité d'une fonction en un point à un certain ordre est unique⁹⁰ et d'autre part, on peut toujours se ramener à un développement limité en 0 .

85. Comme pour les définitions des o , O et $\sim$, on peut simplement supposer que f est définie au voisinage de a .

86. On pourrait aussi dire : soit $a \in \mathbb{R}$ adhérent à I .

87. On emploie aussi l'expression *développement limité* à l'ordre n au voisinage de a .

88. Il existe des fonctions n'admettant de développement limité en aucun réel à n'importe quel ordre : c'est le cas par exemple de la *fonction caractéristique de $\mathbb{Q}$* appelée également *fonction de Dirichlet* (définie dans la note 32 page 42) qui n'est continue en aucun point de $\mathbb{R}$.

89. Le $o((x-a)^n)$ qui désigne une fonction négligeable devant $(x-a)^n$ est parfois appelé *reste* du développement limité de f à l'ordre n en a .

90. On dira même qu'un développement limité d'une fonction en un point à un certain ordre est unique. On parlera ainsi par la suite du développement limité d'une fonction en un point à un certain ordre.

Théorème

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie de I .

1. Si f admet un développement limité à l'ordre n en a , sa partie régulière est unique.
2. f admet un développement limité à l'ordre n en a si, et seulement si, la fonction $g : h \mapsto f(a+h)$ admet un développement limité à l'ordre n en 0.

Théorème

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et a un point de I ou une extrémité finie de I .

Si f admet en a le développement limité à l'ordre n ,

$$f(x) = a_0 + a_1(x-a) + \cdots + a_n(x-a)^n + o((x-a)^n)$$

avec les a_k non tous nuls⁹¹ alors

- $f(x) \underset{x \rightarrow a}{\sim} a_k(x-a)^k$ où $a_k \neq 0$ est le premier coefficient non nul du développement limité⁹².
- f admet en a un développement limité à tout ordre $p \leq n$, appelé *troncature* du développement limité de f à l'ordre p , défini par

$$f(x) = a_0 + a_1(x-a) + \cdots + a_p(x-a)^p + o((x-a)^p).$$

Cette troncature est possible car, au voisinage de a ,

$$a_{p+1}(x-a)^{p+1} + \cdots + a_n(x-a)^n = o((x-a)^p).$$

Les termes après $a_p(x-a)^p$ n'ont donc pas disparu, ils sont tous négligeables devant $(x-a)^p$ au voisinage de a .

91. Il existe des fonctions non nulles admettant en un point un développement limité dont la partie régulière est nulle à tout ordre : c'est le cas par exemple de la fonction

$f : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} e^{-1/x^2} & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases} \end{cases}$ qui vérifie au voisinage de 0, pour tout $n \in \mathbb{N}$, $f(x) = o(x^n)$.

92. En particulier $f(x)$ est du signe de $a_k(x-a)^k$ au voisinage de a .

Théorème

Soit f une fonction d'une variable réelle à valeurs dans $\mathbb{K}$ admettant un développement limité d'ordre $n \in \mathbb{N}$ en 0.

Si f est paire (respectivement impaire), alors la partie régulière de ce développement limité ne contient que des termes d'exposants pairs (respectivement impairs).

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et $a \in I$. Alors⁹³

• f est continue en a si, et seulement si, f admet un développement limité à l'ordre 0 en a .

Dans ce cas, au voisinage de a , $f(x) = f(a) + o(1)$.

• f est dérivable en a si, et seulement si, f admet un développement limité à l'ordre 1 en a .

Dans ce cas, au voisinage de a , $f(x) = f(a) + f'(a)(x - a) + o((x - a))$.

Théorème (primitivation d'un développement limité)

Soient $n \in \mathbb{N}^*$, I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{K}$ continue au moins au voisinage de a de sorte qu'elle admet une primitive F au voisinage de a .

Si f admet en a le développement limité à l'ordre $n-1$

$$f(x) = \sum_{k=0}^{n-1} a_k (x-a)^k + o((x-a)^{n-1})$$

alors F admet en a le développement limité⁹⁴ à l'ordre n

$$F(x) = F(a) + \sum_{k=0}^{n-1} a_k \frac{(x-a)^{k+1}}{k+1} + o((x-a)^n).$$

93. Ce théorème n'est pas vérifié à un ordre supérieur : il existe des fonctions admettant par exemple un développement limité d'ordre 2 en un point sans être deux fois dérivable en ce point.

C'est le cas en 0 de la fonction $f : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} x^3 \sin(\frac{1}{x^2}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases} \end{cases}$ qui vérifie au voisinage

de 0, $f(x) = o(x^2)$ mais qui n'est pas deux fois dérivable en 0 puisque f' n'est même pas continue en 0.

94. obtenu par intégration « terme à terme » sans oublier d'ajouter $F(a)$.

En revanche, si f est une fonction dérivable sur I admettant un développement limité à l'ordre $n \in \mathbb{N}^*$ en $a \in I$, f' n'admet pas nécessairement⁹⁵ un développement limité à l'ordre $n-1$ en a . Il suffit pour s'en convaincre de reprendre l'exemple de la fonction f définie dans la note 93 : f , dérivable sur $\mathbb{R}$, admet un développement limité à l'ordre 2 en 0 mais, en utilisant le théorème page 70, f' n'admet pas de développement limité à l'ordre 1 en 0 car f' n'est pas dérivable (car même pas continue) en 0.

Le théorème qui suit permet d'obtenir, le cas échéant avec l'aide de celui qui précède, les développements limités des fonctions usuelles.

Théorème de Taylor-Young

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et $a \in I$. Si⁹⁶ f est n fois dérivable en a , alors f admet en a le développement limité à l'ordre n

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + o((x-a)^n)$$

c'est-à-dire

$$f(x) = f(a) + \frac{f'(a)}{1!} (x-a) + \cdots + \frac{f^{(n)}(a)}{n!} (x-a)^n + o((x-a)^n).$$

Rassemblons dans les deux théorèmes suivants les développements limités en 0 des fonctions usuelles.

Théorème

Soit $\alpha \in \mathbb{R}$. Au voisinage de 0, on a

$$\frac{1}{1-x} = \sum_{k=0}^n x^k + o(x^n) \qquad e^x = \sum_{k=0}^n \frac{x^k}{k!} + o(x^n)$$

$$\ln(1+x) = \sum_{k=1}^n (-1)^{k-1} \frac{x^k}{k} + o(x^n)$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \cdots + \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n + o(x^n).$$

95. Il faut imposer une hypothèse de régularité plus forte à f , du type de classe $\mathcal{C}^n$, afin que l'implication soit vérifiée.

96. Ce théorème est le plus souvent formulé avec une hypothèse plus forte imposant f de classe $\mathcal{C}^n$ (voire même de classe $\mathcal{C}^{n+1}$) mais celle-ci n'est pas nécessaire et sert simplement à faciliter sa démonstration.

Théorème

Au voisinage de 0, on a ⁹⁷

$$\sin(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+1}) \quad \cos(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2n})$$

$$\tan(x) = x + \frac{x^3}{3} + o(x^3) \quad \arctan(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} + o(x^{2n+1})$$

$$\operatorname{sh}(x) = \sum_{k=0}^n \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+1}) \quad \operatorname{ch}(x) = \sum_{k=0}^n \frac{x^{2k}}{(2k)!} + o(x^{2n}).$$

Ces développements limités permettent de redémontrer en une ligne des limites classiques : $\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{x+o(x)}{x} = \lim_{x \rightarrow 0} (1 + o(1)) = 1$.

Théorème

Soient $\lambda \in \mathbb{K}$, f et g deux fonctions d'une variable réelle à valeurs dans $\mathbb{K}$ admettant les développements limités à l'ordre $n \in \mathbb{N}$ en 0 suivants :

$$f(x) = P(x) + o(x^n) \quad \text{et} \quad g(x) = Q(x) + o(x^n)$$

où $(P, Q) \in (\mathbb{K}_n[X])^2$. Alors

- $\lambda f + g$ admet un développement limité à l'ordre n en 0 dont la partie régulière est $\lambda P(x) + Q(x)$.
- fg admet un développement limité à l'ordre n en 0 dont la partie régulière est la troncature de $P(x)Q(x)$ à l'ordre n .
- Si de plus $f(0) = 0$, alors $g \circ f$ admet un développement limité à l'ordre n en 0 dont la partie régulière est la troncature de $(Q \circ P)(x)$ à l'ordre n .

Si $f(0) \neq 0$, on peut aussi déterminer le développement limité en 0 du quotient ⁹⁸ $\frac{g}{f} = g \times \frac{1}{f}$ en 0 en déterminant le développement limité en 0 de l'inverse $\frac{1}{f}$.

En notant $a = \frac{1}{f(0)}$, on a $\frac{1}{f} = a \times \frac{1}{af(x)} = a \times \frac{1}{1 - (1 - af(x))} = a \times \frac{1}{1 - u(x)}$

où $u : x \mapsto 1 - af(x)$.

Comme $u(0) = 1 - af(0) = 1 - 1 = 0$, on peut utiliser le développement limité de la composée de $x \mapsto \frac{1}{1-x}$ avec u en 0 afin d'obtenir celui de $\frac{1}{f}$ et donc aussi celui de $\frac{g}{f}$ en 0.

⁹⁷. Le développement limité de la fonction tangente en 0 n'est pas d'une forme simple, seul le développement à l'ordre 3 est à connaître.

⁹⁸. Il existe également une autre méthode pour déterminer le développement limité d'un quotient en utilisant la division de polynômes *suivant les puissances croissantes*.

Un développement limité peut également être utile pour positionner le graphe d'une fonction par rapport à ses tangentes comme l'explique le théorème qui suit.

Théorème (position locale par rapport aux tangentes)

Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{R}$ admettant en a le développement limité à l'ordre $k \geq 2$

$$f(x) = a_0 + a_1(x-a) + a_k(x-a)^k + o((x-a)^k)$$

avec⁹⁹ $a_k \neq 0$. Alors

1. La tangente au graphe de f au point d'abscisse a a pour équation $y = a_0 + a_1(x-a)$ et $f(x) - (a_0 + a_1(x-a)) \underset{x \rightarrow a}{\sim} a_k(x-a)^k$.
2. Si k est pair, $(x-a)^k$ est positif et le graphe de f est situé au-dessus ou en dessous de sa tangente au point d'abscisse a selon le signe de a_k .
Si k est impair, $(x-a)^k$ change de signe au voisinage de a et le graphe de f traverse sa tangente au point d'abscisse a : f admet alors un point d'inflexion en a .

Un développement limité d'une fonction en un point peut aussi permettre de préciser si ce point est un extremum local et sa nature (minimum ou maximum) comme le montre le théorème qui suit.

Théorème (condition suffisante d'un extremum local)

Soient I un intervalle de $\mathbb{R}$, a un point intérieur¹⁰⁰ de I et f une fonction de I vers $\mathbb{R}$ admettant en a le développement limité à l'ordre $k \geq 2$

$$f(x) = a_0 + a_1(x-a) + a_k(x-a)^k + o((x-a)^k)$$

avec¹⁰¹ $a_k \neq 0$. Alors

1. Si¹⁰² $a_1 \neq 0$, f n'admet pas d'extremum local en a .
2. Si $a_1 = 0$, f admet un extremum en a si, et seulement si, k est pair.

De plus, dans ce cas :

- si $a_k > 0$, f admet un minimum local en a ;
- si $a_k < 0$, f admet un maximum local en a .

99. On suppose donc pour f dans ce théorème qu'un tel k , le plus petit entier supérieur ou égal à deux tel que $a_k \neq 0$, existe. Rappelons qu'il existe des fonctions non nulles qui ne vérifient pas cette hypothèse (cf. note 91 page 69).

100. On peut aussi prendre I intervalle ouvert de $\mathbb{R}$ et choisir alors a un point quelconque de I .
101. cf. note 99.

102. Par contreposée, si f admet un extremum local en a , $a_1 = f'(a) = 0$: on retrouve la condition nécessaire d'un extremum local vu dans le théorème page 54.

3 | Intégration

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Tout intervalle de $\mathbb{R}$ est supposé non vide et non réduit à un point.

On suppose pour tout segment $[a, b]$, $a < b$.

Dans un souci pédagogique, on utilise le premier théorème du chapitre (cf. note 1) pour énoncer la définition de l'intégrale d'abord d'une fonction continue (cf. page 76) puis d'une fonction continue par morceaux.

La dernière section de ce chapitre est destinée à ceux désirant un approfondissement et une définition équivalente de l'intégrale n'utilisant pas le premier théorème du chapitre.

3.1 Primitives

Définition

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

On appelle *primitive* de f sur I , notée $\int f(t) dt$, toute fonction F de I vers $\mathbb{K}$, dérivable sur I et de dérivée égale à f sur I .

Théorème

Toute fonction continue sur un intervalle de $\mathbb{R}$ admet au moins une primitive sur cet intervalle¹.

1. Ce théorème est généralement présenté comme un corollaire du *théorème fondamental de l'analyse* page 76. Cependant, il est possible de le démontrer sans ce dernier (mais en utilisant la notion de *convergence uniforme* d'une suite de fonctions étudiée en 2^e année). On montre d'abord qu'on peut se restreindre à un segment $[a, b]$, puis on montre que si (f_n) est une suite de fonctions, telle que chaque f_n est continue et admet une primitive sur $[a, b]$, qui converge uniformément vers une fonction f sur $[a, b]$ alors f admet une primitive sur $[a, b]$. Comme tout polynôme possède une primitive sur $[a, b]$, il suffit pour conclure d'appliquer le théorème de Weierstrass : toute fonction continue sur $[a, b]$ est limite uniforme d'une suite de fonctions polynômes.

Théorème

Soient I un intervalle de $\mathbb{R}$, f une fonction de I vers $\mathbb{K}$ et F une primitive de f sur I .

Alors toute primitive de f sur I est de la forme $F + \lambda$ où $\lambda \in \mathbb{K}$.

Corollaire

Soient $\alpha \in \mathbb{R}$, I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{K}$.

Alors il existe une unique primitive F de f sur I telle que $F(a) = \alpha$.

Théorème (primitives usuelles)

Soient $a \in \mathbb{R}^*$ et $\alpha \in \mathbb{R} \setminus \{-1\}$. Alors à une constante près, on a

- $\int t^\alpha dt = \frac{t^{\alpha+1}}{\alpha+1}$ sur $\mathbb{R}_+^*$ et $\int \frac{1}{t} dt = \ln |t|$ sur $\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$.
- $\int e^{at} dt = \frac{1}{a} e^{at}$ sur $\mathbb{R}$ et $\int \ln(t) dt = t \ln(t) - t$ sur $\mathbb{R}_+^*$.
- $\int \sin(t) dt = -\cos(t)$ sur $\mathbb{R}$ et $\int \cos(t) dt = \sin(t)$ sur $\mathbb{R}$.
- $\int \tan(t) dt = -\ln|\cos(t)|$ sur $]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[$ où $k \in \mathbb{Z}$.
- $\int \operatorname{sh}(t) dt = \operatorname{ch}(t)$ sur $\mathbb{R}$ et $\int \operatorname{ch}(t) dt = \operatorname{sh}(t)$ sur $\mathbb{R}$.
- $\int \frac{1}{1+t^2} dt = \arctan(t)$ sur $\mathbb{R}$ et $\int \frac{1}{\sqrt{1-t^2}} dt = \arcsin(t)$ sur $] -1, 1[$.

3.2 Intégration d'une fonction continue

Théorème

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ et f une fonction de I vers $\mathbb{K}$ continue sur I .

Le scalaire $F(b) - F(a)$, noté également $[F(t)]_a^b$, est indépendant de la primitive F de f sur I choisie.

Définition

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ et f une fonction de I vers $\mathbb{K}$ continue sur I . On appelle ^{2, 3} *intégrale de f de a à b* le scalaire noté $\int_a^b f(t) dt$ défini par : $\int_a^b f(t) dt = [F(t)]_a^b = F(b) - F(a)$ où F est une primitive de f sur I .

Remarquons dès à présent que $\int_b^a f(t) dt = -\int_a^b f(t) dt$ et $\int_a^a f(t) dt = 0$.

Théorème fondamental de l'analyse⁴

Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction de I vers $\mathbb{K}$ continue sur I . Alors ⁵ la fonction $F : x \mapsto \int_a^x f(t) dt$ est une primitive de f sur I et l'unique s'annulant en a .

3.3 Intégration d'une fonction continue par morceaux

Définition

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ et f une fonction de I vers $\mathbb{K}$ continue par morceaux sur I . On appelle *intégrale de f de a à b* , toujours notée $\int_a^b f(t) dt$, le scalaire défini par

$$\int_a^b f(t) dt = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(t) dt$$

où $x_0, \dots, x_n$ sont les éventuels points de discontinuité⁶ de f .

2. Lorsque $a < b$, on emploie aussi l'expression *intégrale de f sur $[a, b]$* notée $\int_{[a, b]} f(t) dt$. Cette formulation sera surtout utilisée dans le cadre du chapitre consacré aux intégrales généralisées.

3. Pour ceux désirant une définition « constructive » de l'intégrale (cf. dernière section de ce chapitre), ce résultat est en fait une conséquence du théorème fondamental de l'analyse qui suit.

4. également appelé *théorème fondamental de l'intégration* ou *du calcul intégral*.

5. Toute fonction continue sur I admet des primitives sur I . Mais elles ne sont pas toujours exprimables à l'aide des fonctions usuelles : c'est le cas par exemple des fonctions $t \mapsto e^{t^2}$, $t \mapsto \frac{1}{\ln t}$, $t \mapsto \frac{\sin(t)}{t}$ ou encore $t \mapsto \frac{e^t}{t}$. En revanche, ce théorème est faux pour une fonction seulement continue par morceaux sur I comme la fonction partie entière sur $[0, 1]$. Enfin, la réciproque de la première assertion de cette note est fautive : il existe des fonctions non continues sur un intervalle qui admettent des primitives sur cet intervalle. C'est le cas par exemple de la

fonction $f : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}) & \text{si } x \neq 0, \\ 0 & \text{sinon} \end{cases} \end{cases}$ non continue en 0, qui admet, sur $\mathbb{R}$

tout entier, une primitive $F : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} x^2 \sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases} \end{cases}$.

Théorème

Soient $\lambda \in \mathbb{K}$, I un intervalle de $\mathbb{R}$, $(a, b, c) \in I^3$, f et g deux fonctions de I vers $\mathbb{K}$ continues par morceaux sur I . Alors

- $\int_a^b (\lambda f + g)(t) dt = \lambda \int_a^b f(t) dt + \int_a^b g(t) dt$ (linéarité);
- $\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$ (relation de Chasles);
- $\int_a^b f(t) dt = \int_a^b (\operatorname{Re}(f))(t) dt + i \int_a^b (\operatorname{Im}(f))(t) dt$;
- $\begin{cases} \text{si } a \leq b, & \left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt; \\ \text{si } a > b, & \left| \int_a^b f(t) dt \right| \leq \int_b^a |f(t)| dt. \end{cases}$ (inégalité triangulaire⁷)

Dans le cas particulier des fonctions continues par morceaux à valeurs réelles, on a le théorème suivant.

Théorème

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$, f et g deux fonctions de I vers $\mathbb{R}$ continues par morceaux sur I . Alors

- si f est positive sur I alors $\begin{cases} \text{si } a \leq b, & \int_a^b f(t) dt \geq 0; \\ \text{si } a > b, & \int_a^b f(t) dt \leq 0. \end{cases}$
- si $f \leq g$ alors $\begin{cases} \text{si } a \leq b, & \int_a^b f(t) dt \leq \int_a^b g(t) dt; \\ \text{si } a > b, & \int_a^b f(t) dt \geq \int_a^b g(t) dt. \end{cases}$
- si f est de signe constant et de plus continue⁸ sur I , alors

$$\int_a^b f(t) dt = 0 \implies f \text{ est nulle entre } a \text{ et } b.$$

Donnons à présent l'interprétation de l'intégrale en terme d'aire d'abord pour une fonction à valeurs réelles positives.

6. Comme f est continue par morceaux sur I , f est continue sur $[a, b]$ sauf en un nombre fini de points $x_0, \dots, x_n$ de $[a, b]$ en lesquels f admet une limite à gauche (sauf pour $x_0 = a$) et à droite (sauf pour $x_n = b$). En particulier f est continue sur chaque $]x_i, x_{i+1}[$ et prolongeable par continuité sur $[x_i, x_{i+1}]$. Dans la définition, $\int_{x_i}^{x_{i+1}} f(t) dt$ désigne donc l'intégrale du prolongement par continuité de f restreinte à $]x_i, x_{i+1}[$.

7. Si on utilise la notation $\int_{[a,b]} f(t) dt$ (avec donc $a < b$), on a $\left| \int_{[a,b]} f(t) dt \right| \leq \int_{[a,b]} |f(t)| dt$.

8. Cette propriété devient fausse dans le cas d'une fonction seulement continue par morceaux :

par exemple si $f : \begin{cases} [0, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} 0 & \text{si } x \neq \frac{1}{2} \\ 1 & \text{sinon} \end{cases} \end{cases}$, f n'est pas la fonction nulle sur $[0, 1]$ mais son intégrale sur $[0, 1]$ est nulle.

Théorème

Soient $(O, \vec{i}, \vec{j})$ un repère du plan, I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ et f une fonction de I vers $\mathbb{R}$ continue par morceaux sur I .

Si f est positive et $a \leq b$, alors $\int_a^b f(t) dt$ est l'aire $\mathcal{A}$ de l'ensemble des points $M(x, y)$ du plan⁹ tels que $\begin{cases} a \leq x \leq b \\ 0 \leq y \leq f(x) \end{cases}$.

Si f est négative et $a \leq b$, alors $\int_a^b f(t) dt$ est l'opposée de l'aire $\mathcal{A}$ précédente.

Si f n'est pas de signe constant, $\int_a^b f(t) dt$ est l'aire dite *algébrique*¹⁰ de l'ensemble des points $M(x, y)$ du plan tels que $\begin{cases} a \leq x \leq b \\ y \text{ compris entre } 0 \text{ et } f(x) \end{cases}$.

Si f est positive et $a > b$, comme $\int_a^b f(t) dt = -\int_b^a f(t) dt$, $\int_a^b f(t) dt$ est alors l'opposée de l'aire du théorème précédent.

Théorème

Soient $T \in \mathbb{R}_+^*$ et f une fonction continue par morceaux d'un intervalle I de $\mathbb{R}$ vers $\mathbb{R}$.

- Pour tout $a \in I$, $\int_{-a}^a f(t) dt = \begin{cases} 2 \int_0^a f(t) dt & \text{si } f \text{ est paire;} \\ 0 & \text{si } f \text{ est impaire.} \end{cases}$
- Si f est T -périodique¹¹, alors pour tout $a \in I$, $\int_a^{a+T} f(t) dt = \int_0^T f(t) dt$.

Définition

Soit f une fonction continue par morceaux de $[a, b]$ vers $\mathbb{K}$.

On appelle *valeur moyenne* de f sur $[a, b]$ le scalaire $\frac{1}{b-a} \int_a^b f(t) dt$.

9. Ce qui correspond à la portion du plan délimitée par : le graphe de f , l'axe des abscisses et les droites d'équations $x = a$ et $x = b$.

10. Une aire *algébrique* peut être négative contrairement à une aire *géométrique*.

11. On rappelle qu'une fonction f d'un intervalle I de $\mathbb{R}$ vers $\mathbb{R}$ est T -périodique si pour tout $x \in I : x + T \in I$, $x - T \in I$ et $f(x + T) = f(x)$.

3.4 Sommes de Riemann

Définition

Soient $n \in \mathbb{N}^*$ et f une fonction de $[a, b]$ vers $\mathbb{K}$.

On appelle *somme de Riemann d'ordre n associée à f relativement à une subdivision régulière* de $[a, b]$ l'un des deux scalaires suivants¹² :

$$\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \quad \text{ou} \quad \frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right).$$

Théorème

Soit f une fonction continue par morceaux de $[a, b]$ vers $\mathbb{K}$. Alors

$$\int_a^b f(t) dt = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right).$$

3.5 Intégration par parties et par changement de variable

Théorème (intégration par parties)

Soient I un intervalle de $\mathbb{R}$, f et g deux fonctions de I vers $\mathbb{K}$ de classe $\mathcal{C}^1$ sur I . Alors pour tout $(a, b) \in I^2$,

$$\int_a^b f'(t)g(t) dt = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t) dt.$$

Théorème (intégration par changement de variable)

Soient I un intervalle de $\mathbb{R}$, φ une fonction de I vers un intervalle J de $\mathbb{R}$, de classe $\mathcal{C}^1$ sur I ¹³, et f une fonction de J vers $\mathbb{K}$. Alors pour tout $(a, b) \in I^2$,

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

12. Ces deux scalaires sont parfois appelés somme de Riemann d'ordre n associée à f et à la méthode des rectangles à gauche (respectivement à droite).

13. J est nécessairement un intervalle de $\mathbb{R}$ car l'image d'un intervalle par une fonction continue est un intervalle.

3.6 Formules de Taylor

Théorème (formule de Taylor avec reste intégral)

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$ et f une fonction de I vers $\mathbb{K}$ de classe $\mathcal{C}^{n+1}$ sur I . Alors, pour tout $(a, b) \in I^2$, on a

$$f(b) = \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t) dt.$$

Théorème (formule de Taylor-Lagrange)

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$ et f une application de I vers $\mathbb{R}$ de classe $\mathcal{C}^{n+1}$ sur I . Alors, pour tout $(a, b) \in I^2$ avec $a < b$, il existe $c \in]a, b[$ tel que

$$f(b) = \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c).$$

Théorème (inégalité de Taylor-Lagrange)

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$ et f une application de I vers $\mathbb{K}$ de classe $\mathcal{C}^{n+1}$ sur I . Alors, pour tout $(a, b) \in I^2$,

$$\left| f(b) - \sum_{k=0}^n \frac{(b-a)^k}{k!} f^{(k)}(a) \right| \leq \frac{|b-a|^{n+1}}{(n+1)!} \sup_{t \in [a, b]} |f^{(n+1)}(t)|.$$

3.7 Intégrabilité au sens de Riemann

Cette section s'adresse à ceux désirant un approfondissement et une définition équivalente de l'intégrale n'utilisant pas le premier théorème du chapitre.

Définition

Soit f une fonction de $[a, b]$ vers $\mathbb{K}$.

On dit que f est *en escalier* s'il existe une subdivision¹⁴ $(x_0, \dots, x_n)$ de $[a, b]$ telle que f est constante sur chaque intervalle $]x_i, x_{i+1}[$.

On dit alors que la subdivision est *adaptée* à la fonction en escalier f .

On notera $\mathcal{E}([a, b])$ l'ensemble des fonctions en escalier de $[a, b]$ vers $\mathbb{K}$.

14. Cf. définition page 47.

Si $\lambda \in \mathbb{K}$, f et g sont deux fonctions en escalier de $[a, b]$ vers $\mathbb{K}$, alors $\lambda f + g$, fg , $|f|$, $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$ sont en escalier.

Théorème

Soient f une fonction en escalier de $[a, b]$ vers $\mathbb{K}$, $(x_0, \dots, x_n)$ une subdivision adaptée à f et, pour tout $i \in \llbracket 0, n-1 \rrbracket$, y_i la valeur de f sur $]x_i, x_{i+1}[$. Alors la valeur du scalaire $\sum_{i=0}^{n-1} (x_{i+1} - x_i) y_i$ ne dépend pas de la subdivision adaptée $(x_0, \dots, x_n)$ choisie.

Définition

Soient f une fonction en escalier de $[a, b]$ vers $\mathbb{K}$, $(x_0, \dots, x_n)$ une subdivision adaptée à f et, pour tout $i \in \llbracket 0, n-1 \rrbracket$, y_i la valeur de f sur $]x_i, x_{i+1}[$. Le scalaire $\sum_{i=0}^{n-1} (x_{i+1} - x_i) y_i$ est appelé *intégrale de f sur $[a, b]$* et noté $\int_{[a,b]} f(t) dt$.

Définissons à présent l'intégrale d'une fonction continue par morceaux sur un segment $[a, b]$.

Théorème

Soit f une fonction continue par morceaux de $[a, b]$ vers $\mathbb{K}$.

En notant $I_-(f) = \sup_{\substack{g \in \mathcal{E}([a,b]) \\ g \leq f}} \int_{[a,b]} g(t) dt$ et $I_+(f) = \inf_{\substack{g \in \mathcal{E}([a,b]) \\ g \geq f}} \int_{[a,b]} g(t) dt$, on a $I_-(f) = I_+(f)$.

Cette valeur commune est appelée¹⁵ *intégrale (de Riemann) de f sur $[a, b]$* et notée $\int_{[a,b]} f(t) dt$.

15. On dit aussi que f est *intégrable* (au sens de Riemann) sur $[a, b]$ ou Riemann-intégrable sur $[a, b]$. Une fonction f de $[a, b]$ vers $\mathbb{K}$ est *intégrable* (au sens de Riemann) si $I_+(f)$ et $I_-(f)$ existent (c'est le cas si f est bornée) et sont égales. Toute fonction continue par morceaux de $[a, b]$ vers $\mathbb{K}$ est donc intégrable (au sens de Riemann) sur $[a, b]$. Il existe des fonctions qui ne sont pas intégrables (au sens de Riemann) comme la *fonction de Dirichlet* (définie dans la note 32 page 42). Si f est intégrable (au sens de Riemann) sur $[a, b]$, $|f|$ l'est aussi. En revanche, la

réci-proque est fautive : par exemple, la fonction $f : \begin{cases} [0, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ -1 & \text{sinon} \end{cases} \end{cases}$ n'est pas

intégrable (au sens de Riemann) sur $[0, 1]$ mais la fonction $|f|$, égale à 1, l'est. Cette intégrabilité au sens de Riemann permet de définir l'intégrale d'une classe plus large de fonctions au-delà des fonctions continues par morceaux.

On peut à présent redéfinir l'intégrale de a à b d'une fonction f continue par morceaux sur $[a, b]$.

Définition

Soient I un intervalle de $\mathbb{R}$, $(a, b) \in I^2$ et f une fonction de I vers $\mathbb{K}$ continue par morceaux sur I .

On appelle *intégrale de f de a à b* le scalaire, noté $\int_a^b f(t) dt$, défini par

$$\int_a^b f(t) dt = \begin{cases} \int_{[a,b]} f(t) dt & \text{si } a < b \\ 0 & \text{si } a = b \\ -\int_{[b,a]} f(t) dt & \text{si } a > b \end{cases}.$$

Dans le cas d'une fonction continue sur un intervalle I , on a de plus l'outil des primitives via *le théorème fondamental de l'analyse* (cf. page 76) qui dit, en particulier, que toute fonction continue sur un intervalle I admet des primitives sur I .

Ce théorème a alors pour conséquence : pour toute fonction continue de I vers $\mathbb{K}$,

$$\int_a^b f(t) dt = [F(t)]_a^b = F(b) - F(a)$$

où F est une primitive de f sur I et $(a, b) \in I^2$.

Espaces vectoriels normés

4

$\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Sauf cas particulier, tous les espaces vectoriels de ce chapitre sont des $\mathbb{K}$ -espaces vectoriels.

4.1 Normes

Définition

Soit E un $\mathbb{K}$ -espace vectoriel.

On appelle *norme* sur E toute application $\|\cdot\| : E \longrightarrow \mathbb{R}_+$ telle que pour tout $(x, y) \in E^2$ et tout $\lambda \in \mathbb{K}$,

- $\|x\| = 0 \implies x = 0$ (*séparation*);
- $\|\lambda x\| = |\lambda| \|x\|$ (*homogénéité*);
- $\|x + y\| \leq \|x\| + \|y\|$ (*inégalité triangulaire*).

On appelle *espace vectoriel normé* (sur $\mathbb{K}$)¹ tout couple² $(E, \|\cdot\|)$ où E est un $\mathbb{K}$ -espace vectoriel et $\|\cdot\|$ une norme sur E .

Sur $\mathbb{K}^n$ (où $n \in \mathbb{N}^*$), on utilise les trois normes usuelles $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$, définies pour tout $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ par

$$\|x\|_1 = \sum_{k=1}^n |x_k|, \quad \|x\|_2 = \left(\sum_{k=1}^n |x_k|^2 \right)^{1/2} \quad \text{et} \quad \|x\|_\infty = \max_{k \in \llbracket 1, n \rrbracket} |x_k|.$$

Sur $\mathcal{C}^0([a, b], \mathbb{K})$ l'espace des fonctions continues sur $[a, b]$, on utilise les trois normes usuelles $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$, appelées respectivement norme de la *convergence en moyenne*, de la *convergence en moyenne quadratique* et norme de la *convergence uniforme*, définies pour tout $f \in \mathcal{C}^0([a, b], \mathbb{K})$ par

1. Sans précision, un espace vectoriel normé est un espace vectoriel normé sur $\mathbb{K}$.

2. On présentera souvent dans les définitions ou théorèmes qui suivent des espaces vectoriels normés sans utiliser systématiquement la notation sous forme d'un couple de sorte qu'on emploiera souvent l'expression « Soit E un espace vectoriel normé ».

$$\|f\|_1 = \int_a^b |f(t)| \, dt, \quad \|f\|_2 = \left(\int_a^b |f(t)|^2 \, dt \right)^{1/2} \quad \text{et} \quad \|f\|_\infty = \max_{t \in [a, b]} |f(t)|.$$

Sur un espace préhilbertien réel $(E, \langle \cdot | \cdot \rangle)$, la *norme euclidienne* associée au produit scalaire $\langle \cdot | \cdot \rangle$ est définie pour tout $x \in E$ par

$$\|x\| = \sqrt{\langle x | x \rangle}.$$

Théorème

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Alors pour tout $(x, y) \in E^2$, on a

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

4.2 Distance

Définition

Soit E un ensemble non vide.

On appelle *distance* sur E toute application $d : E^2 \rightarrow \mathbb{R}_+$ telle que pour tout $(x, y) \in E^2$,

- $d(x, y) = 0 \iff x = y$ (*séparation*);
- $d(y, x) = d(x, y)$ (*symétrie*);
- $d(x, z) \leq d(x, y) + d(y, z)$ (*inégalité triangulaire*).

On appelle *espace métrique* tout couple (E, d) où E est un ensemble non vide et d une distance sur E .

Contrairement à la propriété de séparation pour une norme, il n'est pas possible de montrer l'implication $x = y \implies d(x, y) = 0$ à partir des deux autres propriétés.

Théorème

Soit (E, d) un espace métrique. Alors pour tout $(x, y, z) \in E^3$, on a

$$|d(x, y) - d(x, z)| \leq d(y, z).$$

Définition

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. On appelle *distance associée* à $\|\cdot\|$ l'application $d : E^2 \rightarrow \mathbb{R}_+$ définie pour tout $(x, y) \in E^2$ par

$$d(x, y) = \|x - y\|.$$

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé, d la distance associée à $\|\cdot\|$, A une partie non vide de E et $x \in E$.

On appelle *distance de x à A* le réel noté $d(x, A)$ défini par

$$d(x, A) = \inf_{a \in A} d(x, a) = \inf_{a \in A} \|x - a\|.$$

4.3 Produit d'espaces vectoriels normés

Soient $n \in \mathbb{N}^*$ et $(E_1, \|\cdot\|_{E_1}), (E_2, \|\cdot\|_{E_2}), \dots, (E_n, \|\cdot\|_{E_n})$, n espaces vectoriels normés.

Il est possible de définir plusieurs normes sur $E_1 \times E_2 \times \dots \times E_n$ comme $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ définies pour tout $x = (x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ par

$$\|x\|_1 = \sum_{k=1}^n \|x_k\|_{E_k}, \quad \|x\|_2 = \sqrt{\sum_{k=1}^n \|x_k\|_{E_k}^2} \quad \text{et} \quad \|x\|_\infty = \max_{k \in \llbracket 1, n \rrbracket} \|x_k\|_{E_k}.$$

Sans précision, l'espace vectoriel produit sera muni de la norme $\|\cdot\|_\infty$ et l'espace $(E_1 \times \dots \times E_n, \|\cdot\|_\infty)$ sera appelé *espace vectoriel normé produit* des espaces vectoriels normés $E_1, \dots, E_n$.

4.4 Suite d'éléments d'un espace vectoriel normé

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé et (x_n) une suite d'éléments de E . On dit que (x_n) *converge*³ s'il existe $\ell \in E$ tel que

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies \|x_n - \ell\| < \varepsilon.$$

On dit que (x_n) *diverge*⁴ sinon.

La limite éventuelle d'une suite dépend de la norme choisie.

En effet, soient $E = \mathcal{C}^0([0, 1], \mathbb{R})$ et (f_n) la suite définie pour tout $t \in [0, 1]$ par $f_n(t) = t^n$. En reprenant les définitions de $\|\cdot\|_1$ et $\|\cdot\|_\infty$ (cf. p. 83), on a

$$\|f_n\|_1 = \int_0^1 t^n dt = \frac{1}{n+1} \quad \text{et} \quad \|f_n\|_\infty = \sup_{t \in [0, 1]} (t^n) = 1$$

Donc (f_n) converge vers la fonction nulle pour $\|\cdot\|_1$ mais pas pour $\|\cdot\|_\infty$.

3. sous-entendu dans E pour la norme $\|\cdot\|$. On dit aussi (x_n) est *convergente*.

4. ou est *divergente*.

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé et (x_n) une suite d'éléments de E . On dit que (x_n) est *bornée*⁵ si : $\exists K \in \mathbb{R}_+, \forall n \in \mathbb{N}, \|x_n\| \leq K$.

Les propriétés classiques concernant la convergence des suites réelles ou complexes s'étendent aux suites d'éléments d'un espace vectoriel normé quelconque comme le montre le théorème suivant.

Théorème

Soient E un espace vectoriel normé, $\lambda \in \mathbb{K}$, (x_n) et (y_n) deux suites d'éléments de E . Alors

1. la limite éventuelle de (x_n) est unique ;
2. si (x_n) converge vers $\ell \in E$ et (y_n) converge vers $\ell' \in E$, alors $(\lambda x_n + y_n)$ converge vers $\lambda \ell + \ell'$;
3. toute suite convergente d'éléments de E est bornée.

Les notions de suites extraites et de valeur d'adhérence s'étendent également aux suites d'éléments d'un espace vectoriel normé quelconque.

Définition

Soient E un espace vectoriel normé et (x_n) une suite d'éléments de E .

On appelle *suite extraite*⁶ de (x_n) toute suite $(x_{\varphi(n)})$ où $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est strictement croissante.

On appelle *valeur d'adhérence* de (x_n) toute limite d'une suite extraite convergente de (x_n) .

Théorème

Soient E un espace vectoriel normé.

1. Si (x_n) est une suite d'éléments de E convergeant vers $\ell \in E$ alors toute suite extraite de (x_n) converge vers ℓ .
2. La limite d'une suite convergente d'éléments d'un espace vectoriel normé est l'unique valeur d'adhérence de cette suite.

Une suite d'éléments d'un espace vectoriel normé possédant au moins deux valeurs d'adhérence est donc nécessairement divergente.

5. Plus généralement, on dit qu'une fonction f d'un ensemble quelconque X (non nécessairement égal à $\mathbb{N}$) vers E est *bornée* s'il existe $K \in \mathbb{R}_+$ telle que pour tout $x \in X$, $\|f(x)\| \leq K$.

6. ou *sous-suite*.

L'unicité d'une valeur d'adhérence d'une suite d'éléments d'un espace vectoriel normé n'implique nullement la convergence de cette suite : par exemple, la suite $\left(n((-1)^n + 1)\right)$ n'admet que 0 comme valeur d'adhérence et diverge.

Théorème

Soient $p \in \mathbb{N}^*$, $(E_1 \times \cdots \times E_p, \|\cdot\|_\infty)$ espace vectoriel produit de p espaces vectoriels normés $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$, $\dots$, $(E_p, \|\cdot\|_{E_p})$ et (x_n) une suite d'éléments de $E_1 \times \cdots \times E_p$.

Notons pour tout $n \in \mathbb{N}$, $x_n = (x_n^{(1)}, \dots, x_n^{(p)})$.

Alors (x_n) converge vers $\ell = (\ell_1, \dots, \ell_p) \in E_1 \times \cdots \times E_p$ si, et seulement si, pour tout $k \in \llbracket 1, p \rrbracket$, $(x_n^{(k)})$ converge vers ℓ_k .

4.5 Boules et sphères

Définition

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. On appelle *boule ouverte* (respectivement *boule fermée*) de centre $a \in E$ et de rayon $r \in \mathbb{R}_+^*$ l'ensemble noté $B(a, r)$ ou $B_E(a, r)$ (respectivement $\overline{B}(a, r)$ ou $\overline{B}_E(a, r)$) défini par

$$B(a, r) = \{x \in E, \|x - a\| < r\}$$

$$(\text{respectivement } \overline{B}(a, r) = \{x \in E, \|x - a\| \leq r\}).$$

On appelle *sphère* de centre $a \in E$ et de rayon $r \in \mathbb{R}_+^*$ l'ensemble noté $S(a, r)$ défini par

$$S(a, r) = \{x \in E, \|x - a\| = r\}.$$

Dans l'espace vectoriel normé $(\mathbb{R}, |\cdot|)$,

$$B(a, r) =]a - r, a + r[\text{ et } \overline{B}(a, r) = [a - r, a + r].$$

Théorème

Toute boule ouverte ou fermée d'un espace vectoriel normé est convexe⁷.

Aucune sphère d'un espace vectoriel normé non réduit à $\{0\}$ n'est convexe, quelle que soit la norme choisie.

7. On rappelle qu'une partie A d'un $\mathbb{K}$ -espace vectoriel est *convexe* si pour tout $(x, y) \in A^2$, le segment $[x, y] \subset A$ c'est-à-dire si pour tout $(x, y) \in A^2$, pour tout $t \in [0, 1]$, $(1 - t)x + ty \in A$.

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé et $A \subset E$.

On dit que A est *bornée* s'il existe $K \in \mathbb{R}_+$ tel que pour tout $x \in A$, $\|x\| \leq K$.

Les boules d'un espace vectoriel normé sont bornées.

Une partie d'un espace vectoriel normé est bornée si, et seulement si, elle est incluse dans une boule.

4.6 Normes équivalentes

Définition

Soient E un $\mathbb{K}$ -espace vectoriel, $\|\cdot\|$ et $\|\cdot\|'$ deux normes sur E .

On dit que $\|\cdot\|$ et $\|\cdot\|'$ sont *équivalentes*⁸ s'il existe $(\lambda, \mu) \in (\mathbb{R}_+^*)^2$ tel que pour tout $x \in E$,

$$\lambda\|x\| \leq \|x\|' \leq \mu\|x\|.$$

La relation $\mathcal{R}$ définie par

$$\|\cdot\| \mathcal{R} \|\cdot\|' \text{ si } \|\cdot\| \text{ est équivalente à } \|\cdot\|'$$

est une relation d'équivalence sur l'ensemble des normes sur E .

Le théorème qui suit montre l'invariance du caractère borné d'une suite pour deux normes équivalentes et d'autre part que la notion de convergence d'une suite coïncide dans ce même cas de deux normes équivalentes : elles définissent les mêmes suites convergentes, qui ont la même limite pour les deux normes.

Théorème

Soient E un $\mathbb{K}$ -espace vectoriel, $\|\cdot\|$ et $\|\cdot\|'$ deux normes sur E . Alors

1. Si $\|\cdot\|$ et $\|\cdot\|'$ sont équivalentes, toute suite d'éléments de E bornée pour $\|\cdot\|$ est bornée pour $\|\cdot\|'$.
2. Les assertions suivantes sont équivalentes :
 - (i) $\|\cdot\|$ et $\|\cdot\|'$ sont équivalentes.
 - (ii) Toute suite (x_n) d'éléments de E converge vers $\ell \in E$ pour la norme $\|\cdot\|$ si, et seulement si, (x_n) converge vers ℓ pour la norme $\|\cdot\|'$.

8. On écrit parfois $\|\cdot\| \sim \|\cdot\|'$.

Pour montrer que deux normes $\|\cdot\|$ et $\|\cdot\|'$ ne sont pas équivalentes sur un espace vectoriel normé E , il suffit d'exhiber une suite (x_n) d'éléments de E telle que $\frac{\|x_n\|'}{\|x_n\|}$ ne soit pas bornée.

4.7 Voisinage, ouvert, fermé

Cette section aborde les notions fondamentales de la topologie d'un espace vectoriel normé. Commençons par la définition d'un *voisinage*.

Définition

Soient E un espace vectoriel normé, $V \subset E$ et $a \in E$.

On dit que V est un *voisinage* de a (dans E) si V contient une boule ouverte de centre a c'est-à-dire s'il existe $r \in \mathbb{R}_+^*$ tel que $B(a, r) \subset V$.

Théorème

Soient E un espace vectoriel normé et $a \in E$.

1. Une réunion quelconque⁹ de voisinages de a est un voisinage de a .
2. Une intersection finie de voisinages de a est un voisinage de a .

Définissons à présent les notions d'*ouvert* et de *fermé*.

Définition

Soient E un espace vectoriel normé et $U \subset E$. On dit que U est un *ouvert* de E si U est un voisinage de chacun de ses points c'est-à-dire si

$$\forall u \in U, \exists r \in \mathbb{R}_+^*, B(u, r) \subset U$$

On dit que $F \subset E$ est un *fermé* de E si le complémentaire de F dans E , noté $E \setminus F$ ou $\complement_E(F)$, est un ouvert de E .

Il existe des ensembles à la fois ouverts et fermés dans un espace vectoriel normé E : $\emptyset$ et E .

Il existe aussi des ensembles ni ouverts, ni fermés : par exemple dans l'espace vectoriel normé $(\mathbb{R}, |\cdot|)$, $[0, 1[$ n'est ni ouvert, ni fermé.

Dans un espace vectoriel normé E , tout singleton $\{x\}$ est un fermé de E .

Les notions d'ouvert et de fermé sont invariants par passage à une norme équivalente : si deux normes sont équivalentes sur un espace vectoriel normé, tout

9. c'est-à-dire $\bigcup_{i \in I} V_i$ où I un ensemble non vide et $(V_i)_{i \in I}$ une famille de voisinages de a .

ouvert (respectivement fermé) pour l'une est ouvert (respectivement fermé) pour l'autre¹⁰

Théorème

Soit E un espace vectoriel normé. Alors

1. Toute boule ouverte est un ouvert de E .
2. Toute boule fermée est un fermé de E .
3. Une réunion quelconque d'ouverts de E est un ouvert de E .
4. Une intersection finie d'ouverts de E est un ouvert de E .
5. Une réunion finie de fermés de E est un fermé de E .
6. Une intersection quelconque de fermés de E est un fermé de E .

Dans un espace vectoriel normé, toute partie finie est fermée car réunion d'un nombre fini de singletons.

Une intersection quelconque d'ouverts d'un espace vectoriel normé $(E, \|\cdot\|)$ n'est pas nécessairement un ouvert de E . Par exemple dans $(\mathbb{R}, |\cdot|)$, pour tout $n \in \mathbb{N}^*$, $U_n =]-\frac{1}{n}, \frac{1}{n}[$ est un ouvert de $\mathbb{R}$ or $\bigcap_{n \in \mathbb{N}^*} U_n = \{0\}$ n'est pas un ouvert de $\mathbb{R}$ car pour tout $r \in \mathbb{R}_+^*$, $]r, r[\not\subset \{0\}$.

Une réunion quelconque de fermés d'un espace vectoriel normé $(E, \|\cdot\|)$ n'est pas nécessairement un fermé de E . Par exemple dans $(\mathbb{R}, |\cdot|)$, pour tout $n \in \mathbb{N}^*$, $F_n = [\frac{1}{n}, 1]$ est un fermé de $\mathbb{R}$ or $\bigcup_{n \in \mathbb{N}^*} F_n =]0, 1]$ n'est pas un fermé de $\mathbb{R}$.

Théorème

Soient $(E_1 \times \cdots \times E_n, \|\cdot\|_\infty)$ espace vectoriel produit de n espaces vectoriels normés $(E_1, \|\cdot\|_{E_1}), \dots, (E_n, \|\cdot\|_{E_n})$.

1. Soient $U_1, \dots, U_n$, n ouverts respectivement de $E_1, \dots, E_n$.
Alors $U_1 \times \cdots \times U_n$ est un ouvert de $E_1 \times \cdots \times E_n$.
2. Soient $F_1, \dots, F_n$, n fermés respectivement de $E_1, \dots, E_n$.
Alors $F_1 \times \cdots \times F_n$ est un fermé de $E_1 \times \cdots \times E_n$.

Définissons maintenant les notions d'*intérieur*, d'*adhérence* et de *frontière*.

10. On dit parfois que deux normes équivalentes sur un espace vectoriel normé E définissent la même *topologie* de E où la topologie de E désigne l'ensemble des ouverts de E .

Définition

Soient E un espace vectoriel normé, $A \subset E$ et $a \in A$.

1. On dit que a est un *point intérieur* à A si A est un voisinage de a c'est-à-dire s'il existe $r \in \mathbb{R}_+^*$ tel que $B(a, r) \subset A$.

L'ensemble des points intérieurs à A est appelé *intérieur* de A et noté $\overset{\circ}{A}$.

2. On dit que a est un *point adhérent* à A si tout voisinage de a rencontre A c'est-à-dire si pour tout voisinage V de a , $V \cap A \neq \emptyset$.

L'ensemble des points adhérents à A est appelé *adhérence* de A et noté $\overline{A}$.

3. On appelle *frontière* de A , l'ensemble noté $\text{Fr}(A) = \overline{A} \setminus \overset{\circ}{A}$.

Par définition, on a $\overset{\circ}{A} \subset A \subset \overline{A}$.

Théorème

Soient E un espace vectoriel normé et $A \subset E$. Alors

1. $\overset{\circ}{A}$ est le plus grand ouvert de E inclus dans A .
2. A ouvert de $E \iff A = \overset{\circ}{A}$.
3. $E \setminus \overline{A} = \widehat{E \setminus A}$ et $\overline{E \setminus A} = E \setminus \overset{\circ}{A}$.
4. $\overline{A}$ est le plus petit fermé de E contenant A .
5. A fermé de $E \iff A = \overline{A}$.
6. $\text{Fr}(A)$ est un fermé de E .

Théorème

Soient E un espace vectoriel normé, $A \subset E$ et $B \subset E$. Alors

1. $A \subset B \implies \overset{\circ}{A} \subset \overset{\circ}{B}$ et $A \subset B \implies \overline{A} \subset \overline{B}$.
2. $\widehat{A \cap B} = \overset{\circ}{A} \cap \overset{\circ}{B}$ et $\overset{\circ}{A} \cup \overset{\circ}{B} \subset \widehat{A \cup B}$.
3. $\overline{A \cup B} = \overline{A} \cup \overline{B}$ et $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$.

4.8 Caractérisations séquentielles

Théorème (caractérisation séquentielle des points adhérents)

Soient E un espace vectoriel normé, $A \subset E$ et $x \in E$.

Alors $x \in \overline{A}$ si, et seulement si, il existe une suite d'éléments de A convergent vers x .

Corollaire (caractérisation séquentielle des fermés)

Soient E un espace vectoriel normé et $F \subset E$.

Alors F est un fermé de E si, et seulement si, pour toute suite (x_n) d'éléments de F convergeant dans E , $\lim_{n \rightarrow +\infty} x_n \in F$.

Ainsi un fermé contient les limites de ses suites convergentes.

Théorème

Soient $(E, \|\cdot\|)$ un espace vectoriel normé, d la distance associée à $\|\cdot\|$, A une partie non vide de E et $x \in E$.

Alors $d(x, A) = 0 \iff x \in \overline{A}$.

4.9 Partie dense

Définition

Soient E un espace vectoriel normé, $A \subset E$ et $B \subset A$.

On dit que A est *dense* dans E si $\overline{A} = E$.

On dit que B est *dense* dans A si $A \subset \overline{B}$.

Théorème

Soient E un espace vectoriel normé, $A \subset E$ et $B \subset A$. Alors

1. A dense dans E si, et seulement si, tout élément de E est limite d'une suite d'éléments de A .
2. B est dense dans A si, et seulement si, tout élément de A est limite d'une suite d'éléments de B .

Par exemple $\mathbb{Q}$ est dense dans $\mathbb{R}$ pour la norme $|\cdot|$.

De même, l'espace des fonctions polynomiales de $[a, b]$ vers $\mathbb{K}$ est dense dans l'espace des fonctions continues de $[a, b]$ vers $\mathbb{K}$ pour la norme de la convergence uniforme $\|\cdot\|_\infty$ (cf. théorème de Weierstrass page 155).

4.10 Voisinages relatifs, ouverts et fermés relatifs

Si F est un sous-espace vectoriel d'un espace vectoriel normé E , alors F est également un espace vectoriel normé où la norme sur F est la restriction à F de la norme sur E .

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé et F un sous-espace vectoriel de E .

On appelle *boule ouverte* de F de centre $a \in F$ et de rayon $r \in \mathbb{R}_+^*$ l'ensemble noté $B_F(a, r)$ défini par

$$B_F(a, r) = \{x \in F, \|x - a\| < r\}.$$

Remarquons que $B_F(a, r) = F \cap B(a, r)$.

Définition

Soient E un espace vectoriel normé, $A \subset E$, $V \subset A$ et $a \in A$.

On dit que V est un *voisinage relatif*¹¹ de a dans A si V contient une boule ouverte de A de centre a c'est-à-dire s'il existe $r \in \mathbb{R}_+^*$ tel que $B_A(a, r) \subset V$.

Théorème

Soient E un espace vectoriel normé, $A \subset E$, $V \subset A$ et $a \in A$.

Alors V est un voisinage relatif de a dans A si, et seulement si, V est l'intersection de A et d'un voisinage de a (dans E).

Définition

Soient E un espace vectoriel normé, $A \subset E$ et $U \subset A$. On dit que U est un *ouvert relatif* de A si U est un voisinage relatif de chacun de ses points c'est-à-dire si

$$\forall u \in U, \exists r \in \mathbb{R}_+^*, B_A(u, r) \subset U$$

On dit que $F \subset A$ est un *fermé relatif* de A si le complémentaire de F dans A , noté $A \setminus F$ ou $\mathcal{C}_A(F)$, est un ouvert relatif de A .

Théorème

Soient E un espace vectoriel normé, $A \subset E$, $U \subset A$ et $F \subset A$. Alors

1. U est un ouvert relatif de A si, et seulement si, U est l'intersection de A et d'un ouvert de E .
2. F est un fermé relatif de A si, et seulement si, F est l'intersection de A et d'un fermé de E .

11. On dit aussi parfois V est un voisinage de a pour la *topologie induite* sur A ou encore V est un voisinage de a dans A .

Théorème (caractérisation séquentielle des fermés relatifs)

Soient E un espace vectoriel normé, $A \subset E$ et $F \subset A$.

Alors F est un fermé relatif de A si et seulement si, pour toute suite (x_n) d'éléments de F convergeant dans A , $\lim_{n \rightarrow +\infty} x_n \in F$.

4.11 Limites dans les espaces vectoriels normés**Définition**

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $A \subset E$, $a \in \overline{A}$ et f une application de A vers F .

On dit que f admet une *limite* $\ell \in F$ en a si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in A, \|x - a\|_E < \alpha \implies \|f(x) - \ell\|_F < \varepsilon.$$

Réécrivons la définition ci-dessus en utilisant les boules ouvertes :

f admet une limite ℓ en a si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in A, x \in B(a, \alpha) \implies f(x) \in B(\ell, \varepsilon).$$

Réécrivons-la en utilisant à présent les voisinages et en notant $\mathcal{V}_E(a)$ (respectivement $\mathcal{V}_F(a)$) l'ensemble des voisinages de a dans E (respectivement dans F).

f admet une limite ℓ en a si

$$\forall V \in \mathcal{V}_F(\ell), \exists W \in \mathcal{V}_E(a), \forall x \in A, x \in W \implies f(x) \in V.$$

que l'on peut également écrire sous la forme

$$\forall V \in \mathcal{V}_F(\ell), \exists W \in \mathcal{V}_E(a), \forall x \in A \cap W, f(x) \in V.$$

Théorème

Soient E et F deux espaces vectoriels normés, $A \subset E$, $a \in \overline{A}$ et f une application de A vers F . Si f admet une limite $\ell \in F$ en a , alors celle-ci est unique. On écrit alors $\lim_{x \rightarrow a} f(x) = \ell$ ou $\lim_a f = \ell$.

Théorème

Soient E et F deux espaces vectoriels normés, $A \subset E$, $a \in \overline{A}$, $\ell \in F$ et f une application de A vers F .

Si $\lim_{x \rightarrow a} f(x) = \ell$ alors f est bornée au voisinage de a .

Théorème (caractérisation séquentielle de la limite)

Soient E et F deux espaces vectoriels normés, $A \subset E$, $a \in \overline{A}$, $\ell \in F$ et f une application de A vers F .

Alors les assertions suivantes sont équivalentes :

- (i) $\lim_{x \rightarrow a} f(x) = \ell$.
- (ii) Pour toute suite (x_n) de A convergeant vers a , la suite $(f(x_n))$ converge vers ℓ .

Les deux définitions suivantes présentent trois extensions de la notion de limite.

Définition

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $A \subset E$, $a \in \overline{A}$ et f une application de A vers F .

1. Lorsque $E = \mathbb{R}$, on dit que f admet la limite $\ell \in F$ en $+\infty$ si

$$\forall \varepsilon > 0, \exists K > 0, \forall x \in A, x > K \implies \|f(x) - \ell\|_F < \varepsilon.$$

2. Lorsque $F = \mathbb{R}$, on dit que f admet la limite $+\infty$ (respectivement $-\infty$) en a si

$$\forall K > 0, \exists \alpha > 0, \forall x \in A, \|x - a\|_E < \alpha \implies f(x) > K$$

$$(\text{respectivement } \forall K > 0, \exists \alpha > 0, \forall x \in A, \|x - a\|_E < \alpha \implies f(x) < K).$$

Définition

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $\ell \in F$ et f une application de E vers F .

On dit que $f(x)$ admet la limite ℓ lorsque $\|x\|_E$ tend vers $+\infty$ si

$$\forall \varepsilon > 0, \exists K > 0, \forall x \in E, \|x\|_E > K \implies \|f(x) - \ell\|_F < \varepsilon.$$

Théorème

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $A \subset E$, $a \in \overline{A}$, $\lambda \in \mathbb{K}$, $(\ell, \ell') \in F^2$, f et g deux applications de A vers F et φ une application de A vers $\mathbb{K}$. Alors

1. Si $\lim_{x \rightarrow a} f(x) = \ell$ et $\lim_{x \rightarrow a} g(x) = \ell'$ alors $\lim_{x \rightarrow a} (f(x) + g(x)) = \ell + \ell'$.
2. Si $\lim_{x \rightarrow a} \varphi(x) = \lambda$ et $\lim_{x \rightarrow a} f(x) = \ell$ alors $\lim_{x \rightarrow a} \varphi(x)f(x) = \lambda \ell$.
3. Si $\lim_{x \rightarrow a} f(x) = \ell$ alors $\lim_{x \rightarrow a} \|f(x)\|_F = \|\ell\|_F$.

Théorème

Soient E , F et G , trois espaces vectoriels normés, $A \subset E$, $B \subset F$, $a \in \overline{A}$, $b \in \overline{B}$, f une application de A vers F telle que $f(A) \subset B$, g une application de B vers G et $\ell \in G$.

Si $\lim_{x \rightarrow a} f(x) = b$ et $\lim_{x \rightarrow b} g(x) = \ell$ alors $\lim_{x \rightarrow a} (g \circ f)(x) = \ell$.

Étudions à présent le cas d'une application à valeurs dans un produit d'espaces vectoriels normés.

Soient $p \in \mathbb{N}^*$, $(E, \|\cdot\|_E)$ un espace vectoriel normé, $(F_1 \times \cdots \times F_p, \|\cdot\|_\infty)$ espace vectoriel produit de p espaces vectoriels normés $(F_1, \|\cdot\|_{F_1}), \dots, (F_p, \|\cdot\|_{F_p})$ et f une application de A vers $F_1 \times \cdots \times F_p$.

On rappelle que si $y = (y_1, \dots, y_p) \in F_1 \times \cdots \times F_p$, la norme $\|\cdot\|_\infty$ est définie par

$$\|y\|_\infty = \max_{k \in \llbracket 1, p \rrbracket} \|y_k\|_{F_k}.$$

On notera, pour tout $k \in \llbracket 1, p \rrbracket$, f_k les applications définies de A vers F_k par

$$f_k = p_k \circ f \text{ où } p_k : \begin{cases} F_1 \times \cdots \times F_p & \longrightarrow F_k \\ (y_1, \dots, y_p) & \longmapsto y_k \end{cases}.$$

Théorème

Soient $p \in \mathbb{N}^*$, E un espace vectoriel normé, $(F_1 \times \cdots \times F_p, \|\cdot\|_\infty)$ espace vectoriel produit de p espaces vectoriels normés $F_1, \dots, F_p$, $A \subset E$, $a \in \overline{A}$,

$\ell = (\ell_1, \dots, \ell_p) \in F_1 \times \cdots \times F_p$ et $f : \begin{cases} A & \longrightarrow F_1 \times \cdots \times F_p \\ x & \longmapsto (f_1(x), \dots, f_p(x)) \end{cases}.$

Alors $\lim_{x \rightarrow a} f(x) = \ell \iff \forall k \in \llbracket 1, p \rrbracket, \lim_{x \rightarrow a} f_k(x) = \ell_k$.

4.12 Continuité

Définition

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $A \subset E$, $a \in A$ et f une application de A vers F .

On dit que f est *continue* en a si $\lim_{x \rightarrow a} f(x) = f(a)$ c'est-à-dire si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in A, \|x - a\|_E < \alpha \implies \|f(x) - f(a)\|_F < \varepsilon.$$

On dit que f est continue sur A si f est continue en tout point de A .

Théorème (caractérisation séquentielle de la continuité)

Soient E et F deux espaces vectoriels normés, $A \subset E$, $a \in A$ et f une application de A vers F . Alors les assertions suivantes sont équivalentes :

- (i) f continue en a .
- (ii) Pour toute suite (x_n) de A convergeant vers a , la suite $(f(x_n))$ converge vers $f(a)$.

Théorème

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $A \subset E$, $a \in A$, f, g deux applications de A vers F et φ une application de A vers $\mathbb{K}$. Alors

1. Si f et g sont continues en a (respectivement sur A) alors $f + g$ est continue en a (respectivement sur A).
2. Si φ et f sont continues en a (respectivement sur A) alors φf est continue en a (respectivement sur A).
3. Si f est continue en a (respectivement sur A) alors l'application

$$\begin{cases} A & \longrightarrow \mathbb{R} \\ x & \longmapsto \|f(x)\|_F \end{cases}$$
 est continue en a (respectivement sur A).

Théorème

Soient E, F et G , trois espaces vectoriels normés, $A \subset E$, $B \subset F$, $a \in A, b \in \overline{B}$, f une application de A vers F telle que $f(A) \subset B$ et g une application de B vers G .

Si f est continue en a (respectivement sur A) et g continue en $f(a)$ (respectivement sur B), alors $g \circ f$ est continue en a (respectivement sur A).

Théorème

Soient E et F deux espaces vectoriels normés, $A \subset E$ et f une application de A vers F . Alors

1. f continue sur A si, et seulement si, l'image réciproque de tout ouvert de F est un ouvert de A .
2. f continue sur A si et seulement si, l'image réciproque de tout fermé de F est un fermé de A .

Ce théorème permet souvent de montrer aisément qu'une partie est ouverte ou fermée. Par exemple, si E est un espace vectoriel normé et $f : E \rightarrow \mathbb{R}$ est continue sur E , alors $\{x \in E, f(x) = 0\} = f^{-1}(\{0\})$ est un fermé de E .

De même $\{x \in E, f(x) < 0\} = f^{-1}(\mathbb{R}_-^*)$ est un ouvert de E .

Corollaire

Soient E et F deux espaces vectoriels normés, f et g deux applications de E vers F , continues sur E et coïncidant sur une partie dense de E .
Alors $f = g$.

4.13 Uniforme continuité**Définition**

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $A \subset E$ et f une application de A vers F .

On dit que f est *uniformément continue* (sur A) si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall (x, y) \in A^2, \|x - y\|_E < \alpha \implies \|f(x) - f(y)\|_F < \varepsilon.$$

Toute application uniformément continue sur A est continue sur A mais la réciproque est fautive : par exemple, la fonction $x \mapsto x^2$ est continue sur $\mathbb{R}_+$ mais n'est pas uniformément continue sur $\mathbb{R}_+$.

La somme de deux fonctions uniformément continues et le produit d'une fonction uniformément continue par un scalaire dans $\mathbb{K}$ sont uniformément continues.

En revanche, en considérant $(E, \|\cdot\|_E) = (F, \|\cdot\|_F) = (\mathbb{R}, |\cdot|)$, le produit de deux fonctions uniformément continues n'est pas nécessairement uniformément continue : prenons par exemple la fonction $f : x \mapsto x$. Alors f est uniformément continue sur $\mathbb{R}_+$ mais $f^2 : x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}_+$.

Soient E, F et G trois espaces vectoriels normés, $A \subset E$, $B \subset F$, f une application de A vers F et g une application de B vers G tels que $f(A) \subset B$.

Si f et g sont uniformément continues (respectivement sur A et B), $g \circ f$ est uniformément continue (sur A).

4.14 Applications lipschitziennes**Définition**

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés, $A \subset E$, f une application de A vers F et $K \in \mathbb{R}_+$.

On dit que f est *K-lipschitzienne* (sur A) si pour tout $(x, y) \in A^2$,

$$\|f(x) - f(y)\|_F \leq K \|x - y\|_E.$$

On dit que f est *lipschitzienne* (sur A) s'il existe $K \in \mathbb{R}_+$ tel que f est K -lipschitzienne.

On dit que f est *contractante* (sur A) s'il existe $K \in [0, 1[$ tel que f est K -lipschitzienne.

Pour tout espace vectoriel normé $(E, \|\cdot\|)$, l'application $\|\cdot\| : \begin{cases} E & \longrightarrow \mathbb{R} \\ x & \longmapsto \|x\| \end{cases}$ est lipschitzienne. En effet, pour tout $(x, y) \in E^2$, on a

$$|\|x\| - \|y\|| \leq \|x - y\|$$

c'est-à-dire l'application $\|\cdot\|$ est 1-lipschitzienne.

De même on montre que pour tout espace vectoriel normé $(E, \|\cdot\|)$ et toute partie non vide A de E , l'application $d_A : \begin{cases} E & \longrightarrow \mathbb{R} \\ x & \longmapsto d(x, A) = \inf_{a \in A} \|x - a\| \end{cases}$ est 1-lipschitzienne.

La somme de deux applications lipschitziennes et le produit d'une application lipschitzienne par un scalaire de $\mathbb{K}$ sont lipschitziennes.

Soient E, F et G trois espaces vectoriels normés, $A \subset E, B \subset F, f$ une application de A vers F et g une application de B vers G tels que $f(A) \subset B$.

Si f et g sont lipschitziennes (respectivement sur A et B), $g \circ f$ est lipschitzienne (sur A).

Théorème

Soient E et F deux espaces vectoriels normés, $A \subset E$ et f une application de A vers F lipschitzienne. Alors f est uniformément continue.

La réciproque du théorème précédent est fausse.

Par exemple, en prenant $(E, \|\cdot\|_E) = (F, \|\cdot\|_F) = (\mathbb{R}, |\cdot|)$, la fonction $x \mapsto \sqrt{x}$ est uniformément continue sur $\mathbb{R}_+$ mais n'est pas lipschitzienne.

Pour tout espace vectoriel normé $(E, \|\cdot\|)$ et toute partie non vide A de E , les applications

$$\|\cdot\| : \begin{cases} E & \longrightarrow \mathbb{R} \\ x & \longmapsto \|x\| \end{cases} \quad \text{et} \quad d_A : \begin{cases} E & \longrightarrow \mathbb{R} \\ x & \longmapsto d(x, A) = \inf_{a \in A} \|x - a\| \end{cases}$$

sont uniformément continues (donc continues) car lipschitziennes.

4.15 Applications linéaires continues

Le théorème qui suit permet de caractériser la *continuité* d'une application *linéaire* entre deux espaces vectoriels normés.

Théorème

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et $f \in \mathcal{L}(E, F)$.

Alors les assertions suivantes sont équivalentes :

- (i) f est continue sur E .
- (ii) f est continue en 0.
- (iii) f est bornée sur $\overline{B_E}(0, 1)$.
- (iv) Il existe $K \in \mathbb{R}_+$ tel que pour tout $x \in E$, $\|f(x)\|_F \leq K\|x\|_E$.
- (v) f est lipschitzienne.
- (vi) f est uniformément continue.

On note $\mathcal{L}_c(E, F)$ l'ensemble des applications linéaires continues de E vers F .

Comme nous le verrons dans une section suivante, si E est de plus de dimension finie toute application linéaire de E vers F est continue.

En pratique, pour montrer la continuité d'une application linéaire entre deux espaces vectoriels normés, la caractérisation (iv) est la plus utile.

En particulier pour montrer qu'une telle application linéaire n'est pas continue, il suffit par exemple de montrer qu'il existe une suite (x_n) d'éléments de E telle que pour tout $n \in \mathbb{N}$, $\|f(x_n)\|_F > n\|x_n\|_E$.

Définissons à présent la notion de *norme subordonnée* d'une application linéaire continue appelée également *norme d'opérateur*.

Définition

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et $f \in \mathcal{L}_c(E, F)$.

Alors l'application
$$\begin{cases} \mathcal{L}_c(E, F) & \longrightarrow \mathbb{R}_+ \\ f & \longmapsto \sup_{x \in E, x \neq 0} \frac{\|f(x)\|_F}{\|x\|_E} \end{cases} \text{ est une norme}$$
 appelée *norme d'opérateur*¹² et notée¹³ $||| \cdot |||$.

Pour tout $f \in \mathcal{L}_c(E, F)$, on a aussi

$$|||f||| = \sup_{x \in \overline{B}(0,1)} \|f(x)\|_F = \sup_{x \in S(0,1)} \|f(x)\|_F.$$

12. ou *norme* sur $\mathcal{L}_c(E, F)$ *subordonnée* à $\|\cdot\|_E$ et $\|\cdot\|_F$, ou encore *norme linéaire*. Pour tout $f \in \mathcal{L}_c(E, F)$, on dit que $|||f|||$ est la *norme de l'opérateur* f ou la *norme de f subordonnée* à celles de E et F ou simplement la *norme subordonnée* de f .

13. ou $\|\cdot\|_{\text{op}}$.

Théorème

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et $f \in \mathcal{L}_c(E, F)$.

Alors
$$|||f||| = \inf \left\{ K \in \mathbb{R}_+, \forall x \in E, \|f(x)\|_F \leq K \|x\|_E \right\}.$$

En particulier, pour tout $x \in E$, on a

$$\|f(x)\|_F \leq |||f||| \|x\|_E.$$

Donc, si $f \in \mathcal{L}_c(E, F)$, $|||f|||$ est « la meilleure » constante $K \in \mathbb{R}_+$ (c'est-à-dire le plus petit $K \in \mathbb{R}_+$) vérifiant pour tout $x \in E$, $\|f(x)\|_F \leq K \|x\|_E$.

Énonçons à présent la propriété de *sous-multiplicativité* de la norme d'opérateur.

Théorème

Soit E , F et G trois espaces vectoriels normés.

Soient $f \in \mathcal{L}_c(E, F)$ et $g \in \mathcal{L}_c(F, G)$.

Alors $g \circ f \in \mathcal{L}_c(E, G)$ et $|||g \circ f||| \leq |||g||| |||f|||$.

Énonçons à présent la caractérisation de la continuité pour les applications multilinéaires après avoir rappelé la définition de ces dernières.

Définition

Soient $n \in \mathbb{N}^*$, $E_1, \dots, E_n$ et F des espaces vectoriels.

On dit que $f : E_1 \times \dots \times E_n \rightarrow F$ est *multilinéaire*¹⁴ si pour tout $k \in \llbracket 1, n \rrbracket$ et pour tout $(a_1, \dots, a_n) \in E_1 \times \dots \times E_n$, l'application

$$f_k : \begin{cases} E_k & \longrightarrow F \\ x_k & \longmapsto f(a_1, \dots, a_{k-1}, x_k, a_{k+1}, \dots, a_n) \end{cases}$$

est linéaire.

Théorème

Soient $(E_1, \|\cdot\|_{E_1}), \dots, (E_n, \|\cdot\|_{E_n}), (F, \|\cdot\|_F)$ des espaces vectoriels normés et $f : E_1 \times \dots \times E_n \rightarrow F$ multilinéaire. Alors les assertions suivantes sont équivalentes :

- (i) f est continue sur $E_1 \times \dots \times E_n$.
- (ii) Il existe $K \in \mathbb{R}_+$ tel que pour tout $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$,

$$\|f(x_1, \dots, x_n)\|_F \leq K \|x_1\|_{E_1} \cdots \|x_n\|_{E_n}.$$

14. On dit aussi f *n-linéaire*. Dans le cas $n = 2$, on dit que f est *bilinéaire*.

4.16 Parties compactes

Définition

Soient E un espace vectoriel normé et $A \subset E$.

On dit que A est *compact* si¹⁵ de toute suite de A , on peut extraire une sous-suite convergente dont la limite est dans A .

A compact signifie donc que toute suite de A admet au moins une valeur d'adhérence dans A .

Théorème

Soit E un espace vectoriel normé.

Toute partie compacte de E est fermée et bornée.

Énonçons le cas particulier de $\mathbb{R}^n$ qui sera généralisé à tout espace vectoriel normé de dimension finie page 106.

Théorème

Soient $n \in \mathbb{N}^*$ et $A \subset \mathbb{R}^n$.

Alors A compact si, et seulement si, A est fermé et borné.

Théorème

Soient E un espace vectoriel normé et A une partie compacte de E .

Alors tout fermé (relatif) de A est compact.

Théorème

Soient E un espace vectoriel normé et A une partie compacte de E .

Alors toute suite de A converge si, et seulement si, elle admet une unique valeur d'adhérence.

Énonçons maintenant le résultat concernant un produit fini de compacts.

15. Cette définition s'appelle parfois la *propriété de Bolzano-Weierstrass*.

Théorème

Soient $n \in \mathbb{N}^*$, $E_1, \dots, E_n$, n espaces vectoriels normés, $A_1, \dots, A_n$ des parties compactes de $E_1, \dots, E_n$.

Alors $A_1 \times \dots \times A_n$ est une partie compacte de $E_1 \times \dots \times E_n$.

4.17 Continuité sur une partie compacte

Théorème

Soient E et F deux espaces vectoriels normés, A une partie compacte de E et $f : A \rightarrow F$ continue sur A . Alors $f(A)$ est une partie compacte de F .

L'image réciproque d'un compact par une application continue n'est pas nécessairement un compact. Par exemple soit $f : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto 1 \end{cases}$. Alors f est continue, $f^{-1}(\{1\}) = \mathbb{R}$ or $\{1\}$, fermé, borné, est compact via le théorème page 102 et $\mathbb{R}$, non borné, n'est pas compact.

Théorème de Heine

Soient E et F deux espaces vectoriels normés, A une partie compacte de E et $f : A \rightarrow F$ continue sur A . Alors f est uniformément continue sur A .

Théorème des bornes atteintes

Soient E un espace vectoriel normé, A une partie compacte non vide de E et $f : A \rightarrow \mathbb{R}$ continue sur A . Alors f est bornée sur A et atteint ses bornes.

4.18 Connexité par arcs

Définition

Soient E un espace vectoriel normé et A une partie non vide de E et $(a, b) \in A^2$. On appelle *arc*¹⁶ (ou *chemin*) joignant a et b toute application continue γ de $[0, 1]$ vers A vérifiant $\gamma(0) = a$ et $\gamma(1) = b$.

Les points a et b sont alors appelés les *extrémités* de l'arc¹⁷.

16. Certains appellent arc joignant a et b toute application continue γ de $[u, v]$ vers A (où $(u, v) \in \mathbb{R}^2$ tel que $u \neq v$) mais on peut se ramener systématiquement au segment $[0, 1]$: il suffit de considérer alors α de $[0, 1]$ vers A définie pour tout $t \in [0, 1]$ par $\alpha(t) = \gamma(u + (v - u)t)$.

17. Certains appellent $a = \gamma(0)$ l'*origine* de l'arc et $b = \gamma(1)$ son *extrémité*.

Définition

Soient E un espace vectoriel normé et A une partie non vide de E .

On dit que A est *connexe par arcs* si pour tout $(a, b) \in A^2$, il existe un arc joignant a et b .

Théorème

Soient E un espace vectoriel normé et A une partie non vide de E . Alors la relation $\mathcal{R}$ définie pour tout $(a, b) \in A^2$ par

$$a \mathcal{R} b \text{ s'il existe un arc joignant } a \text{ et } b$$

est une relation d'équivalence.

On appelle *composantes connexes par arcs* de A les classes d'équivalence de A relativement à la relation $\mathcal{R}$.

La convexité est un cas particulier de la connexité par arcs comme l'explique le théorème qui suit.

Théorème

Soit E un espace vectoriel normé. Toute partie convexe de E est connexe par arcs.

Énonçons le cas particulier de $\mathbb{R}$.

Théorème

Soit $A \subset \mathbb{R}$.

Alors A connexe par arcs si, et seulement si, A est un intervalle.

Définition

Soient E un espace vectoriel normé.

On dit qu'une partie A de E est *étoilée par rapport à un point* $a \in A$ si pour tout $x \in A$, $[a, x] \subset A$.

On dit que A est *étoilée* si A est étoilée par rapport à un de ses points.

Une partie A de E est convexe si, et seulement si, elle est étoilée par rapport à chacun de ses points.

Théorème

Soit E un espace vectoriel normé.

Toute partie étoilée de E est connexe par arcs.

Théorème

Soient E et F deux espaces vectoriels normés, A une partie connexe par arcs de E et $f : A \rightarrow F$ continue sur A .

Alors $f(A)$ est une partie connexe par arcs de F .

Corollaire (théorème des valeurs intermédiaires)

Soient E un espace vectoriel normé, A une partie connexe par arcs de E et $f : A \rightarrow \mathbb{R}$ continue sur A . Alors $f(A)$ est un intervalle.

4.19 Espaces vectoriels normés de dimension finie

Théorème

Soit E un espace vectoriel normé de dimension finie¹⁸.

Alors toutes les normes sur E sont équivalentes.

Ce théorème permet, dans le cas de la dimension finie, de ne plus préciser quelle est la norme utilisée dans les diverses notions topologiques (voisinages, ouverts, fermés, compacts...).

Soient E un espace vectoriel de dimension finie $p \in \mathbb{N}^*$, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et (u_n) une suite d'éléments de E . Alors, pour tout $n \in \mathbb{N}$,

$$u_n = u_n^{(1)} e_1 + \dots + u_n^{(p)} e_p \quad \text{où les } u_n^{(i)} \in \mathbb{K}.$$

Les suites $(u_n^{(i)})$ à valeurs dans $\mathbb{K}$ sont appelées *suites-coordonnées* de (u_n) dans la base $\mathcal{B}$.

18. Rappelons que dans tout ce chapitre, on ne considère que des $\mathbb{K}$ -espaces vectoriels. Ce théorème peut devenir faux pour un espace vectoriel normé sur un corps différent de $\mathbb{K}$ par exemple le $\mathbb{Q}$ -espace vectoriel $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2}; (a, b) \in \mathbb{Q}^2\}$ de dimension finie 2 : on peut montrer que les deux normes $\|\cdot\|$ et $\|\cdot\|'$ définies pour tout $u = a + b\sqrt{2} \in \mathbb{Q}[\sqrt{2}]$ par $\|u\| = |a + b\sqrt{2}|$ et $\|u\|' = |a| + |b|$ ne sont pas équivalentes.

Théorème

Soit E un espace vectoriel normé de dimension finie $p \in \mathbb{N}^*$ et (u_n) une suite d'éléments de E .

Alors (u_n) converge si, et seulement si, $(u_n^{(1)}), \dots, (u_n^{(p)})$ convergent.

De plus, dans ce cas, on a

$$\lim_{n \rightarrow +\infty} u_n = \left(\lim_{n \rightarrow +\infty} u_n^{(1)} \right) e_1 + \dots + \left(\lim_{n \rightarrow +\infty} u_n^{(p)} \right) e_p.$$

Théorème

Soient E un espace vectoriel normé de dimension finie et $A \subset E$.

Alors A compact si, et seulement si, A est fermé et borné.

Par exemple, en dimension finie, $\overline{B}(0, 1)$ et $S(0, 1)$ sont compacts.

Corollaire (théorème de Bolzano-Weierstrass)

Soit E un espace vectoriel normé de dimension finie.

Alors de toute suite bornée de E , on peut extraire une sous-suite convergente dans E .

Énonçons également un autre corollaire.

Corollaire

Une suite bornée d'un espace vectoriel de dimension finie converge si, et seulement si, elle admet une unique valeur d'adhérence.

Théorème

Soient E un espace vectoriel normé de dimension finie et F un espace vectoriel normé.

Alors toute application linéaire de E vers F est continue sur E .

Donc, si E est dimension finie, $\mathcal{L}(E, F) = \mathcal{L}_c(E, F)$.

Corollaire

Tout sous-espace vectoriel de dimension finie d'un espace vectoriel normé est fermé.

Soient E un $\mathbb{K}$ -espace vectoriel de dimension finie $p \in \mathbb{N}^*$ et $(e_1, \dots, e_p)$ une base de E . On appelle *monôme* définie sur E (à valeurs dans $\mathbb{K}$) toute application

$$\left\{ \begin{array}{ll} E & \longrightarrow \mathbb{K} \\ x = \sum_{k=1}^p x_k e_k & \longmapsto x_1^{\alpha_1} \cdots x_p^{\alpha_p} \end{array} \right.$$

où les $\alpha_i \in \mathbb{N}$.

On appelle *application polynomiale*¹⁹ définie sur E (à valeurs dans $\mathbb{K}$) toute combinaison linéaire de monômes définis sur E (à valeurs dans $\mathbb{K}$).

Théorème

Toute application polynomiale définie sur un espace vectoriel normé de dimension finie est continue.

Étendons enfin le théorème de continuité des applications linéaires en dimension finie aux applications multilinéaires.

Théorème

Soient $E_1, \dots, E_n$ des espaces vectoriels normés de dimension finie et F un espace vectoriel normé.

Alors toute application multilinéaire de $E_1 \times \cdots \times E_n$ vers F est continue.

Soit $n \in \mathbb{N}$ tel que $n \geq 2$. En identifiant $\mathcal{M}_n(\mathbb{K})$, de dimension finie, avec ses n vecteurs colonnes dans $\mathbb{K}^n \times \cdots \times \mathbb{K}^n$, l'application $\det : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \rightarrow \mathbb{K} \\ A & \mapsto \det(A) \end{cases}$ est n -linéaire donc continue via le théorème précédent.

4.20 Complétude

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé et (x_n) une suite d'éléments de E . On dit que (x_n) est une *suite de Cauchy* si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall (p, q) \in \mathbb{N}^2, p \geq N \text{ et } q \geq N \implies \|x_p - x_q\| < \varepsilon.$$

19. ou *fonction polynomiale*.

Théorème

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Alors

1. Toute suite de Cauchy de E est bornée.
2. Toute suite convergente de E est de Cauchy.
3. Toute suite de Cauchy de E , possédant une suite extraite convergente, converge.

Définissons à présent la notion d'espace *complet*.

Définition

Soient E un espace vectoriel normé et $A \subset E$.

On dit que A est *complet* si pour toute suite de Cauchy (x_n) d'éléments de A , (x_n) converge et $\lim_{n \rightarrow +\infty} x_n \in A$.

On appelle *espace de Banach* tout espace vectoriel normé complet.

On appelle *espace de Hilbert* tout espace préhilbertien complet.

Théorème

Soit E un espace vectoriel normé.

1. Si E est complet, toute partie fermée A de E est complète.
2. Toute partie complète de E est fermée.

Théorème

Soient E un espace de Banach et $A \subset E$.

Alors A est complète si, et seulement si, A est fermée

Théorème

Tout espace vectoriel normé de dimension finie est complet.

Le théorème qui suit donne un lien entre compacité et complétude.

Théorème

Soit E un espace vectoriel normé.

Toute partie compacte de E est complète.

La réciproque de ce théorème est fausse : $\mathbb{R}$ est un espace vectoriel normé complet mais non borné donc non compact.

Donnons enfin le *théorème du point fixe* après avoir donné la définition d'un tel point.

Définition

Soient E un ensemble et $f : E \longrightarrow E$.

On dit que $x \in E$ est un *point fixe* de f si $f(x) = x$.

Théorème du point fixe

Soient $(E, \|\cdot\|)$ un espace vectoriel normé, A une partie complète de E et $f : A \longrightarrow A$ contractante.

Alors f admet un unique point fixe dans A et pour tout $a \in A$, la suite (x_n) définie par $x_0 = a$ et pour tout $n \in \mathbb{N}$, $x_{n+1} = f(x_n)$, converge vers le point fixe de f .

5

Séries numériques

$\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

5.1 Convergence et divergence

Définition

Soit (u_n) une suite réelle ou complexe. On appelle *série* de terme général u_n , notée $\sum u_n$, la suite des sommes partielles (S_n) où pour tout $n \in \mathbb{N}$,

$$S_n = \sum_{k=0}^n u_k.$$

On dit que $\sum u_n$ *converge*¹ si (S_n) converge. On dit qu'elle *diverge* sinon.

Cette définition peut être étendue à une série de terme général u_n où la suite (u_n) est à valeurs dans un espace vectoriel normé de *dimension finie*². Rappelons que dans un tel espace toutes les normes sont équivalentes.

Si $(u_n)_{n \geq n_0}$ n'est définie qu'à partir de n_0 , la série numérique associée est alors la suite $(S_n)_{n \geq n_0}$ définie pour tout $n \geq n_0$ par

$$S_n = \sum_{k=n_0}^n u_k.$$

Toutes les définitions et théorèmes figurant dans ce chapitre ne sont énoncés, sauf cas particulier, que pour $n_0 = 0$ mais restent valables si $n_0 \neq 0$.

1. ou est convergente. La série de terme général u_n est parfois notée $\sum_{n \geq 0} u_n$ ou plus généralement $\sum_{n \geq n_0} u_n$ lorsque la suite $(u_n)_{n \geq n_0}$ n'est définie qu'à partir de n_0 . En revanche, la notation

$\sum_{n \in \mathbb{N}} u_n$ est réservée aux familles sommables (cf. note 21 page 124).

2. Elle peut être étendue à une série à valeurs dans un espace vectoriel normé quelconque mais alors l'absolue convergence n'entraîne plus nécessairement la convergence (cf. section 7.3 page 143, théorème page 117 et notes 15 et 16 page 117).

Étudier *la nature* d'une série numérique consiste à déterminer si elle converge ou diverge.

Remarquons que la convergence d'une série $\sum u_n$ ne dépend pas des premiers termes de la suite (u_n) . Plus précisément, si $n_0 \in \mathbb{N}$, $\sum_{n \geq n_0} u_n$ converge si, et seulement si, $\sum u_n$ converge.

Il suffit en effet de remarquer que $\sum_{k=0}^n u_k = \sum_{k=0}^{n_0-1} u_k + \sum_{k=n_0}^n u_k$.

5.2 Série géométrique

Définition

Soit $q \in \mathbb{K}$. On appelle *série géométrique* la série numérique $\sum q^n$.

Théorème

Soit $q \in \mathbb{K}$. Alors $\sum q^n$ converge ssi $|q| < 1$.

De plus si $|q| < 1$, $\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}$.

5.3 Somme et reste d'une série convergente

Définition

Soit $\sum u_n$ une série numérique convergente. On appelle *somme* de la série le scalaire

$$\sum_{n=0}^{+\infty} u_n = \lim_{n \rightarrow +\infty} S_n$$

et on appelle *reste* de la série la suite (R_n) définie pour tout $n \in \mathbb{N}$ par

$$R_n = \sum_{k=n+1}^{+\infty} u_k.$$

Lorsque une série numérique $\sum u_n$ est convergente, elle converge donc vers sa somme³ qui est le nombre réel ou complexe $\sum_{n=0}^{+\infty} u_n$.

3. Ne pas confondre $\sum u_n$ désignant la série de terme général u_n et $\sum_{n=0}^{+\infty} u_n$ désignant la somme de la série convergente $\sum u_n$.

De même, le reste d'une série ne peut être défini que pour une série convergente.

5.4 Somme de deux séries numériques et produit par un scalaire

Théorème

Soient $\sum u_n$, $\sum v_n$ deux séries numériques et $\lambda \in \mathbb{K}$. Alors⁴

1. $(\sum u_n \text{ converge et } \sum v_n \text{ converge}) \implies \sum (u_n + v_n) \text{ converge.}$
2. $\sum u_n \text{ converge} \implies \sum \lambda u_n \text{ converge.}$
3. $(\sum u_n \text{ converge et } \sum v_n \text{ diverge}) \implies \sum (u_n + v_n) \text{ diverge.}$

En particulier, si $\sum u_n$ et $\sum v_n$ convergent, on a⁵

$$\sum_{n=0}^{+\infty} \lambda u_n = \lambda \sum_{n=0}^{+\infty} u_n \quad \text{et} \quad \sum_{n=0}^{+\infty} (u_n + v_n) = \sum_{n=0}^{+\infty} u_n + \sum_{n=0}^{+\infty} v_n$$

5.5 Condition nécessaire de convergence et série harmonique

Théorème

Soit (u_n) une suite réelle ou complexe⁶. Alors

$$\sum u_n \text{ converge} \implies u_n \xrightarrow{n \rightarrow +\infty} 0.$$

Lorsque $u_n \not\xrightarrow{n \rightarrow +\infty} 0$, on dit que $\sum u_n$ *diverge grossièrement*.

Définition

On appelle *série harmonique* la série numérique $\sum \frac{1}{n}$.

4. Les deux premières assertions signifient que l'espace des séries convergentes est un $\mathbb{K}$ -ev.

5. c'est-à-dire l'application $\left\{ \begin{array}{ll} \{\text{séries numériques convergentes}\} & \longrightarrow \mathbb{K} \\ \sum u_n & \longmapsto \sum_{n=0}^{+\infty} u_n \end{array} \right.$ est linéaire.

6. Ce théorème s'étend au cas où (u_n) est à valeurs dans un espace vectoriel normé de dimension finie. Rappelons dans ce cadre que $u_n \xrightarrow{n \rightarrow +\infty} 0 \iff \|u_n\| \xrightarrow{n \rightarrow +\infty} 0$.

La réciproque du dernier théorème est fausse comme l'illustre le théorème suivant.

Théorème

La série harmonique $\sum \frac{1}{n}$ diverge.

Il est parfois délicat de montrer qu'une suite (u_n) est convergente. On peut alors tenter de montrer que la série de terme général⁷ $(u_n - u_{n-1})$ converge comme l'illustre le théorème suivant.

Théorème

Soit $(u_n)_{n \in \mathbb{N}^*}$ une suite réelle ou complexe⁸. Alors

$$\sum (u_n - u_{n-1}) \text{ converge} \iff (u_n) \text{ converge.}$$

5.6 Séries à termes positifs

Définition

On dit qu'une série numérique $\sum u_n$ est à termes positifs si (u_n) est une suite réelle vérifiant pour tout $n \in \mathbb{N}$, $u_n \geq 0$.

Si $\sum u_n$ est une série à termes positifs, la suite de ses sommes partielles (S_n) est croissante d'où le théorème suivant.

Théorème

Soient $\sum u_n$ une série à termes positifs et (S_n) la suite de ses sommes partielles. Alors : $\sum u_n$ converge $\iff (S_n)$ est majorée.

Théorème

Soient (u_n) et (v_n) deux suites réelles telles que pour tout $n \in \mathbb{N}$,

$$0 \leq u_n \leq v_n$$

$$\text{Alors } \begin{cases} \sum v_n \text{ converge} \implies \sum u_n \text{ converge;} \\ \text{et} \\ \sum u_n \text{ diverge} \implies \sum v_n \text{ diverge.} \end{cases}$$

7. Cette série est appelée *série télescopique* associée à la suite (u_n) .

8. Ce théorème s'étend au cas où (u_n) est à valeurs dans un espace vectoriel normé de dimension finie.

Le théorème suivant est très utile car permet de comparer la série étudiée à des séries de références.

Théorème

1. Soient (u_n) et (v_n) deux suites réelles positives telles que $u_n = O(v_n)$.
Alors $\sum v_n$ converge $\implies \sum u_n$ converge⁹.
2. Soient (u_n) et (v_n) deux suites réelles positives telles que $u_n = o(v_n)$.
Alors $\sum v_n$ converge $\implies \sum u_n$ converge¹⁰.
3. Soient (u_n) une suite réelle et (v_n) une suite réelle positive¹¹ que $u_n \sim v_n$.
Alors $\sum u_n$ et $\sum v_n$ sont de même nature¹².

Les séries de Riemann (et les séries géométriques) constituent les principales séries de référence.

Définition

On appelle *série de Riemann* toute série de la forme $\sum \frac{1}{n^\alpha}$ où $\alpha \in \mathbb{R}$.

Théorème

Soit $\alpha \in \mathbb{R}$. Alors $\sum \frac{1}{n^\alpha}$ converge $\iff \alpha > 1$.

Via les deux précédents théorèmes, on en déduit le corollaire suivant.

Corollaire (règle de Riemann)

Soit (u_n) une suite réelle positive.

S'il existe $\alpha > 1$ tel que $n^\alpha u_n \xrightarrow{n \rightarrow +\infty} 0$ alors $\sum u_n$ converge.

On peut généraliser la règle de Riemann de la façon suivante.

9. Le théorème page 117 permet une extension au cas où (u_n) est une suite réelle ou complexe.

10. cf. note précédente

11. La positivité de (v_n) implique celle de (u_n) à partir d'un certain rang.

12. Dire que deux séries sont de *même nature* signifie que si l'une diverge, l'autre diverge également et si l'une converge, l'autre aussi (mais pas nécessairement vers la même valeur).

Théorème (règle de Riemann généralisée)

Soit (u_n) une suite réelle positive.

On suppose qu'il existe $\alpha \in \mathbb{R}$ et $\ell \in \mathbb{R}_+ \cup \{+\infty\}$ tels que $n^\alpha u_n \xrightarrow{n \rightarrow +\infty} \ell$.

Alors $\begin{cases} (\ell \neq +\infty \text{ et } \alpha > 1) \implies \sum u_n \text{ converge;} \\ (\ell \neq 0 \text{ et } \alpha \leq 1) \implies \sum u_n \text{ diverge.} \end{cases}$

Les séries de Bertrand constituent également un réservoir de séries de références.

Définition

On appelle *série de Bertrand* toute série de la forme $\sum \frac{1}{n^\alpha (\ln(n))^\beta}$ où $(\alpha, \beta) \in \mathbb{R}^2$.

Théorème

Soit $(\alpha, \beta) \in \mathbb{R}^2$. Alors

$$\sum \frac{1}{n^\alpha (\ln(n))^\beta} \text{ converge} \iff \alpha > 1 \text{ ou } (\alpha = 1 \text{ et } \beta > 1).$$

Un autre critère de comparaison, très utile via son corollaire (règle de d'Alembert), est le critère de comparaison logarithmique.

Théorème (critère de comparaison logarithmique)

Soient (u_n) et (v_n) deux suites réelles strictement positives telles que pour tout $n \in \mathbb{N}$, $\frac{u_{n+1}}{u_n} \leq \frac{v_{n+1}}{v_n}$. Alors $\sum v_n \text{ converge} \implies \sum u_n \text{ converge}$.

Corollaire (règle de d'Alembert)

Soit (u_n) une suite réelle strictement positive¹³ telle que

$$\frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} \ell \text{ où } \ell \in \mathbb{R}_+ \cup \{+\infty\}.$$

Alors $\begin{cases} \ell < 1 \implies \sum u_n \text{ converge;} \\ \text{et} \\ \ell > 1 \implies \sum u_n \text{ diverge.} \end{cases}$

13. Ce corollaire reste valable dans le cas d'une suite complexe non nulle en remplaçant simplement $\frac{u_{n+1}}{u_n}$ par $\left| \frac{u_{n+1}}{u_n} \right|$.

Cette règle est particulièrement utile lorsque le terme général u_n est sous forme de produits (par exemple sous forme de factorielles).

Mais elle ne permet pas de conclure lorsque la limite ℓ est égale à 1. Il suffit pour s'en assurer de comparer les résultats avec les séries de Riemann $\sum \frac{1}{n}$ et $\sum \frac{1}{n^2}$: dans les deux cas, via la règle de d'Alembert, $\ell = 1$ mais la première série diverge tandis que la seconde converge.

Un autre critère, utile lorsque le terme général d'une série est sous forme de puissance n -ième, est la règle de Cauchy.

Théorème (règle de Cauchy)

Soit (u_n) une suite réelle positive¹⁴ telle que

$$\sqrt[n]{u_n} \xrightarrow{n \rightarrow +\infty} \ell \text{ où } \ell \in \mathbb{R}_+ \cup \{+\infty\}.$$

$$\text{Alors } \begin{cases} \ell < 1 \implies \sum u_n \text{ converge;} \\ \text{et} \\ \ell > 1 \implies \sum u_n \text{ diverge.} \end{cases}$$

La règle suivante, appelée *règle de Raabe-Duhamel* ou *règle de Duhamel* permet d'explorer le cas $\ell = 1$ de la règle de d'Alembert et de conclure dans certains cas.

Théorème (règle de Raabe-Duhamel)

Soient $\alpha \in \mathbb{R}$ et (u_n) une suite réelle strictement positive telle que

$$\frac{u_{n+1}}{u_n} = 1 - \frac{\alpha}{n} + o\left(\frac{1}{n}\right).$$

$$\text{Alors } \begin{cases} \alpha > 1 \implies \sum u_n \text{ converge;} \\ \text{et} \\ \alpha < 1 \implies \sum u_n \text{ diverge.} \end{cases}$$

5.7 Séries à termes quelconques

Dans cette section, on considère des séries à termes quelconques, c'est-à-dire de terme général de signe non constant. Parmi toutes les séries à termes quelconques, il existe deux grandes familles pour lesquelles des théorèmes existent : il s'agit des séries *absolument convergentes* et des séries *alternées*.

Commençons par les séries absolument convergentes.

14. Ce corollaire reste valable dans le cas d'une suite complexe en remplaçant simplement $\sqrt[n]{u_n}$ par $\sqrt[n]{|u_n|}$.

Définition

On dit qu'une série numérique $\sum u_n$ converge absolument ou est absolument convergente¹⁵ si la série $\sum |u_n|$ converge.

Un des intérêts de la convergence absolue est l'implication fournie par le théorème suivant.

Théorème

Soit $\sum u_n$ une série numérique absolument convergente. Alors $\sum u_n$ converge et on a l'inégalité triangulaire¹⁶

$$\left| \sum_{n=0}^{+\infty} u_n \right| \leq \sum_{n=0}^{+\infty} |u_n|.$$

Comme la série $\sum |u_n|$ est à termes positifs, on peut affiner une partie du théorème de comparaison page 114.

Théorème

Soient (u_n) une suite réelle ou complexe et (v_n) une suite

$$u_n = O(v_n) \quad (\text{respectivement } u_n = o(v_n)).$$

Alors $\sum u_n$ converge absolument donc converge.

Définition

Une série convergente mais non absolument convergente est dite *semi-convergente*.

Par exemple la *série harmonique alternée* $\sum \frac{(-1)^n}{n}$ est semi-convergente.

15. Dans le cas d'une série à termes complexes, $|u_n|$ désigne bien entendu le module de u_n . Si (u_n) est à valeurs dans un espace vectoriel normé $(E, \|\cdot\|)$, $\sum u_n$ est absolument convergente si $\sum \|u_n\|$ converge (cf. section 7.3 page 143).

16. Dans le cas d'une suite (u_n) à valeurs dans un espace vectoriel normé de dimension finie $(E, \|\cdot\|)$, l'absolue convergence de $\sum u_n$ entraîne la convergence de $\sum u_n$ et $\left\| \sum_{n=0}^{+\infty} u_n \right\| \leq \sum_{n=0}^{+\infty} \|u_n\|$ (cf. section 7.3 page 143).

Théorème

Pour tout $z \in \mathbb{C}$, la série $\sum \frac{z^n}{n!}$ est absolument convergente donc convergente.

L'application $\exp : \begin{cases} \mathbb{C} & \rightarrow \mathbb{C} \\ z & \mapsto \sum_{n=0}^{+\infty} \frac{z^n}{n!} \end{cases}$ est appelée *fonction exponentielle*.

Pour tout $z \in \mathbb{C}$, l'exponentielle de z , notée e^z est définie par $e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$.

Toutes les propriétés bien connues de l'exponentielle se déduisent de cette définition¹⁷.

Définition

Soit (u_n) une suite réelle.

On dit que (u_n) est *alternée* s'il existe une suite réelle (a_n) positive telle que pour tout $n \in \mathbb{N}$, $u_n = (-1)^n a_n$ (ou pour tout $n \in \mathbb{N}$, $u_n = (-1)^{n+1} a_n$).

On dit qu'une série numérique $\sum u_n$ est *alternée* si la suite (u_n) est alternée.

Si (u_n) ne s'annule pas, (u_n) est alternée lorsque le produit de deux termes consécutifs de la suite est négatif (c'est-à-dire que deux termes consécutifs de la suite sont de signes contraires).

Le critère suivant, appelé *critère spécial des séries alternées* ou *règle de Leibniz*, permet de conclure dans certains cas.

Théorème (critère spécial des séries alternées ou règle de Leibniz)

Soit (u_n) une suite réelle alternée.

Si $(|u_n|)$ est décroissante et converge vers 0 alors

1. $\sum u_n$ converge.
2. $\forall n \in \mathbb{N}$, $|R_n| \leq |u_{n+1}|$ où (R_n) est la suite des restes associée à $\sum u_n$.

Ce critère spécial des séries alternées est un cas particulier de la règle d'Abel suivante.

17. Cf. page 172.

Théorème (règle d'Abel)

Soient (u_n) et (v_n) deux suites réelles telles que

1. (u_n) est décroissante et converge vers 0.
2. La suite $\left(\sum_{k=0}^n v_k\right)$ est bornée.

Alors $\sum u_n v_n$ converge.

5.8 Somme des relations de comparaison

Les deux théorèmes suivants permettent, respectivement, de donner une estimation du reste d'une série convergente et de la somme partielle d'une série divergente.

Théorème

Soient (u_n) et (v_n) deux suites réelles telles que (v_n) est positive et $\sum v_n$ converge.

1. Si $u_n = o(v_n)$ alors $\sum u_n$ converge et $\sum_{k=n+1}^{+\infty} u_k = o\left(\sum_{k=n+1}^{+\infty} v_k\right)$.
2. Si $u_n = O(v_n)$ alors $\sum u_n$ converge et $\sum_{k=n+1}^{+\infty} u_k = O\left(\sum_{k=n+1}^{+\infty} v_k\right)$.
3. Si $u_n \sim v_n$ alors $\sum u_n$ converge et $\sum_{k=n+1}^{+\infty} u_k \sim \sum_{k=n+1}^{+\infty} v_k$.

Théorème

Soient (u_n) et (v_n) deux suites réelles telles que (v_n) est positive et $\sum v_n$ diverge.

1. Si $u_n = o(v_n)$ alors $\sum_{k=0}^n u_k = o\left(\sum_{k=0}^n v_k\right)$.
2. Si $u_n = O(v_n)$ alors $\sum_{k=0}^n u_k = O\left(\sum_{k=0}^n v_k\right)$.
3. Si $u_n \sim v_n$ alors $\sum_{k=0}^n u_k \sim \sum_{k=0}^n v_k$.

5.9 Produit de deux séries numériques

Définition

Soient $\sum u_n$ et $\sum v_n$ deux séries numériques.

On appelle *produit de Cauchy* de $\sum u_n$ et $\sum v_n$ la série $\sum w_n$ où pour tout $n \in \mathbb{N}$,

$$w_n = \sum_{k=0}^n u_k v_{n-k} = \sum_{p+q=n} u_p v_q.$$

Théorème

Soient $\sum u_n$ et $\sum v_n$ deux séries numériques absolument convergentes.

Alors le produit de Cauchy $\sum w_n$ de ces deux séries converge absolument et on a

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{n=0}^{+\infty} u_n \right) \left(\sum_{n=0}^{+\infty} v_n \right)$$

où pour tout $n \in \mathbb{N}$, $w_n = \sum_{p+q=n} u_p v_q$.

5.10 Comparaison série-intégrale

Théorème

Soit $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ continue par morceaux et décroissante.

Alors la série de terme général $\int_{n-1}^n f(t) dt - f(n)$ converge.

5.11 Sommes de séries convergentes classiques

Théorème

$$\sum_{n=0}^{+\infty} \frac{1}{n!} = e \text{ et } \sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = -\ln(2) \text{ (somme de la série harmonique alternée).}$$

$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}, \quad \sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \text{ et } \sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

5.12 Familles sommables

Dans cette section, introduite en vue de l'étude des probabilités, I désigne un ensemble au plus dénombrable¹⁸.

Lorsqu'une série numérique $\sum u_n$ est convergente, sa somme est définie par

$$\sum_{n=0}^{+\infty} u_n = \lim_{n \rightarrow +\infty} \sum_{k=0}^n u_k.$$

Dans le calcul de la somme ci-dessus, on additionne les termes dans un ordre précis. Or, dans le cas d'une série semi-convergente, comme par exemple $\sum \frac{(-1)^n}{n+1}$, on peut, via une permutation des termes, obtenir une somme différente de la valeur classique $\ln(2)$ obtenue dans le théorème précédent.

Cette section a notamment pour but d'introduire une définition de la somme qui exclut cette pathologie en donnant le même résultat quel que soit l'ordre de sommation choisi.

Elle vise plus généralement à définir, dans le cas d'une indexation plus compliquée qui ne possède pas d'ordre naturel, la notion de somme d'une infinité de nombres réels ou complexes ne tenant pas compte de l'ordre de sommation, ni de la façon de les indexer.

Définition

Soit $(u_i)_{i \in I}$ une famille de réels positifs. On dit que $(u_i)_{i \in I}$ est *sommable* si l'ensemble $\left\{ \sum_{i \in J} u_i; J \text{ fini} \subset I \right\}$ est majoré¹⁹.

Dans ce cas, la *somme* de la famille $(u_i)_{i \in I}$, notée $\sum_{i \in I} u_i$, est la borne supérieure de l'ensemble précédent.

Lorsque la famille $(u_i)_{i \in I}$ n'est pas sommable, on écrit $\sum_{i \in I} u_i = +\infty$.

Par définition, la somme d'une famille sommable de réels positifs ne dépend pas de l'ordre des éléments de cette famille : plus précisément, si σ est une permutation de I , $(u_{\sigma(i)})_{i \in I}$ est sommable si, et seulement si, $(u_i)_{i \in I}$ l'est.

Si tel est le cas, on a de plus les deux sommes qui coïncident : $\sum_{i \in I} u_i = \sum_{i \in I} u_{\sigma(i)}$.

Toujours dans le cas de réels positifs, le théorème suivant montre que les notions de famille sommable et de série convergente coïncident lorsque $I = \mathbb{N}$.

18. Via la définition qui suit, on peut montrer que si $(u_i)_{i \in I}$ est sommable, alors I est nécessairement au plus dénombrable d'où cette restriction prise dès le début de cette section.

19. c'est-à-dire s'il existe $K \in \mathbb{R}_+$ tel que pour tout J fini $\subset I$, $\sum_{i \in J} u_i \leq K$.

Théorème

Soit $(u_n)_{n \in \mathbb{N}}$ une suite²⁰ réelle positive.

Alors $(u_n)_{n \in \mathbb{N}}$ est sommable $\iff \sum u_n$ converge.

De plus, dans ce cas, on a $\sum_{n \in \mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n$.

Les deux théorèmes suivants énoncent les premières propriétés élémentaires associées aux familles sommables de réels positifs.

Théorème

Soient $(u_i)_{i \in I}$ et $(v_i)_{i \in I}$ deux familles de réels positifs telles que, pour tout $i \in I$, $u_i \leq v_i$.

Si $(v_i)_{i \in I}$ est sommable alors $(u_i)_{i \in I}$ est sommable et $\sum_{i \in I} u_i \leq \sum_{i \in I} v_i$.

Si $(u_i)_{i \in I}$ n'est pas sommable alors $(v_i)_{i \in I}$ n'est pas sommable.

Théorème

Soient $(u_i)_{i \in I}$ et $(v_i)_{i \in I}$ deux familles sommables de réels positifs et $\lambda \in \mathbb{R}$.

Alors la famille $(\lambda u_i + v_i)_{i \in I}$ est sommable et $\sum_{i \in I} (\lambda u_i + v_i) = \lambda \sum_{i \in I} u_i + \sum_{i \in I} v_i$.

Étendons le concept de famille sommable à une famille de réels ou de complexes.

Pour toute famille (u_i) de réels, on introduit ses parties positive et négative en notant, pour tout $i \in I$, $u_i^+ = \max(u_i, 0)$ et $u_i^- = \max(-u_i, 0)$.

Pour tout $i \in I$, u_i^+ et u_i^- sont donc des réels positifs tels que

$$u_i^+ + u_i^- = |u_i| \quad \text{et} \quad u_i^+ - u_i^- = u_i$$

Définition

Soit $(u_i)_{i \in I}$ une famille de réels ou complexes. On dit que $(u_i)_{i \in I}$ est *sommable* si la famille de réels positifs $(|u_i|)$ l'est.

20. Dans le cas d'une indexation par $\mathbb{N}$, on utilise plutôt le mot *suite* que *famille*.

Théorème

Soit $(u_i)_{i \in I}$ une famille de réels. Alors $(u_i)_{i \in I}$ est sommable si, et seulement si, $(u_i^+)_{i \in I}$ et $(u_i^-)_{i \in I}$ le sont.

Dans ce cas, on appelle *somme* de la famille $(u_i)_{i \in I}$ le réel

$$\sum_{i \in I} u_i = \sum_{i \in I} u_i^+ - \sum_{i \in I} u_i^-.$$

Théorème

Soit $(u_i)_{i \in I}$ une famille de complexes. Alors $(u_i)_{i \in I}$ est sommable si, et seulement si, les deux suites de réels $(\operatorname{Re}(u_i))_{i \in I}$ et $(\operatorname{Im}(u_i))_{i \in I}$ le sont.

Dans ce cas, on appelle *somme* de la famille $(u_i)_{i \in I}$ le complexe

$$\sum_{i \in I} u_i = \sum_{i \in I} \operatorname{Re}(u_i) + i \sum_{i \in I} \operatorname{Im}(u_i).$$

Le théorème de linéarité page 122 reste valable pour une famille de réels ou complexes quelconque.

D'autre part, la remarque page 121 reste valable dans le cas général d'une famille de réels ou complexes comme l'explique le théorème suivant.

Théorème

Soient $(u_i)_{i \in I}$ une famille de réels ou complexes et σ une permutation de I .

$(u_i)_{i \in I}$ est sommable si, et seulement si, $(u_{\sigma(i)})_{i \in I}$ l'est.

De plus, si tel est le cas, les deux sommes coïncident : $\sum_{i \in I} u_i = \sum_{i \in I} u_{\sigma(i)}.$

Le théorème suivant montre que les notions de famille sommable et de série absolument convergente coïncident lorsque $I = \mathbb{N}$.

Théorème

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle ou complexe.

Alors (u_n) est sommable si, et seulement si, la série $\sum u_n$ est absolument convergente.

De plus, dans ce cas, on a $\sum_{n \in \mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n.$

Ne pas confondre les deux notations $\sum_{n \in \mathbb{N}} u_n$ et $\sum_{n=0}^{+\infty} u_n$: dans le premier cas, aucun ordre de sommation n'est suggéré contrairement au deuxième²¹.

Par exemple $\sum_{n=0}^{+\infty} \frac{(-1)^n}{n+1}$ a un sens mais $\sum_{n \in \mathbb{N}} \frac{(-1)^n}{n+1}$ n'a pas de sens car cette dernière notation n'indique pas dans quel ordre ont été sommés les termes : bien que la série $\sum \frac{(-1)^n}{n+1}$ soit convergente, la famille $\left(\frac{(-1)^n}{n+1}\right)_{n \in \mathbb{N}}$ n'est pas sommable.

Le théorème suivant explique que la somme d'une série numérique *absolument convergente*²² n'est pas modifiée par permutation de ses termes.

Théorème (convergence commutative)

Soient $\sum u_n$ une série numérique *absolument convergente* et σ une permutation de $\mathbb{N}$.

Alors la série numérique $\sum u_{\sigma(n)}$ converge²³ et $\sum_{n=0}^{+\infty} u_{\sigma(n)} = \sum_{n=0}^{+\infty} u_n$.

Les deux résultats qui suivent permettent, sous l'hypothèse de sommabilité d'une famille, de calculer sa somme, en réorganisant à sa guise les termes de celle-ci dans n'importe quel ordre.

Théorème (sommation par paquets d'une famille de réels positifs)

Soient $(u_i)_{i \in I}$ une famille de réels positifs et $(I_n)_{n \in \mathbb{N}}$ une partition de I .

Alors $(u_i)_{i \in I}$ est sommable $\iff \begin{cases} \text{pour tout } n \in \mathbb{N}, \text{ la sous-famille } (u_i)_{i \in I_n} \\ \text{est sommable ;} \\ \text{la série numérique } \sum \left(\sum_{i \in I_n} u_i \right) \text{ converge.} \end{cases}$

Dans ce cas, $\sum_{i \in I} u_i = \sum_{n=0}^{+\infty} \left(\sum_{i \in I_n} u_i \right)$.

Voici à présent le théorème de sommation par paquets pour une famille quelconque de réels ou complexes, corollaire du précédent.

21. Ces deux notations ne doivent pas non plus être confondues avec $\sum u_n$ qui désigne la série de terme général u_n .

22. En revanche, on peut modifier la somme d'une série *semi-convergente* via une permutation de ses termes.

23. On dit qu'une série numérique $\sum u_n$ est *commutativement convergente* si pour toute permutation σ de $\mathbb{N}$, $\sum u_{\sigma(n)}$ converge.

Théorème (sommation par paquets d'une famille de réels ou de complexes)

Soient D un ensemble au plus dénombrable, $(u_i)_{i \in I}$ une famille de réels ou complexes et $(I_d)_{d \in D}$ une partition de I .

Si la famille $(u_i)_{i \in I}$ est sommable alors les sous-familles $(u_i)_{i \in I_d}$ sont sommables et la série numérique de leurs sommes, $\sum \left(\sum_{i \in I_d} u_i \right)$, converge.

De plus, on a alors $\sum_{i \in I} u_i = \sum_{d \in D} \left(\sum_{i \in I_d} u_i \right)$.

En appliquant le théorème précédent aux suites doubles $(u_{n,p})_{(n,p) \in \mathbb{N}^2}$, on obtient le corollaire suivant qui permet d'effectuer l'échange de deux sommes infinies.

Théorème de Fubini

Soit $(u_{n,p})_{(n,p) \in \mathbb{N}^2}$ une suite double de réels ou de complexes telle que²⁴

$$\left\{ \begin{array}{l} \text{pour tout } p \in \mathbb{N}, \text{ la série numérique } \sum_{n \geq 0} |u_{n,p}| \text{ converge;} \\ \text{la série numérique } \sum_{p \geq 0} \left(\sum_{n=0}^{+\infty} |u_{n,p}| \right) \text{ converge.} \end{array} \right.$$

Alors

$$\sum_{n=0}^{+\infty} \left(\sum_{p=0}^{+\infty} u_{n,p} \right) = \sum_{p=0}^{+\infty} \left(\sum_{n=0}^{+\infty} u_{n,p} \right)$$

avec absolue convergence des séries ci-dessus.

24. Les deux conditions suivantes sont équivalentes à la sommabilité de $(u_{n,p})_{(n,p) \in \mathbb{N}^2}$.

6

Intégrales généralisées

$\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Tout intervalle de ce chapitre est supposé non vide.

Il s'agit dans ce chapitre de donner un sens à des intégrales de la forme $\int_a^b f(t) dt$ où f est continue par morceaux seulement sur $]a, b]$, $[a, b[$ ou $]a, b[$, avec en particulier des bornes éventuellement infinies : par exemple $\int_1^{+\infty} \frac{dt}{t^2}$, $\int_0^1 \frac{t-1}{\ln(t)} dt$ où dans cette deuxième intégrale, la fonction $t \mapsto \frac{t-1}{\ln(t)}$ n'est ni définie en 0, ni en 1.

6.1 Convergence et divergence

Énonçons successivement les définitions d'une intégrale convergente pour les trois cas $[a, b[$, $]a, b]$ et $]a, b[$.

Définition

Soit $f : [a, b] \rightarrow \mathbb{K}$ continue par morceaux sur $[a, b[$ où $-\infty < a < b \leq +\infty$.

On dit que l'intégrale de f sur $[a, b[$, notée $\int_{[a, b[} f(t) dt$, *converge*² si la

fonction $F : \begin{cases} [a, b[& \longrightarrow \mathbb{R} \\ x & \longmapsto \int_a^x f(t) dt \end{cases}$ admet une limite finie en b .

Dans ce cas, on note $\int_a^b f(t) dt = \lim_{x \rightarrow b} \int_a^x f(t) dt$.

Sinon, on dit que l'intégrale de f sur $[a, b[$ *diverge*.

Si f est *continue* sur $[a, b[$, en notant F une primitive³ de f sur $[a, b[$, on a l'équivalence : $\int_a^b f(t) dt$ converge $\iff F$ admet une limite finie en b .

De plus, dans ce cas, on a alors : $\int_a^b f(t) dt = \lim_{x \rightarrow b} F(x) - F(a) = [F(x)]_a^b$.

1. Cette notation signifie que $a \in \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$.

2. ou est *convergente*. Parfois, on écrira par abus $\int_a^b f(t) dt$ *converge*.

3. Attention, une fonction *continue par morceaux* n'admet pas nécessairement de *primitive* : par exemple la fonction définie sur $[0, 2]$ par $f(1) = 1$ et pour tout $x \in [0, 2] \setminus \{1\}$ par $f(x) = 0$, est continue par morceaux sur $[0, 2]$ mais n'admet pas de primitive sur $[0, 2]$.

Définition

Soit $f :]a, b] \rightarrow \mathbb{K}$ continue par morceaux sur $]a, b]$ où ⁴ $-\infty \leq a < b < +\infty$.

On dit que l'intégrale de f sur $]a, b]$, notée $\int_{]a, b]} f(t) dt$, *converge* ⁵ si la

fonction $F : \begin{cases}]a, b] & \longrightarrow \mathbb{R} \\ x & \longmapsto \int_x^b f(t) dt \end{cases}$ admet une limite finie en a .

Dans ce cas, on note $\int_a^b f(t) dt = \lim_{x \rightarrow a} \int_x^b f(t) dt$.

Sinon, on dit que l'intégrale de f sur $]a, b]$ *diverge*.

Définition

Soit $f :]a, b[\rightarrow \mathbb{K}$ continue par morceaux sur $]a, b[$ où ⁶ $-\infty \leq a < b \leq +\infty$.

On dit que l'intégrale de f sur $]a, b[$, notée $\int_{]a, b[} f(t) dt$, *converge* ⁷ si, pour tout $c \in]a, b[$, $\int_{]a, c]} f(t) dt$ et $\int_{[c, b[} f(t) dt$ convergent ⁸.

Dans ce cas, pour tout $c \in]a, b[$, on a $\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$.

Sinon, on dit que l'intégrale de f sur $]a, b[$ *diverge*.

Si, pour un $c \in]a, b[$, l'intégrale de f sur $]a, c]$ diverge, on peut directement en conclure, via la définition précédente, que l'intégrale de f sur $]a, b[$ diverge ⁹.

On prendra garde de ne pas confondre $\int_{]-\infty, +\infty[} f(t) dt$ avec $\lim_{x \rightarrow +\infty} \int_{-x}^x f(t) dt$.

En effet, $\int_{]-\infty, +\infty[} t dt$ diverge, car $\int_{[0, +\infty[} t dt$ diverge, et pourtant $\lim_{x \rightarrow +\infty} \int_{-x}^x t dt = 0$.

Dans toute la suite du chapitre, les définitions et résultats principaux seront énoncés dans le cas d'une fonction continue par morceaux sur $[a, b[$, les cas $]a, b]$ et $]a, b[$ étant analogues.

4. Cette notation signifie que $a \in \mathbb{R} \cup \{-\infty\}$ et $b \in \mathbb{R}$.

5. ou est *convergente*. Parfois, on écrira par abus $\int_a^b f(t) dt$ *converge*.

6. Cette notation signifie que $a \in \mathbb{R} \cup \{-\infty\}$ et $b \in \mathbb{R} \cup \{+\infty\}$.

7. ou est *convergente*. Parfois, on écrira par abus $\int_a^b f(t) dt$ *converge*.

8. On montre aisément que le choix de c n'influe ni sur la convergence, ni sur l'éventuelle valeur de l'intégrale.

9. idem si pour un $c \in]a, b[$, l'intégrale de f sur $[c, b]$ diverge.

6.2 Linéarité, conjugaison, croissance et stricte positivité

Théorème

Soient f et $g : [a, b[\rightarrow \mathbb{K}$ continues par morceaux sur $[a, b[$ où $-\infty < a < b \leq +\infty$ telles que $\int_{[a, b[} f(t) dt$ et $\int_{[a, b[} g(t) dt$ convergent.

Alors pour tout $\lambda \in \mathbb{K}$, $\int_{[a, b[} (f + \lambda g)(t) dt$ converge et on a de plus

$$\int_a^b (f + \lambda g)(t) dt = \int_a^b f(t) dt + \lambda \int_a^b g(t) dt.$$

En particulier si l'intégrale de f sur $[a, b[$ converge et l'intégrale de g sur $[a, b[$ diverge, on peut en conclure que l'intégrale de $f + g$ sur $[a, b[$ diverge.

Énonçons ensuite deux théorèmes, de conjugaison (pour les fonctions à valeurs complexes) et de croissance (pour celles à valeurs réelles).

Théorème

Soit $f : [a, b[\rightarrow \mathbb{C}$ continue par morceaux sur $[a, b[$ où $-\infty < a < b \leq +\infty$.

Si l'intégrale de f sur $[a, b[$ converge, l'intégrale de $\overline{f}$ sur $[a, b[$ converge et

$$\int_a^b \overline{f(t)} dt = \overline{\int_a^b f(t) dt}.$$

Théorème

Soient f et $g : [a, b[\rightarrow \mathbb{R}$ continue par morceaux sur $[a, b[$ dont les intégrales sur $[a, b[$ convergent où $-\infty < a < b \leq +\infty$. Alors

$$f \leq g \implies \int_a^b f(t) dt \leq \int_a^b g(t) dt.$$

En particulier si f est positive sur $[a, b[$, son intégrale sur $[a, b[$ est positive.

Théorème

Soit $f : [a, b[\rightarrow \mathbb{R}$ continue¹⁰ et positive sur $[a, b[$ dont l'intégrale sur $[a, b[$ converge.

Si $\int_a^b f(t) dt = 0$ alors f est la fonction nulle.

10. Ce théorème devient faux avec une hypothèse de continuité par morceaux.

6.3 Intégrales de Riemann

Introduisons une famille de référence : les *intégrales de Riemann*.

Théorème (intégrales de Riemann)

Soit $\alpha \in \mathbb{R}$. Alors

- $\int_{[0,1[} \frac{dt}{t^\alpha}$ converge si, et seulement si, $\alpha < 1$.
- $\int_{[1,+\infty[} \frac{dt}{t^\alpha}$ converge si, et seulement si, $\alpha > 1$.

On remarquera en particulier que pour tout $\alpha \in \mathbb{R}$, $\int_{]0,+\infty[} \frac{dt}{t^\alpha}$ diverge. Énonçons également un résultat sur une autre famille de référence, utile en probabilités.

Théorème

Soit $a \in \mathbb{R}$. Alors $\int_{[0,+\infty[} e^{-at} dt$ converge si, et seulement si, $a > 0$.

De plus, si tel est le cas, on a $\int_0^{+\infty} e^{-at} dt = \frac{1}{a}$.

6.4 Cas des fonctions positives

On étudie dans cette section le cas des fonctions à valeurs réelles, continues par morceaux et positives sur $[a, b[$.

Théorème

Soit $f : [a, b[\rightarrow \mathbb{R}$ continue par morceaux et positive sur $[a, b[$ où $-\infty < a < b \leq +\infty$.

Soit $F : \begin{cases} [a, b[& \rightarrow \mathbb{R} \\ x & \mapsto \int_a^x f(t) dt \end{cases}$.

Alors $\int_{[a,b[} f(t) dt$ converge si, et seulement si, F est majorée.

6.5 Critères de comparaison

Théorème

Soient f et $g : [a, b[\rightarrow \mathbb{R}$ continues par morceaux et positives sur $[a, b[$ où $-\infty < a < b \leq +\infty$.

1. Supposons que pour tout $t \in [a, b[$, $0 \leq f(t) \leq g(t)$. Alors

$$\int_{[a, b[} g(t) dt \text{ converge} \implies \int_{[a, b[} f(t) dt \text{ converge.}$$

2. Supposons qu'au voisinage de b , $f = o(g)$. Alors

$$\int_{[a, b[} g(t) dt \text{ converge} \implies \int_{[a, b[} f(t) dt \text{ converge.}$$

3. Supposons qu'au voisinage de b , $f = O(g)$. Alors

$$\int_{[a, b[} g(t) dt \text{ converge} \implies \int_{[a, b[} f(t) dt \text{ converge.}$$

4. Supposons que $f \sim_b g$. Alors $\int_{[a, b[} f(t) dt$ et $\int_{[a, b[} g(t) dt$ sont de même nature¹¹.

Dans le 4^e résultat de ce théorème, il suffit que l'une des deux fonctions soit positive¹² car alors l'autre le sera nécessairement au voisinage de b .

Par exemple, l'intégrale de Gauss, $\int_{[0, +\infty[} e^{-t^2} dt$, est convergente¹³ car la fonction $t \mapsto e^{-t^2}$ est continue sur $\mathbb{R}_+$ et, au voisinage de $+\infty$, $e^{-t^2} = o\left(\frac{1}{t^2}\right)$ avec $\int_{[1, +\infty[} \frac{dt}{t^2}$ convergente.

6.6 Comparaison série-intégrale

Le théorème suivant permet, sous certaines hypothèses, de comparer la nature d'une série à celle d'une intégrale généralisée.

Théorème (comparaison série-intégrale¹⁴)

Soient $a \in \mathbb{R}_+$ et $f : [a, +\infty[\rightarrow \mathbb{R}$ continue, positive et décroissante sur $[a, +\infty[$. Alors $\sum f(n)$ et $\int_{[a, +\infty[} f(t) dt$ sont de même nature.

11. c'est-à-dire si l'une converge, l'autre également, et par contraposée, si l'une diverge, l'autre aussi.

12. ou plus généralement de signe constant : en effet, si g est négative, $-g$ est positive et la convergence de l'intégrale de g sur $[a, b[$ équivaut à celle de $-g$ sur $[a, b[$.

13. cf. également théorème page 135.

14. Ce théorème est également appelé *critère intégral de Cauchy*.

Par exemple, pour $\alpha \in \mathbb{R}$, on retrouve, via ce théorème un résultat déjà connu sur les séries de Riemann : $\sum \frac{1}{n^\alpha}$ est de même nature que $\int_{[1, +\infty[} \frac{dt}{t^\alpha}$ qui converge si, et seulement si, $\alpha > 1$.

6.7 Intégration par parties et par changement de variable

Étendons les résultats déjà connus pour les intégrales classiques sur un compact aux intégrales généralisées.

Théorème

Soient f et $g : [a, b[\rightarrow \mathbb{K}$ de classe $\mathcal{C}^1$ sur $[a, b[$ avec $-\infty < a < b \leq +\infty$ telles que fg admette une limite finie en b . Alors $\int_{[a, b[} f'(t)g(t) dt$ et $\int_{[a, b[} f(t)g'(t) dt$ sont de même nature.

Si elles convergent, on a de plus

$$\int_a^b f'(t)g(t) dt = \lim_{t \rightarrow b} (f(t)g(t)) - f(a)g(a) - \int_a^b f(t)g'(t) dt.$$

Théorème

Soient $f : [a, b[\rightarrow \mathbb{K}$ continue sur $[a, b[$ où $-\infty < a < b \leq +\infty$ et

$\varphi : [\alpha, \beta[\rightarrow [a, b[$ bijective, de classe $\mathcal{C}^1$ et strictement croissante sur $[\alpha, \beta[$ où $-\infty < \alpha < \beta \leq +\infty$.

Alors $\int_{[\alpha, \beta[} f(\varphi(u))\varphi'(u) du$ et $\int_{[a, b[} f(t) dt$ sont de même nature.

De plus, en cas de convergence, on a $\int_\alpha^\beta f(\varphi(u))\varphi'(u) du = \int_a^b f(t) dt$.

En particulier, on peut énoncer le théorème suivant, corollaire de celui relatif aux intégrales de Riemann.

Théorème

Soient a , b et α trois réels tels que $a < b$. Alors

- $\int_{[a, b[} \frac{dt}{(t-a)^\alpha}$ converge si, et seulement si, $\alpha < 1$.
- $\int_{[a, b[} \frac{dt}{(b-t)^\alpha}$ converge si, et seulement si, $\alpha < 1$.

6.8 Comportement asymptotique

Rappelons que pour une série numérique $\sum u_n$, on a une condition nécessaire de convergence : $\sum u_n$ converge $\implies u_n \xrightarrow{n \rightarrow +\infty} 0$.

Ce théorème devient faux pour les intégrales généralisées¹⁵ : il existe des fonctions f continues sur $[a, +\infty[$ telles que $\int_{[a, +\infty[} f(t) dt$ converge et $f(t) \not\xrightarrow[t \rightarrow +\infty]{} 0$ ¹⁶.

Par exemple, les *intégrales de Fresnel*¹⁷, $\int_{[0, +\infty[} \cos(t^2) dt$ et $\int_{[0, +\infty[} \sin(t^2) dt$ convergent sans que les fonctions $t \mapsto \cos(t^2)$ et $t \mapsto \sin(t^2)$ n'admettent de limite en $+\infty$.

On a en revanche le théorème suivant.

Théorème

Soient $a \in \mathbb{R}$ et f continue par morceaux sur $[a, +\infty[$ telle que $\int_{[a, +\infty[} f(t) dt$ converge et $\lim_{+\infty} f$ existe. Alors $\lim_{+\infty} f = 0$.

6.9 Fonctions intégrables

Définition

Soit $f : [a, b[\rightarrow \mathbb{K}$ continue par morceaux sur $[a, b[$ où $-\infty < a < b \leq +\infty$.

On dit que f est *intégrable* sur $[a, b[$ si l'intégrale de $|f|$ sur $[a, b[$ converge^{18, 19}.

15. On peut montrer qu'il faut supposer f *uniformément continue* sur $[a, +\infty[$ afin que le théorème soit vérifié pour les intégrales généralisées.

16. et même f non nécessairement bornée au voisinage de $+\infty$. Par exemple $t \mapsto t \cos(t^3)$ n'est pas bornée au voisinage de $+\infty$ mais $\int_{[1, +\infty[} t \cos(t^3) dt$ converge.

17. cf. également théorème page 135.

18. On dit aussi que l'intégrale de f sur $[a, b[$ *converge absolument* ou est *absolument convergente*. On emploiera indifféremment les expressions « f est *intégrable* sur $[a, b[$ » et l'intégrale « $\int_{[a, b[} f(t) dt$ *converge absolument* ».

19. Rappelons que cette définition est valable sur $]a, b]$ ou $]a, b[$ et donc sur tout intervalle I de $\mathbb{R}$ auquel cas la définition devient : f , continue par morceaux sur I , est *intégrable* sur I si l'intégrale de $|f|$ sur I , toujours notée $\int_I |f(t)| dt$, est convergente. Remarquons au passage que f est intégrable sur I si, et seulement si, $|f|$ est intégrable sur I . Il ne faut pas confondre cette notion d'intégrabilité sur I (au sens de Lebesgue) avec la notion d'intégrabilité (au sens de Riemann) d'une fonction f sur un segment $[a, b]$ pour laquelle on peut avoir $|f|$ intégrable (au sens de Riemann) sur I sans que f le soit (cf. note 15 page 81).

Théorème

Soit $f : [a, b[\rightarrow \mathbb{K}$ continue par morceaux sur $[a, b[$ où $-\infty < a < b \leq +\infty$.

Si f est intégrable sur $[a, b[$, alors son intégrale sur $[a, b[$ converge²⁰ et

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt.$$

On dit que l'intégrale d'une fonction $f : I \rightarrow \mathbb{K}$ continue par morceaux sur un intervalle I est *semi-convergente* si elle est convergente mais f est non intégrable sur I .

Par exemple, on peut montrer que l'intégrale de Dirichlet²¹, $\int_{]0, +\infty[} \frac{\sin(t)}{t} dt$ est semi-convergente²².

6.10 Intégrabilité par comparaison et domination

Théorème

Soient $f : [a, b[\rightarrow \mathbb{K}$ et $\varphi : [a, b[\rightarrow \mathbb{R}$ continues par morceaux sur $[a, b[$ avec φ positive et intégrable²³ sur $[a, b[$, où $-\infty < a < b \leq +\infty$, telles que

$$\forall t \in [a, b[, |f(t)| \leq \varphi(t).$$

Alors f est intégrable sur $[a, b[$.

Théorème

Soient f et $g : [a, b[\rightarrow \mathbb{K}$ continues par morceaux sur $[a, b[$ où $-\infty < a < b \leq +\infty$.

Si, au voisinage de b , $f = o(g)$ et g intégrable sur $[a, b[$ alors f est intégrable sur $[a, b[$.

Si, au voisinage de b , $f = O(g)$ et g intégrable sur $[a, b[$, alors f est intégrable sur $[a, b[$.

Si $f \sim_b g$ alors : f intégrable sur $[a, b[\iff g$ intégrable sur $[a, b[$.

20. ce qui peut s'énoncer : si $\int_{[a, b[} f(t) dt$ converge absolument, alors $\int_{[a, b[} f(t) dt$ converge.

21. Le patronyme de ce mathématicien est en réalité Lejeune-Dirichlet, souvent raccourci en Dirichlet.

22. cf. également théorème page 135.

23. ce qui revient à dire ici que l'intégrale de φ sur $[a, b[$ converge puisque φ est positive.

6.11 Intégration des relations de comparaison

Les théorèmes qui suivent sont l'analogie des sommations de relations de comparaison pour les séries numériques. Énonçons d'abord le cas des fonctions intégrables.

Théorème

Soient $f : [a, b[\rightarrow \mathbb{K}$ et $g : [a, b[\rightarrow \mathbb{R}$ continues par morceaux telles que g est positive et intégrable sur $[a, b[$.

Si, au voisinage de b , $f = o(g)$, alors, quand $x \rightarrow b$,

$$\int_x^b f(t) dt = o\left(\int_x^b g(t) dt\right).$$

Si, au voisinage de b , $f = O(g)$, alors, quand $x \rightarrow b$,

$$\int_x^b f(t) dt = O\left(\int_x^b g(t) dt\right).$$

Si $f \sim_b g$ alors f est intégrable sur $[a, b[$ et

$$\int_x^b f(t) dt \sim_b \int_x^b g(t) dt.$$

Énonçons à présent le cas des fonctions non intégrables.

Théorème

Soient $f : [a, b[\rightarrow \mathbb{K}$ et $g : [a, b[\rightarrow \mathbb{R}$ continues par morceaux telles que g est positive et non intégrable sur $[a, b[$.

Si, au voisinage de b , $f = o(g)$, alors, quand $x \rightarrow b$,

$$\int_a^x f(t) dt = o\left(\int_a^x g(t) dt\right).$$

Si, au voisinage de b , $f = O(g)$, alors, quand $x \rightarrow b$,

$$\int_a^x f(t) dt = O\left(\int_a^x g(t) dt\right).$$

Si $f \sim_b g$ alors f n'est pas intégrable sur $[a, b[$ et

$$\int_a^x f(t) dt \sim_b \int_a^x g(t) dt.$$

6.12 Quelques calculs d'intégrales généralisées

Voici réunis dans ce théorème quelques résultats concernant le calcul d'intégrales généralisées célèbres.

Théorème

1. $\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$ (intégrale de Gauss²⁴).
2. $\int_0^{+\infty} \frac{\sin(t)}{t} dt = \frac{\pi}{2}$ (intégrale de Dirichlet).
3. $\int_0^{+\infty} \cos(t^2) dt = \int_0^{+\infty} \sin(t^2) dt = \frac{1}{2}\sqrt{\frac{\pi}{2}}$ (intégrales de Fresnel).
4. $\int_0^{\pi/2} \ln(\sin(t)) dt = \int_0^{\pi/2} \ln(\cos(t)) dt = -\frac{\pi}{2} \ln(2)$ (intégrales d'Euler²⁵).

6.13 Suites et séries d'intégrales généralisées

I désigne un intervalle (non vide) de $\mathbb{R}$. On note $L^1(I, \mathbb{K})$ l'espace des fonctions intégrables de I dans $\mathbb{K}$.

Le théorème de convergence dominée permet, dans le cadre des intégrales généralisées, « l'échange des symboles limite et intégrale » au sens du théorème qui suit.

Théorème de convergence dominée de Lebesgue

Soient (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$ telles que :

- (f_n) converge simplement vers f sur I ;
- les fonctions f_n et f sont continues par morceaux sur I ;
- il existe $\varphi : I \rightarrow \mathbb{R}$ continue par morceaux et intégrable sur I vérifiant²⁶

$$\forall n \in \mathbb{N}, \forall t \in I, |f_n(t)| \leq \varphi(t).$$

Alors les fonctions f_n et f sont intégrables sur I et²⁷

$$\int_I f_n(t) dt \xrightarrow{n \rightarrow +\infty} \int_I f(t) dt.$$

24. et donc $\int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}$. On montre de même pour que pour tout $a \in \mathbb{R}_+^*$, $\int_{-\infty}^{+\infty} e^{-at^2} dt = \sqrt{\frac{\pi}{a}}$.

25. Ces intégrales sont parfois attribuées à tort à Dirichlet.

26. Même si l'hypothèse de positivité de φ est superflue via l'inégalité qui suit, on cherchera toujours en pratique une fonction φ positive et intégrable sur I .

27. ce que l'on peut également écrire : $\lim_{n \rightarrow +\infty} \left(\int_I f_n(t) dt \right) = \int_I \left(\lim_{n \rightarrow +\infty} f_n(t) \right) dt$.

Étendons ce résultat aux familles à paramètres $(f_\lambda)_{\lambda \in J}$ où J est un intervalle de $\mathbb{R}$.

Théorème

Soient J un intervalle de $\mathbb{R}$, $(f_\lambda)_{\lambda \in J}$ une famille de fonctions de I vers $\mathbb{K}$, λ_0 adhérent à J et $f : I \rightarrow \mathbb{K}$ tels que

- pour tout $x \in I$, $f_\lambda(x) \xrightarrow{\lambda \rightarrow \lambda_0} f(x)$;
- les fonctions f_λ et f sont continues par morceaux sur I ;
- il existe $\varphi : I \rightarrow \mathbb{R}$ continue par morceaux et intégrable sur I vérifiant

$$\forall \lambda \in J, \forall t \in I, |f_\lambda(t)| \leq \varphi(t).$$

Alors les fonctions f_λ et f sont intégrables sur I et ²⁸

$$\int_I f_\lambda(t) dt \xrightarrow{\lambda \rightarrow \lambda_0} \int_I f(t) dt.$$

Énonçons à présent le théorème d'intégration terme à terme.

Théorème d'intégration terme à terme

Soit $\sum f_n$ une série de fonctions de I vers $\mathbb{K}$ telle que

- $\sum f_n$ converge simplement sur I ;
- les fonctions f_n et la fonction somme, $\sum_{n=0}^{+\infty} f_n$, sont continues par morceaux sur I ;
- les fonctions f_n sont intégrables sur I ;
- la série numérique $\sum \|f_n\|_1$ converge ²⁹ où $\|f_n\|_1 = \int_I |f_n(t)| dt$.

Alors la fonction somme $\sum_{n=0}^{+\infty} f_n$ est intégrable sur I et

$$\int_I \left(\sum_{n=0}^{+\infty} f_n(t) \right) dt = \sum_{n=0}^{+\infty} \left(\int_I f_n(t) dt \right).$$

28. Ce théorème s'étend au cas $\lambda_0 \in \{-\infty, +\infty\}$.

29. On montre aisément que l'application $\begin{cases} L^1(I, \mathbb{K}) & \rightarrow \mathbb{R} \\ f & \mapsto \int_I |f(t)| dt \end{cases}$ est une norme.

6.14 Intégrales généralisées à paramètre

I désigne toujours un intervalle (non vide) de $\mathbb{R}$.

Dans cette section, on étudie les fonctions du type $F : x \in A \mapsto \int_I f(x, t) dt$ où A est une partie d'un espace vectoriel normé de dimension finie.

Énonçons d'abord le théorème de continuité « sous le signe $\int$ ».

Théorème

Soit $f : A \times I \rightarrow \mathbb{K}$ telle que

- pour tout $t \in I$, $x \mapsto f(x, t)$ est continue sur A ;
- pour tout $x \in A$, $t \mapsto f(x, t)$ est continue par morceaux sur I ;
- il existe $\varphi : I \rightarrow \mathbb{R}$ continue par morceaux et intégrable sur I vérifiant ³⁰

$$\forall (x, t) \in A \times I, |f(x, t)| \leq \varphi(t).$$

Alors la fonction $F : x \mapsto \int_I f(x, t) dt$ est définie et continue sur A .

La troisième hypothèse est appelée *hypothèse de domination*.

Il n'est pas toujours aisé de « dominer » la fonction f pour tout $x \in A$. En pratique, lorsque A est un intervalle de $\mathbb{R}$, il suffit de vérifier l'hypothèse de domination sur tout segment inclus dans A de sorte qu'il est possible dans ce cas particulier de reformuler le théorème précédent comme suit.

Théorème

Soient X un intervalle de $\mathbb{R}$ et $f : X \times I \rightarrow \mathbb{K}$ tels que

- pour tout $t \in I$, $x \mapsto f(x, t)$ est continue sur X ;
- pour tout $x \in X$, $t \mapsto f(x, t)$ est continue par morceaux sur I ;
- pour tout segment $K \subset X$, il existe $\varphi_K : I \rightarrow \mathbb{R}$ continue par morceaux et intégrable sur I vérifiant

$$\forall (x, t) \in K \times I, |f(x, t)| \leq \varphi_K(t).$$

Alors la fonction $F : x \mapsto \int_I f(x, t) dt$ est définie et continue sur X .

Énonçons à présent le théorème de dérivation « sous le signe $\int$ » appelé parfois *théorème de Leibniz* puis son extension à l'ordre $n \in \mathbb{N}^*$.

30. Même si l'hypothèse de positivité de φ est superflue via l'inégalité qui suit, on cherchera toujours en pratique une fonction φ positive et intégrable sur I .

Théorème

Soient X un intervalle de $\mathbb{R}$ non réduit à un point et $f : X \times I \rightarrow \mathbb{K}$ tels que

- pour tout $t \in I$, $x \mapsto f(x, t)$ est de classe $\mathcal{C}^1$ sur X ;
- pour tout $x \in X$, $t \mapsto f(x, t)$ est continue par morceaux et intégrable sur I ;
- pour tout $x \in X$, $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue par morceaux sur I ;
- il existe une fonction $\varphi : I \rightarrow \mathbb{R}$ continue par morceaux et intégrable sur I vérifiant

$$\forall (x, t) \in X \times I, \left| \frac{\partial f}{\partial x}(x, t) \right| \leq \varphi(t).$$

Alors $F : x \mapsto \int_I f(x, t) dt$ est de classe $\mathcal{C}^1$ sur X et pour tout $x \in X$,

$$F'(x) = \int_I \frac{\partial f}{\partial x}(x, t) dt.$$

Théorème

Soient $n \in \mathbb{N}^*$, X un intervalle de $\mathbb{R}$ non réduit à un point et $f : X \times I \rightarrow \mathbb{K}$ tels que

- pour tout $t \in I$, $x \mapsto f(x, t)$ est de classe $\mathcal{C}^n$ sur X ;
- pour tout $k \in \llbracket 0, n-1 \rrbracket$ et tout $x \in X$, $t \mapsto \frac{\partial^k f}{\partial x^k}(x, t)$ est continue par morceaux et intégrable sur I ;
- pour tout $x \in X$, $t \mapsto \frac{\partial^n f}{\partial x^n}(x, t)$ est continue par morceaux sur I ;
- il existe une fonction $\varphi : I \rightarrow \mathbb{R}$ continue par morceaux et intégrable sur I vérifiant

$$\forall (x, t) \in X \times I, \left| \frac{\partial^n f}{\partial x^n}(x, t) \right| \leq \varphi(t).$$

Alors $F : x \mapsto \int_I f(x, t) dt$ est de classe $\mathcal{C}^n$ sur X et pour tout $k \in \llbracket 1, n \rrbracket$ et tout $x \in X$,

$$F^{(k)}(x) = \int_I \frac{\partial^k f}{\partial x^k}(x, t) dt.$$

Équations différentielles linéaires

7

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Tout intervalle de $\mathbb{R}$ est supposé non vide et non réduit à un point.

7.1 Équations différentielles linéaires d'ordre 1

Définition

Soit I un intervalle de $\mathbb{R}$. On appelle *équation différentielle linéaire scalaire d'ordre 1* toute équation du type¹

$$y'(t) + a(t)y(t) = b(t) \quad (E)$$

d'inconnue y , fonction de I vers $\mathbb{K}$, dérivable sur I avec a et b deux fonctions de I vers $\mathbb{K}$, continues sur I .

On appelle *équation homogène* associée à (E) l'équation

$$y'(t) + a(t)y(t) = 0.$$

En reprenant les notations de la définition qui précède, on peut énoncer le théorème suivant.

Théorème

Soient (E) l'équation différentielle $y' + a(t)y = b(t)$ et A une primitive de a sur I . Alors

- les solutions de l'équation homogène $y'(t) + a(t)y(t) = 0$ sont les fonctions

$$y_h : \begin{cases} I & \longrightarrow \mathbb{K} \\ t & \longmapsto \lambda e^{-A(t)} \end{cases} \quad \text{où } \lambda \in \mathbb{K};$$

- les solutions de (E) sont les fonctions $y : \begin{cases} I & \longrightarrow \mathbb{K} \\ t & \longmapsto \lambda e^{-A(t)} + y_p(t) \end{cases}$

où y_p est une solution particulière² de (E) et $\lambda \in \mathbb{K}$.

1. Toute équation de ce type sera notée par la suite en abrégé : $y' + a(t)y = b(t)$.

2. L'usage est d'utiliser le qualificatif « particulière » pour désigner une solution quelconque de (E) sans que cette dernière n'ait quoi que ce soit de particulier.

L'ensemble des solutions de l'équation homogène associée à (E) est donc la droite vectorielle engendrée par la fonction $t \mapsto e^{-A(t)}$.

Théorème (principe de superposition)

Soient I un intervalle de $\mathbb{R}$, a , b_1 et b_2 trois fonctions de I vers $\mathbb{K}$ continues sur I .

Si y_1 est solution de $y' + a(t)y = b_1(t)$ et y_2 est solution de $y' + a(t)y = b_2(t)$ alors, pour tout $\lambda \in \mathbb{K}$, $\lambda y_1 + y_2$ est solution de l'équation

$$y' + a(t)y = \lambda b_1(t) + b_2(t).$$

Explicitons la méthode dite de « variation » de la constante, permettant, le plus souvent³, de déterminer une solution (particulière) de l'équation (E) connaissant les solutions de son équation homogène associée : si $y_h : t \mapsto \lambda e^{-A(t)}$ sont les solutions de l'équation homogène associée à (E) , on cherche une solution particulière de (E) de la forme

$$y_p : t \mapsto \lambda(t)e^{-A(t)}$$

où λ est une fonction de I vers $\mathbb{K}$, dérivable sur I .

Définition

On appelle *problème de Cauchy* associé à l'équation différentielle

$$(E) : y' + a(t)y = b(t)$$

tout problème consistant à chercher parmi les solutions de (E) celles vérifiant une condition initiale $y(t_0) = y_0$ où $t_0 \in I$ et $y_0 \in \mathbb{K}$.

En reprenant les notations de cette définition, on peut énoncer le théorème qui suit.

Théorème

Le problème de Cauchy $\begin{cases} y' + a(t)y = b(t) \\ y(t_0) = y_0 \end{cases}$ admet une unique solution.

3. Comme les primitives d'une fonction continue ne sont pas toujours exprimables à l'aide des fonctions usuelles (cf. note 5 page 76), ce théorème ne permet pas en toute généralité de déterminer une telle solution particulière.

7.2 Équations différentielles linéaires d'ordre 2 à coefficients constants

Définition

Soit I un intervalle de $\mathbb{R}$.

On appelle *équation différentielle linéaire d'ordre 2 à coefficients constants* toute équation du type

$$y'' + ay' + by = f(t) \quad (E)$$

d'inconnue y , fonction de I vers $\mathbb{K}$, deux fois dérivable sur I avec $(a, b) \in \mathbb{K}^2$ et f une fonction de I vers $\mathbb{K}$, continue sur I .

On appelle *équation homogène* associée à (E) l'équation

$$y'' + ay' + by = 0 \quad (E_h).$$

On appelle *équation caractéristique* associée à (E_h) l'équation $r^2 + ar + b = 0$ d'inconnue $r \in \mathbb{K}$.

Théorème (cas réel)

Soit (E_h) l'équation différentielle $y'' + ay' + by = 0$ où a et b sont des réels. Notons Δ le discriminant de l'équation caractéristique associée à (E_h) .

- Si $\Delta > 0$, l'équation caractéristique admet deux racines réelles r_1 et r_2 et les solutions de (E_h) sont les fonctions $y_h : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ t & \longmapsto \lambda e^{r_1 t} + \mu e^{r_2 t} \end{cases}$ où $(\lambda, \mu) \in \mathbb{R}^2$.

- Si $\Delta = 0$, l'équation caractéristique admet une racine réelle r et les solutions de (E_h) sont les fonctions $y_h : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ t & \longmapsto (\lambda t + \mu) e^{rt} \end{cases}$ où $(\lambda, \mu) \in \mathbb{R}^2$.

- Si $\Delta < 0$, l'équation caractéristique admet deux racines complexes conjuguées $\alpha + i\beta$ et $\alpha - i\beta$ et les solutions de (E_h) sont les fonctions $y_h : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ t & \longmapsto (\lambda \cos(\beta t) + \mu \sin(\beta t)) e^{\alpha t} \end{cases}$ où $(\lambda, \mu) \in \mathbb{R}^2$.

Théorème (cas complexe)

Soit (E_h) l'équation différentielle $y'' + ay' + by = 0$ où a et b sont des complexes. Notons Δ le discriminant de l'équation caractéristique associée à (E_h) .

• Si $\Delta \neq 0$, l'équation caractéristique admet deux racines complexes r_1 et r_2 et les solutions de (E_h) sont les fonctions $y_h : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ t & \longmapsto \lambda e^{r_1 t} + \mu e^{r_2 t} \end{cases}$ où $(\lambda, \mu) \in \mathbb{C}^2$.

• Si $\Delta = 0$, l'équation caractéristique admet une racine complexe r et les solutions de (E_h) sont les fonctions $y_h : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ t & \longmapsto (\lambda t + \mu) e^{rt} \end{cases}$ où $(\lambda, \mu) \in \mathbb{C}^2$.

Théorème (principe de superposition)

Soient $(a, b) \in \mathbb{K}^2$, f_1 et f_2 deux fonctions de I vers $\mathbb{K}$ continues sur I . Si y_1 est solution de $y'' + ay' + by = f_1(t)$ et y_2 est solution de $y'' + ay' + by = f_2(t)$ alors, pour tout $\lambda \in \mathbb{K}$, $\lambda y_1 + y_2$ est solution de l'équation

$$y'' + ay' + by = \lambda f_1(t) + f_2(t).$$

Théorème

Soient I un intervalle de $\mathbb{R}$, (E) l'équation différentielle $y'' + ay' + by = f(t)$ où $(a, b) \in \mathbb{K}^2$ et f est une fonction de I vers $\mathbb{K}$, continue sur I . Alors toute solution de E est une fonction $y : I \rightarrow \mathbb{K}$ de la forme $y = y_h + y_p$ où y_h est solution de l'équation homogène associée à (E) et y_p est une solution (particulière) de (E) .

Si $f \in \mathbb{K}[X]$, on cherche une solution (particulière) de (E) sous la forme d'un polynôme de même degré que f .

Si $f : t \mapsto P(t)e^{mt}$ où $P \in \mathbb{K}[X]$, on cherche une solution (particulière) de (E) sous la forme $t \mapsto Q(t)e^{mt}$ où le degré de Q est égal à la somme de k et du degré de P avec k l'ordre de multiplicité de m en tant que racine de l'équation caractéristique associée à l'équation homogène.

Théorème

Soient I un intervalle de $\mathbb{R}$, $t_0 \in I$, $(a, b, y_0) \in \mathbb{K}^3$ et f une fonction de I vers $\mathbb{K}$, continue sur I . Le problème de Cauchy $\begin{cases} y'' + ay' + by = f(t) \\ y(t_0) = y_0 \end{cases}$ admet une unique solution.

7.3 Exponentielle d'une matrice

Étendons la notion de convergence d'une *série numérique* à une *série vectorielle* c'est-à-dire à une série d'éléments d'un espace vectoriel normé.

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé et (u_n) une suite d'éléments de E . On appelle *série* de terme général u_n , notée $\sum u_n$, la suite des sommes partielles (S_n) où pour tout $n \in \mathbb{N}$, $S_n = \sum_{k=0}^n u_k$.

On dit que $\sum u_n$ *converge* si (S_n) converge. On dit qu'elle *diverge* sinon.

Si $\sum u_n$ converge, on appelle *somme* de la série le vecteur

$$\sum_{n=0}^{+\infty} u_n = \lim_{n \rightarrow +\infty} S_n = \lim_{n \rightarrow +\infty} \sum_{k=0}^n u_k.$$

On dit que $\sum u_n$ *converge absolument* ou est *absolument convergente* si $\sum \|u_n\|$ converge.

Théorème

Soient $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie et (u_n) une suite d'éléments de E . Si $\sum u_n$ est absolument convergente, $\sum u_n$ converge et

$$\left\| \sum_{n=0}^{+\infty} u_n \right\| \leq \sum_{n=0}^{+\infty} \|u_n\|.$$

Explicitons à présent une norme matricielle c'est-à-dire une norme sur $\mathcal{M}_n(\mathbb{K})$ où $n \in \mathbb{N}^*$.

Théorème

L'application $(a_{ij}) \mapsto \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ est une norme sur $\mathcal{M}_n(\mathbb{K})$. En la notant $\|\cdot\|$, elle vérifie⁴, pour tout $(A, B) \in \mathcal{M}_n^2(\mathbb{K})$, $\|AB\| \leq \|A\| \|B\|$.

4. Cette propriété dite de *sous-multiplicativité* est également vérifiée pour les normes matricielles $\|\cdot\|_1$ et $\|\cdot\|_2$ définies pour tout $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ par $\|A\|_1 = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|$ et

$\|A\|_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2}$. En revanche, elle n'est pas vérifiée pour la norme matricielle $\|\cdot\|_\infty$ définie par $\|A\|_\infty = \max_{(i,j) \in [1,n]^2} |a_{ij}|$.

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors la série $\sum \frac{A^n}{n!}$ est absolument convergente donc convergente. Sa somme, $\sum_{n=0}^{+\infty} \frac{A^n}{n!} \in \mathcal{M}_n(\mathbb{K})$ est appelée *exponentielle de A* notée $\exp(A)$ ou e^A .

Si E est un espace vectoriel normé de dimension finie, pour tout $u \in \mathcal{L}(E)$, la série $\sum \frac{u^n}{n!}$ est absolument convergente donc convergente et on définit l'*exponentielle de u* par

$$\exp(u) = e^u = \sum_{n=0}^{+\infty} \frac{u^n}{n!}.$$

Théorème

- Pour tout $(A, B) \in \mathcal{M}_n^2(\mathbb{K})$ tel que $AB = BA$, $e^{A+B} = e^A e^B = e^B e^A$.
- L'application $A \mapsto e^A$ est continue sur $\mathcal{M}_n(\mathbb{K})$.
- L'application $t \mapsto e^{tA}$ est dérivable sur $\mathbb{R}$ de dérivée $t \mapsto A e^{tA}$.

On notera $\text{Diag}(\lambda_1, \dots, \lambda_n)$ la matrice diagonale

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$
Théorème

Soient $A \in \mathcal{M}_n(\mathbb{K})$ diagonalisable et $P \in \text{GL}_n(\mathbb{K})$ telles que

$$D = P^{-1}AP = \text{Diag}(\lambda_1, \dots, \lambda_n).$$

Alors e^A est diagonalisable et

$$e^A = P e^D P^{-1} = P \text{Diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) P^{-1}.$$

7.4 Système différentiel linéaire d'ordre 1 à coefficients constants

Dans toute cette section, n désigne un entier naturel non nul.

Définition

Soit I un intervalle de $\mathbb{R}$.

On appelle *système différentiel linéaire d'ordre 1 à coefficients constants*⁵ tout système S de la forme⁶ $X'(t) = AX(t) + B(t)$ d'inconnue $t \mapsto X(t)$, fonction dérivable de I vers $\mathcal{M}_{n,1}(\mathbb{K})$, avec $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ et $t \mapsto B(t)$ une fonction continue de I vers $\mathcal{M}_{n,1}(\mathbb{K})$.

Le système différentiel $X'(t) = AX(t)$ est appelé *système différentiel homogène*⁷ associé à S .

Via l'identification classique entre $\mathbb{K}^n$ et $\mathcal{M}_{n,1}(\mathbb{K})$, tout système différentiel de ce type peut donc s'écrire

$$\begin{cases} x'_1(t) = a_{11}x_1(t) + \cdots + a_{1n}x_n(t) + b_1(t) \\ \vdots \\ x'_n(t) = a_{n1}x_1(t) + \cdots + a_{nn}x_n(t) + b_n(t) \end{cases}.$$

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors, le système différentiel homogène $X'(t) = AX(t)$ admet pour solution générale

$$X : t \mapsto e^{tA}K$$

où $K \in \mathcal{M}_{n,1}(\mathbb{K})$.

Théorème

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $X_0 \in \mathcal{M}_{n,1}(\mathbb{K})$.

Alors le problème de Cauchy $\begin{cases} X'(t) = AX(t) \\ X(t_0) = X_0 \end{cases}$ admet l'unique solution $X : t \mapsto e^{(t-t_0)A}X_0$.

5. On devrait en toute rigueur parler d'*équation différentielle matricielle* mais via l'isomorphisme entre $\mathbb{K}^n$ et $\mathcal{M}_{n,1}(\mathbb{K})$, les deux formulations coïncident (cf. la remarque qui suit la définition).

6. On notera souvent S en abrégé $X' = AX + B(t)$.

7. ou *équation différentielle homogène* associée à l'équation différentielle (matricielle) S .

7.5 Équations différentielles linéaires d'ordre 2

Définition

Soit I un intervalle de $\mathbb{R}$. On appelle *équation différentielle linéaire scalaire d'ordre 2* toute équation du type

$$y'' + a(t)y' + b(t)y = f(t) \quad (E)$$

d'inconnue y , fonction de I vers $\mathbb{K}$, deux fois dérivable sur I avec a , b et f trois fonctions de I vers $\mathbb{K}$, continues sur I .

On appelle *équation homogène* associée à (E) l'équation

$$y'' + a(t)y' + b(t)y = 0 \quad (E_h)$$

Théorème

L'ensemble des solutions de (E_h) est un sous-espace vectoriel de dimension 2 de l'espace des fonctions de classe $\mathcal{C}^2$ de I vers $\mathbb{K}$.

Si y_1 et y_2 sont deux solutions non colinéaires⁸ de (E_h) et y_p une solution (particulière) de (E) , toute solution de (E) est de la forme

$$y : t \mapsto y_p(t) + \lambda y_1(t) + \mu y_2(t) \quad \text{où } (\lambda, \mu) \in \mathbb{K}^2.$$

Théorème

Soient I un intervalle de $\mathbb{R}$, $t_0 \in I$, $(y_0, y_1) \in \mathbb{K}^2$, a , b et f trois fonctions de I vers $\mathbb{K}$, continues sur I . Le problème de Cauchy

$$\begin{cases} y'' + a(t)y' + b(t)y = f(t) \\ y(t_0) = y_0 \text{ et } y'(t_0) = y_1 \end{cases} \quad \text{admet une unique solution.}$$

Définition

Soient I un intervalle de $\mathbb{R}$, y_1 et y_2 deux fonctions de I vers $\mathbb{K}$ dérivables sur I . On appelle *wronskien* de y_1 et y_2 la fonction W de I vers $\mathbb{K}$ définie pour tout $t \in I$ par

$$W(t) = \begin{vmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{vmatrix} = y_1(t)y_2'(t) - y_2(t)y_1'(t).$$

8. Comme l'ensemble des solutions de (E_h) est de dimension 2, (y_1, y_2) forme donc une base de (E_h) .

Théorème

Soient y_1 et y_2 deux solutions de (E_h) et W leur wronskien. Alors W est solution de l'équation différentielle $y' + a(t)y = 0$.

Théorème

Soient y_1 et y_2 deux solutions de (E_h) et W leur wronskien. Alors les assertions sont équivalentes :

- (i) y_1 et y_2 non colinéaires.
- (ii) $\forall t \in I, W(t) \neq 0$.
- (iii) $\exists t_0 \in I, W(t_0) \neq 0$.

Soient I un intervalle de $\mathbb{R}$, a , b et f trois fonctions de I vers $\mathbb{K}$, continues sur I , (E) l'équation différentielle $y'' + a(t)y' + b(t)y = f(t)$, y_1 et y_2 deux solutions non colinéaires de (E_h) .

La méthode dite de « variations » des constantes consiste à chercher une solution (particulière) de (E) de la forme

$$y_p : t \mapsto \lambda(t)y_1(t) + \mu(t)y_2(t)$$

où λ et μ sont deux fonctions de I vers $\mathbb{K}$, dérivables sur I et solutions du système

$$\begin{cases} \lambda'(t)y_1(t) + \mu'(t)y_2(t) = 0 \\ \lambda'(t)y_1'(t) + \mu'(t)y_2'(t) = f(t) \end{cases}.$$

7.6 Équations différentielles linéaires vectorielles d'ordre 1

Définition

Soient I un intervalle de $\mathbb{R}$ et E un $\mathbb{K}$ -espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

On appelle *équation différentielle linéaire vectorielle d'ordre 1* toute équation (E) du type⁹

$$x'(t) = a(t)(x(t)) + b(t)$$

d'inconnue x , fonction de I vers E , dérivable sur I avec a et b deux fonctions de I vers $\mathcal{L}(E)$.

On appelle *équation homogène* associée à (E) l'équation $(E_h) : x' = a(t)(x)$.

9. On notera par la suite (E) en abrégé $x' = a(t)(x) + b(t)$. On rencontre également parfois l'écriture $x' = a(t) \cdot x + b(t)$.

En introduisant une base $\mathcal{B}$ de E et en notant $A(t)$ et $B(t)$ les matrices de $a(t)$ et $b(t)$ relativement à $\mathcal{B}$, l'équation (E) équivaut à l'équation différentielle matricielle

$$X'(t) = A(t)X(t) + B(t) \text{ ou en abrégé } X' = A(t)X + B(t).$$

Théorème

L'ensemble des solutions de (E_h) est un sous-espace vectoriel de n de l'espace des fonctions de classe $\mathcal{C}^1$ de I vers E .

Si $(x_1, \dots, x_n)$ est une base de l'ensemble des solutions de (E_h) et x_p une solution (particulière) de (E) alors toute solution de (E) est de la forme

$$x : t \mapsto x_p(t) + \lambda_1 x_1(t) + \dots + \lambda_n x_n(t) \quad \text{où } (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n.$$

Théorème

Soient I un intervalle de $\mathbb{R}$, $t_0 \in I$, E un espace vectoriel de dimension finie, $x_0 \in E$, a une fonction de I vers $\mathcal{L}(E)$, continue sur I et b une fonction de I vers E , continue sur I .

Alors le problème de Cauchy $\begin{cases} x' = a(t)(x) + b(t) \\ x(t_0) = x_0 \end{cases}$ admet une unique solution.

7.7 Équations différentielles linéaires d'ordre n

Dans cette section, n est un entier naturel supérieur ou égal à 2.

$y^{(k)}$ désigne la dérivée k -ième de la fonction y d'un intervalle I de $\mathbb{R}$ vers $\mathbb{K}$, k fois dérivable sur I avec $k \in \mathbb{N}$.

Définition

Soient I un intervalle de $\mathbb{R}$. On appelle *équation différentielle linéaire scalaire d'ordre n* toute équation du type

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = f(t) \quad (E)$$

d'inconnue y , fonction de I vers $\mathbb{K}$, n fois dérivable sur I avec $a_0, \dots, a_{n-1}$ et f des fonctions de I vers $\mathbb{K}$, continues sur I .

On appelle *équation homogène* associée à (E) l'équation

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = 0 \quad (E_h)$$

En posant $Y = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$, (E_h) est équivalente au système différentiel

$$Y' = A(t)Y \text{ où } A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ -a_0(t) & -a_1(t) & \cdots & \cdots & -a_{n-1}(t) \end{pmatrix}.$$

Théorème

L'ensemble des solutions de (E_h) est un sous-espace vectoriel de dimension n de l'espace des fonctions de classe $\mathcal{C}^n$ de I vers $\mathbb{K}$.

Si $(y_1, \dots, y_n)$ est une base de l'espace des solutions de (E_h) et y_p une solution (particulière) de (E) , toute solution de (E) est de la forme

$$y : t \mapsto y_p(t) + \lambda_1 y_1(t) + \cdots + \lambda_n y_n(t) \quad \text{où } (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n.$$

Théorème

Soient I un intervalle de $\mathbb{R}$, $t_0 \in I$, $(y_0, \dots, y_{n-1}) \in \mathbb{K}^n$, $a_0, \dots, a_{n-1}$ et f des fonctions de I vers $\mathbb{K}$, continues sur I .

Le problème de Cauchy
$$\begin{cases} y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y' + a_0(t)y = f(t) \\ y(t_0) = y_0, \dots, y^{(n-1)}(t_0) = y_{n-1} \end{cases}$$

admet une unique solution.

8

Suites de fonctions

Dans ce chapitre, I désigne, sauf cas particulier, une partie non vide de $\mathbb{R}$.
 $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

8.1 Convergence simple et convergence uniforme

Définition

Soient (f_n) une suite de fonctions¹ de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$.

On dit que (f_n) *converge simplement*² vers f sur I si

$$\forall x \in I, f_n(x) \xrightarrow{n \rightarrow +\infty} f(x)$$

autrement dit si

$$\forall x \in I, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies |f_n(x) - f(x)| < \varepsilon.$$

Si (f_n) est une suite de fonctions de I vers $\mathbb{K}$ et $x \in I$, $(f_n(x))$ est une suite numérique. Ainsi, étudier la convergence simple de (f_n) sur I revient, après avoir fixé x dans I , à étudier la limite de la suite $(f_n(x))$.

Remarquons également dans la définition ci-dessus que le rang N dépend à la fois de x et de ε contrairement à la définition de la convergence uniforme suivante.

Définition

Soient (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$.

On dit que (f_n) *converge uniformément*³ vers f sur I si :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \implies \forall x \in I, |f_n(x) - f(x)| < \varepsilon.$$

1. On peut étendre les définitions et certains théorèmes de ce chapitre à des fonctions d'une partie A d'un espace vectoriel E de dimension finie à valeurs dans un espace vectoriel normé F de dimension finie.

2. On dit aussi f est la *limite simple* de (f_n) sur I . On montre aisément l'unicité de la fonction vers laquelle une suite de fonctions peut converger simplement.

3. On dit aussi f est la *limite uniforme* de (f_n) sur I . Si cette limite uniforme existe, elle coïncide avec la limite simple.

Contrairement à la définition de la convergence simple, le rang N ne dépend que de ε et non de x c'est-à-dire que le rang N est le même⁴ pour chaque valeur de x .

Théorème

Soient (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$.

Si (f_n) converge uniformément vers f sur I , alors (f_n) converge simplement vers f sur I .

Ainsi, avant d'étudier une convergence uniforme de (f_n) sur I , on commence par déterminer l'éventuelle limite simple de (f_n) sur I .

Le théorème suivant donne une condition nécessaire et suffisante de convergence uniforme.

Théorème

Soient (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$. Alors

(f_n) converge uniformément vers f sur I si, et seulement si,

- les fonctions $f_n - f$ sont bornées à partir d'un certain rang sur I ;
- $\sup_{x \in I} |f_n(x) - f(x)| \xrightarrow{n \rightarrow +\infty} 0$.

Dans ce théorème, notons que $\left(\sup_{x \in I} |f_n(x) - f(x)| \right)$ est bien une suite numérique.

Dans le cas de fonctions à valeurs réelles, un tableau de variations permet souvent de déterminer cette borne supérieure. Néanmoins, si cette détermination s'avère non aisée, le théorème suivant permet également de montrer la convergence uniforme.

Théorème

Soient (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$.

S'il existe une suite réelle (ε_n) convergeant vers 0 telle qu'à partir d'un certain rang,

$$\forall x \in I, |f_n(x) - f(x)| \leq \varepsilon_n$$

alors (f_n) converge uniformément vers f sur I .

Sur l'espace des fonctions bornées de I vers $\mathbb{K}$, on peut définir la norme⁵

$$\|\cdot\|_\infty = \sup_{x \in I} |f(x)|$$

4. On dit que N est *uniforme* en x .

5. appelée *norme de la convergence uniforme*.

d'où le théorème suivant, qui est une reformulation du second théorème page 151.

Théorème

Soient (f_n) une suite de fonctions bornées de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$ bornée.

Alors (f_n) converge uniformément vers f sur $I \iff \|f_n - f\|_\infty \xrightarrow{n \rightarrow +\infty} 0$.

Lorsque (f_n) converge simplement mais non uniformément vers f sur I , (f_n) peut néanmoins converger uniformément vers f sur $J \subset I$.

Par exemple, si (f_n) est définie pour tout $n \in \mathbb{N}$ par $f_n : \begin{cases} [0, 1[& \longrightarrow \mathbb{R} \\ x & \longmapsto x^n \end{cases}$.

Alors (f_n) converge simplement mais non uniformément vers la fonction nulle sur $[0, 1[$ mais pour tout $a \in [0, 1[$, (f_n) converge uniformément vers la fonction nulle sur $[0, a]$.

8.2 Convergence uniforme et continuité

Énonçons le théorème principal et son corollaire immédiat.

Théorème

Soient $x_0 \in I$, (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$ tels que chaque f_n est continue en x_0 et (f_n) converge uniformément vers f sur I .

Alors ⁶ f est continue en x_0 .

Corollaire

Soient (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$ tels que chaque f_n est continue sur I et (f_n) converge uniformément vers f sur I .

Alors ⁷ f est continue sur I .

Ce théorème est faux avec une hypothèse de convergence simple.

Par exemple, soit (f_n) définie pour tout $n \in \mathbb{N}$ par $f_n : \begin{cases} [0, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto x^n \end{cases}$.

6. ce qui peut également s'écrire sous la forme d'une interversion de limites : $\lim_{n \rightarrow +\infty} \lim_{x \rightarrow x_0} f_n(x) = \lim_{x \rightarrow x_0} \lim_{n \rightarrow +\infty} f_n(x)$.

7. ce qui peut s'énoncer ainsi : « Toute limite uniforme d'une suite de fonctions continues sur I est continue sur I . »

Alors chaque f_n est continue sur $[0, 1]$ mais (f_n) converge simplement sur $[0, 1]$

$$\text{vers la fonction discontinue } f : \begin{cases} [0, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} 0 & \text{si } x \neq 1 \\ 1 & \text{si } x = 1 \end{cases} \end{cases}.$$

Il n'est pas toujours possible d'établir la continuité de f sur I via la convergence uniforme de (f_n) vers f sur I tout entier. Comme la continuité est une propriété locale, il suffit, par exemple dans le cas d'un intervalle I , d'établir la convergence uniforme sur tout segment inclus dans I : reprenons par exemple (f_n) définie pour

$$\text{tout } n \in \mathbb{N} \text{ par } f_n : \begin{cases} [0, 1[& \longrightarrow \mathbb{R} \\ x & \longmapsto x^n \end{cases}.$$

Alors, bien que (f_n) ne converge pas uniformément vers f (qui est ici la fonction nulle) sur $[0, 1[$, elle converge uniformément vers f sur $[0, a]$ pour tout $a \in]0, 1[$. On peut donc en déduire que f est continue sur $[0, a]$ pour tout $a \in [0, 1[$ donc sur $[0, 1[$.

8.3 Convergence uniforme et double limite

Le théorème suivant propose une extension du théorème page 152 traitant de la continuité en un point d'une suite de fonctions convergeant uniformément.

Théorème de la double limite

Soient x_0 adhérent à I , (f_n) une suite de fonctions de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$ tels que chaque f_n admet une limite finie ℓ_n en x_0 et (f_n) converge uniformément vers f sur I alors⁸ (ℓ_n) admet une limite finie ℓ et f tend vers ℓ en x_0 autrement dit

$$\lim_{n \rightarrow +\infty} \lim_{x \rightarrow x_0} f_n(x) = \lim_{x \rightarrow x_0} \lim_{n \rightarrow +\infty} f_n(x)$$

La condition de convergence uniforme est indispensable pour assurer la conclusion du théorème : il suffit pour s'en assurer de reprendre l'exemple de (f_n) définie pour

$$\text{tout } n \in \mathbb{N} \text{ par } f_n : \begin{cases} [0, 1[& \longrightarrow \mathbb{R} \\ x & \longmapsto x^n \end{cases}.$$

On a en effet : $\lim_{n \rightarrow +\infty} \lim_{x \rightarrow 1} f_n(x) = 1$ alors que $\lim_{x \rightarrow 1} \lim_{n \rightarrow +\infty} f_n(x) = 0$.

8.4 Intégration d'une suite de fonctions

Énonçons d'abord le théorème d'interversion des symboles limite et intégrale dans le cas d'un segment $I = [a, b]$.

8. Ce théorème s'étend au cas $x_0 \in \{-\infty, +\infty\}$.

Théorème

Soient $(a, b) \in \mathbb{R}^2$ avec $a \leq b$, (f_n) une suite de fonctions continues de $[a, b]$ vers $\mathbb{K}$ et $f : [a, b] \rightarrow \mathbb{K}$ tels que (f_n) converge uniformément vers f sur $[a, b]$. Alors

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow +\infty} f_n(x) dx = \int_a^b f(x) dx.$$

Remarquons que l'hypothèse de *convergence uniforme* sur un *segment* est primordiale afin d'assurer la conclusion du théorème : soit (f_n) définie pour tout $n \in \mathbb{N}$

$$\text{par } f_n : \begin{cases} [0, 1[& \rightarrow \mathbb{R} \\ x & \mapsto n^2 x^n \end{cases}.$$

Alors (f_n) converge simplement vers la fonction nulle sur $[0, 1[$ mais

$$\lim_{n \rightarrow +\infty} \int_0^1 f_n(x) dx = +\infty \text{ alors que } \int_0^1 \lim_{n \rightarrow +\infty} f_n(x) dx = 0.$$

Dans le cas plus général d'un intervalle I de $\mathbb{R}$, on peut néanmoins énoncer le théorème suivant sur les primitives.

Théorème

Soient I un intervalle de $\mathbb{R}$, $x_0 \in I$, (f_n) une suite de fonctions continues de I vers $\mathbb{K}$ et $f : I \rightarrow \mathbb{K}$ tels que (f_n) converge uniformément vers f sur tout segment de I . Pour tout $n \in \mathbb{N}$ et tout $x \in I$, notons

$$F_n(x) = \int_{x_0}^x f_n(t) dt \quad \text{et} \quad F(x) = \int_{x_0}^x f(t) dt.$$

Alors (F_n) converge uniformément vers F sur tout segment de I .

8.5 Dérivation d'une suite de fonctions

Théorème

Soient I un intervalle de $\mathbb{R}$ non réduit à un point et (f_n) une suite de fonctions de classe $\mathcal{C}^1$ de I vers $\mathbb{K}$ convergeant simplement vers une fonction f sur I avec (f'_n) convergeant uniformément vers une fonction g sur tout segment de I .

Alors (f_n) converge uniformément vers f sur tout segment de I , f est de classe $\mathcal{C}^1$ sur I et $f' = g$ c'est-à-dire $\left(\lim_{n \rightarrow +\infty} f_n \right)' = \lim_{n \rightarrow +\infty} f'_n$.

Remarquons qu'il est suffisant d'avoir une convergence uniforme de (f'_n) vers g sur tout segment de I et non sur tout I afin d'assurer la conclusion du théorème.

Sans cette hypothèse de convergence uniforme de (f'_n) (l'éventuelle convergence uniforme de (f_n) ne suffit pas), non seulement la dérivabilité de f n'est pas assurée mais la convergence même simple de (f'_n) vers f' non plus : par exemple, soit (f_n)

définie pour tout $n \in \mathbb{N}^*$ par $f_n : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \frac{1}{n} \sin(nx) \end{cases}$. Alors (f_n) converge uniformément vers la fonction nulle sur $\mathbb{R}$. Or, pour tout $x \in \mathbb{R}$, $f'_n(x) = \cos(nx)$ donc (f'_n) ne converge pas simplement sur $\mathbb{R}$.

Étendons à présent le théorème précédent à un ordre de dérivation supérieur.

Théorème

Soient $p \in \mathbb{N}^*$, I un intervalle de $\mathbb{R}$ non réduit à un point et (f_n) une suite de fonctions de classe $\mathcal{C}^p$ de I vers $\mathbb{K}$ tels que

- $\forall k \in \llbracket 0, p-1 \rrbracket$, $(f_n^{(k)})$ converge simplement vers une fonction f_k sur I ;
- $(f_n^{(p)})$ converge uniformément vers une fonction f_p sur tout segment de I .

Alors (f_n) converge uniformément vers $f = f_0$ sur tout segment de I , f est de classe $\mathcal{C}^p$ sur I et, pour tout $k \in \llbracket 1, p \rrbracket$, $f^{(k)} = f_k$.

8.6 Approximation uniforme

Théorème

Toute fonction continue par morceaux sur $[a, b]$ est la limite uniforme sur $[a, b]$ d'une suite de fonctions en escalier sur $[a, b]$.

Donc, si $f : [a, b] \rightarrow \mathbb{K}$ est continue par morceaux sur $[a, b]$, pour tout $\varepsilon > 0$, il existe une fonction en escalier $g : [a, b] \rightarrow \mathbb{K}$ telle que⁹ $\|f - g\|_\infty < \varepsilon$ c'est-à-dire telle que pour tout $x \in [a, b]$, $|f(x) - g(x)| < \varepsilon$.

Théorème de Weierstrass

Toute fonction continue sur $[a, b]$ est la limite uniforme sur $[a, b]$ d'une suite de fonctions polynomiales.

Donc, si $f : [a, b] \rightarrow \mathbb{K}$ est continue sur $[a, b]$, pour tout $\varepsilon > 0$, il existe une fonction polynomiale $P : [a, b] \rightarrow \mathbb{K}$ telle que $\|f - P\|_\infty < \varepsilon$ c'est-à-dire telle que pour tout $x \in [a, b]$, $|f(x) - P(x)| < \varepsilon$.

9. $\|\cdot\|_\infty$ désigne la norme de la convergence uniforme définie page 151.

9

Séries de fonctions

Dans ce chapitre, I désigne, sauf cas particulier, une partie non vide de $\mathbb{R}$.
 $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

9.1 Convergence simple et convergence uniforme

Définition

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$.

On appelle *série de fonctions*¹ de terme général f_n , notée² $\sum f_n$, la suite de fonctions des sommes partielles (S_n) où pour tout $n \in \mathbb{N}$, $S_n = \sum_{k=0}^n f_k$.

Tout comme les séries numériques sont des suites « particulières », les séries de fonctions sont des suites de fonctions « particulières ».

Si $(f_n)_{n \geq n_0}$ n'est définie qu'à partir de n_0 , la série de fonctions associée est alors la suite de fonctions $(S_n)_{n \geq n_0}$ définie pour tout $n \geq n_0$ par $S_n = \sum_{k=n_0}^n f_k$.

Toutes les définitions et théorèmes figurant dans ce chapitre ne sont énoncés, sauf cas particulier, que pour $n_0 = 0$ mais restent valables si $n_0 \neq 0$.

Définition

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$.

On dit que $\sum f_n$ converge simplement sur I si la suite de fonctions de ses sommes partielles (S_n) converge simplement sur I .

1. Comme pour les suites de fonctions, on peut étendre les définitions et certains théorèmes de ce chapitre à des fonctions d'une partie A d'un espace vectoriel E de dimension finie à valeurs dans un espace vectoriel normé F de dimension finie.

2. ou parfois $\sum_{n \geq 0} f_n$.

Ainsi $\sum f_n$ converge simplement sur I si, pour tout $x \in I$, la série numérique $\sum f_n(x)$ converge c'est-à-dire si, pour tout $x \in I$, la suite numérique $\left(\sum_{k=0}^n f_k(x)\right)$ converge.

Définition

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$ telle que $\sum f_n$ converge simplement sur I . On appelle *somme* (ou *fonction somme*) de la série $\sum f_n$, la fonction notée S ou $\sum_{n=0}^{+\infty} f_n$, définie pour tout $x \in I$ par

$$S(x) = \sum_{n=0}^{+\infty} f_n(x) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n f_k(x).$$

On appelle *reste* de la série $\sum f_n$ la suite de fonctions de I vers $\mathbb{K}$, notée (R_n) , définie pour tout $n \in \mathbb{N}$ et tout $x \in I$ par

$$R_n(x) = \sum_{k=n+1}^{+\infty} f_k(x).$$

Remarquons que si $\sum f_n$ converge simplement vers la fonction S sur I , alors, pour tout $n \in \mathbb{N}$, $R_n = S - S_n$ donc (R_n) converge simplement vers la fonction nulle sur I .

Définition

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$.

On dit que $\sum f_n$ *converge uniformément* sur I si la suite de fonctions de ses sommes partielles (S_n) converge uniformément sur I .

On en déduit que la convergence uniforme d'une série de fonctions sur I entraîne sa convergence simple sur I . Cette définition n'apporte pas beaucoup d'éléments pour démontrer la convergence uniforme d'une série de fonctions.

Commençons par donner un théorème utile par sa contraposée.

Théorème

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$ telle que $\sum f_n$ converge uniformément sur I .

Alors ³ (f_n) converge uniformément vers la fonction nulle sur I .

3. La réciproque est fautive : la convergence uniforme de (f_n) vers la fonction nulle n'entraîne pas nécessairement la convergence uniforme de $\sum f_n$. Pour s'en convaincre, on pourra considérer (f_n) définie pour tout $n \in \mathbb{N}^*$ et tout $x \in \mathbb{R}_+$ par $f_n(x) = \frac{1}{n+n^3x^2}$.

Ainsi, via la contraposée de ce théorème, si (f_n) ne converge pas uniformément sur I vers la fonction nulle⁴ alors $\sum f_n$ ne converge pas uniformément sur I .

Le théorème suivant donne une condition nécessaire et suffisante afin qu'une série de fonctions converge uniformément.

Théorème

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$.

Alors $\sum f_n$ converge uniformément sur I si, et seulement si, $\sum f_n$ converge simplement sur I et la suite de fonctions de ses restes (R_n) converge uniformément vers la fonction nulle sur I .

Pour l'étude de la convergence uniforme de (R_n) vers la fonction nulle, on peut :

- étudier si $\sup_{x \in I} |R_n(x)| \xrightarrow{n \rightarrow +\infty} 0$;
- déterminer une suite numérique (ε_n) , convergeant vers 0, telle que

$$\forall x \in I, |R_n(x)| \leq \varepsilon_n.$$

Ce théorème est particulièrement adapté dans le cas d'une série numérique alternée vérifiant les hypothèses du critère spécial page 118, car alors ce dernier fournit une majoration des restes.

Par exemple, soit (f_n) définie pour tout $n \in \mathbb{N}^*$ par $f_n : \begin{cases} [0, 1] & \longrightarrow \mathbb{R} \\ x & \longmapsto \frac{(-1)^n x^n}{n} \end{cases}$.

Pour tout $x \in [0, 1]$, la série numérique $\sum f_n(x)$ est alternée et vérifie les hypothèses du critère spécial, donc $\sum f_n$ converge simplement sur I .

De plus, pour tout $x \in [0, 1]$, $|R_n(x)| \leq \left| \frac{(-1)^{n+1} x^{n+1}}{n+1} \right| \leq \frac{1}{n+1} \xrightarrow{n \rightarrow +\infty} 0$.

Donc (R_n) converge uniformément vers la fonction nulle sur $[0, 1]$ d'où la convergence uniforme de $\sum f_n$ sur $[0, 1]$.

9.2 Convergence normale et convergence absolue

Un autre outil, qui va s'avérer parfois très utile pour montrer la convergence uniforme, est la *convergence normale*.

4. c'est-à-dire si $\sup_{x \in I} |f_n(x)|$ ne tend pas vers 0 lorsque n tend vers l'infini.

Définition

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$.

On dit que $\sum f_n$ *converge normalement* sur I si

- les fonctions f_n sont bornées sur I ;
- la série numérique⁵ $\sum \|f_n\|_\infty$ converge.

Pour l'étude de la convergence normale de $\sum f_n$ sur I , on peut :

- déterminer explicitement⁶ $\|f_n\|_\infty$ puis étudier la nature de $\sum \|f_n\|_\infty$;
- déterminer une suite numérique (α_n) telle que $\sum \alpha_n$ converge et vérifiant

$$\forall x \in I, |f_n(x)| \leq \alpha_n.$$

Théorème

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$ telle que $\sum f_n$ converge normalement sur I . Alors $\sum f_n$ converge uniformément sur I .

Ainsi, le premier test à effectuer lors de l'étude de la convergence uniforme est de voir si la série converge normalement.

La réciproque du théorème est fausse : soit (f_n) définie pour tout $n \in \mathbb{N}^*$ par

$$f_n : \begin{cases} \mathbb{R}_+ & \longrightarrow \mathbb{R} \\ x & \longmapsto \frac{(-1)^n}{n+x} \end{cases}.$$

La série $\sum f_n$ converge simplement sur $\mathbb{R}_+$ via le critère spécial des séries alternées.

De plus, via ce même critère, pour tout $x \in \mathbb{R}_+$, on a

$$|R_n(x)| \leq \frac{1}{n+1+x} \leq \frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0$$

donc $\sum f_n$ converge uniformément sur $\mathbb{R}_+$.

Or, $\sup_{x \in \mathbb{R}_+} \left| \frac{(-1)^n}{n+x} \right| = \frac{1}{n}$ donc $\sum f_n$ ne converge pas normalement sur $\mathbb{R}_+$.

Introduisons à présent un dernier type de convergence.

Définition

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$.

On dit que $\sum f_n$ *converge absolument* sur I si, pour tout $x \in I$, la série numérique $\sum |f_n(x)|$ converge.

5. On rappelle que $\|\cdot\|_\infty$ désigne la norme de la convergence uniforme c'est-à-dire $\|f_n\|_\infty = \sup_{x \in I} |f_n(x)|$.

6. via un tableau de variations dans le cas de fonctions à valeurs réelles.

Comme la convergence absolue d'une série numérique entraîne sa convergence, la convergence absolue d'une série de fonctions entraîne sa convergence simple.

Théorème

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$ telle que $\sum f_n$ converge normalement sur I . Alors $\sum f_n$ converge absolument sur I .

Il n'y a en revanche aucun lien entre la convergence uniforme et la convergence absolue⁷.

9.3 Convergence uniforme et continuité

Les deux théorèmes page 152 peuvent être adaptés aux séries de fonctions et résumés en un seul.

Théorème

Soient $x_0 \in I$ et $\sum f_n$ une série de fonctions de I vers $\mathbb{K}$ convergeant uniformément sur⁸ I .

Si les fonctions f_n sont continues en x_0 (respectivement sur I), alors $\sum_{n=0}^{+\infty} f_n$ est continue en x_0 (respectivement sur I).

9.4 Convergence uniforme et double limite

Adaptons à nouveau le théorème déjà vu sur les suites de fonctions.

Théorème de la double limite

Soient x_0 adhérent à I et $\sum f_n$ une série de fonctions de I vers $\mathbb{K}$ tels que chaque f_n admet une limite finie ℓ_n en x_0 et $\sum f_n$ converge uniformément sur I alors⁹ $\sum \ell_n$ converge, $\sum_{n=0}^{+\infty} f_n$ admet une limite en x_0 et

$$\lim_{x \rightarrow x_0} \sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} \ell_n \quad \text{c'est-à-dire} \quad \lim_{x \rightarrow x_0} \sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} \lim_{x \rightarrow x_0} f_n(x).$$

7. En considérant (f_n) définie pour tout $n \in \mathbb{N}^*$ et tout $x \in [0, 1]$ par $f_n(x) = \frac{(-1)^n x^n}{n}$, on montre que la convergence uniforme n'entraîne pas la convergence absolue. Puis en considérant (f_n) définie pour tout $n \in \mathbb{N}$ et tout $x \in]1, +\infty[$ par $f_n(x) = \frac{1}{n^x}$, on montre que la convergence absolue n'entraîne pas la convergence uniforme.

8. Comme pour le théorème relatif aux suites de fonctions, il suffit dans le cas d'un intervalle I , de vérifier la convergence uniforme sur tout segment inclus dans I .

9. Ce théorème s'étend au cas $x_0 \in \{-\infty, +\infty\}$.

9.5 Intégration terme à terme sur un segment

Théorème d'intégration terme à terme sur un segment

Soit $\sum f_n$ une série de fonctions de $[a, b]$ vers $\mathbb{K}$, convergeant uniformément sur $[a, b]$ et telle que chaque f_n est continue sur $[a, b]$. Alors ¹⁰

$$\int_a^b \sum_{n=0}^{+\infty} f_n(x) \, dx = \sum_{n=0}^{+\infty} \int_a^b f_n(x) \, dx.$$

9.6 Dérivation terme à terme

Théorème de dérivation terme à terme

Soient I un intervalle de $\mathbb{R}$ non réduit à un point, $\sum f_n$ une série de fonctions de classe $\mathcal{C}^1$ de I vers $\mathbb{K}$ convergeant simplement sur I et tels que $\sum f'_n$ converge uniformément sur tout segment de I .

Alors $\sum_{n=0}^{+\infty} f_n$ converge uniformément sur tout segment de I , $\sum_{n=0}^{+\infty} f_n$ est de classe $\mathcal{C}^1$ sur I et $\left(\sum_{n=0}^{+\infty} f_n\right)' = \sum_{n=0}^{+\infty} f'_n$.

Comme pour les suites de fonctions, on peut étendre ce théorème à un ordre de dérivation supérieur.

Théorème

Soient $p \in \mathbb{N}^*$, I un intervalle de $\mathbb{R}$ non réduit à un point et (f_n) une suite de fonctions de classe $\mathcal{C}^p$ de I vers $\mathbb{K}$ tels que

- pour tout $k \in \llbracket 0, p-1 \rrbracket$, $(f_n^{(k)})$ converge simplement sur I ;
- $(f_n^{(p)})$ converge uniformément sur tout segment de I .

Alors $\sum_{n=0}^{+\infty} f_n$ est de classe $\mathcal{C}^p$ sur I et, pour tout $k \in \llbracket 1, p \rrbracket$,

$$\left(\sum_{n=0}^{+\infty} f_n\right)^{(k)} = \sum_{n=0}^{+\infty} f_n^{(k)}.$$

10. Dans le cas plus général d'une série de fonctions d'un intervalle I quelconque vers $\mathbb{K}$, cf. théorème d'intégration terme à terme page 136.

10

Séries entières

10.1 Rayon de convergence

Définition

Soit (a_n) est une suite complexe. On appelle *série entière* (d'une variable complexe)¹ toute série² $\sum f_n$ où pour tout $n \in \mathbb{N}$,

$$f_n : \begin{cases} \mathbb{C} & \longrightarrow \mathbb{C} \\ z & \longmapsto a_n z^n \end{cases}$$

Les séries entières sont donc des séries de fonctions « particulières ».

Étant donné une série entière $\sum a_n z^n$, il est naturel de déterminer non seulement l'ensemble des $z \in \mathbb{C}$ tels que la série converge mais aussi les propriétés de sa fonction somme lorsqu'elle est définie.

Par exemple la série entière $\sum z^n$ (série géométrique) converge pour tout $z \in \mathbb{C}$ tel que $|z| < 1$ et dans ce cas

$$\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}.$$

Lemme d'Abel

Soient $(a_n) \in \mathbb{C}^{\mathbb{N}}$ et $z_0 \in \mathbb{C}$ tels que la suite $(a_n z_0^n)_{n \in \mathbb{N}}$ est bornée³.

Alors, pour tout $z \in \mathbb{C}$ tel que $|z| < |z_0|$, la série numérique $\sum a_n z^n$ est absolument convergente.

1. Par défaut, une série entière sera toujours d'une variable complexe. On parlera ponctuellement de série entière d'une variable réelle c'est-à-dire de série $\sum f_n$ où pour tout $n \in \mathbb{N}$,

$f_n : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto a_n x^n \end{cases}$ et (a_n) est une suite complexe.

2. Une série entière est une série de fonctions et sa somme, si elle est définie, est une fonction de la variable complexe z . Par abus, si $z \in \mathbb{C}$ est fixé, la série numérique $\sum a_n z^n$ sera encore appelée série entière.

3. C'est en particulier le cas si la série numérique $\sum a_n z_0^n$ converge car alors $a_n z_0^n \xrightarrow{n \rightarrow +\infty} 0$.

Si $\sum a_n z_0^n$ converge pour un certain $z_0 \in \mathbb{C}$, la série ne converge pas nécessairement absolument pour tout $z \in \mathbb{C}$ tel que $|z| = |z_0|$: en effet, en choisissant par exemple $(a_n) = \left(\frac{1}{n}\right)$ et $z_0 = -1$, la série alternée $\sum \frac{(-1)^n}{n}$ converge alors que la série $\sum \frac{1}{n}$ diverge.

Définissons à présent le *rayon de convergence* d'une série entière.

Définition

On appelle *rayon de convergence* d'une série entière $\sum a_n z^n$ la borne supérieure⁴ de l'ensemble $I = \{r \in \mathbb{R}_+ \text{ tels que la suite } (a_n r^n) \text{ est bornée}\}$.

Théorème

Soit $\sum a_n z^n$ de rayon de convergence R .

- Si $|z| < R$ alors $\sum a_n z^n$ converge absolument.
- Si $|z| > R$ alors $\sum a_n z^n$ diverge grossièrement⁵.

Dire que le rayon de convergence d'une série entière $\sum a_n z^n$ est égal à 0 signifie qu'elle ne converge que pour $z = 0$.

Dire que le rayon de convergence d'une série entière $\sum a_n z^n$ est égal à $+\infty$ signifie que pour tout $z \in \mathbb{C}$, $\sum a_n z^n$ converge.

Soit $\sum a_n z^n$ une série entière de rayon de convergence R . Ce théorème ne permet pas de conclure quant à sa convergence ou sa divergence pour les $z \in \mathbb{C}$ tel que $|z| = R$.

Plus précisément, en notant $\mathcal{C} = \{z \in \mathbb{C}, |z| = R\}$, $\sum a_n z^n$ peut converger en tout point⁶ de $\mathcal{C}$, diverger en tout point⁷ de $\mathcal{C}$ ou seulement converger en certains points⁸ de $\mathcal{C}$.

Le corollaire suivant est souvent très utile pour déterminer un rayon de convergence.

4. Si I n'est pas majorée, le rayon de convergence de la série entière est égale à $+\infty$.

5. On rappelle que $\sum a_n z^n$ *diverge grossièrement* signifie que son terme général ne tend pas vers 0.

6. C'est le cas de la série entière $\sum \frac{z^n}{n^2}$ de rayon de convergence $R = 1$.

7. C'est le cas de la série entière géométrique $\sum z^n$ de rayon de convergence $R = 1$.

8. Par exemple, la série entière $\sum \frac{z^n}{n}$ de rayon de convergence $R = 1$, converge en tout point $z \in \mathbb{C}$ tel que $|z| = 1$ et $z \neq 1$.

Corollaire

Soit $\sum a_n z^n$ de rayon de convergence R et $z_0 \in \mathbb{C}$.

- Si $\sum a_n z_0^n$ converge alors $|z_0| \leq R$.
- Si $\sum a_n z_0^n$ diverge alors $|z_0| \geq R$.

Définition

Soit $\sum a_n z^n$ une série entière de rayon de convergence R .

Le disque $\{z \in \mathbb{C}, |z| < R\}$ est appelé *disque ouvert de convergence* de la série entière⁹.

Ne pas confondre le *disque ouvert de convergence* et le *domaine de convergence*¹⁰ d'une série entière $\sum a_n z^n$ qui est l'ensemble $\mathcal{D}$ des $z \in \mathbb{C}$ tels que la série numérique $\sum a_n z^n$ converge. Via ce qui précède,

- si $R = 0$ alors $\mathcal{D} = \{0\}$;
- si $R = +\infty$ alors $\mathcal{D} = \mathbb{C}$;
- si $R \in]0, +\infty[$ alors $\{z \in \mathbb{C}, |z| < R\} \subset \mathcal{D} \subset \{z \in \mathbb{C}, |z| \leq R\}$.

La fonction somme de la série entière $\sum a_n z^n$ est alors la fonction $f : \mathcal{D} \rightarrow \mathbb{C}$ définie pour tout $z \in \mathcal{D}$ par

$$f(z) = \sum_{n=0}^{+\infty} a_n z^n.$$

10.2 Règles de d'Alembert et de Cauchy

Les règles de d'Alembert et de Cauchy évoquées dans le cadre des séries numériques page 115 peuvent être adaptées aux séries entières afin de déterminer, dans certains cas, le rayon de convergence.

Théorème (règle de d'Alembert)

Soit (a_n) une suite complexe à termes non nuls¹¹ telle que

$$\left| \frac{a_{n+1}}{a_n} \right| \xrightarrow{n \rightarrow +\infty} \ell \in \mathbb{R}_+^*.$$

Alors le rayon de convergence de $\sum a_n z^n$ est¹² $R = \frac{1}{\ell}$.

9. Dans le cas d'une série entière d'une variable réelle $\sum a_n x^n$, l'intervalle $\{x \in \mathbb{R}, |x| < R\} =]-R, R[$ est appelé *intervalle ouvert de convergence*.

10. Dans le cas d'une série entière d'une variable réelle $\sum a_n x^n$, on parle de l'*intervalle de convergence* qui est l'ensemble I des $x \in \mathbb{R}$ tels que la série numérique $\sum a_n x^n$ converge.

11. au moins à partir d'un certain rang.

12. Le théorème est encore vérifié avec $\ell = 0$ et $\ell = +\infty$ en convenant que $R = 0$ si $\ell = +\infty$ et $R = +\infty$ si $\ell = 0$.

Corollaire

Soit $\alpha \in \mathbb{R}$. Alors le rayon de convergence de $\sum n^\alpha z^n$ est $R = 1$.

Théorème (règle de Cauchy)

Soit (a_n) une suite complexe telle que

$$\sqrt[n]{|a_n|} \xrightarrow{n \rightarrow +\infty} \ell \in \mathbb{R}_+^*.$$

Alors le rayon de convergence de $\sum a_n z^n$ est¹³ $R = \frac{1}{\ell}$.

10.3 Comparaisons des rayons de convergence

Énonçons les critères de comparaison pour les séries entières.

Théorème

Soient $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons de convergence respectifs R_a et R_b .

1. Si pour tout $n \in \mathbb{N}$, $|a_n| \leq |b_n|$, alors $R_a \geq R_b$.
2. Si $a_n = O(b_n)$, alors $R_a \geq R_b$.
3. Si $a_n = o(b_n)$, alors $R_a \geq R_b$.
4. Si $a_n \sim b_n$, $R_a = R_b$.

On en déduit le corollaire suivant.

Corollaire

Soit (a_n) une suite complexe telle qu'il existe $(k, K) \in (\mathbb{R}_+^*)^2$ vérifiant

$$\forall n \in \mathbb{N}, k \leq |a_n| \leq K.$$

Alors le rayon de convergence de la série entière $\sum a_n z^n$ est égal à 1.

Énonçons à présent un résultat souvent utile.

13. Cf. note précédente.

Théorème

Les séries entières $\sum a_n z^n$ et $\sum n a_n z^n$ ont même rayon de convergence.

10.4 Opérations sur les séries entières

Énonçons d'abord le théorème relatif à la somme de deux séries entières.

Théorème

Soient $\sum a_n z^n$, $\sum b_n z^n$ deux séries entières de rayons de convergence R_a et R_b et R le rayon de convergence de la série entière somme $\sum (a_n + b_n) z^n$. Alors

1. Si $R_a \neq R_b$, $R = \min(R_a, R_b)$.
2. Si $R_a = R_b$, $R \geq R_a$ et $R \geq R_b$.
3. Pour tout $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$,

$$\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n.$$

Si $R_a = R_b$, on peut avoir $R > R_a$.

En effet, considérons les deux séries entières $\sum z^n$ et $\sum -z^n$ de rayon de convergence égale à 1. Alors le rayon de convergence de la série somme est $+\infty$.

Énonçons à présent le théorème relatif au produit par un scalaire.

Théorème

Soient $\sum a_n z^n$ une série entière de rayon de convergence R et $\lambda \in \mathbb{C}^*$. Alors

1. La série entière $\sum \lambda a_n z^n$ a pour rayon de convergence R .
2. Pour tout $z \in \mathbb{C}$ tel que $|z| < R$, $\sum_{n=0}^{+\infty} \lambda a_n z^n = \lambda \sum_{n=0}^{+\infty} a_n z^n$.

Donnons la définition du *produit de Cauchy* adaptée aux séries entières puis énonçons le théorème relatif au produit de Cauchy de deux séries entières.

Définition

On appelle *produit de Cauchy* de deux séries entières $\sum a_n z^n$ et $\sum b_n z^n$ la série entière $\sum c_n z^n$ où pour tout $n \in \mathbb{N}$,

$$c_n = \sum_{k=0}^n a_k b_{n-k}.$$

Théorème

Soient $\sum a_n z^n$ et $\sum b_n z^n$ deux séries entières de rayons de convergence R_a et R_b et R le rayon de convergence de la série entière produit de Cauchy $\sum c_n z^n$. Alors

1. $R \geq \min(R_a, R_b)$.
2. Pour tout $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$,

$$\sum_{n=0}^{+\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) z^n = \left(\sum_{n=0}^{+\infty} a_n z^n \right) \left(\sum_{n=0}^{+\infty} b_n z^n \right).$$

10.5 Dérivée et primitive d'une série entière

Définition

On appelle *série dérivée* de $\sum a_n z^n$ la série entière $\sum (n+1)a_{n+1} z^n$.

Théorème

Les rayons de convergence d'une série entière et de sa série dérivée sont égaux.

Définition

On appelle *série primitive* de $\sum a_n z^n$ la série entière $\sum \frac{a_{n-1}}{n} z^n$.

Théorème

Les rayons de convergence d'une série entière et de sa série primitive sont égaux.

10.6 Convergence normale

Théorème

Une série entière de rayon de convergence non nul converge normalement (donc uniformément) sur tout disque fermé de centre 0 inclus dans le disque ouvert de convergence¹⁴.

De plus sa fonction somme est continue sur le disque ouvert de convergence $\{z \in \mathbb{C}, |z| < R\}$.

A priori, une série entière ne converge pas normalement sur tout le disque ouvert de convergence : par exemple, le rayon de convergence de la série géométrique $\sum z^n$ est $R = 1$ et $\sup_{z \in \mathcal{D}(0, R)} |z^n| = 1$ où $\mathcal{D}(0, R) = \{z \in \mathbb{C}, |z| < R\}$.

10.7 Séries entières d'une variable réelle

On s'intéresse dans cette section aux séries entières d'une variable réelle x c'est-à-dire aux séries de fonctions¹⁵ $\sum f_n$ où pour tout $n \in \mathbb{N}$,

$$f_n : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto a_n x^n \end{cases}$$

avec (a_n) une suite réelle ou complexe.

Commençons par adapter le théorème page 163.

Théorème

Soit $\sum a_n x^n$ une série entière d'une variable réelle de rayon de convergence R .

- Si $x \in]-R, R[$ alors $\sum a_n x^n$ converge absolument.
- Si $|x| > R$ alors $\sum a_n x^n$ diverge grossièrement.

On en peut rien conclure *a priori* quant à la convergence de la série entière en R ou $-R$.

Étant donné une série entière d'une variable réelle $\sum a_n x^n$, rappelons que I , son intervalle de convergence, est l'ensemble des $x \in \mathbb{R}$ tels que la série numérique $\sum a_n x^n$ converge.

14. c'est-à-dire si R est le rayon de convergence non nul de la série entière $\sum a_n z^n$, cette dernière converge normalement sur tout disque fermé de centre 0 et de rayon $r < R$.

15. que l'on notera comme d'habitude abusivement $\sum a_n x^n$.

Ainsi

- si $R = 0$ alors $I = \{0\}$;
- si $R = +\infty$ alors $I = \mathbb{R}$;
- si $R \in]0, +\infty[$ alors $] -R, R[\subset I \subset [-R, R]$.

I est donc l'un des quatre intervalles $] -R, R[$, $] -R, R]$, $[-R, R[$ ou $[-R, R]$.

La fonction somme est alors la fonction $f : I \rightarrow \mathbb{C}$ définie pour tout $x \in I$ par

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Adaptons à présent le théorème de convergence normale page 168.

Théorème

Une série entière d'une variable réelle $\sum a_n x^n$, de rayon de convergence R non nul, converge normalement (donc uniformément) sur tout segment inclus dans l'intervalle ouvert de convergence $] -R, R[$.

De plus, sa fonction somme est continue sur $] -R, R[$.

Étudions à présent la continuité de la somme aux extrémités de l'intervalle $] -R, R[$. Soit $\sum a_n x^n$ une série entière de rayon de convergence R non nul. Alors si $|x| > R$, la série numérique $\sum a_n x^n$ diverge. Mais, comme nous l'avons vu précédemment, il se peut également que $\sum a_n R^n$ converge. Le *théorème radial d'Abel* qui suit affirme que dans ce cas, la fonction somme (qui est continue sur $] -R, R[$ via le théorème précédent) est également continue à gauche en R .

Théorème radial d'Abel

Soit $\sum a_n x^n$ une série entière d'une variable réelle de rayon de convergence R non nul telle que la série numérique $\sum a_n R^n$ converge.

Alors la fonction somme $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est continue à gauche en R c'est-à-dire ¹⁶

$$\lim_{x \rightarrow R^-} f(x) = f(R) = \sum_{n=0}^{+\infty} a_n R^n.$$

Énonçons à présent le théorème relatif à la primitive de la fonction somme s'annulant en 0.

16. Ce théorème est également vérifié en $-R$: si $\sum a_n (-1)^n R^n$ converge alors S est continue à droite en $-R$.

Théorème d'intégration terme à terme

Soit $\sum a_n x^n$ une série entière d'une variable réelle de rayon de convergence R non nul et de fonction somme f . Alors pour tout $x \in]-R, R[$,

$$\int_0^x \left(\sum_{n=0}^{+\infty} a_n t^n \right) dt = \sum_{n=0}^{+\infty} \int_0^x a_n t^n dt = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

Ainsi la fonction F définie pour tout $x \in]-R, R[$ par $F(x) = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$ est la primitive de la fonction somme f s'annulant en 0.

Énonçons ensuite le théorème de dérivation terme à terme de la fonction somme : on peut dériver terme à terme une série entière sur son intervalle ouvert de convergence.

Théorème de dérivation terme à terme

Soit $\sum a_n x^n$ une série entière d'une variable réelle de rayon de convergence R non nul. Alors sa fonction somme f est de classe $\mathcal{C}^\infty$ sur $]-R, R[$ et vérifie pour tout $k \in \mathbb{N}^*$ et tout $x \in]-R, R[$,

$$f^{(k)}(x) = \sum_{n=k}^{+\infty} n(n-1) \cdots (n-k+1) a_n x^{n-k}.$$

Le théorème suivant permet de donner l'expression des coefficients d'une série entière à l'aide des dérivées en 0 de sa somme.

Théorème

Soit $\sum a_n x^n$ une série entière d'une variable réelle de rayon de convergence R non nul et de somme f . Alors, pour tout $n \in \mathbb{N}$,

$$a_n = \frac{f^{(n)}(0)}{n!}.$$

On en déduit le corollaire d'unicité suivant.

Corollaire

Soient $\sum a_n x^n$ et $\sum b_n x^n$ deux séries entières d'une variable réelle dont les sommes sont égales sur un voisinage de 0.

Alors, pour tout $n \in \mathbb{N}$, $a_n = b_n$.

10.8 Fonctions développables en série entière

Dans cette section, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On notera, pour tout $a > 0$, $\mathcal{D}(0, a)$ le disque ouvert de centre 0 et de rayon a c'est-à-dire

$$\mathcal{D}(0, a) = \{z \in \mathbb{C}, |z| < a\}.$$

Précédemment ont été étudiées les propriétés de la fonction somme d'une série entière de rayon de convergence non nul. Nous proposons dans cette section d'étudier la réciproque : soient U un voisinage de 0 dans $\mathbb{R}$ et une fonction $f : U \rightarrow \mathbb{C}$. Existe-t-il une suite complexe (a_n) telle que pour tout $x \in U$,

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n ?$$

De même, dans le cas d'une variable complexe, si V est un voisinage de 0 dans $\mathbb{C}$ et $f : V \rightarrow \mathbb{C}$. Existe-t-il une suite complexe (a_n) telle que pour tout $z \in V$,

$$f(z) = \sum_{n=0}^{+\infty} a_n z^n ?$$

Définition

Soit $f : V \rightarrow \mathbb{C}$ où V est un voisinage de 0 dans $\mathbb{K}$.

On dit que f est *développable en série entière*¹⁷ en 0 si f coïncide sur un voisinage de 0 avec la somme d'une série entière.

Si V est un voisinage de 0 dans $\mathbb{C}$, $f : V \rightarrow \mathbb{C}$ développable en série entière en 0 signifie donc^{18, 19} qu'il existe $a > 0$ avec $\mathcal{D}(0, a) \subset V$ et (a_n) une suite complexe tels que

$$\forall z \in \mathcal{D}(0, a), f(z) = \sum_{n=0}^{+\infty} a_n z^n.$$

Si V est un voisinage de 0 dans $\mathbb{R}$, $f : V \rightarrow \mathbb{C}$ développable en série entière en 0 signifie donc^{20, 21} qu'il existe $a > 0$ avec $] -a, a[\subset V$ et (a_n) une suite complexe tels que

$$\forall x \in] -a, a[, f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

17. ou au voisinage de 0.

18. On peut aussi écrire qu'il existe $a > 0$ et (a_n) une suite complexe tels que pour tout $z \in \mathcal{D}(0, a) \cap V$, $f(z) = \sum_{n=0}^{+\infty} a_n z^n$.

19. On emploiera parfois dans ce cas l'expression « f est développable en série entière sur $\mathcal{D}(0, a)$ » étant sous-entendu qu'il s'agit d'un développement en 0.

20. On peut aussi écrire qu'il existe $a > 0$ et (a_n) une suite complexe tels que pour tout $x \in] -a, a[\cap V$, $f(x) = \sum_{n=0}^{+\infty} a_n x^n$.

21. On emploiera parfois dans ce cas l'expression « f est développable en série entière sur $] -a, a[$ » étant sous-entendu qu'il s'agit d'un développement en 0.

Dans les deux cas, on remarquera que le rayon de convergence de la série entière est alors nécessairement supérieur ou égal à a .

Théorème

La fonction²² $z \mapsto \frac{1}{1-z}$ est développable en série entière en 0 et, pour tout $z \in \mathcal{D}(0, 1)$,

$$\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n.$$

La somme et le produit de fonctions développables en série entière en 0 sont également développables en série entière en 0.

Les primitives et les dérivées successives d'une fonction développable en série entière en 0 le sont également.

Rappelons également la définition de la fonction exponentielle sur $\mathbb{C}$, définie page 118, qui apparaît comme un développement en série entière.

Définition

L'application $\exp : \begin{cases} \mathbb{C} & \longrightarrow \mathbb{C} \\ z & \longmapsto \sum_{n=0}^{+\infty} \frac{z^n}{n!} \end{cases}$ est appelée *fonction exponentielle*.

Pour tout $z \in \mathbb{C}$, l'exponentielle de z , notée e^z est définie par $e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$.

Le théorème suivant exhibe les propriétés de la fonction exponentielle complexe.

Théorème

1. Pour tout $(z_1, z_2) \in \mathbb{C}^2$, $e^{z_1+z_2} = e^{z_1}e^{z_2}$.
2. Pour tout $z \in \mathbb{C}$, $e^z \neq 0$, $\frac{1}{e^z} = e^{-z}$, $\overline{e^z} = e^{\bar{z}}$ et $|e^z| = e^{\operatorname{Re}(z)}$.
3. Pour tout $z \in \mathbb{C}$, $|e^z| = 1 \iff z \in i\mathbb{R}$.

10.9 Fonctions d'une variable réelle développables en série entière

On ne considère dans cette section que des fonctions d'une variable réelle.

²² Ne pas confondre le domaine de définition d'une fonction et le voisinage de 0 sur lequel elle est éventuellement somme d'une série entière : la fonction $z \mapsto \frac{1}{1-z}$ est définie pour tout $z \neq 1$ mais n'est développable en série entière en 0 que sur $\mathcal{D}(0, 1)$.

Définition

Soient V un voisinage de 0 dans $\mathbb{R}$ et $f : V \rightarrow \mathbb{C}$ de classe $\mathcal{C}^\infty$ sur V .
On appelle *série de Taylor* de f en 0 la série entière $\sum \frac{f^{(n)}(0)}{n!} x^n$.

Théorème

Soient V un voisinage de 0 dans $\mathbb{R}$ et $f : V \rightarrow \mathbb{C}$ développable²³ en série entière en 0. Alors

- f est de classe $\mathcal{C}^\infty$ au voisinage²⁴ de 0 ;
- la série entière dont f est la somme est unique : f est²⁵ la somme de sa série de Taylor en 0.

Insistons sur les deux résultats fondamentaux délivrés par ce théorème :

- le développement en série entière d'une fonction f de classe $\mathcal{C}^\infty$ au voisinage de 0 est unique ;
- si f est développable en série entière en 0 sur $] -a, a[$, sa série de Taylor en 0 converge et f coïncide avec la somme de sa série de Taylor sur $] -a, a[$.

Notons aussi qu'une fonction qui n'est pas de classe $\mathcal{C}^\infty$ au voisinage de 0, ne peut pas être développable en série entière en 0.

En revanche, il existe des fonctions de classe $\mathcal{C}^\infty$ qui ne sont pas développables en série entière : c'est le cas de la fonction

$$f : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} e^{-1/x^2} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases} \end{cases}$$

On peut montrer que f est de classe $\mathcal{C}^\infty$, que pour tout $n \in \mathbb{N}$, $f^{(n)}(0) = 0$ et donc que f n'est pas développable en série entière en 0.

Bien que souvent difficile à appliquer, énonçons pour information, un théorème donnant une condition nécessaire et suffisante pour qu'une fonction soit développable en série entière en 0. On adoptera dans ce théorème l'expression développée dans la note 21 page 171.

23. Donc il existe $a > 0$ avec $] -a, a[\subset V$ et (a_n) une suite complexe tels que pour tout $x \in] -a, a[$, $f(x) = \sum_{n=0}^{+\infty} a_n x^n$.

24. c'est-à-dire sur $] -a, a[$ en reprenant les notations de la note précédente.

25. c'est-à-dire, en reprenant les notations de la note 23 ci-dessus, pour tout $x \in] -a, a[$,

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

Théorème

Soient V un voisinage de 0 dans $\mathbb{R}$, $a > 0$ avec $] -a, a[\subset V$ et $f : V \rightarrow \mathbb{C}$. Alors f est développable en série entière en 0 sur $] -a, a[$ si, et seulement si

- f est de classe $\mathcal{C}^\infty$ sur $] -a, a[$;
- pour tout $x \in] -a, a[$, $\int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt \xrightarrow{n \rightarrow +\infty} 0$.

Énonçons à présent un théorème, plus facile d'emploi, mais n'exhibant qu'une condition suffisante pour qu'une fonction soit développable en série entière.

Théorème

Soient V un voisinage de 0 dans $\mathbb{R}$, $a > 0$ avec $] -a, a[\subset V$ et $f : V \rightarrow \mathbb{C}$ de classe $\mathcal{C}^\infty$ sur $] -a, a[$.

S'il existe $M \in \mathbb{R}_+^*$ tel que pour tout $n \in \mathbb{N}^*$ et pour tout $x \in] -a, a[$, $|f^{(n)}(x)| \leq M$, alors f est développable en série entière en 0 sur $] -a, a[$.

Présentons le lien entre développement en série entière et développement limité via le théorème suivant.

Théorème

Soient V un voisinage de 0 dans $\mathbb{R}$ et $f : V \rightarrow \mathbb{C}$ développable en série entière en 0 : il existe donc $a > 0$ avec $] -a, a[\subset V$ et (a_n) une suite complexe tels que pour tout $x \in] -a, a[$, $f(x) = \sum_{n=0}^{+\infty} a_n x^n$.

Alors le développement limité de f à l'ordre $p \in \mathbb{N}$ en 0 s'obtient en tronquant le développement en série entière de f en 0 à l'ordre p c'est-à-dire

$$f(x) = a_0 + a_1 x + \cdots + a_p x^p + o(x^p).$$

Théorème

Les fonctions sinus, cosinus, sinus hyperbolique et cosinus hyperbolique sont développables en série entière en 0 et, pour tout $x \in \mathbb{R}$, on a

$$\begin{aligned} \sin(x) &= \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} & \cos(x) &= \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n} \\ \operatorname{sh}(x) &= \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!} & \operatorname{ch}(x) &= \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!}. \end{aligned}$$

Théorème

1. La fonction $x \mapsto \ln(1+x)$ est développable en série entière en 0 et, pour tout $x \in]-1, 1]$, on a ²⁶

$$\ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n.$$

2. La fonction $x \mapsto \arctan(x)$ est développable en série entière en 0 et, pour tout $x \in [-1, 1]$, on a ²⁷

$$\arctan(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}.$$

3. Soit ²⁸ $\alpha \in \mathbb{R} \setminus \mathbb{N}$. Alors la fonction $x \mapsto (1+x)^\alpha$ est développable en série entière en 0 et pour tout $x \in]-1, 1[$, on a

$$(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n.$$

26. L'égalité est bien vérifiée pour $x = 1$ via le théorème de convergence radiale d'Abel.

27. L'égalité est bien vérifiée pour $x = 1$ et $x = -1$ via le théorème de convergence radiale d'Abel. On notera au passage que la fonction $\arctan$ est définie sur $\mathbb{R}$ mais que le développement en série entière en 0 n'est valable que sur $[-1, 1]$.

28. Si $\alpha \in \mathbb{N}$, pour tout $x \in \mathbb{R}$, on a $(1+x)^\alpha = 1 + \sum_{n=1}^{\alpha} \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n$.

11

Séries de Fourier

Pour toute fonction $f : I \subset \mathbb{R} \rightarrow \mathbb{C}$, continue par morceaux sur I , on notera en tout point de discontinuité éventuel a de f ,

$$f(a^+) = \lim_{x \rightarrow a^+} f(x) = \lim_{\substack{x \rightarrow a \\ x > a}} f(x) \quad \text{et} \quad f(a^-) = \lim_{x \rightarrow a^-} f(x) = \lim_{\substack{x \rightarrow a \\ x < a}} f(x).$$

Pour tout $T \in \mathbb{R}_+^*$, on notera

- $\mathcal{C}_{m,T}^0(\mathbb{R}, \mathbb{C})$ l'ensemble des fonctions de $\mathbb{R}$ vers $\mathbb{C}$, continues par morceaux et T -périodiques ;
- $\mathcal{C}_T^0(\mathbb{R}, \mathbb{C})$ l'ensemble des fonctions de $\mathbb{R}$ vers $\mathbb{C}$, continues et T -périodiques.

Après les séries entières, nous étudions dans ce chapitre d'autres séries de fonctions « particulières » : les *séries de Fourier*.

11.1 Séries trigonométriques

Définition

On appelle *série trigonométrique* toute série de fonctions $\sum f_n$ où pour tout $n \in \mathbb{N}$,

$$f_n : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto a_n \cos(n\omega x) + b_n \sin(n\omega x) \end{cases}$$

avec $\omega \in \mathbb{R}^*$, $(a_n) \in \mathbb{C}^{\mathbb{N}}$ et $(b_n) \in \mathbb{C}^{\mathbb{N}}$.

Si $\sum f_n$ est une série trigonométrique, alors que chaque f_n est $\frac{2\pi}{\omega}$ -périodique.

Si $\sum f_n$ est une série trigonométrique convergeant simplement vers $f = \sum_{n=0}^{+\infty} f_n$ sur $I \subset \mathbb{R}$, alors f est $\frac{2\pi}{\omega}$ -périodique.

Les séries de Fourier étudient en particulier la réciproque de la remarque ci-dessus : étant donné $T \in \mathbb{R}_+^*$ et une fonction $f : \mathbb{R} \rightarrow \mathbb{C}$, T -périodique, existe-t-il $\omega \in \mathbb{R}^*$, (a_n) et (b_n) deux suites complexes tels que pour tout $x \in \mathbb{R}$,

$$f(x) = \sum_{n=0}^{+\infty} (a_n \cos(n\omega x) + b_n \sin(n\omega x)) ?$$

11.2 Coefficients et séries de Fourier

Définition

Soient $T \in \mathbb{R}_+^*$ et $f \in \mathcal{C}_{m,T}^0(\mathbb{R}, \mathbb{C})$.

On appelle *coefficients de Fourier* de f les nombres complexes $a_n(f)$ et $b_n(f)$ définis pour tout $n \in \mathbb{N}$ par

$$\begin{cases} a_n(f) = \frac{2}{T} \int_0^T f(x) \cos(n\omega x) dx \\ b_n(f) = \frac{2}{T} \int_0^T f(x) \sin(n\omega x) dx \end{cases} \quad \text{où } \omega = \frac{2\pi}{T}.$$

On appelle *série de Fourier* de f la série de fonctions $\frac{a_0(f)}{2} + \sum_{n \geq 1} F_n(f)$ où pour tout $n \in \mathbb{N}^*$,

$$F_n(f) : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto a_n(f) \cos(n\omega x) + b_n(f) \sin(n\omega x) \end{cases}.$$

La série de Fourier de f sera souvent notée abusivement

$$\frac{a_0(f)}{2} + \sum_{n \geq 1} (a_n(f) \cos(n\omega x) + b_n(f) \sin(n\omega x)).$$

Certains définissent à part le coefficient $a_0(f)$ comme la valeur moyenne de f sur $[0, T]$, c'est-à-dire

$$a_0(f) = \frac{1}{T} \int_0^T f(x) dx.$$

La série de Fourier de f est alors définie, avec l'abus ci-dessus, par

$$a_0(f) + \sum_{n \geq 1} (a_n(f) \cos(n\omega x) + b_n(f) \sin(n\omega x)).$$

Nous nous restreindrons par la suite à des fonctions 2π -périodiques, et ce, sans perte de généralité : il suffit de remarquer que si $T \in \mathbb{R}_+^*$ et f est T -périodique, alors la fonction g définie pour tout $x \in \mathbb{R}$ par $g(x) = f(\frac{T}{2\pi}x)$ est 2π -périodique.

Dans le cas d'une fonction $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$, les coefficients de Fourier de f sont donc définis pour tout $n \in \mathbb{N}$ par

$$\begin{cases} a_n(f) = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \\ b_n(f) = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx \end{cases}$$

et la série de Fourier de f est la série de fonctions définie par

$$\frac{a_0(f)}{2} + \sum_{n \geq 1} (a_n(f) \cos(nx) + b_n(f) \sin(nx)).$$

Les définitions des coefficients de Fourier peuvent paraître à première vue « parachutées ». On peut en fait montrer que si une série trigonométrique

$$\frac{a_0}{2} + \sum (a_n \cos(nx) + b_n \sin(nx))$$

converge *uniformément* vers une fonction f sur $\mathbb{R}$ alors nécessairement, pour tout $n \in \mathbb{N}$,

$$\begin{cases} a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) \, dx \\ b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) \, dx \end{cases}$$

Remarquons que si $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$, via la 2π -périodicité des fonctions f , sinus et cosinus, on a pour tout $\alpha \in \mathbb{R}$,

$$\begin{cases} a_n(f) = \frac{1}{\pi} \int_{\alpha}^{\alpha+2\pi} f(x) \cos(nx) \, dx \\ b_n(f) = \frac{1}{\pi} \int_{\alpha}^{\alpha+2\pi} f(x) \sin(nx) \, dx \end{cases}$$

d'où le théorème suivant.

Théorème

Soit $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$. Alors $\begin{cases} f \text{ paire} \implies \forall n \in \mathbb{N}, b_n(f) = 0; \\ f \text{ impaire} \implies \forall n \in \mathbb{N}, a_n(f) = 0. \end{cases}$

Définissons à présent les *coefficients de Fourier exponentiels*, fondamentaux dans la théorie des séries de Fourier et en particulier dans les théorèmes de convergence.

Définition

Soit $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$. On appelle *coefficients de Fourier exponentiels* de f les nombres complexes notés $c_n(f)$ définis pour tout $n \in \mathbb{Z}$ par

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} \, dx.$$

On en déduit le théorème suivant.

Théorème

Soit $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$. Alors la série de Fourier de f s'écrit ^{1, 2}

$$\sum_{n \in \mathbb{Z}} c_n(f) e^{inx} \quad \text{c'est-à-dire} \quad c_0(f) + \sum_{n \geq 1} (c_n(f) e^{inx} + c_{-n}(f) e^{-inx}).$$

Dans le cas d'une fonction $f \in \mathcal{C}_{m,T}^0(\mathbb{R}, \mathbb{C})$ avec $T \in \mathbb{R}_+^*$, la série de Fourier de f s'écrit

$$\sum_{n \in \mathbb{Z}} c_n(f) e^{in\omega x} \quad \text{où, pour tout } n \in \mathbb{Z}, \quad c_n(f) = \frac{1}{T} \int_0^T f(x) e^{-in\omega x} dx.$$

11.3 Espace de Dirichlet

Définition

On appelle *espace de Dirichlet*³, noté $\mathcal{D}$, l'ensemble des fonctions $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$ telles que⁴ pour tout $x \in \mathbb{R}$,

$$f(x) = \frac{1}{2}(f(x^+) + f(x^-)).$$

Remarquons que $\mathcal{C}_{2\pi}^0(\mathbb{R}, \mathbb{C}) \subset \mathcal{D} \subset \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$.

Si $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$ n'est pas dans $\mathcal{D}$, il est facile de « transformer » f pour qu'elle soit dans $\mathcal{D}$ comme l'illustre la définition qui suit.

Définition

Soit $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$.

On appelle *régularisée* de f la fonction $\tilde{f} : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto \frac{1}{2}(f(x^+) + f(x^-)) \end{cases}$.

1. en utilisant l'abus habituel. En toute rigueur, on devrait écrire la série de Fourier sous la forme $\sum_{n \in \mathbb{Z}} c_n(f) e_n$ où, pour tout $n \in \mathbb{N}$, $e_n : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto e^{inx} \end{cases}$.

2. Si $(u_n)_{n \in \mathbb{N}}$ est une suite complexe, $\sum_{n \in \mathbb{Z}} u_n$ converge signifie que $\sum_{n \geq 1} (u_n + u_{-n})$ converge. Si $\sum_{n \in \mathbb{Z}} u_n$ converge, on a alors $\sum_{n=-\infty}^{+\infty} u_n = u_0 + \sum_{n=1}^{+\infty} (u_n + u_{-n})$.

3. Le patronyme complet de ce mathématicien est en réalité Lejeune Dirichlet, mais est souvent raccourci en Dirichlet.

4. Cette condition est parfois appelée *condition de Dirichlet*.

Si $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$, f et $\tilde{f}$ coïncident sur $[0, 2\pi]$ sauf en un nombre fini de points donc

$$\int_0^{2\pi} f(x) dx = \int_0^{2\pi} \tilde{f}(x) dx.$$

En particulier les coefficients de Fourier de f et $\tilde{f}$ sont égaux.

Il est possible de définir sur le $\mathbb{C}$ -espace vectoriel $\mathcal{D}$ un produit scalaire de sorte que $\mathcal{D}$ devient un espace préhilbertien complexe comme le montre le théorème suivant.

Théorème

L'application $\varphi : (f, g) \mapsto \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) dx$ est un produit scalaire sur $\mathcal{D}$.

Le produit scalaire sur $\mathcal{D}$ sera dorénavant noté $\langle \cdot | \cdot \rangle$ et la norme associée $\| \cdot \|_2$ de sorte que pour tout $(f, g) \in \mathcal{D}^2$,

$$\langle f | g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) dx \quad \text{et} \quad \|f\|_2^2 = \langle f | f \rangle = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx.$$

Les fonctions e_n définies dans la note 1 page 179 possèdent une propriété fondamentale dans l'espace préhilbertien $\mathcal{D}$ comme l'explique le théorème suivant.

Théorème

La famille⁵ $(e_n)_{n \in \mathbb{Z}}$ est orthonormée dans $(\mathcal{D}, \langle \cdot | \cdot \rangle)$.

11.4 Inégalité de Bessel et lemme de Riemann-Lebesgue

Théorème (inégalité de Bessel)

Soit $f \in \mathcal{D}$.

Alors⁶ pour tout $n \in \mathbb{N}$, $\sum_{k=-n}^n |c_k(f)|^2 \leq \|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx$.

5. Chaque e_n est continue sur $\mathbb{R}$ (donc dans $\mathcal{D}$) car les fonctions à valeurs réelles $x \mapsto \cos(nx)$ et $x \mapsto \sin(nx)$ sont continues sur $\mathbb{R}$.

6. Cette inégalité est encore vérifiée si $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$ car alors $\tilde{f} \in \mathcal{D}$.

Corollaire (lemme de Riemann-Lebesgue)

Soit $f \in \mathcal{D}$. Alors ⁷

1. $c_n(f) \xrightarrow{n \rightarrow \pm\infty} 0$ c'est-à-dire $\begin{cases} c_n(f) \xrightarrow{n \rightarrow +\infty} 0. \\ c_{-n}(f) \xrightarrow{n \rightarrow +\infty} 0. \end{cases}$
2. $a_n(f) \xrightarrow{n \rightarrow +\infty} 0$ et $b_n(f) \xrightarrow{n \rightarrow +\infty} 0$.

Théorème

Soit $f \in \mathcal{C}_{2\pi}^0(\mathbb{R}, \mathbb{C})$. Alors $(\forall n \in \mathbb{Z}, c_n(f) = 0) \implies f = 0$.

Ce théorème est faux dans le cas d'une fonction $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$ comme l'illustre le contre-exemple suivant.

Soit $f : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto \begin{cases} 1 & \text{si } x \in 2\pi\mathbb{Z} \\ 0 & \text{sinon} \end{cases} \end{cases}$. Alors $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$.

Mais $c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx = 0$ et pourtant f n'est pas nulle.

11.5 Théorèmes de convergence

Plusieurs questions importantes se posent, notamment les suivantes : la série de Fourier d'une fonction $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$ converge-t-elle simplement sur $\mathbb{R}$? Si c'est le cas, sa somme est-elle égale à f ?

Nous verrons dans cette section qu'une hypothèse plus forte est nécessaire afin d'obtenir une convergence simple sur $\mathbb{R}$.

Pour tout $f \in \mathcal{C}_{m,2\pi}^0(\mathbb{R}, \mathbb{C})$, on notera $(S_n(f))$ la suite de fonctions des sommes partielles associée à la série de Fourier de f c'est-à-dire pour tout $n \in \mathbb{N}$,

$$S_n(f) = \frac{a_0(f)}{2} + \sum_{k=1}^n F_k(f)$$

en convenant que $S_0(f) = \frac{a_0(f)}{2}$ et, avec pour tout $k \in \mathbb{N}^*$,

$$F_k(f) : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{C} \\ x & \longmapsto a_k(f) \cos(kx) + b_k(f) \sin(kx) \end{cases}.$$

Définition

On dit qu'une suite de fonctions $(f_n) \in \mathcal{D}^{\mathbb{N}}$ converge *en moyenne quadratique* vers $f \in \mathcal{D}$ si

$$\|f_n - f\|_2 \xrightarrow{n \rightarrow +\infty} 0.$$

7. Cf. note précédente.

Théorème

Soit $f \in \mathcal{D}$. Alors la série de Fourier de f converge en moyenne quadratique vers f c'est-à-dire

$$\|S_n(f) - f\|_2 \xrightarrow{n \rightarrow +\infty} 0.$$

Corollaire (égalité de Parseval)

Soit $f \in \mathcal{D}$. Alors ⁸ $\sum_{n=-\infty}^{+\infty} |c_n(f)|^2 = \|f\|_2^2$.

Si f est à valeurs réelles, on a $\frac{a_0^2(f)}{4} + \frac{1}{2} \sum_{n=1}^{+\infty} (a_n^2(f) + b_n^2(f)) = \|f\|_2^2$.

Définition

On dit que $f : [a, b] \rightarrow \mathbb{C}$ est de classe $\mathcal{C}^1$ par morceaux sur $[a, b]$ si f est de classe $\mathcal{C}^1$ sur $[a, b]$ privé d'un nombre fini (éventuellement nul) de points x_i en lesquels f' admet une limite finie à gauche et à droite c'est-à-dire s'il existe $n \in \mathbb{N}^*$ et $(x_0, \dots, x_n) \in [a, b]^{n+1}$ tels que :

- $x_0 = a < x_1 < \dots < x_n = b$;
- f de classe $\mathcal{C}^1$ sur chaque intervalle $]x_i, x_{i+1}[$ et f' admet une limite finie à gauche et à droite ^{9, 10, 11} en chaque x_i .

Si $f : \mathbb{R} \rightarrow \mathbb{C}$, on dit que f est de classe $\mathcal{C}^1$ par morceaux sur $\mathbb{R}$ si f est de classe $\mathcal{C}^1$ par morceaux sur tout segment $[a, b]$ de $\mathbb{R}$.

Donc, si $f : \mathbb{R} \rightarrow \mathbb{C}$ est 2π -périodique, il suffit bien entendu que f soit de classe $\mathcal{C}^1$ par morceaux sur $[0, 2\pi]$ pour que f soit de classe $\mathcal{C}^1$ par morceaux sur $\mathbb{R}$.

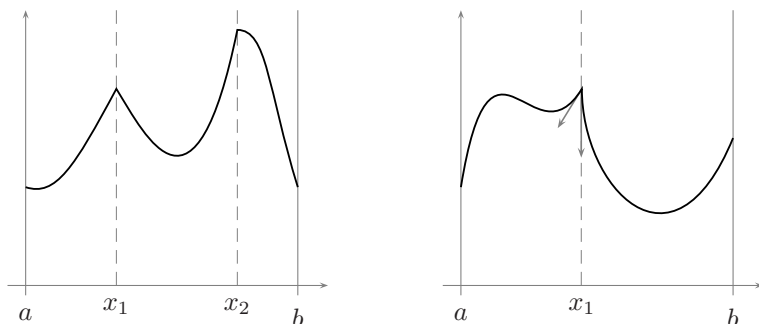
On notera $\mathcal{C}_{m, 2\pi}^1(\mathbb{R}, \mathbb{C})$ l'ensemble des fonctions de $\mathbb{R}$ vers $\mathbb{C}$ de classe $\mathcal{C}^1$ par morceaux sur $\mathbb{R}$.

8. Cf. note 6 page 180.

9. sauf en $x_0 = a$ et $x_n = b$ en lesquels f' n'admet respectivement qu'une limite à droite et à gauche.

10. On peut montrer que si f' admet une limite finie à gauche et à droite en chaque x_i alors f admet une limite finie à gauche et à droite en chaque x_i de sorte qu'une fonction de classe $\mathcal{C}^1$ par morceaux est nécessairement continue par morceaux.

11. On peut montrer que la deuxième condition de cette définition équivaut à dire que pour tout $i \in \llbracket 0, n-1 \rrbracket$, la restriction de f à $]x_i, x_{i+1}[$ se prolonge en une fonction de classe $\mathcal{C}^1$ sur $[x_i, x_{i+1}]$.



Une fonction de classe $\mathcal{C}^1$ par morceaux sur $[a, b]$... une qui ne l'est pas.

Théorème de Dirichlet

Soit $f \in \mathcal{C}_{m, 2\pi}^1(\mathbb{R}, \mathbb{C})$. Alors la série de Fourier de f converge simplement sur $\mathbb{R}$ vers la régularisée $\tilde{f}$ de f c'est-à-dire, pour tout $x \in \mathbb{R}$,

$$\sum_{n=-\infty}^{+\infty} c_n(f) e^{inx} = \tilde{f}(x)$$

ou encore¹², pour tout $x \in \mathbb{R}$,

$$\frac{a_0(f)}{2} + \sum_{n=1}^{+\infty} (a_n(f) \cos(nx) + b_n(f) \sin(nx)) = \tilde{f}(x).$$

Non seulement la continuité par morceaux d'une fonction f ne suffit pas à assurer la convergence simple de la série de Fourier de f mais la continuité non plus. On peut en effet exhiber des fonctions continues sur $\mathbb{R}$ et 2π -périodique dont la série de Fourier diverge en 0 c'est-à-dire telle que la suite $(S_n(f)(0))$ diverge : c'est le cas de la fonction f , 2π -périodique, définie pour tout $x \in [0, \pi]$ par

$$f(x) = \sum_{n=1}^{+\infty} \frac{1}{n^2} \sin\left(\left(2^{n^3} + 1\right) \frac{x}{2}\right)$$

et prolongée par parité¹³.

En revanche, on peut montrer que si f est continue sur $\mathbb{R}$ et 2π -périodique telle que sa série de Fourier converge simplement sur $\mathbb{R}$, alors cette dernière converge simplement vers f sur $\mathbb{R}$.

12. En tout point x sur lequel f est continue, $\frac{a_0(f)}{2} + \sum_{n=1}^{+\infty} (a_n(f) \cos(nx) + b_n(f) \sin(nx)) = f(x)$ car $\tilde{f}(x)$ coïncide alors avec $f(x)$.

13. c'est-à-dire, on pose $f(x) = f(-x)$ pour $x \in [-\pi, 0]$ de sorte que f est une fonction paire sur $\mathbb{R}$.

12

Fonctions vectorielles

Dans tout ce chapitre, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, les espaces vectoriels sont des $\mathbb{K}$ -espaces vectoriels non réduits au vecteur nul et tout intervalle de $\mathbb{R}$ est supposé non vide et non réduit à un point.

Le but de ce chapitre est d'étendre aux fonctions à valeurs dans un espace vectoriel normé de dimension finie¹ les résultats sur la dérivation et l'intégration connus pour les fonctions de $\mathbb{R}$ vers $\mathbb{K}$.

12.1 Limite, continuité et dérivabilité

Soient E un espace vectoriel normé de dimension finie $p \in \mathbb{N}^*$ et $\mathcal{B} = (e_1, \dots, e_p)$ une base de E . Si f est une application d'un intervalle I de $\mathbb{R}$ vers E , on note $f_1, \dots, f_p$ les applications-coordonnées de f relativement à $\mathcal{B}$. On a donc

$$f = \sum_{k=1}^p f_k e_k.$$

Comme les f_k sont des applications de I vers $\mathbb{K}$, beaucoup de résultats obtenus pour les fonctions de $\mathbb{R}$ dans $\mathbb{K}$ se généralisent aisément aux fonctions vectorielles c'est-à-dire aux applications de I vers E .

En somme, les extensions aux fonctions vectorielles se feront naturellement *composante par composante*.

Commençons par étendre les définitions de la limite et de la continuité aux fonctions vectorielles.

1. On rappelle que dans un espace vectoriel normé de dimension finie, toutes les normes sont équivalentes.

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie, I un intervalle de $\mathbb{R}$, f une application de I vers E et a adhérent à I .

On dit que f admet la² limite $\ell \in E$ en a si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall t \in I, |t - a| < \alpha \implies \|f(t) - \ell\| < \varepsilon.$$

On écrit alors $\lim_{x \rightarrow a} f(x) = \ell$ ou $\lim_a f = \ell$.

Définition

Soient $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie, I un intervalle de $\mathbb{R}$, f une application de I vers E et $a \in I$.

On dit que f est *continue* en a si $\lim_{t \rightarrow a} f(t) = f(a)$ c'est-à-dire si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall t \in I, |t - a| < \alpha \implies \|f(t) - f(a)\| < \varepsilon.$$

On dit que f est continue sur I si f est continue en tout point de I .

La caractérisation séquentielle de la limite s'étend également aisément :

$\lim_{t \rightarrow a} f(t) = \ell \in E$ si, et seulement si, pour toute suite (t_n) de I convergeant vers a , la suite $(f(t_n))$ converge vers ℓ .

En notant $f_1, \dots, f_p$ les applications-coordonnées de f relativement à une base $(e_1, \dots, e_p)$ de E , alors $\lim_{t \rightarrow a} f(t) = \ell = \sum_{k=1}^p \ell_k e_k \in E$ si, et seulement si, pour tout $k \in \llbracket 1, p \rrbracket$, $\lim_{t \rightarrow a} f_k(t) = \ell_k$.

Étendons à présent la notion de dérivabilité.

Définition

Soient E un espace vectoriel normé de dimension finie, I un intervalle de $\mathbb{R}$ et f une application de I vers E .

On dit que f est *dérivable* en $a \in I$ si $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$ existe dans E .

Si elle existe, elle est notée $f'(a)$ et appelée *dérivée* de f en a .

On dit que f est *dérivable sur* I si f est dérivable en tout point de I .

Si $n \in \mathbb{N}$, on dit que f est de *classe* $\mathcal{C}^n$ (respectivement $\mathcal{C}^\infty$) si $f^{(n)}$ existe et est continue sur I (respectivement si f est indéfiniment dérivable sur I).

2. L'unicité de la limite se démontre de la même façon que pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$.

Si f est dérivable sur I , sa *fonction dérivée* f' est alors définie par

$$f' : \begin{cases} I & \longrightarrow E \\ t & \longmapsto f'(t) \end{cases}.$$

De la même manière que pour les notions précédentes, si $f_1, \dots, f_p$ sont les applications-coordonnées de f relativement à une base $(e_1, \dots, e_p)$ de E , f dérivable en $a \in I$ si, et seulement si, chaque f_k est dérivable en a , et dans ce cas

$$f'(a) = \sum_{k=1}^p f'_k(a) e_k.$$

De même f de classe $\mathcal{C}^n$ sur I si, et seulement si, chaque f_k est de classe $\mathcal{C}^n$ sur I .

Une combinaison linéaire d'applications de classe $\mathcal{C}^n$ sur I est de classe $\mathcal{C}^n$ sur I .

Théorème

Soient E un espace vectoriel normé de dimension finie, I un intervalle de $\mathbb{R}$, $a \in I$ et f une application de I vers E .

Alors f dérivable en a si, et seulement si,

$$\exists \ell \in E, \forall t \in I, f(t) = f(a) + (t - a)\ell + o(t - a).$$

Si f est dérivable en a , le développement limité à l'ordre 1 de f en a est

$$f(t) = f(a) + (t - a)f'(a) + o(t - a).$$

Étendons à présent les notions de dérivabilité à gauche et à droite.

Définition

Soient E un espace vectoriel normé de dimension finie, I un intervalle de $\mathbb{R}$, $a \in I$ et f une application de I vers E .

On dit que f est *dérivable à droite* en a si $\lim_{\substack{h \rightarrow 0 \\ h > 0}} \frac{f(a+h) - f(a)}{h}$ existe dans E .

On dit que f est *dérivable à gauche* en a si $\lim_{\substack{h \rightarrow 0 \\ h < 0}} \frac{f(a+h) - f(a)}{h}$ existe dans E .

Si f est dérivable à droite en a , on note $f'_d(a) = \lim_{\substack{h \rightarrow 0 \\ h > 0}} \frac{f(a+h) - f(a)}{h}$.

Si f est dérivable à gauche en a , on note $f'_g(a) = \lim_{\substack{h \rightarrow 0 \\ h < 0}} \frac{f(a+h) - f(a)}{h}$.

Théorème

Soient E un espace vectoriel normé de dimension finie, I un intervalle de $\mathbb{R}$, $a \in I$ et f une application de I vers E . Alors f dérivable en a si, et seulement si, f dérivable à droite et à gauche en a et $f'_d(a) = f'_g(a)$.

De plus, dans ce cas, $f'(a) = f'_d(a) = f'_g(a)$.

12.2 Opérations sur les fonctions dérivables

Théorème

Soient E un espace vectoriel normé de dimension finie, I un intervalle de $\mathbb{R}$, f et g deux applications de I vers E dérivables sur I .

1. Si $\lambda \in \mathbb{K}$, $\lambda f + g$ est dérivable sur I et $(\lambda f + g)' = \lambda f' + g'$.
2. Soient J est un intervalle de $\mathbb{R}$, φ une application de J vers $\mathbb{R}$, dérivable sur J avec $\varphi(J) \subset I$. Alors $f \circ \varphi$ est dérivable sur J et

$$(f \circ \varphi)' = (f' \circ \varphi)\varphi'.$$

Toujours dans les propriétés des fonctions dérivables, introduisons une application linéaire dans le théorème suivant.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, I un intervalle de $\mathbb{R}$, f une application de I vers E et $L \in \mathcal{L}(E, F)$. Si f est dérivable sur I alors ³ $L \circ f$ est dérivable sur I et

$$(L \circ f)' = L \circ f'.$$

Introduisons à présent une application bilinéaire.

Théorème

Soient E , F et G trois espaces vectoriels normés de dimension finie, I un intervalle de $\mathbb{R}$, f une application de I vers F , g une application de I vers G et B une application bilinéaire de $F \times G$ vers E . Si f et g sont dérivables sur I alors ⁴ $B(f, g)$ est dérivable sur I et

$$(B(f, g))' = B(f', g) + B(f, g').$$

3. L'application $L \circ f$ est parfois notée $L(f)$.

4. L'écriture abusive $B(f, g)$ désigne l'application $t \mapsto B(f(t), g(t))$.

Illustrons ce théorème par trois exemples :

$$1. \text{ Soient } \lambda : \begin{cases} I & \longrightarrow \mathbb{K} \\ t & \longmapsto \lambda(t) \end{cases} \quad \text{et } X : \begin{cases} I & \longrightarrow \mathcal{M}_{n,1}(\mathbb{K}) \\ t & \longmapsto X(t) \end{cases}.$$

Si X et λ sont dérivables sur I alors λX est dérivable sur I et

$$(\lambda X)' = \lambda' X + \lambda X'.$$

$$2. \text{ Soient } A : \begin{cases} I & \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ t & \longmapsto A(t) \end{cases} \quad \text{et } X : \begin{cases} I & \longrightarrow \mathcal{M}_{n,1}(\mathbb{K}) \\ t & \longmapsto X(t) \end{cases}.$$

Si A et X sont dérivables sur I alors AX est dérivable sur I et

$$(AX)' = A'X + AX'.$$

3. Soient $(E, \langle \cdot | \cdot \rangle)$ un espace euclidien, u et v deux applications de I vers E dérivables sur I . Alors l'application $t \mapsto \langle u(t) | v(t) \rangle$ est dérivable sur I et

$$(\langle u | v \rangle)' = \langle u' | v \rangle + \langle u | v' \rangle.$$

Théorème

Soient E un espace vectoriel normé de dimension finie, $n \in \mathbb{N}^*$, $E_1, \dots, E_n$ des espaces vectoriels normés de dimension finie, $f_1, \dots, f_n$ des applications de I vers $E_1, \dots, E_n$ et M une application multilinéaire de $E_1 \times \dots \times E_n$ vers E .

Si les f_i sont dérivables sur I alors M est dérivable sur I et

$$(M(f_1, \dots, f_n))' = \sum_{i=1}^n M(f_1, \dots, f_{i-1}, f'_i, f_{i+1}, \dots, f_n).$$

Illustrons ce théorème avec le déterminant.

Soit $A : \begin{cases} I & \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ t & \longmapsto A(t) \end{cases}$. Pour tout $t \in I$, notons $C_1(t), \dots, C_n(t)$ les colonnes

de $A(t)$ et $D(t) = \det(A(t)) = \det(C_1(t), \dots, C_n(t))$.

Alors $D'(t) = \sum_{i=1}^n \det(C_1(t), \dots, C_{i-1}(t), C'_i(t), C_{i+1}(t), \dots, C_n(t))$.

12.3 Intégration sur un segment

Dans cette section, lorsqu'on considère le segment $[a, b]$ de $\mathbb{R}$, $a < b$.

Définition

Soient E un espace vectoriel normé de dimension finie, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E , f une application de $[a, b]$ vers E et $f_1, \dots, f_p$ ses applications-coordonnées relativement à la base $\mathcal{B}$.

On dit que f est *continue par morceaux* sur $[a, b]$ si les f_i le sont.

Théorème

Soient E un espace vectoriel normé de dimension finie, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E , f une application de $[a, b]$ vers E continue par morceaux sur $[a, b]$ et $f_1, \dots, f_p$ ses applications-coordonnées relativement à la base $\mathcal{B}$.

Alors $\sum_{k=1}^p \left(\int_a^b f_k(t) dt \right) e_k$, indépendante de la base de E choisie, est appelée *intégrale de f sur $[a, b]$* et notée $\int_a^b f(t) dt$.

Comme le calcul d'une intégrale d'une fonction continue sur $[a, b]$ à valeurs dans un espace vectoriel normé de dimension finie se ramène donc aux calculs des intégrales des fonctions coordonnées, la linéarité et la relation de Chasles sont naturellement étendues aux fonctions vectorielles.

Présentons à présent une propriété de linéarité de l'intégrale d'une fonction vectorielle.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, f une application de $[a, b]$ vers E , continue par morceaux sur $[a, b]$ et $L \in \mathcal{L}(E, F)$.

Alors l'application $L \circ f$ est continue par morceaux sur $[a, b]$ et

$$L \left(\int_a^b f(t) dt \right) = \int_a^b (L \circ f)(t) dt.$$

Illustrons ce théorème par un exemple.

Soient $A : \begin{cases} [a, b] & \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ t & \longmapsto A(t) = a_{ij}(t) \end{cases}$ continue par morceaux⁵ sur $[a, b]$ et l'endomorphisme $L : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ X & \longmapsto P^{-1}XP \end{cases}$ où $P \in \text{GL}_n(\mathbb{K})$.

Alors $P^{-1} \left(\int_a^b A(t) dt \right) P = \int_a^b P^{-1}A(t)P dt$.

5. c'est-à-dire que les fonctions $t \mapsto a_{ij}(t)$ sont continues par morceaux sur $[a, b]$.

12.4 Sommes de Riemann

Étendons la notion de sommes de Riemann aux fonctions vectorielles.

Définition

Soient $n \in \mathbb{N}^*$, E un espace vectoriel normé de dimension finie et f une application de $[a, b]$ vers E .

On appelle *somme de Riemann d'ordre n associée à f relativement à une subdivision régulière de $[a, b]$* l'une des deux sommes suivantes

$$\frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right) \quad \text{ou} \quad \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right).$$

Théorème

Soient E un espace vectoriel normé de dimension finie, f une application de $[a, b]$ vers E continue par morceaux sur $[a, b]$. Alors

$$\int_a^b f(t) dt = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right).$$

Ce théorème permet d'étendre l'*inégalité triangulaire* aux fonctions vectorielles.

Théorème (inégalité triangulaire)

Soient $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie et f une application de $[a, b]$ vers E continue par morceaux sur $[a, b]$. Alors

$$\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt.$$

12.5 Intégrale fonction de sa borne supérieure

Définition

Soient I un intervalle de $\mathbb{R}$, E un espace vectoriel normé de dimension finie, f et F deux applications de I vers E . On dit que F est une *primitive* de f sur I si F est dérivable sur I de dérivée égale à f .

En revenant à nouveau aux applications coordonnées, si F est une primitive de f sur I , toute primitive de f sur I est de la forme $F + u$ où $u \in E$.

Théorème

Soient I un intervalle de $\mathbb{R}$, $a \in I$, E un espace vectoriel normé de dimension finie, f une application de I vers E continue sur I .

Alors ⁶ $F : \begin{cases} I & \longrightarrow E \\ x & \longmapsto \int_a^x f(t) dt \end{cases}$ est une primitive ⁷ de f sur I et l'unique s'annulant en a .

Corollaire

Soient I un intervalle de $\mathbb{R}$, E un espace vectoriel normé de dimension finie, f une application de I vers E continue sur I et F une primitive de f sur I . Alors pour tout $(a, b) \in I^2$, $\int_a^b f(t) dt = F(b) - F(a)$.

Énonçons à présent l'extension de l'inégalité des accroissements finis aux fonctions vectorielles.

Théorème (inégalité des accroissements finis)

Soient I un intervalle de $\mathbb{R}$, $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie et f une application de I vers E de classe $\mathcal{C}^1$ sur I . S'il existe $K \in \mathbb{R}_+$ tels que pour tout $t \in I$, $\|f'(t)\| \leq K$, alors pour tout $(a, b) \in I^2$,

$$\|f(b) - f(a)\| \leq K|b - a|.$$

12.6 Formules de Taylor

Théorème (formule de Taylor avec reste intégral)

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, E un espace vectoriel normé de dimension finie et f une application de I vers E de classe $\mathcal{C}^{n+1}$ sur I .

Alors pour tout $(a, x) \in I^2$,

$$f(x) = \sum_{k=0}^n \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt.$$

6. Ce résultat porte parfois le nom de *théorème fondamental de l'analyse* ou *théorème fondamental de l'intégration*.

7. ce qui signifie que F est de classe $\mathcal{C}^1$ sur I avec $F' = f$.

Théorème (inégalité de Taylor-Lagrange)

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie et f une application de I vers E de classe $\mathcal{C}^{n+1}$ sur I . Alors, pour tout $(a, x) \in I^2$,

$$\left\| f(x) - \sum_{k=0}^n \frac{(x-a)^k}{k!} f^{(k)}(a) \right\| \leq \frac{|x-a|^{n+1}}{(n+1)!} \sup_{t \in [a, x]} \|f^{(n+1)}(t)\|.$$

Théorème (formule de Taylor-Young)

Soient $n \in \mathbb{N}$, I un intervalle de $\mathbb{R}$, E un espace vectoriel normé de dimension finie et f une application de I vers E de classe $\mathcal{C}^n$ sur I . Alors, pour tout $(a, x) \in I^2$,

$$f(x) = \sum_{k=0}^n \frac{(x-a)^k}{k!} f^{(k)}(a) + o((x-a)^n).$$

Cette égalité est appelée *développement limité* de f à l'ordre n en a .

12.7 Arcs paramétrés du plan

Définition

Soient I un intervalle de $\mathbb{R}$ et $k \in \mathbb{N} \cup \{+\infty\}$. On appelle *arc paramétré* du plan⁸ toute application $\gamma : \begin{cases} I \subset \mathbb{R} & \longrightarrow \mathbb{R}^2 \\ t & \longmapsto \gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \end{cases}$ où x et y sont des applications de I vers $\mathbb{R}$.

Pour tout $t \in I$, on note M_t le point de coordonnées $(x(t), y(t))$ dans un repère orthonormé du plan cartésien.

L'ensemble $\gamma(I) = \{M_t; t \in I\}$ est appelé *trajectoire* de l'arc⁹ et noté Γ .

L'arc paramétré γ est dit *de classe* $\mathcal{C}^k$ sur I si x et y sont de classe $\mathcal{C}^k$ sur I .

Un arc paramétré du plan est donc un cas particulier d'une fonction vectorielle. Il décrit le mouvement d'un mobile ponctuel M dans le plan. La variable de départ t représente alors le temps et l'ensemble d'arrivée est dans le plan de sa trajectoire. Pour tout $t \in I$, M_t donne la position du mobile à l'instant t . La trajectoire d'un arc γ modélise donc l'ensemble des positions successives $\gamma(t) = M_t$ prises par le mobile au cours du temps.

8. ou *courbe paramétrée*.

9. ou *support* de l'arc qui est donc une courbe du plan.

Lorsque les vecteurs $\gamma'(t)$ et $\gamma''(t)$ existent en un instant donné t , ils donnent respectivement la vitesse et l'accélération du point M à l'instant t .

Définition

Soient γ un arc de classe $\mathcal{C}^1$, Γ son support et $M_{t_0} \in \Gamma$.

On dit que M_{t_0} est un point *régulier* si $\gamma'(t_0) \neq (0, 0)$.

On dit que M_{t_0} est un point *stationnaire*¹⁰ si $\gamma'(t_0) = (0, 0)$.

Soient $k \in \mathbb{N}^*$, γ un arc paramétré du plan de classe $\mathcal{C}^k$, Γ son support et $M_{t_0} \in \Gamma$.

On suppose qu'il existe $i \in \llbracket 1, k \rrbracket$ tel que $\gamma^{(i)}(t_0)$ soit non nul¹¹.

Γ admet une tangente en M_{t_0} si la droite $(M_{t_0}M_t)$ admet une position limite quand t tend vers t_0 .

Théorème

La tangente à Γ en M_{t_0} est la droite passant par M_{t_0} dirigée par $\gamma^{(p)}(t_0)$
où
$$p = \min\{i \in \llbracket 1, k \rrbracket \text{ tel que } \gamma^{(i)}(t_0) \neq (0, 0)\}.$$

Si M_{t_0} est un point régulier, $p = 1$ donc la tangente à Γ en M_{t_0} est dirigée par $\gamma'(t_0)$.

Intéressons-nous à présent à l'étude locale d'un arc paramétré.

Soient $k \in \mathbb{N}^*$, γ un arc paramétré de classe $\mathcal{C}^k$, Γ son support et $M_{t_0} \in \Gamma$.

On suppose qu'il existe $i \in \llbracket 1, k \rrbracket$ tel que $\gamma^{(i)}(t_0)$ soit non nul.

On note comme précédemment $p = \min\{i \in \llbracket 1, k \rrbracket \text{ tel que } \gamma^{(i)}(t_0) \neq (0, 0)\}$ puis notons q l'entier, défini, s'il existe, par

$$q = \min\{i \in \llbracket p, k \rrbracket \text{ tel que } (\gamma^{(p)}(t_0), \gamma^{(i)}(t_0)) \text{ soit libre}\}.$$

$(\gamma^{(p)}(t_0), \gamma^{(q)}(t_0))$, famille libre de deux vecteurs de $\mathbb{R}^2$ de dimension 2, est une base de $\mathbb{R}^2$.

Le nouveau repère $(M_{t_0}, \gamma^{(p)}(t_0), \gamma^{(q)}(t_0))$ va nous permettre de déterminer le comportement de la courbe au voisinage de M_{t_0} .

10. ou *singulier*.

11. Même un arc de classe $\mathcal{C}^\infty$ ne vérifie pas toujours cette condition. Choisissons par exemple

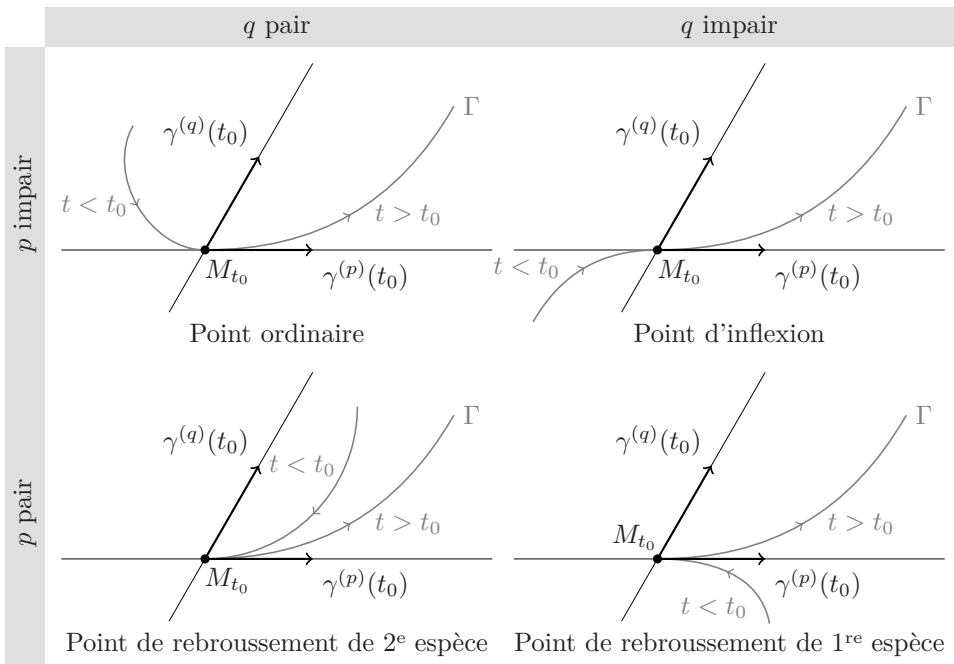
$$x : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ t & \longmapsto \begin{cases} e^{-1/t^2} & \text{si } t \neq 0 \\ 0 & \text{si } t = 0 \end{cases} \end{cases} \quad \text{et } \gamma : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R}^2 \\ t & \longmapsto \gamma(t) = \begin{pmatrix} x(t) \\ 0 \end{pmatrix} \end{cases}, \text{ alors pour tout } n \in \mathbb{N},$$

 $\gamma^{(n)}(0) = (0, 0).$

Définition

Soient $k \in \mathbb{N}^*$, γ un arc paramétré de classe $\mathcal{C}^k$, Γ son support et $M_{t_0} \in \Gamma$. On suppose que les entiers p et q définis précédemment existent.

- Si p est impair et q pair, on dit que M_{t_0} est un *point ordinaire*.
- Si p et q sont impairs, on dit que M_{t_0} est un *point d'inflexion*.
- Si p est pair et q impair, on dit que M_{t_0} est un *point de rebroussement de première espèce*.
- Si p et q sont pairs, on dit que M_{t_0} est un *point de rebroussement de deuxième espèce*.



Un point de rebroussement (de première ou de deuxième espèce) M_{t_0} est nécessairement un point stationnaire car p pair implique $\gamma'(t_0) = (0, 0)$.

Ainsi, un point régulier est soit un point ordinaire, soit un point d'inflexion.

Pour qu'un point régulier M_{t_0} soit un point d'inflexion, il faut nécessairement que $\gamma'(t_0)$ et $\gamma''(t_0)$ soient colinéaires c'est-à-dire $x'(t_0)y''(t_0) = x''(t_0)y'(t_0)$.

Définition

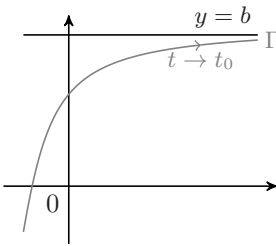
Soient $\gamma : \begin{cases} I \subset \mathbb{R} & \longrightarrow \mathbb{R}^2 \\ t & \longmapsto \gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \end{cases}$ un arc paramétré du plan, Γ sa

trajectoire et t_0 l'une des extrémités de I ($+\infty$ ou $-\infty$ le cas échéant).

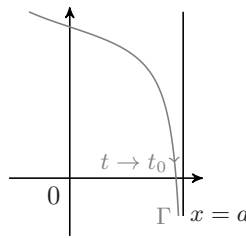
On dit que Γ admet une *branche infinie* en t_0 si ¹² : $\lim_{t \rightarrow t_0} \|\gamma(t)\| = +\infty$.

Si $\lim_{t \rightarrow t_0} x(t) = \pm\infty$ ou si $\lim_{t \rightarrow t_0} y(t) = \pm\infty$, alors $\lim_{t \rightarrow t_0} \sqrt{x^2(t) + y^2(t)} = +\infty$ donc Γ admet une branche infinie en t_0 .

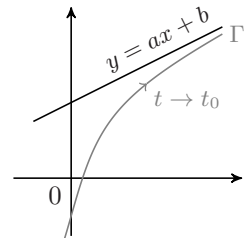
La réciproque est fausse. Pour s'en convaincre, il suffit par exemple de choisir $x(t) = t \cos(t)$ et $y(t) = t \sin(t)$ et $t_0 = +\infty$: $x(t)$ et $y(t)$ n'ont pas de limite quand t tend vers $+\infty$ et pourtant $\sqrt{x^2(t) + y^2(t)} = |t| \xrightarrow{t \rightarrow +\infty} +\infty$.



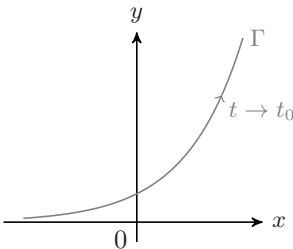
La droite $y = b$ est asymptote à Γ lorsque $t \rightarrow t_0$.



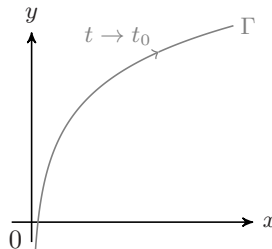
La droite $x = a$ est asymptote à Γ lorsque $t \rightarrow t_0$.



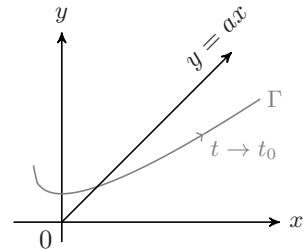
La droite $y = ax + b$ est asymptote à Γ lorsque $t \rightarrow t_0$.



Branche parabolique de direction asymptotique Oy



Branche parabolique de direction asymptotique Ox



Branche parabolique de direction asymptotique $y = ax$

12. On peut choisir par exemple la norme euclidienne : $\|\gamma(t)\| = \sqrt{x^2(t) + y^2(t)}$, la définition restant valable pour une autre norme.

Définition

Soient $\gamma : \begin{cases} I \subset \mathbb{R} & \longrightarrow \mathbb{R}^2 \\ t & \longmapsto \gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \end{cases}$ un arc paramétré du plan et Γ

sa trajectoire admettant une branche infinie en t_0 .

- Si $\lim_{t \rightarrow t_0} x(t) = a \in \mathbb{R}$ et $\lim_{t \rightarrow t_0} y(t) = \pm\infty$, Γ admet une *asymptote verticale* d'équation $x = a$.
- Si $\lim_{t \rightarrow t_0} x(t) = \pm\infty$ et $\lim_{t \rightarrow t_0} y(t) = b \in \mathbb{R}$, Γ admet une *asymptote horizontale* d'équation $y = b$.
- Si $\lim_{t \rightarrow t_0} x(t) = \pm\infty$ et $\lim_{t \rightarrow t_0} y(t) = \pm\infty$ alors
 - Si $\lim_{t \rightarrow t_0} \frac{y(t)}{x(t)} = 0$, Γ admet une *branche parabolique* de direction asymptotique l'axe des abscisses.
 - Si $\lim_{t \rightarrow t_0} \frac{y(t)}{x(t)} = \pm\infty$, Γ admet une *branche parabolique* de direction asymptotique l'axe des ordonnées.
 - Si $\lim_{t \rightarrow t_0} \frac{y(t)}{x(t)} = c \in \mathbb{R}^*$, alors
 - Si $\lim_{t \rightarrow t_0} y(t) - cx(t) = \pm\infty$, alors Γ admet une *branche parabolique* de direction asymptotique la droite d'équation $y = cx$.
 - Si $\lim_{t \rightarrow t_0} y(t) - cx(t) = d$, alors Γ admet une *asymptote* (oblique) d'équation $y = cx + d$.

Dans tout ce chapitre, les espaces vectoriels sont des $\mathbb{R}$ -espaces vectoriels non réduits au vecteur nul et tout ouvert est supposé non vide.

13.1 Application différentiable et différentielle

Les notions de différentiabilité et de différentielle pour des applications dans des espaces vectoriels normés généralisent celles de dérivabilité et de dérivée pour des applications de $\mathbb{R}$ vers $\mathbb{R}$.

Rappelons la définition de la dérivabilité de f en un point $a \in I$ pour une application f d'un intervalle I de $\mathbb{R}$ vers $\mathbb{R}$:

f est dérivable en a si $\frac{f(a+h)-f(a)}{h}$ tend vers une limite finie $\ell \in \mathbb{R}$ lorsque h tend vers 0.

Cette définition est équivalente à l'existence d'un réel ℓ et d'une fonction ε de I vers $\mathbb{R}$ tels que pour tout $h \in \mathbb{R}$ (vérifiant $a + h \in I$), on ait

$$f(a + h) = f(a) + h\ell + h\varepsilon(h) \quad \text{où } \varepsilon(h) \xrightarrow{h \rightarrow 0} 0$$

ou encore

$$f(a + h) = f(a) + h\ell + o(h).$$

Nous étendons donc dans ce chapitre les notions de dérivabilité d'une fonction d'une variable réelle aux applications d'un espace vectoriel normé E de dimension finie vers un espace vectoriel normé F de dimension finie.

Bien qu'il y ait un isomorphisme isométrique^{1, 2} entre un espace vectoriel normé E de dimension finie $n \geq 1$ et $\mathbb{R}^n$ les notions abordées sont adaptées à des espaces de natures bien différentes d'un espace $\mathbb{R}^k$ comme $\mathcal{L}(E, F)$ ou $\mathcal{M}_n(\mathbb{R})$.

Il n'est cependant pas inutile, afin de mieux assimiler ce cours, d'appliquer d'abord les définitions à $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ et même d'abord $E = \mathbb{R}^n$ (voire $\mathbb{R}^2$) et $F = \mathbb{R}$ puis $E = \mathbb{R}$ et $F = \mathbb{R}^n$.

1. c'est-à-dire qu'il existe une application linéaire bijective f de $\mathbb{R}^n$ vers E telle que pour tout $x \in \mathbb{R}^n$, $\|f(x)\| = \|x\|$.

2. Car si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E , l'application $f : \mathbb{R}^n \longrightarrow E$ définie pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ par $f(x) = \sum_{k=1}^n x_k e_k$ est linéaire, bijective et isométrique.

Étendons tout d'abord une notion de négligeabilité aux fonctions définies et à valeurs dans des espaces vectoriels normés.

Définition

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés³, $A \subset E$, f une application de A vers F et N une application de A vers $\mathbb{R}_+$.

On dit que le vecteur $f(x)$ est *négligeable* devant le réel $N(x)$ au voisinage du vecteur nul⁴, et on écrit⁵ $f(x) = o(N(x))$ si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in A, \|x\|_E < \alpha \implies \|f(x)\|_F \leq \varepsilon N(x).$$

Généralisons à présent la notion de dérivabilité.

Définition

Soient E et F deux espaces vectoriels normés de dimension finie⁶, U un ouvert de E , $a \in U$ et f une application de U vers F . On dit que f est *différentiable* en a s'il existe une application linéaire⁷ $L \in \mathcal{L}(E, F)$ telle que, pour tout^{8, 9} $h \in E$,

$$f(a+h) = f(a) + L(h) + o(\|h\|). \quad (*)$$

f différentiable en a signifie donc que l'on peut « approcher », au voisinage de a , l'« accroissement » $h \mapsto f(a+h) - f(a)$ par une application linéaire L .

Plus précisément le reste $f(a+h) - f(a) - L(h)$ est négligeable devant $\|h\|$ lorsque h tend vers le vecteur nul c'est-à-dire

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0$$

ce qui peut encore s'écrire sous la forme

3. Tous les espaces vectoriels normés de ce chapitre sont des espaces vectoriels sur $\mathbb{R}$.
4. Sans précision, la négligeabilité sera toujours au voisinage de 0.
5. Se lit « $f(x)$ est un petit o de $N(x)$ ».
6. Tous les espaces vectoriels normés de ce chapitre sont dorénavant de dimension finie. On rappelle que toutes les normes d'un espace de dimension finie sont équivalentes. Afin de ne pas alourdir les notations, les normes sur E et F seront notées, non pas $\|\cdot\|_E$ et $\|\cdot\|_F$, mais simplement $\|\cdot\|$.
7. La définition est encore valable pour des espaces de dimension infinie en remplaçant « linéaire » par « linéaire continue ». Rappelons que toute application linéaire de E vers F est continue lorsque E est de dimension finie.
8. Afin que $f(a+h)$ soit définie, il est sous-entendu (comme dans le reste du chapitre) que $a+h$ doit appartenir à U ce qui est possible si $\|h\|$ est suffisamment petite. En effet, comme U est un ouvert de E , U est un voisinage de a donc il existe $r \in \mathbb{R}_+^*$ tel que $B(a, r) \subset U$. Donc si $\|h\| < r$, $a+h \in B(a, r) \subset U$.
9. On peut aussi écrire, pour tout h voisin du vecteur nul de E c'est-à-dire appartenant à un voisinage du vecteur nul de E .

$$\forall \varepsilon > 0, \exists \alpha > 0, \|h\| < \alpha \implies \|f(a+h) - f(a) - L(h)\| \leq \varepsilon \|h\|.$$

L'expression (*) de la définition précédente peut également s'écrire

$$f(a+h) = f(a) + L(h) + \|h\|\varepsilon(h)$$

où le vecteur $\varepsilon(h) \in F$ tend vers le vecteur nul lorsque le vecteur $h \in E$ tend vers le vecteur nul c'est-à-dire $\lim_{h \rightarrow 0} \|\varepsilon(h)\| = 0$.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $a \in U$ et f une application de U vers F .

S'il existe $L \in \mathcal{L}(E, F)$ telle que pour tout $h \in E$,

$$f(a+h) = f(a) + L(h) + o(\|h\|)$$

alors L est unique.

L est appelée *différentielle de f en a* ¹⁰ et est notée¹¹ $df(a)$.

L'image de tout $h \in E$ par $df(a)$ sera notée $df(a)(h)$.

Définition

Soient E et F deux espaces vectoriels normés¹² de dimension finie, U un ouvert de E et f une application de U vers F . On dit que f est *différentiable* sur U si f est différentiable en tout point de U .

On appelle alors *différentielle* de f l'application notée df définie par

$$df : \begin{cases} U & \longrightarrow \mathcal{L}(E, F) \\ x & \longmapsto df(x) \end{cases}.$$

Soient $(E, \langle \cdot | \cdot \rangle)$ un espace euclidien et $f : \begin{cases} E & \longrightarrow \mathbb{R} \\ x & \longmapsto \langle x | x \rangle = \|x\|^2 \end{cases}$

Alors pour tout $x \in E$ et tout $h \in E$, on a $f(x+h) = \|x\|^2 + 2\langle x | h \rangle + \|h\|^2$.

Comme $\|h\|^2$ est un $o(\|h\|)$, on a directement $df(x)(h) = 2\langle x | h \rangle$.

10. ou *application linéaire tangente à f en a* .

11. La notation pour la différentielle de f en a n'est malheureusement pas figée : certains la notent df_a ou $d_a f$ ou $Df(a)$ ou Df_a ou f'_a voire même $f'(a)$!

12. On pourrait supposer simplement que E et F sont deux espaces vectoriels de dimension finie car la notion de différentielle ne dépend pas des normes choisies sur E et F . On rappelle d'autre part que toutes les normes d'un espace vectoriel de dimension finie sont équivalentes.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E et f une application de U vers F .

1. Si f est constante, alors pour tout $a \in U$, $df(a)$ est l'application linéaire nulle.
2. Si $E = \mathbb{R}$, I est un intervalle ouvert de E et $a \in I$, on a

$$f \text{ différentiable en } a \iff f \text{ dérivable en } a$$

et pour tout $a \in I$, $f'(a) = df(a)(1)$.

3. Si f est linéaire, alors f est différentiable sur U et, pour tout $a \in U$, $df(a) = f$.

Énonçons dans les trois théorèmes qui suivent quelques propriétés des applications différentiables.

Théorème

Soient E , F deux espaces vectoriels normés de dimension finie et U un ouvert de E . Alors ¹³ toute application de U vers F différentiable en $a \in U$ est continue en a .

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $\lambda \in \mathbb{R}$, f et g deux applications de U vers F .

Si f et g sont différentiables sur U , $\lambda f + g$ est différentiable sur U et pour tout $a \in U$, $d(\lambda f + g)(a) = \lambda df(a) + dg(a)$.

Théorème (règle de la chaîne)

Soient E , F et G trois espaces vectoriels normés de dimension finie, U un ouvert de E , $a \in U$, V un ouvert de F , f une application de U vers F telle que $f(U) \subset V$ et g une application de V vers G . Si f est différentiable en a et g différentiable en $f(a)$ alors $g \circ f$ est différentiable en a et

$$d(g \circ f)(a) = dg(f(a)) \circ df(a).$$

13. Donc toute application de U vers F différentiable sur U est continue sur U .

Corollaire (dérivée le long d'un arc)

Soient E et F , deux espaces vectoriels normés de dimension finie, I un intervalle de $\mathbb{R}$, $a \in I$, V un ouvert de E , γ une application¹⁴ de I vers E dérivable en a tel que $\gamma(I) \subset V$ et f une application de V vers F différentiable en $\gamma(a)$. Alors $f \circ \gamma$ est dérivable en a et

$$(f \circ \gamma)'(a) = df(\gamma(a))(\gamma'(a)).$$

L'application $(f \circ \gamma)'$ est appelée *dérivée de f le long de l'arc γ* .

Énonçons le théorème relatif à la différentielle d'une application multilinéaire.

Théorème

Soient $n \in \mathbb{N}^*$, $E_1, \dots, E_n$ et F des espaces vectoriels normés de dimensions finie. Toute application multilinéaire $f : E_1 \times \dots \times E_n \rightarrow F$ est différentiable sur $E_1 \times \dots \times E_n$ et pour tout $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ et tout $(h_1, \dots, h_n) \in E_1 \times \dots \times E_n$, on a¹⁵

$$df(x_1, \dots, x_n)(h_1, \dots, h_n) = \sum_{k=1}^n f(x_1, \dots, x_{k-1}, h_k, x_{k+1}, \dots, x_n).$$

13.2 Dérivées partielles

Définition

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $a \in U$, v un vecteur non nul de E et f une application de U vers F . On dit que f admet une¹⁶ *dérivée en a* (ou est *dérivable en a*) *selon le vecteur v* si l'application $g : t \mapsto f(a + tv)$, définie¹⁷ sur un voisinage de 0, est dérivable en 0 c'est-à-dire si $\lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t}$ existe.

Cette dérivée en a selon le vecteur v se note alors $D_v f(a)$ ou $\frac{\partial f}{\partial v}(a)$.

Si f admet une dérivée selon le vecteur v en tout point $x \in U$, on appelle

dérivée de f selon v l'application notée $D_v f : \begin{cases} U & \longrightarrow F \\ x & \longmapsto D_v f(x) \end{cases}$.

14. Toute application d'un intervalle I de $\mathbb{R}$ vers un espace vectoriel E de dimension finie est appelée *arc* (tracé ou inscrit dans E).

15. via la convention suivante : le premier terme de la somme est $f(h_1, x_2, \dots, x_n)$ et le dernier est $f(x_1, \dots, x_{n-1}, h_n)$.

16. On dit aussi *dérivée en a suivant v* ou *dérivée en a dans la direction de v* ou encore *dérivée directionnelle selon v en a* .

$$\text{Soit } f : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x, y) & \longmapsto \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases} \end{cases}.$$

On peut montrer que f admet une dérivée en $(0, 0)$ selon tout vecteur non nul (a, b) définie par $D_{(a,b)}f(0, 0) = \begin{cases} \frac{a^2}{b} & \text{si } b \neq 0 \\ 0 & \text{sinon} \end{cases}$.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $a \in U$ et f une application de U vers F .

Si f est différentiable en a , alors, pour tout vecteur v non nul de E , f admet une dérivée en a selon v et on a $D_v f(a) = df(a)(v)$.

La réciproque de ce théorème est fautive : en reprenant l'exemple précédent, f admet une dérivée en $(0, 0)$ selon tout vecteur non nul (a, b) mais f n'est pas continue en $(0, 0)$ donc non différentiable en $(0, 0)$.

Définition

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , $a \in U$ et f une application de U vers F . On dit que f admet une *dérivée partielle en a par rapport à la i -ème place*¹⁸ si f admet une dérivée en a selon le vecteur e_i .

Elle se note alors $D_i f(a)$ ou $\frac{\partial f}{\partial x_i}(a)$.

Si f admet une dérivée partielle par rapport à la i -ème place en tout point $a \in U$, on appelle *dérivée partielle de f par rapport à la i -ème place* l'application notée $D_i f$ ou $\frac{\partial f}{\partial x_i} : \begin{cases} U & \longrightarrow F \\ a & \longmapsto \frac{\partial f}{\partial x_i}(a) \end{cases}$.

Si f admet pour tout $i \in \llbracket 1, n \rrbracket$ une dérivée partielle par rapport à la i -ème place en tout point $a \in U$, on appelle *dérivées partielles de f* les n dérivées partielles¹⁹ $\frac{\partial f}{\partial x_i}$ pour $i \in \llbracket 1, n \rrbracket$.

17. Comme $a \in U$ et U est un ouvert de E , il existe $r \in \mathbb{R}_+^*$ tel que $B(a, r) \subset U$. Il existe $\delta \in \mathbb{R}_+^*$ tel que $\delta \|v\| < r$. Ainsi pour tout $t \in]-\delta, \delta[$, $a + tv \in B(a, r) \subset U$ et g est ainsi définie sur $]-\delta, \delta[$.

18. On dit aussi *i -ème dérivée partielle en a* ou encore *dérivée partielle en a par rapport à la i -ème variable*.

19. Lorsque $E = \mathbb{R}^2$, $\frac{\partial f}{\partial x_1}$ et $\frac{\partial f}{\partial x_2}$ seront généralement notées $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$. Elles seront appelées dérivées partielles de f par rapport à x et y .

Via cette définition f admet une dérivée partielle en $a = (a_1, \dots, a_n) \in U$ par rapport à la i -ème place si $\lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t}$ existe.

En particulier, si $E = \mathbb{R}^n$ et $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$, alors les assertions suivantes sont équivalentes :

(i) f admet une dérivée partielle en $a = (a_1, \dots, a_n) \in U$ par rapport à la i -ème place.

(ii) $\lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{t}$ existe.

(iii) L'application d'une variable réelle : $\begin{cases} \mathbb{R} & \longrightarrow F \\ t & \longmapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n) \end{cases}$ est dérivable en a_i .

Le théorème suivant montre le lien entre différentielle et dérivées partielles.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , $a \in U$ et f une application de U vers F .

Si f est différentiable en a alors les n dérivées partielles de f en a existent et pour tout $h = \sum_{k=1}^n h_k e_k \in E$, on a

$$df(a)(h) = \sum_{k=1}^n h_k \frac{\partial f}{\partial x_k}(a).$$

La réciproque de ce théorème est fautive : en reprenant l'exemple page 201, les dérivées partielles de f en $(0, 0)$ existent

$$\frac{\partial f}{\partial x}(0, 0) = D_{(1,0)}f(0, 0) = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(0, 0) = D_{(0,1)}f(0, 0) = 0.$$

Or, on a vu précédemment que f n'est pas différentiable en $(0, 0)$ car non continue en $(0, 0)$.

Notons que, contrairement aux fonctions d'une variable réelle où la dérivabilité implique la continuité, les dérivées partielles de f en $(0, 0)$ existent et pourtant f n'est pas continue en $(0, 0)$.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , $\mathcal{B}' = (e'_1, \dots, e'_p)$ une base de F , $f = \sum_{k=1}^p f_k e'_k$ une application²⁰ de U vers F où pour tout $k \in \llbracket 1, p \rrbracket$, les applications f_k de U vers $\mathbb{R}$ sont les applications-coordonnées de f relativement à $\mathcal{B}'$. Alors f différentiable en a si, et seulement si, chaque f_k est différentiable en a .

On a alors²¹ pour tout $h \in E$, $df(a)(h) = \sum_{k=1}^p df_k(a)(h) e'_k$.

En particulier, on a pour tout $i \in \llbracket 1, n \rrbracket$, $df(a)(e_i) = \sum_{k=1}^p \frac{\partial f_k}{\partial x_i}(a) e'_k$.

13.3 Matrice jacobienne et jacobien

Définition

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , $\mathcal{B}' = (e'_1, \dots, e'_p)$ une base de F , $a \in U$ et $f = \sum_{k=1}^p f_k e'_k$ une application de U vers F où pour tout $k \in \llbracket 1, p \rrbracket$, les f_k sont les applications de U vers $\mathbb{R}$, coordonnées de f relativement à $\mathcal{B}'$.

Si f est différentiable en²² a , on appelle *matrice jacobienne* de f en a relativement au couple de bases²³ $(\mathcal{B}, \mathcal{B}')$ la matrice de l'application linéaire $df(a)$ relativement à $(\mathcal{B}, \mathcal{B}')$. Elle est notée²⁴ $\text{Jac } f(a)$ et, via le théorème précédent,

$$\text{Jac } f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{R}).$$

Dans le cas où $\dim(E) = \dim(F)$, on appelle *jacobien* de f en a le déterminant de la matrice jacobienne de f en a .

20. Cette écriture abusive signifie que pour tout $x \in U$, $f(x) = \sum_{k=1}^p f_k(x) e'_k$.

21. Dans le cas où $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ munis de leur base canonique respective, on pourra écrire $df(a)(h)$ sous la forme $df(a)(h) = (df_1(a)(h), \dots, df_p(a)(h))$.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, $\mathcal{B}$ et $\mathcal{B}'$ des bases respectives de E et F , U un ouvert de E , $a \in U$, $\lambda \in \mathbb{R}$, f et g deux applications de U vers F différentiables en a .

Alors $\text{Jac}(\lambda f + g)(a) = \lambda \text{Jac} f(a) + \text{Jac} g(a)$.

Théorème

Soient E , F et G trois espaces vectoriels normés de dimension finie, $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ des bases de E , F et G , U un ouvert de E , $a \in U$, V un ouvert de F , f une application de U vers F telle que $f(U) \subset V$, différentiable en a et g une application de V vers G différentiable en $f(a)$.

Alors $\text{Jac}(g \circ f)(a) = \text{Jac} g(f(a)) \text{Jac} f(a)$.

On en déduit le théorème de composition des dérivées partielles suivant.

Corollaire

Soient $(n, p) \in (\mathbb{N}^*)^2$, U un ouvert de $\mathbb{R}^n$, V un ouvert de $\mathbb{R}^p$,

$$f : \begin{cases} U & \longrightarrow \mathbb{R}^p \\ (x_1, \dots, x_n) & \longmapsto (f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n)) \end{cases}$$

telle que $f(U) \subset V$, différentiable en $(a_1, \dots, a_n) \in U$ et

$$g : \begin{cases} V & \longrightarrow \mathbb{R} \\ (y_1, \dots, y_p) & \longmapsto g(y_1, \dots, y_p) \end{cases}$$

différentiable en $(b_1, \dots, b_p) = (f_1(a_1, \dots, a_n), \dots, f_p(a_1, \dots, a_n))$.

Alors $g \circ f$ est différentiable en a et pour tout $i \in \llbracket 1, n \rrbracket$,

$$\frac{\partial (g \circ f)}{\partial x_i}(a_1, \dots, a_n) = \sum_{k=1}^p \frac{\partial g}{\partial y_k}(b_1, \dots, b_p) \frac{\partial f_k}{\partial x_i}(a_1, \dots, a_n).$$

22. L'existence des n dérivées partielles de chaque f_k en a est suffisante pour définir la matrice jacobienne. Rappelons que via le théorème précédent, f différentiable en a implique que chaque f_k est différentiable en a et donc que les n dérivées partielles de chaque f_k existent.

23. Dans le cas de $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$, on appelle matrice jacobienne de f en a la matrice de $df(a)$ relativement aux bases canoniques de E et F .

24. Elle est également parfois notée $J_f(a)$ ou plus rarement $J(f, a)$.

13.4 Gradient

Commençons par donner la définition du gradient d'une application numérique différentiable via le théorème suivant.

Théorème

Soient $(E, \langle \cdot | \cdot \rangle)$ un espace euclidien, U un ouvert de E et f une application de U vers $\mathbb{R}$ différentiable sur U .

Alors pour tout $a \in U$, il existe une unique vecteur de E , appelé *gradient de f en a* et noté $\nabla f(a)$ ou $\text{grad } f(a)$, tel que pour tout $h \in E$,

$$df(a)(h) = \langle \nabla f(a) | h \rangle.$$

Donnons maintenant une expression du gradient dans le cas d'une base orthonormée de E .

Théorème

Soient $(E, \langle \cdot | \cdot \rangle)$ un espace euclidien, $(e_1, \dots, e_n)$ une base orthonormée de E , U un ouvert de E et f une application de U vers $\mathbb{R}$ différentiable sur U .

Alors pour tout $a \in U$, on a

$$\nabla f(a) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(a) e_k.$$

Dans le cas particulier de $E = \mathbb{R}^n$, muni de sa base canonique et de son produit

scalaire canonique défini pour tout $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ et tout $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$

par $\langle x | y \rangle = \sum_{k=1}^n x_k y_k$, le vecteur gradient de f en a sera noté²⁵

$$\nabla f(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

Donnons maintenant l'interprétation géométrique du gradient.

25. On peut considérer le gradient comme un cas particulier de matrice jacobienne : plus précisément, $\nabla f(a)$ est la transposée de la matrice-ligne $\text{Jac } f(a) = \left(\frac{\partial f}{\partial x_1}(a) \quad \dots \quad \frac{\partial f}{\partial x_n}(a) \right)$.

Théorème

Soient U un ouvert de $\mathbb{R}^2$, f une application de U vers $\mathbb{R}$, différentiable sur U . Alors la valeur maximale de la dérivée de f en $a \in U$ dans une direction d'un vecteur unitaire (c'est-à-dire de norme égale à 1) est atteinte dans la direction du vecteur gradient de f en a et est égale à $\|\nabla f(a)\|$ ²⁶.

13.5 Applications de classe $\mathcal{C}^1$

Définition

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E et f une application de U vers F . On dit que f est²⁷ de classe $\mathcal{C}^1$ sur U si f est différentiable sur U et $df : U \rightarrow \mathcal{L}(E, F)$ est continue sur U .

Énonçons maintenant un des théorèmes fondamentaux de ce chapitre.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E et f une application de U vers F .

Alors²⁸ f de classe $\mathcal{C}^1$ sur U si, et seulement si, les dérivées partielles de f relativement à une base de E existent et sont continues sur U .

Les trois théorèmes suivants donnent les propriétés classiques des fonctions de classe $\mathcal{C}^1$.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $\lambda \in \mathbb{R}$, f et g deux applications de U vers F .

Si f et g sont de classe $\mathcal{C}^1$ sur U , $\lambda f + g$ est de classe $\mathcal{C}^1$ sur U et on a

$$d(\lambda f + g) = \lambda df + dg.$$

26. Ainsi, si on se place en un point a de U , $\nabla f(a)$ indique la direction de plus grande pente à partir de ce point (c'est-à-dire la direction pour laquelle f croît le plus vite) et la pente maximale est égale à $\|\nabla f(a)\|$.

27. ou *continûment différentiable*.

28. En particulier si les dérivées partielles de f existent et sont continues sur U , alors f est continue sur U .

Soient E, F, G trois espaces vectoriels normés de dimension finie, U un ouvert de E , V un ouvert de F , f une application de U vers F telle que $f(U) \subset V$ et g une application de V vers G .

Si f est de classe $\mathcal{C}^1$ sur U et g de classe $\mathcal{C}^1$ sur V alors $g \circ f$ est de classe $\mathcal{C}^1$ sur U et pour tout $a \in U$, $d(g \circ f)(a) = dg(f(a)) \circ df(a)$.

Théorème

Soient a et b deux réels, E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , f une application de U vers F de classe $\mathcal{C}^1$ sur U et γ un arc de $[0, 1]$ vers E tel que $\gamma([0, 1]) \subset U$, de classe $\mathcal{C}^1$ sur $[0, 1]$. Alors

$$f(\gamma(1)) - f(\gamma(0)) = \int_0^1 df(\gamma(t))(\gamma'(t)) dt.$$

En particulier, si $\gamma(1) = b$ et $\gamma(0) = a$, on a

$$f(b) - f(a) = \int_0^1 df(\gamma(t))(\gamma'(t)) dt.$$

Le théorème suivant donne la caractérisation des applications de différentielle nulle sur un ouvert connexe par arcs.

Théorème

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert connexe par arcs²⁹ de E et f une application de U vers F différentiable sur U .

Alors f est constante si, et seulement si, df est nulle.

13.6 Vecteurs tangents

Définition

Soient E un espace vectoriel normé de dimension finie, X une partie de E et $x \in X$. On dit qu'un vecteur v de E est *tangent* à X en x s'il existe $\varepsilon > 0$ et un arc γ défini sur $]-\varepsilon, \varepsilon[$, à valeurs dans X , dérivable en 0, tels que $\gamma(0) = x$ et $\gamma'(0) = v$.

On note $T_x X$ l'ensemble des vecteurs tangents à X en x .

Soient U un ouvert de $\mathbb{R}^2$ et f une application de U vers $\mathbb{R}$. Pour représenter graphiquement cette application, on se place dans $\mathbb{R}^3$ et on trace l'ensemble des points $(x, y, z) \in \mathbb{R}^3$ tel que $z = f(x, y)$ autrement dit f peut être représentée

29. Ce théorème est en particulier vérifié si U est un ouvert convexe de E .

graphiquement par une *surface* S d'équation $z = f(x, y)$ tout comme le graphe d'une fonction d'une variable est une *courbe* d'équation $y = f(x)$.

Soit $a \in \mathbb{R}$. On appelle *courbe de niveau* a , notée C_a , la courbe obtenue en sectionnant la surface S par le plan d'équation $z = a$. Une courbe de niveau³⁰ de f est donc une courbe C_a d'équation $f(x, y) = a$ où $a \in \mathbb{R}$. Plus précisément $C_a = \{(x, y) \in \mathbb{R}^2, f(x, y) = a\}$. En considérant le paramètre a comme l'altitude, les courbes de niveau a de f sont donc des courbes d'altitude constante a et montrent graphiquement où f passe à l'altitude a . Concrètement, en coupant la surface S de $\mathbb{R}^3 = (O, \vec{i}, \vec{j}, \vec{k})$ par le plan (horizontal) $z = a$, on obtient une courbe que l'on projette dans le plan $(O, \vec{i}, \vec{j})$. La superposition dans ce plan de plusieurs courbes de niveau permettent donc de donner une visualisation de f dans le plan. Plus la pente de la surface est raide, plus les courbes de niveau sont rapprochées.

Théorème

Soient $(E, \langle \cdot | \cdot \rangle)$ un espace euclidien, U un ouvert de E , f une application de U vers $\mathbb{R}$ différentiable sur U et X une ligne de niveau de f . Alors les vecteurs tangents à X au point $x \in X$ sont orthogonaux au vecteur gradient de f en x .

Pour une fonction f d'une variable, la *tangente* de f en un point x_0 donne une approximation de f au voisinage de ce point. Pour une fonction de deux variables représentée par une surface S d'équation $z = f(x, y)$, le *plan tangent* à S en un point $M_0 = (x_0, y_0, z_0)$ de S donne l'allure approximative de la surface au voisinage de ce point.

L'équation $z = f(x, y)$ peut également s'écrire $g(x, y, z) = 0$ où

$$g(x, y, z) = z - f(x, y).$$

Le *plan tangent* à S en M_0 est l'unique plan passant par M_0 et de vecteur normal

$$\nabla g(x_0, y_0, z_0) = \begin{pmatrix} -\frac{\partial f}{\partial x}(x_0, y_0, z_0) \\ -\frac{\partial f}{\partial y}(x_0, y_0, z_0) \\ 1 \end{pmatrix}.$$

On en déduit le théorème suivant.

30. Dans le cas plus général d'un $\mathbb{R}$ -espace vectoriel normé E de dimension finie, d'un ouvert U de E et d'une application $f : U \rightarrow \mathbb{R}$, on parlera plutôt de *lignes de niveau* de f pour les ensembles $\{x \in U, f(x) = a\}$ où $a \in \mathbb{R}$.

Théorème

Soit $M_0 = (x_0, y_0, z_0)$ un point d'une surface S d'équation $z = f(x, y)$ où f est une application d'un ouvert U de $\mathbb{R}^2$ vers $\mathbb{R}$, différentiable sur U telle que $\nabla f(x_0, y_0) \neq 0$. Alors le plan tangent à S en M_0 est le plan d'équation

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

13.7 Point critique et extremum

Définition

Soient E un espace vectoriel normé de dimension finie, U un ouvert de E , $a \in U$ et f une application de U vers $\mathbb{R}$ différentiable en a . On dit que a est un *point critique* de f si ³¹ $df(a) = 0$.

Définition

Soient E un espace vectoriel normé de dimension finie, U un ouvert de E , $a \in U$ et f une application de U vers $\mathbb{R}$.

On dit que f admet un *maximum local* (respectivement *minimum local*) en a s'il existe $r \in \mathbb{R}_+^*$ tel que pour tout $x \in U$,

$$x \in B(a, r) \implies f(x) \leq f(a) \quad (\text{respectivement } x \in B(a, r) \implies f(x) \geq f(a)).$$

On dit que f admet un *extremum local* en a si f admet un maximum ou un minimum local en a .

Théorème

Soient E un espace vectoriel normé de dimension finie, U un ouvert de E , $a \in U$ et f une application de U vers $\mathbb{R}$ différentiable en a . Si f admet un extremum local en a , alors ³² $df(a) = 0$.

Les extrema locaux de f se trouvent donc *parmi* les points critiques de f .

31. En particulier, $\mathcal{B} = (e_1, \dots, e_n)$ étant une base de E , si a est un point critique de f , alors les dérivées partielles $\frac{\partial f}{\partial x_i}$ sont nulles en a pour tout $i \in \llbracket 1, n \rrbracket$.

32. et donc, $\mathcal{B} = (e_1, \dots, e_n)$ étant une base de E , pour tout $i \in \llbracket 1, n \rrbracket$, $\frac{\partial f}{\partial x_i}(a) = 0$.

Théorème

Soient E un espace vectoriel normé de dimension finie, U un ouvert de E , X une partie de U , $x \in X$ et $f : U \rightarrow \mathbb{R}$, différentiable en x et dont la restriction à X admet un extremum local en x .

Alors $df(x)$ s'annule en tout vecteur tangent à X en x .

Théorème d'optimisation sous une contrainte

Soient E un espace vectoriel normé de dimension finie, U un ouvert de E , f et g deux applications de U vers $\mathbb{R}$, de classe $\mathcal{C}^1$ sur U , X l'ensemble des zéros de g et $x \in X$ tels que la restriction de f à X admet un extremum local en x et la différentielle de g en x est non nulle.

Alors $df(x)$ est colinéaire à $dg(x)$.

13.8 Applications de classe $\mathcal{C}^k$

Définition

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert de E , $a \in U$, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , et f une application de U vers F différentiable³³ sur U .

Soit $(i, j) \in \llbracket 1, n \rrbracket^2$. On dit que f admet une *dérivée partielle seconde* (ou d'ordre 2) en a par rapport aux places (j, i) , notée³⁴ $\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$, si l'application $\frac{\partial f}{\partial x_i}$ admet une dérivée partielle en a par rapport à la j -ème place.

On a alors $\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right)(a)$.

Soient un entier $k \geq 2$ et $(i_1, \dots, i_k) \in \llbracket 1, n \rrbracket^n$. On dit que f admet une *dérivée partielle d'ordre k* en a par rapport aux places $(i_k, \dots, i_1)$, notée $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a)$, si $\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \dots \partial x_{i_1}}$ existe au voisinage de a et l'application $\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \dots \partial x_{i_1}}$ admet une dérivée partielle par rapport à la i_k -ième place en a .

On a alors $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a) = \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \dots \partial x_{i_1}} \right)(a)$.

33. En particulier les dérivées partielles $\frac{\partial f}{\partial x_i}$ sont définies en tout point de U pour tout $i \in \llbracket 1, n \rrbracket$.

34. Pour $i = j$, $\frac{\partial^2 f}{\partial x_i \partial x_i}(a)$ est notée $\frac{\partial^2 f}{\partial x_i^2}(a)$.

Définition

Soient E, F deux espaces vectoriels de dimension finie, U un ouvert de E , f une application de U vers F différentiable sur U et un entier $k \geq 1$.

On dit que f est de classe $\mathcal{C}^k$ sur U si les dérivées partielles de f d'ordre k existent en tout point de U et sont continues sur U .

On dit que f est de classe $\mathcal{C}^\infty$ sur U si pour tout ³⁵ $k \in \mathbb{N}$, f est de classe $\mathcal{C}^k$ sur U .

Si $\mathcal{B}' = (e'_1, \dots, e'_p)$ est une base de F , $f = (f_1, \dots, f_p)$ de classe $\mathcal{C}^k$ sur U si et seulement si chaque f_i est de classe $\mathcal{C}^k$ sur U .

Via une récurrence sur $k \geq 1$, la somme de deux applications de classe $\mathcal{C}^k$ et le produit par un scalaire d'une fonction de classe $\mathcal{C}^k$ sont de classe $\mathcal{C}^k$ de même que la composée de deux applications de classe $\mathcal{C}^k$.

Soient U un ouvert de $\mathbb{R}^2$ et f une application de $\mathbb{R}^2$ vers $\mathbb{R}$.

Y a-t-il un lien entre les dérivées croisées $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$? Le théorème qui suit affirme que si f est de classe $\mathcal{C}^2$ sur U alors les deux dérivées croisées coïncident.

Théorème de Schwarz

Soient k un entier supérieur ou égal à 2, E et F deux espaces vectoriels de dimension finie, U un ouvert de E et f une application de U vers F .

Si f est de classe $\mathcal{C}^k$ sur U alors pour tout $(i_1, \dots, i_k) \in \llbracket 1, n \rrbracket^k$ et toute permutation σ de $\llbracket 1, k \rrbracket$, on a
$$\frac{\partial^k f}{\partial x_{\sigma(i_k)} \cdots \partial x_{\sigma(i_1)}} = \frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}.$$

13.9 Formule de Taylor-Young

Définition

Soient E un espace vectoriel normé de dimension finie, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , U un ouvert de E , $a \in U$ et f une application de U vers $\mathbb{R}$ de classe $\mathcal{C}^2$ sur U . On appelle *matrice hessienne* de f en a la matrice carrée d'ordre n , notée $H_f(a)$ ou $\text{Hess } f(a)$, définie par

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}.$$

35. Pour $k = 0$, f de classe $\mathcal{C}^0$ signifie simplement f continue sur U .

Théorème (formule de Taylor-Young à l'ordre 2)

Soient U un ouvert de $\mathbb{R}^n$, $a \in U$ et f une application de U vers $\mathbb{R}$ de classe $\mathcal{C}^2$ sur U . Alors pour tout $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$, on a³⁶

$$f(a+h) = f(a) + \langle \nabla f(a) \mid h \rangle + \frac{1}{2} \langle H_f(a) h \mid h \rangle + o(\|h\|^2)$$

ce qui peut également s'écrire sous la forme

$$f(a+h) = f(a) + \nabla f(a)^T h + \frac{1}{2} h^T H_f(a) h + o(\|h\|^2).$$

Cette formule s'écrit donc

$$f(a+h) = f(a) + \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + \frac{1}{2} (h_1 \ \cdots \ h_n) \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + o(\|h\|^2).$$

Énonçons le théorème relatif à la condition nécessaire d'un extremum local.

Théorème (condition nécessaire d'optimalité)

Soient U un ouvert de $\mathbb{R}^n$, $a \in U$ et f une application de U vers $\mathbb{R}$ de classe $\mathcal{C}^2$ sur U admettant un minimum (respectivement maximum) local en a .

Alors a est un point critique de f et la matrice hessienne de f en a est positive³⁷ (respectivement négative).

Énonçons le théorème relatif à la condition suffisante d'un extremum local strict.

36. Il est sous-entendu que h est tel que $a+h \in U$ (cf. note 8 p. 198).

37. f étant de classe $\mathcal{C}^2$, via le théorème de Schwarz, la matrice hessienne de f en a est symétrique réelle. Dire que cette dernière est positive signifie que ses valeurs propres (qui sont réelles) sont positives ou nulles (ou encore pour tout $x \in \mathbb{R}^n$, $x^T H_f(a) x \geq 0$).

Théorème (condition suffisante d'optimalité)

Soient U un ouvert de $\mathbb{R}^n$, f une application de U vers $\mathbb{R}$ de classe $\mathcal{C}^2$ sur U et $a \in U$ un point critique de f .

Si la matrice hessienne de f en a est définie positive³⁸ (respectivement définie négative) alors f atteint un minimum (respectivement maximum) local strict en a .

Si la matrice hessienne de f en a admet au moins une valeur propre strictement positive et une valeur propre strictement négative, on dit que a est un *point-col*.

Étudions le cas particulier de $n = 2$.

Soient U un ouvert de $\mathbb{R}^2$, f une application de U vers $\mathbb{R}$ de classe $\mathcal{C}^2$ sur U et $a \in U$ un point critique. On note³⁹

$$r = \frac{\partial^2 f}{\partial x^2}(a), \quad s = \frac{\partial^2 f}{\partial x \partial y}(a) \quad \text{et} \quad t = \frac{\partial^2 f}{\partial y^2}(a).$$

La matrice hessienne de f en a est donc la matrice $H = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$.

Le polynôme caractéristique de H est

$$X^2 - \text{tr}(H)X + \det(H).$$

Si λ_1 et λ_2 désignent les valeurs propres réelles de la matrice symétrique H , on a

$$\lambda_1 + \lambda_2 = \text{tr}(H) = r + t \quad \text{et} \quad \lambda_1 \lambda_2 = \det(H) = rt - s^2.$$

On peut ainsi reformuler le théorème précédent dans le cas $n = 2$ sous la forme suivante.

Théorème

Soient U un ouvert de $\mathbb{R}^2$, f une application de U vers $\mathbb{R}$ de classe $\mathcal{C}^2$ sur U et $a \in U$ un point critique de f .

- Si $rt - s^2 > 0$ et r (ou t) > 0 alors f atteint un minimum local strict en a .
- Si $rt - s^2 > 0$ et r (ou t) < 0 alors f atteint un maximum local strict en a .
- Si $rt - s^2 < 0$, f n'admet pas d'extremum en a ; f admet un point-col en a .

38. Dire que la matrice hessienne de f en a est définie positive signifie que ses valeurs propres sont strictement positives (ou encore pour tout $x \in \mathbb{R}^n$, $x^T H_f(a) x > 0$).

39. Ces notations sont appelées *notations de Monge*.

ALGÈBRE

par Jean-Étienne ROMBALDI

14.1 Quelques notions de logique

De manière intuitive, une *assertion* est un énoncé mathématique qui ne peut prendre que deux valeurs de vérité à savoir « vrai » ou « faux » mais jamais entre les deux comme dans le langage courant. Une assertion qui est toujours vraie est une *tautologie*. Deux assertions sont dites *logiquement équivalentes*, ou plus simplement équivalentes, si elles sont toutes deux vraies ou toutes deux fausses.

À la base de toute théorie mathématique, on dispose d'un petit nombre d'assertions qui sont supposées vraies *a priori* et que l'on nomme *axiomes* ou *postulats*. Ces axiomes sont élaborés par abstraction à partir de l'intuition et ne sont pas déduits d'autres relations. Un exemple important est donné par la construction de l'ensemble noté $\mathbb{N}$ des entiers naturels. Cette construction peut se faire en utilisant les axiomes suivants :

- 0 est un entier naturel ;
- tout entier naturel n a un unique successeur noté $n + 1$;
- deux entiers naturels ayant même successeur sont égaux ;
- toute partie non vide de $\mathbb{N}$ admet un plus petit élément, ce qui signifie que si A est une partie non vide de $\mathbb{N}$, il existe alors un entier $m \in A$ tel que $m \leq n$ pour tout $n \in A$ (axiome du bon ordre).

Cette construction avec les opérations usuelles d'addition et de multiplication est admise. On note $\mathbb{N}^*$ l'ensemble $\mathbb{N}$ privé de 0. De l'axiome du bon ordre, on déduit les deux théorèmes fondamentaux qui suivent.

Théorème de récurrence simple

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une propriété portant sur les entiers $n \geq n_0$. La propriété $\mathcal{P}(n)$ est vraie pour tout entier $n \geq n_0$ si, et seulement si :

- $\mathcal{P}(n_0)$ est vraie ;
- pour tout $n \geq n_0$, si $\mathcal{P}(n)$ est vraie alors $\mathcal{P}(n+1)$ est vraie.

Théorème de récurrence double

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une propriété portant sur les entiers $n \geq n_0$. La propriété $\mathcal{P}(n)$ est vraie pour tout entier $n \geq n_0$ si, et seulement si :

- $\mathcal{P}(n_0)$ et $\mathcal{P}(n_1)$ sont vraies ;
- pour tout $n \geq n_0$, si $\mathcal{P}(n)$ et $\mathcal{P}(n+1)$ sont vraies alors $\mathcal{P}(n+2)$ est vraie.

Théorème de récurrence forte

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une propriété portant sur les entiers $n \geq n_0$. La propriété $\mathcal{P}(n)$ est vraie pour tout entier $n \geq n_0$ si, et seulement si :

- $\mathcal{P}(n_0)$ est vraie ;
- pour tout $n \geq n_0$, si $\mathcal{P}(k)$ est vraie pour tout entier k compris entre n_0 et n , alors $\mathcal{P}(n+1)$ est vraie.

La notion de définition nous permet de décrire un objet ou une situation précise à l'aide du langage courant.

Les énoncés qui se démontrent sont classés en fonction de leur importance dans une théorie comme suit :

- un théorème est une assertion vraie déduite d'autres assertions, il s'agit en général d'un résultat important à retenir ;
- un lemme est un résultat préliminaire utilisé pour démontrer un théorème ;
- un corollaire est une conséquence importante d'un théorème ;
- une proposition est de manière générale un résultat auquel on peut attribuer la valeur vraie ou fausse sans ambiguïté.

Pour rédiger un énoncé mathématique, on utilise le langage courant et les objets manipulés sont représentés en général par des lettres de l'alphabet latin ou grec. Usuellement, on utilise :

- les lettres minuscules $a, b, c, \dots$ pour des objets fixés ;
- les lettres minuscules $x, y, z, t, \dots$ pour des objets inconnus à déterminer ;

- les lettres majuscules $E, F, G, H, \dots$ pour des ensembles ;
- des lettres de l'alphabet grec minuscules ou majuscules $\alpha, \beta, \varepsilon, \dots \Lambda, \Gamma, \dots$

L'élaboration de nouvelles assertions à partir d'autres se fait en utilisant les connecteurs logiques de négation, de conjonction, de disjonction, d'implication et d'équivalence définis comme suit, où P et Q désignent des assertions.

- La *négation* de P , notée $\neg P$, ou non P ou $\overline{P}$, est l'assertion qui est vraie si P est fausse et fausse si P est vraie. En théorie des ensembles on admet qu'il n'existe pas d'assertion P telle que P et $\overline{P}$ soient toutes deux vraies. On dit que cette théorie est non contradictoire.
- La *conjonction* de P et Q , notée $P \wedge Q$ (lire P et Q), est l'assertion qui est vraie uniquement si P et Q sont toutes deux vraies.
- La *disjonction* de P et Q , notée $P \vee Q$ (lire P ou Q), est l'assertion qui est vraie uniquement si l'une des deux assertions P ou Q est vraie. On peut aussi introduire le « ou exclusif », qui est vrai uniquement lorsque l'une des deux assertions, mais pas les deux simultanément, est vraie.
- L'*implication*, notée $P \Rightarrow Q$, est l'assertion qui est fausse uniquement si P est vraie et Q fausse.
- L'*équivalence* de P et Q , notée $P \Leftrightarrow Q$, est l'assertion qui est vraie uniquement si $P \Rightarrow Q$ et $Q \Rightarrow P$ sont toutes deux vraies.

Dans ce cas, on dit que Q est une condition nécessaire et suffisante de P .

Théorème

Soient P, Q, R des propositions. On a les équivalences :

- $(P \wedge Q) \Leftrightarrow (Q \wedge P)$ et $(P \vee Q) \Leftrightarrow (Q \vee P)$ (commutativité) ;
- $(P \wedge (Q \wedge R)) \Leftrightarrow ((P \wedge Q) \wedge R)$ et $(P \vee (Q \vee R)) \Leftrightarrow ((P \vee Q) \vee R)$ (associativité) ;
- $\begin{cases} (P \wedge (Q \vee R)) \Leftrightarrow ((P \wedge Q) \vee (P \wedge R)) \\ (P \vee (Q \wedge R)) \Leftrightarrow ((P \vee Q) \wedge (P \vee R)) \end{cases}$ (distributivité) ;
- $(\overline{\overline{P}}) \Leftrightarrow (P) ; (\overline{P \wedge Q}) \Leftrightarrow (\overline{P} \vee \overline{Q}) ; (\overline{P \vee Q}) \Leftrightarrow (\overline{P} \wedge \overline{Q})$ (relations de De Morgan) ;
- $(P \Rightarrow Q) \Leftrightarrow (\overline{Q} \Rightarrow \overline{P}) ; (P \Rightarrow Q) \Leftrightarrow (\overline{P} \vee Q) ; (\overline{P \Rightarrow Q}) \Leftrightarrow (P \wedge \overline{Q})$.

En général l'énoncé d'une proposition à démontrer est formé d'une ou plusieurs hypothèses qui constituent l'assertion H et d'une ou plusieurs conclusions qui constituent l'assertion C . Il s'agit donc de montrer l'implication $H \Rightarrow C$. Si de plus, on peut montrer que $C \Rightarrow H$, on dira alors que la réciproque de la proposition est vraie. Les idées de base que l'on peut utiliser sont les suivantes.

- Une assertion peut toujours être remplacée par n'importe quelle assertion qui lui est équivalente.
- On peut effectuer une démonstration directe, c'est-à-dire de déduire logiquement C de H .
- L'implication étant transitive, on peut essayer de montrer que $C \Rightarrow C'$ sachant par ailleurs que $C' \Rightarrow H$.
- Dans le cas où une démonstration directe semble difficile, on peut essayer une *démonstration par l'absurde* qui consiste à étudier l'assertion $H \wedge \overline{C}$ équivalente à $\overline{H \Rightarrow C}$ et on montre qu'on aboutit à une impossibilité si cette dernière assertion est vraie.
- On peut aussi essayer une *démonstration par contraposition* qui consiste à montrer la *contraposée* $\overline{C} \Rightarrow \overline{H}$ puisque les implications $H \Rightarrow C$ et $\overline{C} \Rightarrow \overline{H}$ sont équivalentes.
- La *démonstration par contre-exemple* permet de montrer qu'une implication $H \Rightarrow C$, où H et C sont des propriétés portant sur des variables x , est fausse. Pour ce faire on cherche une ou des valeurs de x pour lesquels $H(x)$ est vraie et $C(x)$ est fausse.

14.2 Notions de base sur les ensembles

Nous nous contenterons d'une définition intuitive de la notion d'ensemble.

Un ensemble est une collection d'objets possédant des propriétés communes, ces objets sont les éléments de l'ensemble. On admet l'existence d'un ensemble qui ne contient aucun élément. Cet ensemble est noté $\emptyset$ et on dit que c'est l'*ensemble vide*. On utilisera les notations suivantes, pour les ensembles de nombres usuels :

- $\mathbb{N}$ est l'ensemble des entiers naturels ; $\mathbb{Z}$ celui des entiers relatifs ;
- $\mathbb{Q}$ est l'ensemble des nombres rationnels
- $\mathbb{R}$ est l'ensemble des nombres réels ; $\mathbb{C}$ celui des nombres complexes.

Nous serons souvent amenés à décrire un ensemble en précisant les propriétés que doivent vérifier tous ses éléments, ce que nous noterons de la façon suivante :

$$E = \{\text{description des propriétés des éléments de } E\}$$

On peut aussi décrire un ensemble en donnant la liste finie ou infinie de tous ses éléments, quand cela est possible, ce qui se note $E = \{x_1, \dots, x_n\}$ s'il s'agit d'un ensemble fini ou $E = \{x_1, \dots, x_n, \dots\}$ s'il s'agit d'un ensemble infini pour lequel on peut numérotter les éléments (un tel ensemble est dit *dénombrable*¹). On dit alors que l'ensemble E est défini en extension.

Un *singleton* est un ensemble qui ne contient qu'un élément, soit $E = \{a\}$.

1. Cf. définition page 365.

Si E est un ensemble, on notera alors $a \in E$ pour signifier que a est un élément de E , ce qui se lit « a appartient à E ». La négation « a n'appartient pas à E » se note $a \notin E$.

Pour signifier qu'un ensemble F est contenu dans un ensemble E , ce qui signifie que tout élément de F est dans E , nous noterons $F \subset E$ qui se lit « F est contenu dans E ». On peut écrire de manière équivalente que $E \supset F$ pour dire que E contient F . La négation de cette assertion est notée $F \not\subset E$.

Deux ensembles E et F sont égaux si, et seulement si, ils ont les mêmes éléments, ce qui se traduit par $E \subset F$ et $F \subset E$.

Si E est un ensemble, il existe un ensemble dont tous les éléments sont formés de tous les sous-ensembles (ou parties) de E . On note $\mathcal{P}(E)$ cet ensemble et on dit que c'est l'ensemble des parties de E . Ainsi $F \subset E$ est équivalent à $F \in \mathcal{P}(E)$. L'ensemble vide et E sont des éléments de $\mathcal{P}(E)$. Pour décrire des ensembles, ou faire des raisonnements, nous utiliserons les deux *quantificateurs* suivants ;

- le *quantificateur universel* « quel que soit » ou « pour tout » noté $\forall$ utilisé pour signifier que tout élément x d'un ensemble E vérifie une propriété $P(x)$, la syntaxe étant : $\forall x \in E, P(x)$;
- le *quantificateur existentiel* « il existe » noté $\exists$ pour signifier qu'il existe au moins un élément x de E vérifiant la propriété $P(x)$, la syntaxe étant : $\exists x \in E, P(x)$.
- Pour signifier qu'il existe un et un seul x dans E vérifiant la propriété $P(x)$, on utilisera la syntaxe : $\exists! x \in E, P(x)$.

14.3 Les symboles Σ et Π

Si n est un entier naturel non nul et $x_1, \dots, x_n$ des entiers, rationnels, réels ou complexes, on notera $\sum_{k=1}^n x_k = x_1 + \dots + x_n$ et $\prod_{k=1}^n x_k = x_1 \cdot \dots \cdot x_n$ la somme et le produit des x_k . Dans une telle somme ou produit l'indice est « muet », ce qui signifie que $\sum_{k=1}^n x_k = \sum_{i=1}^n x_i$ et $\prod_{k=1}^n x_k = \prod_{i=1}^n x_i$.

On peut également effectuer des changements d'indice. Par exemple, en posant $i = k + 1$, on a $\sum_{k=1}^n x_k = \sum_{i=2}^{n+1} x_{i-1} = \sum_{k=2}^{n+1} x_{k-1}$.

On peut ajouter ou multiplier de telles sommes (ou produits). Par exemple, on a :

$$\sum_{k=1}^n x_k + \sum_{k=1}^n y_k = \sum_{k=1}^n (x_k + y_k); \quad \lambda \sum_{k=1}^n x_k = \sum_{k=1}^n \lambda x_k$$

$$\left(\sum_{k=1}^n x_k \right) \left(\sum_{k=1}^m y_k \right) = \left(\sum_{j=1}^n x_j \right) \left(\sum_{k=1}^m y_k \right) = \sum_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}} x_j y_k.$$

14.4 L'algèbre des parties d'un ensemble

On note $\mathcal{P}(E)$ l'ensemble des parties d'un ensemble E . En désignant par A , B et C des parties de E , on définit les ensembles suivants :

- le *complémentaire* de A dans E est l'ensemble noté $E \setminus A$ ou $\overline{A}$ des éléments de E qui ne sont pas dans A , ce qui se traduit par :

$$(x \in E \setminus A) \Leftrightarrow ((x \in E) \text{ et } (x \notin A))$$

ou encore par $E \setminus A = \{x \in E, x \notin A\}$.

- L'*intersection* de A et B , notée $A \cap B$, est l'ensemble des éléments de E qui sont dans A et dans B , soit : $(x \in A \cap B) \Leftrightarrow ((x \in A) \text{ et } (x \in B))$ ou encore $A \cap B = \{x \in E, x \in A \text{ et } x \in B\}$. Si $A \cap B = \emptyset$, on dit alors que A et B sont *disjointes*.

- La *réunion* de A et B , notée $A \cup B$, est l'ensemble des éléments de E qui sont soit dans A , soit dans B (éventuellement dans A et B) soit :

$$(x \in A \cup B) \Leftrightarrow ((x \in A) \text{ ou } (x \in B))$$

ou encore $A \cup B = \{x \in E, x \in A \text{ ou } x \in B\}$.

- La *différence* de A et B , notée $A \setminus B$, est l'ensemble des éléments de E qui sont dans A et qui ne sont pas dans B , soit :

$$(x \in A \setminus B) \Leftrightarrow ((x \in A) \text{ et } (x \notin B))$$

ou encore $A \setminus B = \{x \in A, x \notin B\}$.

- La *différence symétrique* de A et B , notée $A \Delta B$, est l'ensemble des éléments de E qui sont soit dans A et pas dans B , soit dans B et pas dans A , donc :

$$(x \in A \Delta B) \Leftrightarrow ((x \in A) \text{ et } (x \notin B)) \text{ ou } ((x \in B) \text{ et } (x \notin A)).$$

Théorème

Soient E un ensemble et A, B, C des sous-ensembles de E . On a :

- $A \cap B = B \cap A$, $A \cup B = B \cup A$ et $A \Delta B = B \Delta A$ (commutativité) ;
- $A \cap (B \cap C) = (A \cap B) \cap C$, $A \cup (B \cup C) = (A \cup B) \cup C$ et $A \Delta (B \Delta C) = (A \Delta B) \Delta C$ (associativité) ;
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$ et $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
- $A \Delta A = \emptyset$, $A \Delta B = (A \setminus B) \cup (B \setminus A)$, $A \Delta B = (A \cap \overline{B}) \cup (B \cap \overline{A})$ et $A \Delta B = (A \cup B) \setminus (A \cap B)$;
- $\overline{\overline{A}} = A$, $(A \subset B) \Leftrightarrow (\overline{B} \subset \overline{A})$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$ et $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

De manière plus générale, on définit l'intersection ou la réunion d'une famille $(A_i)_{i \in I}$ de sous-ensembles de E par :

$$\left(x \in \bigcap_{i \in I} A_i\right) \Leftrightarrow (\forall i \in I, x \in A_i) \text{ et } \left(x \in \bigcup_{i \in I} A_i\right) \Leftrightarrow (\exists i \in I, x \in A_i).$$

Pour tout $j \in I$, on a

$$E \setminus \bigcap_{i \in I} A_i = \bigcup_{i \in I} (E \setminus A_i), \quad E \setminus \bigcup_{i \in I} A_i = \bigcap_{i \in I} (E \setminus A_i) \text{ et } \bigcap_{i \in I} A_i \subset A_j \subset \bigcup_{i \in I} A_i.$$

Définition

On appelle *recouvrement disjoint* d'un ensemble E toute famille $(A_i)_{i \in I}$ de parties de E de réunion égale à E avec les A_i deux à deux disjoints. Si de plus les A_i sont non vides, on dit que la famille $(A_i)_{i \in I}$ est une *partition* de E .

Définition

Étant donné deux ensembles E et F , on appelle *produit cartésien* de E par F l'ensemble $E \times F$ des couples (x, y) formés d'un élément x de E et d'un élément y de F .

Dans le cas où $F = E$, on note E^2 pour $E \times E$.

On peut itérer le procédé et définir le produit cartésien $E_1 \times \cdots \times E_n$ de n ensembles comme l'ensemble des listes (ordonnées) $(x_1, \dots, x_n)$ formées d'un élément x_1 de E_1 suivi d'un élément x_2 de $E_2, \dots$, suivi d'un élément x_n de E_n . On notera de façon condensée $\prod_{k=1}^n E_k = E_1 \times \cdots \times E_n$. Dans le cas où tous les E_k sont égaux à un même ensemble E , on note E^n pour $E \times \cdots \times E$ (n fois).

14.5 Fonctions, injectivité, surjectivité, bijectivité

E, F et G désignent des ensembles.

Définition

On appelle *application*, ou *fonction*², de E dans F (ou de E vers F) toute partie Γ du produit cartésien $E \times F$ telle que :

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

2. On ne distinguera pas dans cet ouvrage les notions d'*application* et de *fonction* de sorte qu'il sera sous-entendu qu'une fonction f de E dans F est *définie* sur tout E , autrement dit le domaine de définition d'une telle fonction f est E .

On note $f : \begin{cases} E & \rightarrow F \\ x & \mapsto f(x) \end{cases}$.

On utilisera aussi les notations $f : E \rightarrow F$ ou $f : x \mapsto f(x)$. Pour $x \in E$, $y = f(x)$ est l'*image* de x par f et x est un *antécédent* de y par f . On dit que E est l'*ensemble de départ* (ou le *domaine de définition*), F l'*ensemble d'arrivée* et Γ le *graphe* de l'application f .

On note $\mathcal{F}(E, F)$ ou F^E l'ensemble de toutes les applications de E dans F .

Deux applications f et g sont égales si, et seulement si, elles ont même ensemble de départ E , même ensemble d'arrivée F et si $g(x) = f(x)$ pour tout $x \in E$.

L'application qui associe à tout x d'un ensemble E le même x est l'application identique notée Id_E ou Id si l'ensemble E est fixé.

Si f est une fonction de E dans F et D un sous-ensemble non vide de E , on définit alors une application g de D dans F en posant $g(x) = f(x)$ pour tout $x \in D$ et on dit que g est la *restriction* de f à D , ce qui se note $g = f|_D$.

Définition

Soient f une application de E dans F , A une partie de E et B une partie de F . L'*image* de A par f ³ est⁴ $f(A) = \{f(x); x \in A\}$ et l'*image réciproque* de B par f est $f^{-1}(B) = \{x \in E, f(x) \in B\}$. L'ensemble $f(E)$ est appelé l'*image* de f .

Théorème

Soit f une application de E dans F . Pour toutes parties A, B de E et C, D de F , on a :

- $A \subset B \Rightarrow f(A) \subset f(B)$;
- $f(A \cup B) = f(A) \cup f(B)$; $f(A \cap B) \subset f(A) \cap f(B)$;
- $C \subset D \Rightarrow f^{-1}(C) \subset f^{-1}(D)$;
- $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$; $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$;
- $f^{-1}(\overline{C}) = \overline{f^{-1}(C)}$.

Définition

Soient f une application de E dans F et g une application de F dans G . La *composée* de f par g est l'application de E dans G notée $g \circ f$ définie par :

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

3. également appelée *l'image directe* de A par f .

4. Cet ensemble peut également s'écrire $f(A) = \{y \in F, \exists x \in A, y = f(x)\}$.

Dans le cas où $E = F$, on dit que f et g *commutent*, si $f \circ g = g \circ f$.

La suite des itérées $f : E \rightarrow E$ est définie par $f^0 = \text{Id}_E$ et $f^{n+1} = f^n \circ f$ pour tout $n \in \mathbb{N}$. On a $f^p \circ f^q = f^q \circ f^p = f^{p+q}$ pour tous p, q dans $\mathbb{N}$.

Définition

Soient E, F deux ensembles et $f : E \rightarrow F$. On dit que f est :

- *injective* (ou que c'est une *injection*), si deux éléments distincts de E ont deux images distinctes dans F , soit :

$$\forall (x_1, x_2) \in E^2, (x_1 \neq x_2) \Rightarrow (f(x_1) \neq f(x_2)).$$

- *surjective* (ou que c'est une *surjection*), si tout élément de F a au moins un antécédent dans E , soit :

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

- *bijective* (ou que c'est une *bijection*), si elle est injective et surjective.

Une bijection peut se caractériser en disant que tout élément de y a un unique antécédent par f , ce qui est encore équivalent à dire que pour que pour tout $y \in F$ l'équation $y = f(x)$ a une et une seule solution x dans E , ce qui permet de définir la *bijection réciproque* (ou l'*application réciproque*) de f , notée f^{-1} , de F dans E par :

$$(y \in F \text{ et } x = f^{-1}(y)) \Leftrightarrow (x \in E \text{ et } y = f(x)).$$

Cette application f^{-1} est une bijection de F dans E . L'application $f \circ f^{-1}$ est alors l'application identité $y \mapsto y$ de F dans F et l'application $f^{-1} \circ f$ est l'application identité $x \mapsto x$ de E dans E , ce qui se note $f \circ f^{-1} = \text{Id}_F$ et $f^{-1} \circ f = \text{Id}_E$.

Définition

Une *permutation* d'un ensemble E est une bijection de E dans lui même.

On note $\mathcal{S}(E)$ l'ensemble des permutations de E .

Théorème

Soient E, F, G des ensembles, f une application de E dans F et g une application de F dans G .

- Si f et g sont injectives [resp. surjectives], alors $g \circ f$ est injective [resp. surjective] (la composée de deux injections [resp. surjections] est une injection [resp. surjection]).
- Si f et g sont bijectives, alors $g \circ f$ est bijective (la composée de deux bijections est une bijection) et on a $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Théorème

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

- S'il existe $g : F \rightarrow E$ telle que $g \circ f = \text{Id}_E$, alors f est injective.
- S'il existe $h : F \rightarrow E$ telle que $f \circ h = \text{Id}_F$, alors f est surjective.
- S'il existe deux applications g et h de F dans E telles que $g \circ f = \text{Id}_E$ et $f \circ h = \text{Id}_F$, alors f est bijective et $g = h = f^{-1}$.

14.6 Relations d'ordre et d'équivalence

On se donne un ensemble non vide E et une propriété $\mathcal{P}$ sur les éléments (x, y) de E^2 . On dit que x et y dans E sont en relation, ce que l'on notera $x \mathcal{R} y$, si le couple (x, y) vérifie la propriété $\mathcal{P}$. On dit que $\mathcal{R}$ est une *relation binaire* sur E .

Définition

On dit que $\mathcal{R}$ est une *relation d'ordre* sur E , si elle est :

- *réflexive* : pour tout $x \in E$, on a $x \mathcal{R} x$;
- *antisymétrique* : pour tout $(x, y) \in E^2$, si $x \mathcal{R} y$ et $y \mathcal{R} x$, on a $x = y$;
- *transitive* : pour tout $(x, y, z) \in E^3$, si $x \mathcal{R} y$ et $y \mathcal{R} z$, on a alors $x \mathcal{R} z$.

Un ensemble E muni d'une relation d'ordre est dit *ordonné*. Si deux éléments quelconques de E sont comparables, c'est-à-dire que pour tous x, y dans E , on a $x \mathcal{R} y$ ou $y \mathcal{R} x$, on dit alors que l'ensemble E est *totalelement ordonné* ou que $\mathcal{R}$ est une *relation d'ordre total* sur E .

Définition

On dit que $\mathcal{R}$ est une *relation d'équivalence* sur E , si elle est :

- *réflexive* : pour tout $x \in E$, on a $x \mathcal{R} x$;
- *symétrique* : pour tout $(x, y) \in E^2$, si $x \mathcal{R} y$, on a alors $y \mathcal{R} x$;
- *transitive* : pour tout $(x, y, z) \in E^3$, si $x \mathcal{R} y$ et $y \mathcal{R} z$, on a alors $x \mathcal{R} z$.

Si $\mathcal{R}$ est une relation d'équivalence sur E , on note $\bar{x} = \{y \in E, x \mathcal{R} y\}$ pour tout $x \in E$. Cet ensemble est non vide et on dit que c'est la *classe d'équivalence* de x . L'ensemble de toutes ces classes d'équivalence est appelé l'*ensemble quotient* de E par $\mathcal{R}$ et noté $E/\mathcal{R}$.

L'application $\pi : x \in E \mapsto \bar{x} \in E/\mathcal{R}$ est la surjection canonique de E sur $E/\mathcal{R}$ et les classes d'équivalence forment une partition de E .

Groupe, anneaux et corps

15

Définition

On appelle *loi de composition interne* sur un ensemble non vide G toute application φ définie sur $G \times G$ et à valeurs dans G .

Si φ est une loi de composition interne sur G , on notera plus simplement, pour tous a, b dans G , $a \star b = \varphi(a, b)$. Il sera parfois commode de noter une telle loi sous la forme additive $(a, b) \mapsto a + b$ ou sous la forme multiplicative $(a, b) \mapsto a \cdot b$ ou plus simplement $(a, b) \mapsto ab$. On note $(G, \star)$ l'ensemble G muni de la loi de composition interne $\star$.

Définition

Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne. On dit que :

- cette loi est *associative* si $(a \star b) \star c = a \star (b \star c)$ pour tous a, b, c dans G ;
- cette loi est *commutative* si $a \star b = b \star a$ pour tous a, b dans G ;
- e est un *élément neutre* pour cette loi si $a \star e = e \star a = a$ pour tout $a \in G$;
- un élément a de G est dit *régulier* (ou *simplifiable*) si :

$$\forall (b, c) \in G^2, \begin{cases} a \star b = a \star c \Rightarrow b = c \\ b \star a = c \star a \Rightarrow b = c \end{cases}.$$

Dans le cas d'une loi associative, on peut effectuer les opérations $a_1 \star \cdots \star a_n$ où les a_j sont des éléments de G , ce que l'on notera $\prod_{j=1}^n a_j$ dans le cas d'une loi

multiplicative ou $\sum_{j=1}^n a_j$ dans le cas d'une loi additive. Ce produit est donc défini par $a_1 \in G$ et supposant $\prod_{j=1}^{n-1} a_j$ construit pour $n \geq 2$, on a $\prod_{j=1}^n a_j = \prod_{j=1}^{n-1} a_j \star a_n$, le parenthésage étant sans importance du fait de l'associativité¹.

Dans le cas où tous les a_j sont égaux à un même élément a , ce produit est noté a^n et on dit que c'est la puissance n -ième de a . Dans le cas où la loi est notée additivement, on note na au lieu de a^n .

Théorème

Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne. Si G admet un élément neutre, ce dernier est alors unique.

Définition

Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne et admettant un élément neutre e . On dit qu'un élément a de G est *inversible* s'il existe un élément a' dans G tel que $a \star a' = a' \star a = e$. On dit alors que a' est un *inverse* (ou un *symétrique*) de a dans G .

Théorème

Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne associative et admettant un élément neutre e . Si $a \in G$ admet un inverse dans G , ce dernier est alors unique.

En cas d'existence, on notera a^{-1} l'inverse de a dans $(G, \star)$, la loi $\star$ étant associative. Dans le cas d'une loi de composition interne notée de façon additive, on notera plutôt $-a$ l'inverse de a et on l'appellera *opposé*.

1. Pour $n = 0$, il sera commode de noter $\prod_{j=1}^n a_j = 1$ [resp. $\sum_{j=1}^n a_j = 0$ dans le cas d'une loi additive], où 1 [resp. 0 pour une loi additive] est l'élément neutre.

Définition

Un *groupe* est un ensemble non vide G muni d'une loi de composition interne $\star$ possédant les propriétés suivantes :

- la loi $\star$ est associative ;
- il existe un élément neutre e pour la loi $\star$;
- tout élément de G admet un symétrique.

Si de plus la loi $\star$ est commutative, on dit que le groupe G est *commutatif* ou *abélien*.

En général, s'il n'y pas de confusion possible, on dira tout simplement que G est un groupe et on notera ab ou $a + b$ le résultat de l'opération $a \star b$. Pour un groupe multiplicatif $(G, \cdot)$, on notera 1 l'élément neutre, a^{-1} le symétrique d'un élément a de G et pour un groupe additif $(G, +)$, on notera 0 l'élément neutre, $-a$ le symétrique qu'on appelle *opposé*.

Théorème

Dans un groupe $(G, \star)$ tout élément est simplifiable.

Si $(G, \star)$ est un groupe tel que G a un nombre fini $n \geq 1$ d'éléments, on dira alors que G est un groupe fini d'ordre (ou de cardinal) n .

Théorème

Soit $(G, \star)$ un groupe. Pour tous $a_1, \dots, a_p$ dans G , on a :

$$(a_1 \star \dots \star a_p)^{-1} = a_p^{-1} \star \dots \star a_1^{-1}.$$

Définition

Soit $(G, \star)$ un groupe. Un *sous-groupe* de G est un sous-ensemble H de G tel que :

- H est non vide ;
- pour tous a, b dans H , $a \star b^{-1}$ est dans H .

Théorème

Soit $(G, \star)$ un groupe. Un sous-ensemble H de G est sous-groupe si, et seulement si :

- H contient l'élément neutre e de G ;
- H est stable pour la loi $\star$, c'est-à-dire que $a \star b \in H$ pour tous a, b dans H ;
- H est stable par passage à l'inverse, c'est-à-dire que $a^{-1} \in H$ pour tout $a \in H$.

Un sous-groupe H d'un groupe G est lui-même un groupe et il est commutatif si G l'est.

Théorème

Si G est un sous-groupe de $(\mathbb{Z}, +)$, il existe alors un unique entier $n \in \mathbb{N}$ tel que $G = n\mathbb{Z} = \{qn ; q \in \mathbb{Z}\}$.

Théorème

Soit $(G, \star)$ un groupe. L'intersection d'une famille quelconque $(H_i)_{i \in I}$ de sous-groupes de G est un sous-groupe de G .

Corollaire

Si X est une partie d'un groupe $(G, \star)$, l'intersection de tous les sous-groupes de G qui contiennent X est un sous-groupe de G .

Définition

Si X est une partie d'un groupe $(G, \star)$, le *sous-groupe de G engendré par X* est l'intersection de tous les sous-groupes de G qui contiennent X .

On note $\langle X \rangle$ le sous-groupe de G engendré par X et ce sous-groupe est le plus petit (pour l'ordre de l'inclusion) des sous-groupes de G qui contiennent X . Dans le cas où X est l'ensemble vide, on a $\langle X \rangle = \{e\}$.

Définition

Si X est une partie d'un groupe $(G, \star)$, on dit alors que X *engendre* G , si on a $G = \langle X \rangle$.

Théorème

Soient $(G, \star)$ un groupe, X et Y deux parties de G .

- On a $X \subset \langle X \rangle$ et l'égalité est réalisée si, et seulement si, X est un sous-groupe de G .
- Si $X \subset Y$, on a alors $\langle X \rangle \subset \langle Y \rangle$.
- En notant, pour X non vide, $X^{-1} = \{x^{-1} ; x \in X\}$ l'ensemble formé des symétriques des éléments de X , les éléments de $\langle X \rangle$ sont de la forme $x_1 \star \cdots \star x_n$ où $n \in \mathbb{N}^*$ et les x_k sont dans $X \cup X^{-1}$ pour tout k compris entre 1 et n .

Théorème

Soient $(G_1, \star_1)$ et $(G_2, \star_2)$ deux groupes. Alors $(G_1 \times G_2, \star)$ est un groupe appelé *groupe produit* où $\star$ est défini pour tout (x_1, x_2) et (y_1, y_2) dans $G_1 \times G_2$ par

$$(x_1, x_2) \star (y_1, y_2) = (x_1 \star_1 y_1, x_2 \star_2 y_2).$$

15.1 Groupes monogènes

On se donne un groupe multiplicatif $(G, \cdot)$. Si X est une partie de G formée d'un nombre fini d'éléments, $x_1, \dots, x_n$, on note alors $\langle X \rangle = \langle x_1, \dots, x_n \rangle$.

Définition

On dit que G est un *groupe monogène*, s'il existe $x_1 \in G$ tel que $G = \langle x_1 \rangle$. Si de plus, G est fini, on dit alors qu'il est *cyclique*.

Théorème

Pour a dans G et n, m dans $\mathbb{Z}$, on a $a^n a^m = a^{n+m}$ et pour $b \in G$ qui commute avec a , on a $(ab)^n = a^n b^n = b^n a^n$.

Théorème

Pour tout $a \in G$, le sous-groupe engendré par a est $\langle a \rangle = \{a^n ; n \in \mathbb{Z}\}$.

Pour une loi de groupe notée additivement, on a dans le cas où G est commutatif,

$$\langle x_1, \dots, x_n \rangle = \left\{ \sum_{k=1}^n \alpha_k x_k ; (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n \right\}.$$

15.2 Groupes finis, théorème de Lagrange

Pour ce paragraphe, on se donne un groupe multiplicatif $(G, \cdot)$.

Théorème

Pour tout sous-groupe H de G , la relation $\sim$ définie sur G par :

$$(x \sim y) \Leftrightarrow (x^{-1}y \in H)$$

est une relation d'équivalence.

Avec les notations du théorème précédent, on note, pour tout $g \in G$, $\bar{g}$ la classe d'équivalence de g et on dit que $\bar{g}$ est la *classe à gauche* modulo H de g .

On a, pour tout $g \in G$:

$$(h \in \bar{g}) \Leftrightarrow (g \sim h) \Leftrightarrow (g^{-1}h \in H) \Leftrightarrow (\exists k \in H, h = gk) \Leftrightarrow (h \in gH)$$

soit $\bar{g} = gH$. L'ensemble de ces classes d'équivalence est noté G/H et on l'appelle l'*ensemble des classes à gauche* modulo H . On a donc :

$$G/H = \{\bar{g} ; g \in G\} = \{gH ; g \in G\}.$$

L'application $\pi : \begin{cases} G & \rightarrow G/H \\ g & \mapsto \bar{g} = gH \end{cases}$ est surjective.

C'est la surjection canonique de G sur G/H .

Théorème de Lagrange

Soient G un groupe fini d'ordre $n \geq 2$ et H un sous-groupe de G .

- Les classes à gauche modulo H forment une partition de G .
- Pour tout $g \in G$, on a $\text{Card}(\bar{g}) = \text{Card}(H)$.
- L'ordre du sous-groupe H divise l'ordre du groupe G .

Le cardinal de l'ensemble G/H est noté $[G : H]$ et on l'appelle l'*indice* de H dans G . Le théorème de Lagrange peut aussi se traduire par :

$$[G : H] = \text{Card}(G/H) = \frac{\text{Card}(G)}{\text{Card}(H)}.$$

15.3 Morphismes de groupes

On désigne par $(G, \star)$ et $(H, \cdot)$ deux groupes et on note respectivement e et 1 les éléments neutres de $(G, \star)$ et $(H, \cdot)$.

Définition

On dit que φ est un *morphisme de groupes* de G dans H si φ est une application de G dans H telle que :

$$\forall (a, b) \in G^2, \varphi(a \star b) = \varphi(a) \cdot \varphi(b).$$

Dans le cas où φ est de plus bijective, on dit que φ est un *isomorphisme* du groupe G sur le groupe H . Dans le cas où $H = G$, on dit que φ est un *endomorphisme* du groupe $(G, \star)$ et que c'est un *automorphisme* du groupe $(G, \star)$ si φ est de plus bijective.

Théorème

Soient G, H, K trois groupes, φ un morphisme de groupes de G dans H et ψ un morphisme de groupes de H dans K . L'application $\psi \circ \varphi$ est un morphisme de groupes de G dans K . Si φ est un automorphisme de G , alors φ^{-1} est également un automorphisme de G .

L'ensemble $(\text{Aut}(G), \circ)$ des automorphismes de G dans G est un sous-groupe du groupe symétrique $(\mathcal{S}(G), \circ)$ formé des bijections (ou permutations) de G .

Théorème

Si φ est un morphisme de groupes de G dans H , on alors $\varphi(e) = 1$ et $\varphi(a)^{-1} = \varphi(a^{-1})$ pour tout $a \in G$.

Définition

Soit φ un morphisme de groupes de G dans H .

- Le *noyau* de φ est l'ensemble $\text{Ker}(\varphi) = \{x \in G, \varphi(x) = 1\}$.
- L'*image* de φ est l'ensemble $\text{Im}(\varphi) = \{\varphi(x); x \in G\}$.

Théorème

Si φ est un morphisme de groupes de G dans H , alors :

- $\text{Ker}(\varphi)$ est un sous-groupe de G et $\text{Im}(\varphi)$ est un sous-groupe de H ;
- φ est injectif si, et seulement si, on a $\text{Ker}(\varphi) = \{e\}$ et φ est surjectif si, et seulement si, on a $\text{Im}(\varphi) = H$;
- pour tout sous-groupe G' de G , $\varphi(G')$ est un sous-groupe de H et pour tout sous-groupe H' de H , $\varphi^{-1}(H')$ est un sous-groupe de G .

15.4 Ordre d'un élément dans un groupe

On se donne un groupe multiplicatif $(G, \cdot)$ et pour tout $a \in G$, $\langle a \rangle = \{a^n ; n \in \mathbb{Z}\}$ est le sous-groupe de G engendré par a .

Définition

L'ordre de $a \in G$ est $\theta(a) = \text{Card}(\langle a \rangle)$. Si $\theta(a)$ est dans $\mathbb{N}^*$, on dit alors que a est d'ordre fini, sinon on dit qu'il est d'ordre infini.

Pour $a \in G$, le sous-groupe de G engendré par a peut être vu comme l'image du morphisme de groupes $\varphi_a : k \in \mathbb{Z} \mapsto a^k \in G$.

Théorème

Pour $a \in G$, on a $\theta(a) = +\infty$ si, et seulement si, φ_a est injective et pour a d'ordre fini, on a $\text{Ker}(\varphi_a) = \theta(a)\mathbb{Z}$.

Corollaire

Soit $a \in G$. Les assertions suivantes sont équivalentes :

- a est d'ordre fini $n \geq 1$;
- $a^n = 1$ et $a^k \neq 1$ pour tout entier k compris entre 1 et $n-1$;
- pour $k \in \mathbb{Z}$, on a $a^k = 1$ si, et seulement si, k est multiple de n .

On retiendra que $((\theta(a) = +\infty)) \Leftrightarrow (\varphi_a \text{ injective}) \Leftrightarrow (\forall k \in \mathbb{Z}^*, a^k \neq 1)$
 $\Leftrightarrow (\langle a \rangle \text{ est infini isomorphe à } \mathbb{Z})$

et $(\theta(a) = n \in \mathbb{N}^*) \Leftrightarrow (\text{Ker}(\varphi_a) = n\mathbb{Z}) \Leftrightarrow (\langle a \rangle = \{a^r, 0 \leq r \leq n-1\})$
 $\Leftrightarrow (k \in \mathbb{Z} \text{ et } a^k = 1 \text{ équivaut à } k \equiv 0 \pmod{n})$
 $\Leftrightarrow (n \text{ est le plus petit entier naturel non nul tel que } a^n = 1).$

Pour a d'ordre fini, le groupe $\langle a \rangle$ est cyclique.

Corollaire

Si G est fini d'ordre m , on a alors $a^m = 1$ pour tout $a \in G$.

Théorème

Soient $a, b \in G$ d'ordre fini et $k \in \mathbb{Z}^*$.

- On a $\theta(a^k) = \frac{\theta(a)}{\theta(a) \wedge k}$ (en particulier $\theta(a^{-1}) = \theta(a)$).
- Si k divise $\theta(a)$, on a alors $\theta(a^k) = \frac{\theta(a)}{|k|}$.
- Si k est premier avec $\theta(a)$, on a alors $\theta(a^k) = \theta(a)$.
- Si $ab = ba$, alors ab est d'ordre fini divisant $\theta(a) \vee \theta(b)$. Dans le cas où $\langle a \rangle \cap \langle b \rangle = \{1\}$, on a $\theta(ab) = \theta(a) \vee \theta(b)$. Pour $\theta(a)$ et $\theta(b)$ premiers entre eux, on a $\langle a \rangle \cap \langle b \rangle = \{1\}$ et $\theta(ab) = \theta(a) \vee \theta(b) = \theta(a)\theta(b)$.

Théorème

Pour tout diviseur d de n , le groupe cyclique $\langle a^{\frac{n}{d}} \rangle$ est l'unique sous-groupe d'ordre d de $\langle a \rangle$.

15.5 Groupe des permutations d'un ensemble fini

E est un ensemble fini ayant au moins deux éléments et Id_E est l'application identité sur E . On note $\text{card}(E)$ le cardinal de E et $\mathcal{S}(E)$ le groupe des bijections de E sur lui même.

Définition

Le groupe $\mathcal{S}(E)$ est appelé *groupe des permutations* de E .

Pour $E = \{1, \dots, n\} \subset \mathbb{N}$, on note $\mathcal{S}_n$ le groupe $\mathcal{S}(E)$ et on l'appelle *groupe symétrique* d'indice n .

Pour toute permutation $\sigma \in \mathcal{S}_n$, on note : $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$ pour signifier que σ est la bijection $\sigma : k \in E \mapsto \sigma(k)$.

Pour toute permutation $\sigma \in \mathcal{S}(E)$ et tout entier relatif r , σ^r est la permutation de E définie par :

$$\sigma^r = \begin{cases} \text{Id}_E & \text{si } r = 0 \\ \sigma \circ \cdots \circ \sigma & (r \text{ fois}) \text{ si } r \geq 1 \\ (\sigma^{-r})^{-1} & \text{si } r \leq -1 \end{cases}$$

Définition

Soit r un entier compris entre 2 et $\text{card}(E)$. On appelle *cycle* d'ordre r (ou r -cycle), toute permutation $\sigma \in \mathcal{S}(E)$ qui permute circulairement r éléments de E et laisse fixe les autres, c'est-à-dire qu'il existe une partie $\{x_1, \dots, x_r\}$ de E telle que :

$$\begin{cases} \forall k \in \{1, \dots, r-1\}, \sigma(x_k) = x_{k+1} \\ \sigma(x_r) = x_1 \\ \forall x \in E \setminus \{x_1, \dots, x_r\}, \sigma(x) = x \end{cases}.$$

On notera $\sigma = (x_1, \dots, x_r)$ un tel cycle et on dit que $\{x_1, \dots, x_r\}$ est le *support* de σ et on le note $\text{Supp}(\sigma)$.

L'inverse d'un r -cycle est un r -cycle de même support. Précisément, on a :

$$(x_1, x_2, \dots, x_r)^{-1} = (x_r, x_{r-1}, \dots, x_1).$$

Si $\sigma = (x_1, \dots, x_r)$ est un r -cycle, on a alors $x_k = \sigma^{k-1}(x_1)$ pour tout entier k compris entre 1 et r .

Définition

Une *transposition* est un 2-cycle.

Théorème

Si E, F sont deux ensembles non vides et φ est une bijection de E sur F , les groupes $\mathcal{S}(E)$ et $\mathcal{S}(F)$ sont alors isomorphes.

Donc tout groupe de permutations d'un ensemble E à n éléments est isomorphe au groupe symétrique $\mathcal{S}_n$ des permutations de $\{1, \dots, n\}$.

Théorème

Pour E de cardinal $n \geq 1$, on a $\text{card}(\mathcal{S}(E)) = n!$

Théorème

Le *support* d'une permutation $\sigma \in \mathcal{S}(E)$ est le complémentaire dans E de l'ensemble de ses points fixes, soit l'ensemble :

$$\text{Supp}(\sigma) = \{x \in E, \sigma(x) \neq x\}.$$

Le support d'un cycle $\sigma = (x_1, \dots, x_r)$ est $\{x_1, \dots, x_r\}$.

Théorème

Soient σ, σ' dans $\mathcal{S}(E)$.

- $\sigma(\text{Supp}(\sigma)) = \text{Supp}(\sigma)$; $\text{Supp}(\sigma) = \text{Supp}(\sigma^{-1})$.
- Pour tout $r \in \mathbb{Z}$, on a $\text{Supp}(\sigma^r) \subset \text{Supp}(\sigma)$.
- Si $\text{Supp}(\sigma) \cap \text{Supp}(\sigma') = \emptyset$, on a alors $\sigma \circ \sigma' = \sigma' \circ \sigma$.

Définition

On dit que deux cycles σ et σ' dans $\mathcal{S}(E)$ sont *disjoints*, si leurs supports sont disjoints dans E .

Théorème

Toute permutation $\sigma \in \mathcal{S}(E) \setminus \{\text{Id}_E\}$ se décompose en produit de cycles deux à deux disjoints (le groupe $\mathcal{S}(E)$ est engendré par les cycles). Cette décomposition est unique à l'ordre près.

Si $\sigma = \gamma_1 \cdots \gamma_p$ est une telle décomposition, on a alors la partition

$$\text{Supp}(\sigma) = \bigcup_{k=1}^p \text{Supp}(\gamma_k) \text{ et } \theta(\sigma) = \text{ppcm}(\theta(\gamma_1), \dots, \theta(\gamma_p)).$$

Théorème

Toute permutation $\sigma \in \mathcal{S}(E)$ se décompose en produit de transpositions (le groupe $\mathcal{S}(E)$ est engendré par les transpositions).

Lemme

$\mathcal{S}_n$ est engendré par les $n - 1$ transpositions $(1, k)$ où $2 \leq k \leq n$.

Lemme

$\mathcal{S}_n$ est engendré par les $n - 1$ transpositions $(k, k + 1)$ où k est compris entre 1 et $n - 1$.

Lemme

$\mathcal{S}_n$ est engendré par $(1, 2)$ et $(1, \dots, n)$.

Définition

Soit $\sigma \in \mathcal{S}(E)$. Les *orbites* de σ sont les classes d'équivalence pour la relation d'équivalence définie sur E par :

$$(x \mathcal{R}_\sigma y) \Leftrightarrow (\exists k \in \mathbb{Z}, y = \sigma^k(x))$$

soit les ensembles $\text{Orb}_\sigma(x) = \{\sigma^k(x) ; k \in \mathbb{Z}\}$.

Les orbites deux à deux distinctes forment une partition de E .

Une σ -orbite $\text{Orb}_\sigma(x)$ est réduite à un point si, et seulement si, $\sigma(x) = x$ et les orbites non réduites à un point forment une partition du support de σ .

Théorème

Une permutation $\sigma \in \mathcal{S}(E)$ est un cycle d'ordre $r \geq 2$ si, et seulement si, il n'y a qu'une seule σ -orbite non réduite à un point.

Pour toute permutation $\sigma \in \mathcal{S}(E)$, on note $\mu(\sigma)$ le nombre de σ -orbites distinctes.

Définition

La *signature* d'une permutation $\sigma \in \mathcal{S}(E)$ est l'élément $\varepsilon(\sigma)$ de $\{-1, 1\}$ défini par $\varepsilon(\sigma) = (-1)^{n-\mu(\sigma)}$.

Théorème

Si $\sigma \in \mathcal{S}(E)$ est produit de p transpositions, on a alors $\varepsilon(\sigma) = (-1)^p$ (la parité de p est donc uniquement déterminée par σ).

Théorème

Les seuls morphismes de groupes de $(\mathcal{S}(E), \circ)$ dans $(\mathbb{R}^*, \cdot)$ sont l'application constante égale à 1 et la signature ε . La signature étant surjective de $\mathcal{S}(E)$ sur $\{-1, 1\}$.

Théorème

Pour toute permutation $\sigma \in \mathcal{S}_n$, on a $\varepsilon(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$.

Du théorème précédent, on déduit que $\varepsilon(\sigma) = (-1)^{\nu(\sigma)}$, où :

$$\nu(\sigma) = \text{card} \{(i, j) \in \mathbb{N}^2 ; 1 \leq i < j \leq n \text{ et } \sigma(j) < \sigma(i)\}$$

est le nombre d'inversions de σ . Ce qui nous donne une définition supplémentaire de la signature.

Définition

On dit qu'une permutation $\sigma \in \mathcal{S}(E)$ est *paire* [resp. est *impaire*] si $\varepsilon(\sigma) = 1$ [resp. $\varepsilon(\sigma) = -1$].

Définition

Le *groupe alterné* est le sous-ensemble de $\mathcal{S}(E)$ formé des permutations paires. On le note $\mathcal{A}(E)$.

Pour $E = \{1, \dots, n\}$, on note $\mathcal{A}_n$ le groupe alterné.

Théorème

Pour $n \geq 3$, $\mathcal{A}(E)$ est engendré par les 3-cycles.

15.6 Anneaux

Définition

Soit A un ensemble non vide muni de deux lois de composition internes notées $+$ (une addition) et $\cdot$ (une multiplication). On dit que $(A, +, \cdot)$ est un *anneau* si :

- $(A, +)$ est un groupe commutatif ;
- la loi $\cdot$ est associative, admet un élément neutre et est distributive par rapport à la loi $+$.

Si la loi $\cdot$ est commutative, on dit alors que l'anneau A est *commutatif*.

Si $(A, +, \cdot)$ est un anneau, on note 0 le neutre pour l'addition et 1 le neutre pour la multiplication. L'opposé d'un élément a (soit le symétrique pour l'addition) sera noté $-a$ et on notera $a - b$ pour $a + (-b)$. On écrira souvent ab pour $a \cdot b$ dans un anneau. On supposera que $0 \neq 1$ (sans quoi l'anneau est réduit à $\{0\}$).

Si $A_1, \dots, A_p$ sont des anneaux, on peut alors munir le produit $\prod_{k=1}^p A_k$ d'une structure d'anneau. Si $A_k = A$ pour tout k compris entre 1 et p , on note alors A^p cet anneau produit.

Théorème

Dans un anneau $(A, +, \cdot)$, on a les règles de calcul suivantes, où $a, b, c, a_1, \dots, a_p$ sont des éléments de A , p un entier naturel non nul et n un entier relatif :

- $a \cdot 0 = 0 \cdot a = 0$; $(-a) \cdot b = a \cdot (-b) = -a \cdot b$;
- $(-a) \cdot (-b) = a \cdot b$; $(a - b) \cdot c = a \cdot c - b \cdot c$;
- $a \cdot (b - c) = a \cdot b - a \cdot c$; $n(a \cdot b) = (na) \cdot b = a \cdot (nb)$;
- $a \cdot \left(\sum_{k=1}^p b_k \right) = \sum_{k=1}^p a \cdot b_k$; $\left(\sum_{k=1}^p b_k \right) \cdot a = \sum_{k=1}^p b_k \cdot a$.

Théorème (formule du binôme de Newton)

Soit $(A, +, \cdot)$ un anneau. Dans le cas où a et b commutent, on a pour tout $n \in \mathbb{N}$:

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \text{ et } b^{n+1} - a^{n+1} = (b - a) \sum_{k=0}^n a^k b^{n-k}.$$

15.7 Éléments inversibles dans un anneau

$(A, +, \cdot)$ est un anneau et on note $A^\times$ l'ensemble des éléments de A inversibles pour la multiplication. Comme $1 \neq 0$ dans A , on a $A^\times \subset A \setminus \{0\}$ et cette inclusion peut être stricte.

Théorème

Soit $(A, +, \cdot)$ un anneau. L'ensemble $A^\times$ des éléments inversibles de A est un groupe pour le produit.

$A^\times$ est le groupe des unités de A .

Définition

On dit que $a \in A$ est un *diviseur de 0*, si $a \neq 0$ et s'il existe $b \neq 0$ dans A tel que $a \cdot b = 0$.

Définition

Un anneau est dit *intègre*, s'il est commutatif et n'admet pas de diviseur de 0.

Dans un anneau intègre, on a $a \cdot b = 0$ si, et seulement si, $a = 0$ ou $b = 0$.

15.8 Sous-anneaux

Définition

Soit $(A, +, \cdot)$ un anneau. Un *sous-anneau* de A est une partie non vide B de A telle que $1 \in B$, $(B, +)$ est un sous-groupe de A et B est stable pour la multiplication, c'est-à-dire que pour tous a, b dans B , $a \cdot b$ est aussi dans B .

Un sous-anneau d'un anneau est lui-même un anneau.

Théorème

Soient $(A, +, \cdot)$ un anneau et B une partie non vide de A .
 B est un sous-anneau de A si, et seulement si, $1 \in B$ et :

$$\forall (a, b) \in B^2, a - b \in B \text{ et } a \cdot b \in B.$$

15.9 Morphismes d'anneaux

On désigne par $(A, +, \cdot)$ et $(B, +, \cdot)$ deux anneaux.

Définition

On dit que φ est un *morphisme d'anneaux* de A dans B , si φ est une application de A dans B telle que $\varphi(1) = 1$ et :

$$\forall (a, b) \in A^2, \varphi(a + b) = \varphi(a) + \varphi(b) \text{ et } \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$$

Dans le cas où φ est bijective, on dit que φ est un *isomorphisme* d'anneaux de A sur B . Dans le cas où $A = B$, on dit que φ est un *endomorphisme* de l'anneau A et que c'est un *automorphisme* de l'anneau A si φ est bijective.

Définition

Soit φ un morphisme d'anneaux de A dans B . Le *noyau* de φ est l'ensemble $\text{Ker}(\varphi) = \{x \in A, \varphi(x) = 0\}$ et son *image* est $\text{Im}(\varphi) = \{\varphi(x); x \in A\}$.

$\text{Ker}(\varphi)$ est un sous-anneau de A et $\text{Im}(\varphi)$ un sous-anneau de B .

Définition

La *caractéristique* de l'anneau A est l'entier naturel p tel que $\text{Ker}(\varphi) = p\mathbb{Z}$, où φ est le morphisme d'anneaux de $\mathbb{Z}$ dans A défini par $\varphi(n) = n \cdot 1$ pour tout $n \in \mathbb{Z}$.

Théorème

La caractéristique d'un anneau commutatif et intègre est soit nulle soit un nombre premier.

Dire que la caractéristique d'un anneau commutatif A est nulle équivaut à dire que l'application φ est injective et dans ce cas $\varphi(\mathbb{Z})$ est un sous-anneau de A isomorphe à $\mathbb{Z}$, il est donc en particulier infini. On identifie ce sous-anneau $\varphi(\mathbb{Z})$ à $\mathbb{Z}$. Un anneau de caractéristique nul est donc infini et contient $\mathbb{Z}$ comme sous-anneau.

Si $\varphi : \mathbb{A} \rightarrow \mathbb{B}$ est un morphisme d'anneaux commutatifs, $\text{Ker}(\varphi)$ est alors un sous-groupe de $\mathbb{A}$ et pour tout $a \in \text{Ker}(\varphi)$, tout $b \in \mathbb{A}$, on a $\varphi(ab) = \varphi(a)\varphi(b) = 0_{\mathbb{B}}$, donc $ab \in \text{Ker}(\varphi)$. Ces propriétés se traduisent en disant que $\text{Ker}(\varphi)$ est un idéal de $\mathbb{A}$.

Définition

Un *idéal* d'un anneau commutatif $\mathbb{A}$ est un sous-groupe I de $(\mathbb{A}, +)$ tel que :

$$\forall (a, b) \in I \times \mathbb{A}, ab \in I$$

(on dit que I est *absorbant* pour le produit).

Définition

Soit I un idéal de $\mathbb{A}$. On dit que a est *congru* à b modulo I dans $\mathbb{A}$ si $b - a \in I$. On note alors $a \equiv b \pmod{I}$.

Cette relation de congruence modulo I est une relation d'équivalence sur $\mathbb{A}$ et pour tout $a \in \mathbb{A}$, on note $\bar{a} = \{b \in \mathbb{A}, b \equiv a \pmod{I}\} = a + I$ la classe d'équivalence correspondante. L'ensemble de ces classes d'équivalence modulo I est noté $\frac{\mathbb{A}}{I}$.

Théorème

Il existe une unique structure d'anneau commutatif sur $\frac{\mathbb{A}}{I}$ telle que la surjection canonique $\pi_I : a \in \mathbb{A} \rightarrow \bar{a} = a + I \in \frac{\mathbb{A}}{I}$ soit un morphisme d'anneaux.

Théorème

Les idéaux de $\mathbb{A}$ sont les noyaux de morphismes d'anneaux $\varphi : \mathbb{A} \rightarrow \mathbb{B}$ où $\mathbb{B}$ est un anneau commutatif.

15.10 Corps

Définition

Soit $\mathbb{K}$ un ensemble non vide muni de deux lois de composition internes notées $+$ et $\cdot$. On dit que $(\mathbb{K}, +, \cdot)$ est un *corps* si $(\mathbb{K}, +, \cdot)$ est un anneau commutatif et tous les éléments de $\mathbb{K} \setminus \{0\}$ sont inversibles pour la multiplication.

Pour un corps on note aussi $\mathbb{K}^*$ l'ensemble $\mathbb{K}^\times = \mathbb{K} \setminus \{0\}$.

Dans un corps, on notera $-a$ l'opposé d'un élément a et a^{-1} ou $\frac{1}{a}$ l'inverse d'un élément non nul a .

Dans un corps tout élément non nul est simplifiable et il n'y a pas de diviseurs de 0. Un corps est donc en particulier un anneau intègre.

Les règles de calcul valables dans un anneau le sont aussi dans un corps avec de plus l'équivalence : $(ab = 0) \Leftrightarrow (a = 0 \text{ ou } b = 0)$.

Dans un corps, pour $(a, b) \in \mathbb{K}^* \times \mathbb{K}$, on écrira $a^{-1} \cdot b = \frac{b}{a}$.

Définition

Soit $(\mathbb{K}, +, \cdot)$ un corps. On dit qu'une partie $\mathbb{L}$ de $\mathbb{K}$ est un *sous-corps* de $\mathbb{K}$ si $\mathbb{L}$ est un sous-anneau de $\mathbb{K}$ et $\mathbb{L}^* = \mathbb{L} \setminus \{0\}$ est stable par passage à l'inverse, c'est-à-dire que pour tout $x \in \mathbb{L}^*$, x^{-1} est dans $\mathbb{L}^*$.

On vérifie facilement qu'un sous-corps d'un corps est un corps.

Théorème

Soit $(\mathbb{K}, +, \cdot)$ un corps et $\mathbb{L}$ une partie non vide de $\mathbb{K}$. $\mathbb{L}$ est un sous-corps de $\mathbb{K}$ si, et seulement si $1 \in \mathbb{L}$ et :

$$\forall (x, y) \in \mathbb{L} \times \mathbb{L}^*, x - y \in \mathbb{L} \text{ et } xy^{-1} \in \mathbb{L}.$$

Si $\mathbb{L}$ est un sous-corps d'un corps $\mathbb{K}$, on dit alors que $\mathbb{K}$ est une *extension* de $\mathbb{L}$.

15.11 Morphismes de corps

On désigne par $(\mathbb{K}, +, \cdot)$ et $(\mathbb{L}, +, \cdot)$ deux corps.

Définition

On dit que $\varphi : \mathbb{K} \rightarrow \mathbb{L}$ est un *morphisme de corps*² si $\varphi(1_{\mathbb{K}}) = 1_{\mathbb{L}}$ et :

$$\forall (a, b) \in \mathbb{K}^2, \varphi(a + b) = \varphi(a) + \varphi(b) \text{ et } \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b).$$

Pour φ bijective, on dit que φ est un *isomorphisme*. Pour $\mathbb{K} = \mathbb{L}$, on dit que φ est un *endomorphisme* du corps $\mathbb{K}$ et que c'est un *automorphisme* si φ est bijective.

On a, pour un tel morphisme, $\varphi(0) = 0$, $\varphi(1) = 1$, $\varphi(-a) = -\varphi(a)$ pour tout $a \in \mathbb{K}$ et $\varphi(a^{-1}) = \varphi(a)^{-1}$ pour tout $a \in \mathbb{K}^*$.

2. Un morphisme de corps est en fait un morphisme d'anneaux.

Arithmétique dans $\mathbb{Z}$

16

16.1 Théorème de division euclidienne dans $\mathbb{Z}$

Définition

On dit que l'entier relatif d *divise* l'entier relatif a , ou que d est un *diviseur* de a , ou que a est *divisible* par d , ou que a est un *multiple* de d , s'il existe $q \in \mathbb{Z}$ tel que $a = qd$.

La relation de divisibilité est une relation d'ordre non totale sur $\mathbb{N}$.

Pour tout entier relatif n , on note $n\mathbb{Z} = \{nq ; q \in \mathbb{Z}\}$ l'ensemble de tous les multiples de n .

Définition

Soient $n \in \mathbb{N}$ et a, b deux entiers relatifs. On dit que a est *congru* à b modulo n , si n divise $a - b$. On note $a \equiv b [n]$ ou $a \equiv b \pmod{n}$.

Théorème

Soit $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$. Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que $a = bq + r$ et $0 \leq r < |b|$.

a est le *dividende*, b le *diviseur*, q le *quotient* et r le *reste* dans la division euclidienne de a par b .

Corollaire

Soit $b \geq 2$ un entier. Pour tout $n \in \mathbb{N}^*$, il existe un unique entier p et un unique $(p+1)$ -uplet $(n_0, \dots, n_p) \in \mathbb{N}^{p+1}$ tels que $n_p \neq 0$, $0 \leq n_k \leq b-1$ pour tout $k \in \{0, 1, \dots, p\}$ et $n = \sum_{k=0}^p n_k b^k$.

Avec les notations du théorème précédent, on dit que $\sum_{k=0}^p n_k b^k$ est la représentation en base b de l'entier n .

16.2 Plus grand commun diviseur et plus petit commun multiple

Théorème

Soient a, b deux entiers relatifs non tous deux nuls. Il existe un unique entier naturel δ tel que $a\mathbb{Z} + b\mathbb{Z} = \delta\mathbb{Z}$. Cet entier s'écrit $\delta = au + bv$ avec $(u, v) \in \mathbb{Z}^2$ et c'est le plus grand entier naturel qui divise a et b .

Définition

Soient a, b deux entiers relatifs non tous deux nuls. On appelle *plus grand commun diviseur* de a et b le plus grand entier naturel qui divise a et b . On le note $\text{pgcd}(a, b)$ ou $a \wedge b$.

On définit de manière analogue le pgcd d'une famille $a_1, \dots, a_p$ formée de p entiers non tous nuls comme le plus grand des diviseurs communs à $a_1, \dots, a_p$. On le note $\text{pgcd}(a_1, \dots, a_p)$ ou $a_1 \wedge \dots \wedge a_p$ et c'est un entier supérieur ou égal à 1.

Théorème

Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls. Il existe un unique entier naturel δ tel que $a_1\mathbb{Z} + \dots + a_p\mathbb{Z} = \delta\mathbb{Z}$. Cet entier s'écrit $\delta = \sum_{k=1}^p u_k a_k$ avec $(u_1, \dots, u_p) \in \mathbb{Z}^p$ et c'est le plus grand entier naturel qui divise $a_1, \dots, a_p$.

L'égalité $\delta = \sum_{k=1}^p u_k a_k$ s'appelle *identité de Bézout*.

Théorème

Pour a, b, c entiers relatifs, on a :

- pour $(a, b) \neq (0, 0)$, $a \wedge b \in \mathbb{N}^*$ et pour $a \neq 0$, $a \wedge 0 = a \wedge a = |a|$;
- $a \wedge 1 = 1$;
- pour $(a, b) \neq (0, 0)$, $a \wedge b = b \wedge a = |a| \wedge |b| = |a| \wedge b = a \wedge |b|$;
- pour $b \neq 0$, on a $a \wedge b = |b|$ si, et seulement si, b divise a ;
- pour $(a, b) \neq (0, 0)$ et $c \neq 0$, $(ac) \wedge (bc) = |c| (b \wedge a)$;
- si $d \neq 0$ est un diviseur commun de a et b non tous deux nuls, alors $\frac{a}{d} \wedge \frac{b}{d} = \frac{a \wedge b}{|d|}$;
- pour $(a, b, c) \neq (0, 0, 0)$, on a $a \wedge (b \wedge c) = (a \wedge b) \wedge c$ (associativité).

Définition

Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls. On dit qu'ils sont *premiers entre eux dans leur ensemble*¹, si leur pgcd est égal à 1.

Théorème

Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls et $\delta = a_1 \wedge \dots \wedge a_p$. Il existe des entiers $a'_1, \dots, a'_p$ premiers entre eux dans leur ensemble tels que $a_k = \delta a'_k$ pour tout k compris entre 1 et p .

Théorème de Bézout

Des entiers relatifs $a_1, \dots, a_p$ non tous nuls sont premiers entre eux dans leur ensemble si, et seulement si, il existe deux entiers relatifs $u_1, \dots, u_p$ tels que $\sum_{k=1}^p u_k a_k = 1$.

Corollaire

Soient a, b et c des entiers relatifs non nuls. Si c est premier avec a , on a alors $a \wedge b = a \wedge (bc)$.

1. pour $p = 2$, on dit simplement que a_1 et a_2 sont premiers entre eux.

Corollaire

Soient $a_1, \dots, a_p$ et c des entiers relatifs non nuls. Si c est premier avec chacun des a_k , il est alors premier avec leur produit $\prod_{k=1}^p a_k$.

Théorème de Gauss

Soient a , b et c des entiers relatifs non nuls. Si a divise bc et a est premier avec b , alors a divise c .

Corollaire

Tout nombre rationnel non nul r s'écrit de manière unique $r = \frac{p}{q}$ avec $p \in \mathbb{Z}^*$ et $q \in \mathbb{N}^*$ premiers entre eux.

Corollaire

Si un entier relatif non nul n est divisible par des entiers $a_1, \dots, a_p$ deux à deux premiers entre eux, il est alors divisible par leur produit.

Théorème

Soient a et b deux entiers relatifs. Il existe un unique entier naturel μ tel que $a\mathbb{Z} \cap b\mathbb{Z} = \mu\mathbb{Z}$. Si $a = 0$ ou $b = 0$, alors $\mu = 0$. Si $a \neq 0$ et $b \neq 0$, alors μ est le plus petit entier naturel non nul multiple de a et de b .

Définition

Soient a, b deux entiers relatifs. On appelle *plus petit commun multiple* de a et b le plus petit entier naturel multiple de a et b . On le note $\text{ppcm}(a, b)$ ou $a \vee b$.

Théorème

Pour a, b, c entiers relatifs, on a :

- pour a, b non nuls, $a \vee b \in \mathbb{N}^*$ et pour a non nul, $a \vee 1 = a \vee a = |a|$;
- $a \vee b = b \vee a = |a| \vee |b| = |a| \vee b = a \vee |b|$;
- $a \vee b = |b|$ si, et seulement si, b est multiple de a ;
- $(ac) \vee (bc) = |c| (b \vee a)$;
- si $d \in \mathbb{Z}^*$ est un diviseur commun de a et b , alors $\frac{a}{d} \vee \frac{b}{d} = \frac{a \vee b}{|d|}$;
- pour $(a, b, c) \neq (0, 0, 0)$, $a \vee (b \vee c) = (a \vee b) \vee c$ (associativité).

Théorème

Pour tout couple (a, b) d'entiers relatifs, on a $(a \wedge b) (a \vee b) = |ab|$.

On définit de manière analogue le ppcm d'une famille $a_1, \dots, a_p$ formée de p entiers non tous nuls comme le plus petit des multiples communs à $a_1, \dots, a_p$. On le note $\text{ppcm}(a_1, \dots, a_p)$ ou $a_1 \vee \dots \vee a_p$ et c'est un entier supérieur ou égal à 1.

Théorème

Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls. Il existe un unique entier naturel μ tel que $a_1\mathbb{Z} \cap \dots \cap a_p\mathbb{Z} = \mu\mathbb{Z}$. Cet entier μ est le plus petit entier naturel divisible par $a_1, \dots, a_p$.

16.3 L'algorithme d'Euclide

Théorème

Soient a, b deux entiers naturels non nuls et r le reste dans la division euclidienne de a par b . On a alors $a \wedge b = b \wedge r$.

Le principe de l'algorithme d'Euclide est le suivant pour $a > b$ dans $\mathbb{N}^*$ (par symétrie, on peut supposer que $a \geq b$ et pour $a = b$, on a $a \wedge a = a$).

On note $r_0 = b$ et on désigne par r_1 le reste dans la division euclidienne de a par b . On a alors $0 \leq r_1 < r_0$ et $a \wedge b = r_0 \wedge r_1$.

Si $r_1 = 0$ alors $r_0 \wedge r_1 = r_0 = b$ et c'est terminé, sinon on désigne par r_2 le reste dans la division euclidienne de r_0 par r_1 et on a $0 \leq r_2 < r_1$ et $a \wedge b = r_0 \wedge r_1 = r_1 \wedge r_2$.

Si $r_2 = 0$ alors $r_1 \wedge r_2 = r_1$ et c'est terminé. Sinon on continue.

On définit donc ainsi une suite d'entiers $(r_n)_{n \geq 0}$ par :

- $r_0 = b$;
- r_1 est le reste dans la division euclidienne de a par b ; on a donc $0 \leq r_1 < b$;
- pour $n \geq 2$, si $r_{n-1} = 0$ alors $r_n = 0$, sinon r_n est le reste dans la division euclidienne de r_{n-2} par r_{n-1} et on a $0 \leq r_n < r_{n-1}$. On a $r_n \leq r_{n-1}$ l'égalité étant réalisée si, et seulement si, les deux termes sont nuls.

Il existe un entier $p \geq 1$ tel que $r_p = 0 < r_{p-1} < \dots < r_1 < r_0$ et :

$$a \wedge b = r_0 \wedge r_1 = \dots = r_{p-1} \wedge r_p = r_{p-1}$$

c'est-à-dire que $a \wedge b$ est le dernier reste non nul dans cette suite de divisions. On a donc construit deux suites d'entiers $(r_n)_{0 \leq n \leq p}$ et $(q_n)_{1 \leq n \leq p}$ de la manière suivante :

$$\begin{cases} a = q_1 r_0 + r_1 & (0 < r_1 < r_0 = b) \\ r_0 = q_2 r_1 + r_2 & (0 < r_2 < r_1) \\ r_1 = q_3 r_2 + r_3 & (0 < r_3 < r_2) \\ \vdots \\ r_{p-3} = q_{p-1} r_{p-2} + r_{p-1} & (0 < r_{p-1} < r_{p-2}) \\ r_{p-2} = q_p r_{p-1} + r_p & (r_p = 0) \end{cases}$$

Il existe des entiers u_n et v_n tels que $r_n = au_n + bv_n$. Les suites $(r_n)_{0 \leq n \leq p-1}$ $(u_n)_{0 \leq n \leq p-1}$ et $(v_n)_{0 \leq n \leq p-1}$ vérifient la relation de récurrence $x_n = x_{n-2} - q_n x_{n-1}$ pour $2 \leq n \leq p-1$ avec les conditions initiales $(r_0, r_1) = (b, 1)$, $(u_0, u_1) = (0, 1)$, $(v_0, v_1) = (1, -q_1)$, où q_1, r_1 sont, respectivement, le quotient et le reste dans la division euclidienne de a par b .

16.4 L'ensemble $\mathcal{P}$ des nombres premiers

Définition

On dit qu'un entier naturel p est *premier*, s'il est supérieur ou égal à 2 et si les seuls diviseurs positifs de p sont 1 et p .

On note $\mathcal{P}$ l'ensemble de tous les nombres premiers.

Théorème

Tout entier relatif $n \in \mathbb{Z} \setminus \{-1, 0, 1\}$ a au moins un diviseur premier.

Un entier naturel non premier s'écrit donc $n = pq$ avec $p \geq 2$ premier et $q \geq 2$. On dit alors qu'il est composé.

Théorème

Tout entier $n \geq 2$ qui est composé a au moins un diviseur premier p tel que $2 \leq p \leq \sqrt{n}$.

Théorème

Soit p un nombre premier et $r \geq 2$ un entier. Si p divise le produit $n_1 \cdots n_r$ de r entiers naturels non nuls, il divise alors l'un des n_k .

Théorème d'Euclide

L'ensemble $\mathcal{P}$ des nombres premiers est infini.

16.5 Décomposition en facteurs premiers

Définition

Soient p un nombre premier et $n \in \mathbb{Z}^*$. La *valuation p -adique* de n est l'entier $\nu_p(n) = \max\{k \in \mathbb{N}, p^k \mid n\}$.

La valuation p -adique d'un entier n est l'exposant de p dans la décomposition en facteurs premiers du théorème suivant.

Théorème (décomposition en facteurs premiers d'un entier)

Soit $n \in \mathbb{N}^*$. Alors $n = \prod_{p \in \mathcal{P}} p^{\nu_p(n)}$. Ce produit est appelé *décomposition en facteurs premiers* ou *décomposition primaire* de n .

Dans le produit ci-dessus, tous les $\nu_p(n)$ sont nuls sauf un nombre fini d'entre eux de sorte que ce produit ne contient qu'un nombre fini de facteurs différents de 1.

Théorème

Soient a et b deux entiers relatifs non nuls. Alors

- si p est premier, $\nu_p(ab) = \nu_p(a) + \nu_p(b)$;
- a divise b si, et seulement si, pour tout p premier, $\nu_p(a) \leq \nu_p(b)$;
- $a \wedge b = \prod_{p \in \mathcal{P}} p^{\min\{\nu_p(a), \nu_p(b)\}}$ et $a \vee b = \prod_{p \in \mathcal{P}} p^{\max\{\nu_p(a), \nu_p(b)\}}$ où $\mathcal{P}$ est l'ensemble des nombres premiers.

16.6 Les théorèmes de Fermat et de Wilson

Lemme

Un entier naturel $p \geq 2$ est premier si, et seulement si, pour tout entier k compris entre 1 et $p-1$, p divise $\binom{p}{k}$.

Théorème (petit théorème de Fermat)

Soit p un nombre premier. Pour tout entier relatif n , on a $n^p \equiv n \pmod{p}$ et pour tout entier relatif n non multiple de p , on a $n^{p-1} \equiv 1 \pmod{p}$.

Théorème de Wilson

Un entier p supérieur ou égal à 2 est premier si, et seulement si, $(p-1)!$ est congru à -1 modulo p .

16.7 Les anneaux $\mathbb{Z}/n\mathbb{Z}$ et l'indicatrice d'Euler

La relation de congruence est une relation d'équivalence compatible avec l'addition et la multiplication sur $\mathbb{Z}$, ce qui signifie que :

$$(a \equiv b \pmod{n}, c \equiv d \pmod{n}) \Rightarrow (a + c \equiv b + d \pmod{n}, ac \equiv bd \pmod{n}).$$

Cette compatibilité permet de munir l'ensemble $\mathbb{Z}/n\mathbb{Z}$ des classes d'équivalence modulo n d'une structure d'anneau. Pour $n = 0$, $\mathbb{Z}/0\mathbb{Z} = \{\{a\} ; a \in \mathbb{Z}\}$ est identifié à $\mathbb{Z}$. Pour $n = 1$, on a $\mathbb{Z}/1\mathbb{Z} = \{\mathbb{Z}\} = \{\bar{0}\}$.

Théorème

Pour tout $n \in \mathbb{N}^*$, $\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$ est de cardinal égal à n et en bijection avec l'ensemble de tous les restes modulo n .

On désigne par π_n la surjection canonique de $\mathbb{Z}$ sur $\mathbb{Z}/n\mathbb{Z}$, à savoir l'application qui associe à tout entier relatif sa classe modulo n . Tout antécédent par π_n d'un élément x de $\mathbb{Z}/n\mathbb{Z}$ est appelé un *représentant* de x .

Théorème

Il existe une unique structure d'anneau commutatif sur $\mathbb{Z}/n\mathbb{Z}$ telle que la surjection canonique π_n soit un morphisme d'anneaux.

Définition

Deux polynômes P et Q à coefficients entiers sont *congrus* modulo un nombre premier p , s'ils sont de même degré et tous leurs coefficients sont égaux modulo p .

Théorème

p est premier si, et seulement si, il existe un entier relatif n premier avec p tel que $(X + n)^p$ soit congru à $X^p + n$ modulo p .

Définition

On dit que $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ est *inversible* s'il existe $\bar{b} \in \mathbb{Z}/n\mathbb{Z}$ tel que $\bar{a}\bar{b} = \bar{1}$.

L'ensemble $(\mathbb{Z}/n\mathbb{Z})^\times$ des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$ est un groupe pour la loi multiplicative.

Théorème

Soit a un entier relatif. Les propriétés suivantes sont équivalentes :

$$\begin{aligned} (\bar{a} \text{ est inversible dans } \mathbb{Z}/n\mathbb{Z}) &\Leftrightarrow (a \text{ est premier avec } n) \\ &\Leftrightarrow (\bar{a} \text{ est un générateur de } \mathbb{Z}/n\mathbb{Z}). \end{aligned}$$

Définition

On appelle *fonction indicatrice d'Euler* la fonction qui associe à tout entier naturel non nul n , le nombre, noté $\varphi(n)$, d'entiers compris entre 1 et n qui sont premiers avec n .

Théorème d'Euler

Pour tout entier relatif a premier avec n , on a $a^{\varphi(n)} \equiv 1 \pmod{n}$.

Pour n premier, on a $\varphi(n) = n - 1$ et le théorème d'Euler devient le petit théorème de Fermat.

Théorème

Pour $n \geq 2$, on a les équivalences :

$$(n \text{ est premier}) \Leftrightarrow (\mathbb{Z}/n\mathbb{Z} \text{ est un corps}) \Leftrightarrow (\mathbb{Z}/n\mathbb{Z} \text{ est intègre}).$$

Théorème des restes chinois

Les entiers n et m sont premiers entre eux si, et seulement si, les anneaux $\mathbb{Z}/nm\mathbb{Z}$ et $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ sont isomorphes. Dans ce cas, on a l'égalité $\varphi(nm) = \varphi(n)\varphi(m)$.

Théorème

Si l'entier $n \geq 2$ a pour décomposition en facteurs premiers $n = \prod_{i=1}^r p_i^{\alpha_i}$ avec $2 \leq p_1 < \dots < p_r$ premiers et les α_i entiers naturels non nuls, on a alors :

$$\varphi(n) = \prod_{i=1}^r p_i^{\alpha_i-1} (p_i - 1) = n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right).$$

17.1 Construction de $\mathbb{C}$

On définit sur $\mathbb{R}^2$ les opérations d'addition et de multiplication suivantes, en désignant par $z = (x, y)$ et $z' = (x', y')$ deux éléments quelconques de $\mathbb{R}^2$: $z + z' = (x + x', y + y')$ et $z \cdot z' = (xx' - yy', xy' + yx')$. On note $\mathbb{C}$ l'ensemble $\mathbb{R}^2$ muni de ces deux opérations et on l'appelle *ensemble des nombres complexes*. C'est un corps.

L'égalité de deux nombres complexes $z = (x, y)$ et $z' = (x', y')$ est réalisée si, et seulement si, on a $x = x'$ et $y = y'$. En particulier, pour $y = y' = 0$, on a $(x, 0) = (x', 0)$ si, et seulement si, $x = x'$, ce qui signifie que l'application $x \in \mathbb{R} \mapsto (x, 0) \in \mathbb{C}$ est injective, ce qui permet de réaliser une bijection de $\mathbb{R}$ sur le sous ensemble $\mathbb{R}'$ de $\mathbb{C}$ formé des couples $(x, 0)$. Cette bijection permet d'identifier $\mathbb{R}$ à $\mathbb{R}'$, ce qui signifie qu'un nombre réel x est identifié à son image $(x, 0)$ dans $\mathbb{C}$. Cette identification est compatible avec les opérations d'addition et de multiplication des réels dans le sens où :

$$x + x' = (x, 0) + (x', 0) = (x + x', 0) = x + x' \text{ et } xx' = (x, 0)(x', 0) = (xx', 0) = xx'.$$

On note $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ et $(\mathbb{C}^*, \cdot)$ est un groupe commutatif.

En notant $i = (0, 1)$ on a $i^2 = -1$, soit $\frac{1}{i} = -i$. Le nombre complexe $-i$ est aussi solution de l'équation $x^2 + 1 = 0$ et $i, -i$ sont les seules solutions de cette équation.

Théorème

Tout nombre complexe s'écrit de manière unique $z = x + iy$, où $(x, y) \in \mathbb{R}^2$.

Définition

Avec les notations du théorème qui précède, x est la *partie réelle* de z et y sa *partie imaginaire*, ce qui se note $x = \operatorname{Re}(z)$ et $y = \operatorname{Im}(z)$. On dit que z est un *imaginaire pur*, si $\operatorname{Re}(z) = 0$.

17.2 Conjugué et module d'un nombre complexe**Définition**

Le *conjugué* de $z = x + iy \in \mathbb{C}$ est le nombre complexe $\bar{z} = x - iy$ et son *module* est le réel $|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$.

Théorème

Pour tous nombres complexes z et z' , on a :

- $\bar{\bar{z}} = z$; $\overline{z + z'} = \bar{z} + \bar{z'}$; $\overline{zz'} = \bar{z}\bar{z'}$; $z + \bar{z} = 2\operatorname{Re}(z)$; $z - \bar{z} = 2i\operatorname{Im}(z)$;
- $(z \in \mathbb{R}) \Leftrightarrow (z = \bar{z})$; z est imaginaire pur si, et seulement si, $\bar{z} = -z$;
- $z\bar{z} = (\operatorname{Re}(z))^2 + (\operatorname{Im}(z))^2 = |z|^2 \in \mathbb{R}_+$.

Théorème

Pour tous nombres complexes z et z' , on a :

- $|z| \geq 0$ et $|z| = 0$ si, et seulement si, $z = 0$;
- $|z| = |\bar{z}|$; $|zz'| = |z||z'|$; $|z + z'|^2 = |z|^2 + 2\operatorname{Re}(z\bar{z'}) + |z'|^2$;
- pour tous nombres complexes $z_1, \dots, z_n$, on a $\left| \prod_{k=1}^n z_k \right| = \prod_{k=1}^n |z_k|$;
- si $z \neq 0$, alors $\frac{1}{z} = \frac{\bar{z}}{|z|^2} = \frac{x-iy}{x^2+y^2}$; si $z' \neq 0$, alors $\left| \frac{z}{z'} \right| = \frac{|z|}{|z'|}$;
- $|z| = 1$ si, et seulement si, $\frac{1}{z} = \bar{z}$;
- $|\operatorname{Re}(z)| \leq |z|$, l'égalité étant réalisée si, et seulement si, z est réel ;
- $|\operatorname{Im}(z)| \leq |z|$, l'égalité étant réalisée si, et seulement si, z est imaginaire pur.

Théorème (inégalité triangulaire)

Pour tous nombres complexes z et z' , on a $|z + z'| \leq |z| + |z'|$, l'égalité étant réalisée si, et seulement si, z et z' sont positivement liées sur $\mathbb{R}$ (c'est-à-dire tels que $z' = 0$ ou $z' \neq 0$ et $\frac{z}{z'} \in \mathbb{R}_+$), ce qui équivaut à dire que $z\bar{z'} \in \mathbb{R}_+$.

Corollaire

Pour tous nombres complexes z et z' , on a $||z| - |z'||| \leq |z - z'| \leq |z| + |z'|$.

Théorème

Pour toute suite $(z_k)_{1 \leq k \leq n}$ de nombres complexes, on a $\left| \sum_{k=1}^n z_k \right| \leq \sum_{k=1}^n |z_k|$,
l'égalité étant réalisée si, et seulement si, il existe des réels λ_k tels que $z_k = \lambda_k z_1$ pour $k = 1, \dots, n$.

17.3 Les équations de degré 2

Théorème

Pour tout nombre réel non nul a , l'équation $x^2 + a = 0$ a deux solutions données par $-\sqrt{-a}$ et $\sqrt{-a}$ pour $a < 0$ et par $-i\sqrt{a}$ et $i\sqrt{a}$ pour $a > 0$.

Définition

Si α est un nombre complexe, on dit alors que le nombre complexe u est une *racine carrée* de α , si $u^2 = \alpha$.

Théorème

Tout complexe non nul a exactement deux racines carrées qui sont opposées.

Théorème

Toute équation de degré 2, $az^2 + bz + c = 0$ ($a \neq 0$) a deux solutions complexes distinctes ou confondues.

Avec les notations du théorème la quantité $\delta = b^2 - 4ac$ est appelé *discriminant* de l'équation $az^2 + bz + c = 0$.

Théorème

Étant donné des nombres complexes a , b et c avec $a \neq 0$, les nombres complexes z_1 et z_2 sont les racines de l'équation $az^2 + bz + c = 0$ si, et seulement si, on a $z_1 + z_2 = -\frac{b}{a}$ et $z_1 z_2 = \frac{c}{a}$.

Pour a, b, c réels, il en est de même de δ et on distingue trois cas de figure :

- soit $\delta = 0$ et $-\frac{b}{2a}$ est l'unique solution réelle de cette équation ;
- soit $\delta > 0$ et $\frac{-b \pm \sqrt{\delta}}{2a}$ sont les deux solutions réelles de cette équation ;
- soit $\delta < 0$ et $\frac{-b \pm i\sqrt{-\delta}}{2a}$ sont les deux solutions complexes non réelles de cette équation.

17.4 Arguments d'un nombre complexe

Définition

Pour tout $z \in \mathbb{C}^*$, le réel $\theta \in [-\pi, \pi[$ tel que $\frac{z}{|z|} = \cos(\theta) + i \sin(\theta)$ est son *argument principal*. On dit qu'un réel θ est un *argument* du nombre complexe non nul z , si $\frac{z}{|z|} = \cos(\theta) + i \sin(\theta)$.

Théorème

En désignant par z et z' des nombres complexes non nuls, λ un réel non nul et n un entier relatif, on a :

- $\arg(\bar{z}) \equiv -\arg(z) \pmod{2\pi}$; $\arg(zz') \equiv \arg(z) + \arg(z') \pmod{2\pi}$;
- $\arg\left(\frac{z}{z'}\right) \equiv \arg(z) - \arg(z') \pmod{2\pi}$; $\arg(z^n) \equiv n \arg(z) \pmod{2\pi}$;
- pour $\lambda > 0$, on a $\arg(\lambda z) \equiv \arg(z) \pmod{2\pi}$ et pour $\lambda < 0$, on a $\arg(\lambda z) \equiv \arg(z) + \pi \pmod{2\pi}$;
- z est réel si, et seulement si, $\arg(z) \equiv 0 \pmod{\pi}$;
- z est imaginaire pur si, et seulement si, $\arg(z) \equiv \frac{\pi}{2} \pmod{\pi}$.

En désignant par ψ l'application qui associe à tout réel θ le nombre complexe $\psi(\theta) = \cos(\theta) + i \sin(\theta)$ on réalise une application surjective de $\mathbb{R}$ sur l'ensemble des nombres complexes de module 1.

Théorème

Avec les notations qui précèdent, on a $\psi(0) = 1$ et $\psi(\theta + \theta') = \psi(\theta)\psi(\theta')$ pour tous réels θ et θ' .

La fonction ψ vérifie donc la même équation fonctionnelle que la fonction exponentielle réelle. Cette remarque justifie la notation $\psi(\theta) = e^{i\theta}$. On a donc en résumé la notation $e^{i\theta} = \cos(\theta) + i \sin(\theta)$ pour tout réel θ , ce qui définit une fonction

2π -périodique surjective de $\mathbb{R}$ sur l'ensemble Γ des nombres complexes de module 1 avec les propriétés suivantes, où θ, θ' sont des réels et n un entier relatif¹ :

$$\left\{ \begin{array}{l} e^{i \cdot 0} = e^0 = 1; \quad e^{i(\theta+\theta')} = e^{i\theta} e^{i\theta'}; \quad \frac{1}{e^{i\theta}} = e^{-i\theta} = \overline{e^{i\theta}}; \quad (e^{i\theta})^n = e^{in\theta} \\ (e^{i\theta} = e^{i\theta'}) \Leftrightarrow (\exists k \in \mathbb{Z}, \theta' = \theta + 2k\pi) \\ \cos(\theta) = \operatorname{Re}(e^{i\theta}) = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ et } \sin(\theta) = \operatorname{Im}(e^{i\theta}) = \frac{e^{i\theta} - e^{-i\theta}}{2i}. \end{array} \right.$$

On a en particulier les valeurs suivantes $e^{i\pi} = -1$, $e^{i\frac{\pi}{2}} = i$, les égalités $e^{i\theta} = 1$, $e^{i\theta} = -1$ et $e^{i\theta} = i$ étant réalisées respectivement si, et seulement si, $\theta = 2k\pi$, $\theta = (2k+1)\pi$ et $\theta = \frac{\pi}{2} + 2k\pi$, où k est un entier relatif.

Un nombre complexe non nul peut donc s'écrire sous la forme $z = \rho e^{i\theta}$ où ρ est un réel strictement positif uniquement déterminé, c'est le module de z , et θ est un argument de z . Cette écriture est l'*écriture polaire* (ou trigonométrique, ou la forme exponentielle) de z .

Théorème

Soient a, b des réels non tous deux nuls. Il existe un unique couple (ρ, θ) de réels dans $\mathbb{R}_+^* \times [-\pi, \pi[$ tel que $a \cos(x) + b \sin(x) = \rho \cos(x - \theta)$ pour tout réel x .

Avec les notations du théorème précédent, ρ est le module de $a + ib$ et θ son argument principal.

Corollaire

Soient a et b des réels non tous deux nuls. Les solutions de l'équation $a \cos(x) + b \sin(x) = 0$ sont les réels $x = \theta + \pi/2 + k\pi$, où $\theta \in [-\pi, \pi[$ est l'argument principal de $a + ib$ et k un entier relatif.

17.5 Racines n-ièmes d'un nombre complexe

Définition

Étant donné $\alpha \in \mathbb{C}$ et $n \in \mathbb{N}^*$, on appelle *racine n-ième* de α tout nombre complexe z tel que $z^n = \alpha$.

1. La formule $(e^{i\theta})^n = e^{in\theta}$ qui s'écrit aussi $(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta)$ est appelée *formule de Moivre*.

Définition

Étant donné $n \in \mathbb{N}^*$, on appelle *racine n -ième de l'unité* toute racine n -ième de 1.

Théorème

Pour tout $n \in \mathbb{N}^*$, il y a exactement n racines n -ièmes de l'unité données par $e^{\frac{2ik\pi}{n}} = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right)$ où k est compris entre 0 et $n - 1$.

Théorème

Pour tout $n \in \mathbb{N}^*$ et tout $z \in \mathbb{C}$, on a $z^n - 1 = \prod_{k=0}^{n-1} (z - \omega_k)$, où les $\omega_k = e^{\frac{2ik\pi}{n}}$, pour k compris entre 0 et $n - 1$, sont les racines n -ièmes de l'unité.

Les racines cubiques de l'unité sont 1, $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ et $\bar{j} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$.

Corollaire

Soit $n \in \mathbb{N}^*$. Tout nombre complexe non nul $\alpha = \rho e^{i\theta}$ a exactement n racines n -ièmes données par $\sqrt[n]{\rho} e^{i\frac{\theta}{n}} e^{\frac{2ik\pi}{n}}$ où k est compris entre 0 et $n - 1$.

$\mathbb{K}$ est un corps dont les éléments sont appelés *scalaires*.

Définition

Un $\mathbb{K}$ -*espace vectoriel* est un ensemble E muni d'une addition interne et d'une multiplication externe :

$$(x, y) \in E^2 \mapsto x + y \in E; (\lambda, x) \in \mathbb{K} \times E \mapsto \lambda x \in E.$$

telles que $(E, +)$ soit un groupe commutatif et pour tous scalaires λ, μ et tous vecteurs x, y on ait :

$$\lambda(x + y) = \lambda x + \lambda y; (\lambda + \mu)x = \lambda x + \mu x; \lambda(\mu x) = (\lambda\mu)x; 1 \cdot x = x.$$

Pour simplifier, on dira *espace vectoriel* pour $\mathbb{K}$ -espace vectoriel. Les éléments d'un espace vectoriel sont appelés *vecteurs*.

Dans un espace vectoriel l'élément neutre pour l'addition est noté 0 et on dit que c'est le *vecteur nul* et le symétrique d'un vecteur x est noté $-x$ et on dit que c'est l'opposé de x . Pour x et y dans E , la somme $x + (-y)$ est simplement notée $x - y$.

Si $(E_i)_{i \in I}$ est une famille de $\mathbb{K}$ -espaces vectoriels, alors le produit cartésien $\prod_{i \in I} E_i$ muni des lois : $(x_i)_{i \in I} + (y_i)_{i \in I} = (x_i + y_i)_{i \in I}$ et $\lambda(x_i)_{i \in I} = (\lambda x_i)_{i \in I}$ est aussi un $\mathbb{K}$ -espace vectoriel.

Définition

Soient E un espace vectoriel, n un entier naturel non nul et $x, y, x_1, \dots, x_n$ des éléments de E . On dit que y est *colinéaire* à x s'il existe un scalaire λ tel que $y = \lambda x$. Plus généralement, on dit que y est *combinaison linéaire* de $x_1, \dots, x_n$ s'il existe des scalaires $\lambda_1, \dots, \lambda_n$ tels que $y = \sum_{k=1}^n \lambda_k x_k$.

18.1 Sous-espaces vectoriels

Définition

Soit E un espace vectoriel. On dit qu'une partie non vide F de E est un *sous-espace vectoriel* de E si $0 \in F$ et pour tous vecteurs x, y dans F et tout scalaire λ , le vecteur $\lambda x + y$ est dans F .

Théorème

Tout sous-espace vectoriel d'un espace vectoriel est un espace vectoriel.

L'intersection $\bigcap_{i \in I} E_i$ d'une famille $(E_i)_{i \in I}$ de sous-espaces vectoriels de E est un sous-espace vectoriel de E .

Théorème

Soient E un espace vectoriel et $x_1, \dots, x_n$ des éléments de E .

L'ensemble F de toutes les combinaisons linéaires de $x_1, \dots, x_n$ est un sous-espace vectoriel de E .

Définition

Avec les notations du théorème précédent, on dit que F est le sous-espace vectoriel de E *engendré* par $x_1, \dots, x_n$ et on le note $F = \langle x_1, \dots, x_n \rangle$, ou $F = \text{Vect} \{x_1, \dots, x_n\}$, ou encore $F = \sum_{k=1}^n \mathbb{K}x_k$.

De manière plus générale, on peut définir le sous-espace vectoriel de E engendré par une famille X non vide de E comme l'ensemble $F = \text{Vect}(X)$ (ou $F = \langle X \rangle$) de toutes les combinaisons linéaires d'éléments de X . Un vecteur x de E est donc

dans $\text{Vect}(X)$ si, et seulement si, il existe un entier $p \geq 1$, des vecteurs $x_1, \dots, x_p$ dans X et des scalaires $\lambda_1, \dots, \lambda_p$ tels que $x = \sum_{k=1}^p \lambda_k x_k$.

Théorème

Si X est une partie non vide d'un espace vectoriel E , $\text{Vect}(X)$ est l'intersection de tous les sous-espaces vectoriels de E qui contiennent X . C'est aussi le plus petit (pour l'ordre défini par l'inclusion) sous-espace vectoriel de E qui contient X , c'est-à-dire que $\text{Vect}(X)$ contient X et est contenu dans tout sous-espace vectoriel de E qui contient X .

Théorème

Étant donné un entier naturel non nul n et n scalaires non tous nuls $a_1, \dots, a_n$, l'ensemble $H = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n, \sum_{k=1}^n a_k x_k = 0 \right\}$ est un sous-espace vectoriel de $\mathbb{K}^n$.

Définition

Étant donné un entier naturel non nul n , on appelle *hyperplan vectoriel* de $\mathbb{K}^n$ tout sous-espace vectoriel de la forme :

$$H = \left\{ (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n, \sum_{k=1}^n a_k x_k = 0 \right\}$$

où les scalaires $a_1, \dots, a_n$ ne sont pas tous nuls.

On dit aussi que H est l'hyperplan d'équation $\sum_{k=1}^n a_k x_k = 0$.

Théorème

Étant donné deux entiers naturels non nuls n et p , des scalaires $a_{1,1}, \dots, a_{1,n}, \dots, a_{p,1}, \dots, a_{p,n}$, l'ensemble :

$$F = \left\{ (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n, \sum_{k=1}^n a_{i,k} x_k = 0 \text{ pour } 1 \leq i \leq p \right\}$$

est un sous-espace vectoriel de $\mathbb{K}^n$.

On dit que F est le sous-espace vectoriel de $\mathbb{K}^n$ d'équations linéaires :

$$\sum_{k=1}^n a_{1,k} x_k = 0, \dots, \sum_{k=1}^n a_{p,k} x_k = 0.$$

18.2 Applications linéaires

On désigne par E et F deux espaces vectoriels sur un même corps $\mathbb{K}$.

Définition

On dit qu'une application u de E dans F est *linéaire* (ou que c'est un morphisme d'espaces vectoriels) si pour tous vecteurs x, y de E et tout scalaire λ , on a $u(x + y) = u(x) + u(y)$ et $u(\lambda x) = \lambda u(x)$.

On note $\mathcal{L}(E, F)$ l'ensemble de toutes les applications linéaires de E dans F . Dans le cas où $F = E$, on note plus simplement $\mathcal{L}(E)$ pour $\mathcal{L}(E, E)$. Les éléments de $\mathcal{L}(E)$ sont aussi appelés *endomorphismes* de E .

Si u et v sont deux applications linéaires de E dans F , $u + v$ est l'application définie sur E par :

$$\forall x \in E, (u + v)(x) = u(x) + v(x)$$

et pour tout scalaire λ , λu est l'application définie sur E par :

$$\forall x \in E, (\lambda u)(x) = \lambda u(x).$$

Théorème

L'ensemble $\mathcal{L}(E, F)$ de toutes les applications linéaires de E dans F muni de ces deux opérations est un espace vectoriel.

Théorème

Soient E , F et G des espaces vectoriels, u une application linéaire de E dans F et v une application linéaire de F dans G . La composée $v \circ u$ est alors une application linéaire de E dans G .

Pour $F = E$, $\mathcal{L}(E)$ est un anneau non commutatif pour $n \geq 2$. Le fait que $\mathcal{L}(E)$ soit un espace vectoriel et un anneau, la composition étant bilinéaire, se traduit en disant que c'est une algèbre. Sur cette algèbre, on peut définir les puissances entières u^n par $u^0 = \text{Id}$ et $u^{n+1} = u^n \circ u$ pour tout $n \in \mathbb{N}$.

De manière générale, on donne les définitions suivantes.

Définition

Une $\mathbb{K}$ -algèbre est un ensemble $\mathbb{A}$ muni d'une addition interne, d'une multiplication interne et d'une multiplication externe :

$$(x, y) \mapsto x + y, (x, y) \mapsto x \times y, (\lambda, x) \mapsto \lambda \cdot x$$

telles que $(\mathbb{A}, +, \times)$ est un anneau, $(\mathbb{A}, +, \cdot)$ est un $\mathbb{K}$ -espace vectoriel et pour tous $(\lambda, x, y) \in \mathbb{K} \times E^2$ on a :

$$(\lambda \cdot x) \times y = \lambda \cdot (x \times y), x \times (\lambda \cdot y) = \lambda \cdot (x \times y)$$

Une *sous-algèbre* de $\mathbb{A}$ est une partie de $\mathbb{A}$ qui est à la fois un sous-anneau et un sous-espace vectoriel.

Définition

Soient $\mathbb{A}$ et $\mathbb{B}$ deux $\mathbb{K}$ -algèbres. Un *morphisme d'algèbres* de $\mathbb{A}$ dans $\mathbb{B}$ est une application $u : \mathbb{A} \rightarrow \mathbb{B}$ qui est à la fois un morphisme d'anneaux et d'espaces vectoriels.

Théorème

Soit $u : E \rightarrow F$ une application linéaire. Pour tout $n \in \mathbb{N}^*$, tous vecteurs $x_1, \dots, x_n$ de E et tous scalaires $\lambda_1, \dots, \lambda_n$, on a :

$$u \left(\sum_{k=1}^n \lambda_k x_k \right) = \sum_{k=1}^n \lambda_k u(x_k).$$

Définition

Soit $u : E \rightarrow F$ une application linéaire. Le *noyau* de u est l'ensemble $\text{Ker}(u) = \{x \in E, u(x) = 0\}$ et son *image* est $\text{Im}(u) = \{u(x); x \in E\}$.

Théorème

Le noyau d'une application linéaire u de E dans F est un sous-espace vectoriel de E et son image un sous-espace vectoriel de F .

Plus généralement, pour tout sous-espace vectoriel H de E [resp. G de F] l'image $u(H)$ [resp. l'image réciproque $u^{-1}(G)$] est un sous-espace vectoriel de F [resp. de E].

Théorème

Soit u une application linéaire de E dans F .

- L'application u est injective si, et seulement si, on a $\text{Ker}(u) = \{0\}$.
- L'application u est surjective si, et seulement si, on a $\text{Im}(u) = F$.

Définition

On appelle *isomorphisme* de E sur F toute application linéaire bijective de E sur F .

Théorème

Soit u un isomorphisme de E sur F .
Alors sa bijection réciproque est linéaire.

Dans le cas où $E = F$, un isomorphisme de E sur E est un *automorphisme* de E . On note $\text{GL}(E)$ l'ensemble de tous les automorphismes de E . C'est un sous-groupe du groupe des permutations de E et on dit que c'est le *groupe linéaire* de E .

18.3 Dualité

E, F sont deux $\mathbb{K}$ -espaces vectoriels.

Définition

On appelle *forme linéaire* sur E toute application linéaire de E dans $\mathbb{K}$.

Une forme linéaire non nulle est surjective.

On note E^* l'espace vectoriel de toutes les formes linéaires sur E . Cet espace est l'*espace dual* de E .

Définition

On appelle *hyperplan vectoriel* de E , le noyau d'une forme linéaire non nulle sur E .

Si $H = \text{Ker}(\varphi)$ est un hyperplan de E , on dit alors que φ (ou $\varphi(x) = 0$) est une équation de H .

Définition

On dit qu'une forme linéaire $\varphi \in E^*$ et un vecteur $x \in E$ sont *orthogonaux* si $\varphi(x) = 0$.

Définition

L'orthogonal dans E^* d'une partie non vide X de E est l'ensemble :

$$X^\perp = \{\varphi \in E^*, \forall x \in X, \varphi(x) = 0\}.$$

L'orthogonal dans E d'une partie non vide Y de E^* est l'ensemble :

$$Y^\circ = \{x \in E, \forall \varphi \in Y, \varphi(x) = 0\}.$$

$X^\perp$ est un sous-espace vectoriel de E^* et Y° est un sous-espace vectoriel de E .

Pour $X = \emptyset$, on pose $X^\perp = E^*$ et pour $Y = \emptyset$, $Y^\circ = E$.

Théorème

Soient A, B des parties non vides de E et U, V des parties non vides de E^* .

- Si $A \subset B$, on a alors $B^\perp \subset A^\perp$; si $U \subset V$, on a alors $V^\circ \subset U^\circ$.
- $A \subset (A^\perp)^\circ$; $U \subset (U^\circ)^\perp$, les égalités n'étant pas réalisées en général.
- $A^\perp = (\text{Vect}(A))^\perp$ et $U^\circ = (\text{Vect}(U))^\circ$.
- $\{0\}^\perp = E^*$; $E^\perp = \{0\}$; $\{0\}^\circ = E$ et $(E^*)^\circ = \{0\}$.

Définition

La *transposée* de l'application linéaire $u \in \mathcal{L}(E, F)$ est l'application notée $u^\top$ (ou ${}^t u$) de F^* dans E^* définie par : $\forall \varphi \in F^*, u^\top(\varphi) = \varphi \circ u$.

Théorème

L'application de transposition $u \mapsto u^\top$ est linéaire et injective de $\mathcal{L}(E, F)$ dans $\mathcal{L}(F^*, E^*)$.

Théorème

Soient u dans $\mathcal{L}(E, F)$ et v dans $\mathcal{L}(F, G)$. On a :

- $(v \circ u)^T = u^T \circ v^T$;
- pour $F = E$, $(\text{Id}_E)^T = \text{Id}_{E^*}$;
- si u est un isomorphisme de E sur F , alors u^T est un isomorphisme de F^* sur E^* et $(u^T)^{-1} = (u^{-1})^T$;
- $\text{Ker}(u^T) = (\text{Im}(u))^\perp$;
- u est surjective si, et seulement si, u^T est injective ;
- $\text{Im}(u^T) = (\text{Ker}(u))^\perp$;
- u est injective si, et seulement si, u^T est surjective.

18.4 Sommes et sommes directes de sous-espaces vectoriels

On se donne un espace vectoriel E .

Définition

Soient F et G deux sous-espaces vectoriels de E . On dit que E est *somme* des espaces F et G si l'application :

$$\begin{aligned} \varphi : F \times G &\rightarrow E \\ (x, y) &\mapsto x + y \end{aligned}$$

est surjective et on note alors $E = F + G$. Si cette application φ est bijective, on dit que E est *somme directe* des espaces F et G et on note $E = F \oplus G$.

Théorème

Soient F et G deux sous-espaces vectoriels de E . On a $E = F \oplus G$ si, et seulement si, $E = F + G$ et $F \cap G = \{0\}$.

Définition

On dit que deux sous-espaces vectoriels F et G de E sont *supplémentaires*, si $E = F \oplus G$. On dit aussi que F est un supplémentaire de G ou que G est un supplémentaire de F dans E .

Dire que $E = F \oplus G$ revient à dire que tout x de E s'écrit de manière unique $x = x' + x''$ avec $x' \in F$ et $x'' \in G$. Les applications $p_F : x \mapsto x'$ et $p_G : x \mapsto x''$

sont linéaires et surjectives de E sur F et sur G respectivement. On dit que p_F est le *projecteur* sur F parallèlement à G et p_G est le projecteur sur G parallèlement à F . On a $p_F^2 = p_F$, $\text{Im}(p_F) = F$, $\text{Ker}(p_F) = G$, $p_G^2 = p_G$, $\text{Im}(p_G) = G$ et $\text{Ker}(p_G) = F$.

Réciproquement si $u \in \mathcal{L}(E)$ est tel que $u^2 = u$, on a alors $E = \text{Im}(u) \oplus \text{Ker}(u)$ et u est le projecteur sur $\text{Im}(u)$ parallèlement à $\text{Ker}(u)$. Pour cette raison, on dit que u est un *projecteur*.

L'application $s : x \mapsto p_F(x) - p_G(x) = 2p_F(x) - x$ est la symétrie par rapport à F et parallèlement à G . On a $s \circ s = \text{Id}$, donc s est un automorphisme de E égal à son inverse, $\text{Ker}(s - \text{Id}) = F$, $\text{Ker}(s + \text{Id}) = F$. Réciproquement si $u \in \mathcal{L}(E)$ est tel que $u^2 = \text{Id}$, on a alors $E = \text{Im}(u - \text{Id}) \oplus \text{Ker}(u + \text{Id})$ et u est la symétrie par rapport à $\text{Im}(u - \text{Id})$ parallèlement à $\text{Ker}(u + \text{Id})$.

Définition

Soient $p \geq 2$ un entier et $F_1, \dots, F_p$ des sous-espaces vectoriels de E . On dit que E est somme des espaces $F_1, \dots, F_p$ si l'application :

$$\begin{aligned} \varphi : F_1 \times \dots \times F_p &\rightarrow E \\ (x_1, \dots, x_p) &\mapsto x_1 + \dots + x_p \end{aligned}$$

est surjective et on note alors $E = F_1 + \dots + F_p$ ou de manière plus compacte

$$E = \sum_{k=1}^p F_k. \text{ Si cette application } \varphi \text{ est bijective, on dit que } E \text{ est } \textit{somme}$$

$$\textit{directe} \text{ des espaces } F_1, \dots, F_p \text{ et on note } E = F_1 \oplus \dots \oplus F_p \text{ ou } E = \bigoplus_{k=1}^p F_k.$$

Si $E = \bigoplus_{k=1}^p F_k$, alors chaque application $p_k : x \mapsto x_k$ est un projecteur d'image F_k

et de noyau $\bigoplus_{\substack{j=1 \\ j \neq k}}^p F_j$.

Théorème

Si $F_1, \dots, F_p$ sont des sous-espaces vectoriels de E , alors la somme $\sum_{k=1}^p F_k$ est le sous-espace vectoriel de E engendré par $\bigcup_{k=1}^p F_k$.

Théorème

Un hyperplan de E est un sous-espace de E supplémentaire d'une droite.

18.5 Familles libres, familles génératrices et bases

On se donne un $\mathbb{K}$ -espace vectoriel E .

Définition

Soient I un ensemble non vide et $\mathcal{B} = (e_i)_{i \in I}$ une famille de vecteurs de E . On dit que $\mathcal{B}$ est :

- une *famille libre*, ou que les vecteurs e_i sont *linéairement indépendants*, si pour toute partie finie non vide J de I et toute famille $(\lambda_j)_{j \in J}$ l'égalité $\sum_{j \in J} \lambda_j e_j = 0$ est réalisée si, et seulement si, tous les λ_j sont nuls ;
- une *famille liée*, si ce n'est pas une famille libre (il existe une partie finie non vide J de I et une famille $(\lambda_j)_{j \in J}$ de scalaires non tous nuls tels que $\sum_{j \in J} \lambda_j e_j = 0$) ;
- une *famille génératrice*, si pour tout vecteur $x \in E$, il existe une partie finie non vide J de I et une famille $(\lambda_j)_{j \in J}$ de scalaires tels que $x = \sum_{j \in J} \lambda_j e_j$, ce qui revient à dire que $E = \text{Vect}(\mathcal{B})$;
- une *base* de E , si elle est libre et génératrice.

$\mathcal{B} = (e_i)_{i \in I}$ est une base de E si, et seulement si, tout $x \in E$ s'écrit de manière unique $x = \sum_{i \in I} \lambda_i e_i$, où $(\lambda_i)_{i \in I}$ est une famille de scalaires à support fini (ce qui signifie que seul un nombre fini de λ_i sont non nuls). Les λ_i sont les *composantes* ou *coordonnées* de x dans la base $\mathcal{B}$.

Théorème

Soient I un ensemble non vide et $\mathcal{B} = (e_i)_{i \in I}$ une famille de vecteurs de E .

- Si $\mathcal{B}$ est libre, alors tous les vecteurs e_i sont non nuls.
- Si l'un des e_i est nul, alors $\mathcal{B}$ est liée.
- Si $\mathcal{B}$ contient une famille liée, elle est alors liée.
- Si $\mathcal{B}$ est contenue dans une partie libre, elle est alors libre.
- Si $\mathcal{B}$ est liée, l'un des vecteurs e_j est alors combinaison linéaire d'un nombre fini d'autres e_i .

18.6 Espaces vectoriels de dimension finie

Théorème

Si un espace vectoriel E admet une famille génératrice $\mathcal{B}$ formée de $n \geq 1$ éléments, alors toute famille libre dans E a au plus n éléments¹.

Définition

On dit qu'un espace vectoriel est de *dimension finie* s'il est réduit à $\{0\}$, ou s'il est différent de $\{0\}$ et admet une base formée d'un nombre fini de vecteurs. Dans le cas contraire, on dit qu'il est de *dimension infinie*.

Théorème

Si E un espace vectoriel non réduit à $\{0\}$ et de dimension finie, alors toutes les bases ont le même nombre d'éléments.

Définition

Si E est un espace vectoriel non réduit à $\{0\}$ et de dimension finie, alors sa *dimension* est le nombre de vecteurs de l'une quelconque de ses bases. On note $\dim_{\mathbb{K}}(E)$ (ou simplement $\dim(E)$) cette dimension.

Par convention, on dit que l'espace vectoriel $\{0\}$ est de dimension 0.

On appelle *droite vectorielle* tout espace vectoriel de dimension 1 et *plan vectoriel* tout espace vectoriel de dimension 2. Une droite vectorielle est donc de la forme $D = \{\lambda a ; \lambda \in \mathbb{K}\}$ où a est un vecteur non nul donné, ce que l'on note $D = \mathbb{K}a$ et un plan vectoriel de la forme $P = \{\lambda a + \mu b ; (\lambda, \mu) \in \mathbb{K}^2\}$, où a et b sont deux vecteurs non colinéaires donnés, ce que l'on note $P = \mathbb{K}a \oplus \mathbb{K}b$.

Théorème

Un espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

1. ce qui équivaut à dire qu'un système de plus de $n + 1$ vecteurs est lié.

Théorème

Soit E un espace vectoriel de dimension $n \geq 1$.

- Une famille libre dans E a au plus n éléments et c'est une base si, et seulement si, elle a exactement n éléments.
- Une famille génératrice dans E a au moins n élément et c'est une base si, et seulement si, elle a exactement n éléments.

Théorème

Soit E un espace vectoriel admettant une famille génératrice finie. De cette famille on peut extraire une base et E est de dimension finie.

Théorème de la base incomplète

Soit E un espace vectoriel de dimension $n \geq 1$. Toute famille libre à p éléments dans E (nécessairement $1 \leq p \leq n$) peut être complétée en une base.

Corollaire

Soit E un espace vectoriel de dimension n . Tout sous-espace vectoriel F de E est de dimension finie $m \leq n$ et on a $m = n$ si, et seulement si, $F = E$.

Corollaire

Tout sous-espace-vectoriel F d'un espace vectoriel E de dimension finie admet des supplémentaires et pour tout supplémentaire G de F dans E , on a $\dim(E) = \dim(F) + \dim(G)$.

Théorème

Soit E un espace vectoriel de dimension n . Si E est somme directe de $p \geq 2$

sous-espaces stricts $F_1, \dots, F_p$, soit $E = \bigoplus_{k=1}^p F_k$, alors en désignant, pour tout k compris entre 1 et p , par $\mathcal{B}_k$ une base de F_k , la famille $\bigcup_{k=1}^p \mathcal{B}_k$ est une base de E et on a $\dim(E) = \sum_{k=1}^p \dim(F_k)$.

Théorème (formule de Grassmann)

Soient E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E . On a $\dim(F + G) + \dim(F \cap G) = \dim(F) + \dim(G)$.

Cette formule permet de montrer que $F \cap G \neq \{0\}$ si $\dim(F) + \dim(G) > n$ où F et G sont deux sous-espaces vectoriels d'un espace vectoriel E de dimension $n \in \mathbb{N}^*$.

Elle permet aussi de démontrer par exemple que si F et G sont deux sous-espaces vectoriels distincts de dimension $n - 1$ d'un espace vectoriel de dimension $n \geq 2$, alors $\dim(F \cap G) = n - 2$.

Théorème

Soient E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E . On a :

$$\begin{aligned}(E = F \oplus G) &\Leftrightarrow (E = F + G \text{ et } \dim(E) = \dim(F) + \dim(G)) \\ &\Leftrightarrow (F \cap G = \{0\} \text{ et } \dim(E) = \dim(F) + \dim(G)).\end{aligned}$$

Théorème

Si $F_1, \dots, F_p$ sont des espaces vectoriels de dimension finie, il en est de même de l'espace produit $F = F_1 \times \dots \times F_p$ et on a $\dim(F) = \sum_{k=1}^p \dim(F_k)$.

Théorème

Si $F_1, \dots, F_p$ sont des sous-espaces vectoriels d'un espace vectoriel E de dimension finie, on a alors $E = \bigoplus_{k=1}^n F_k$ si, et seulement si, $E = \sum_{k=1}^p F_k$ et $\dim(E) = \sum_{k=1}^p \dim(F_k)$.

Ce théorème met en relief un résultat un peu surprenant : si F et G sont deux sous-espaces vectoriels d'un espace vectoriel E de dimension finie, alors

$$\dim(F \times G) = \dim(F \oplus G).$$

Théorème

Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F .

- Si u est injective, elle transforme alors toute famille libre de E en une famille libre de F et $\dim(E) \leq \dim(F)$.
- Si u est surjective, elle transforme alors toute famille génératrice de E en une famille génératrice de F et $\dim(F) \leq \dim(E)$.
- Si u est bijective, elle transforme alors toute base de E en une base de F et $\dim(E) = \dim(F)$ (deux espaces vectoriels de dimension finie isomorphes ont la même dimension).
- Si $\dim(E) = \dim(F)$, alors :

$$(u \text{ bijective}) \Leftrightarrow (u \text{ injective}) \Leftrightarrow (u \text{ surjective}).$$

- L'espace vectoriel $\mathcal{L}(E, F)$ de toutes les applications linéaires de E dans F est de dimension finie égale à $\dim(E) \cdot \dim(F)$.

Si $(e_j)_{1 \leq j \leq n}$ est une base de E et $(f_i)_{1 \leq i \leq m}$ une base de F , la famille $(u_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ d'applications linéaires définie par $u_{ij}(e_k) = 0$ pour $k \neq j$ et $u_{ij}(e_j) = f_i$ est alors une base de $\mathcal{L}(E, F)$.

L'espace E^* de toutes les formes linéaires sur E est de dimension n , de base $\mathcal{B}^* = (e_j^*)_{1 \leq j \leq n}$, où les e_j^* sont les projections relativement à la base $\mathcal{B}$ définies par :

$$e_j^*(e_i) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (1 \leq i, j \leq n).$$

On dit que $\mathcal{B}^*$ est la base duale de $\mathcal{B}$. Toute forme linéaire ℓ sur E s'écrit de manière unique $\ell = \sum_{j=1}^n \ell(e_j) e_j^*$.

Théorème

Soit E un espace vectoriel de dimension n . Si $\mathcal{B}' = (\ell_i)_{1 \leq i \leq n}$ est une base de E^* , il existe alors une base $\mathcal{B} = (f_i)_{1 \leq i \leq n}$ de E telle que $\mathcal{B}'$ soit la base duale de $\mathcal{B}$.

Définition

Avec les notations du théorème précédent, on dit que $\mathcal{B}$ est la *base anté-duale* de $\mathcal{B}'$.

Théorème

Dans un espace vectoriel E de dimension n un hyperplan est un sous-espace de E de dimension $n - 1$.

Théorème

On suppose que E est de dimension finie égale à $n \geq 1$.

— Pour tout sous-espace vectoriel F de E , on a :

$$\dim(F) + \dim(F^\perp) = \dim(E)$$

où $F^\perp$ est l'orthogonal de F dans l'espace dual E^* .

— Pour tout sous-espace vectoriel G de l'espace dual E^* , on a :

$$\dim(G) + \dim(G^\circ) = \dim(E)$$

où G° est l'orthogonal de G dans E .

— Pour tout sous-espace vectoriel F de E et tout sous-espace vectoriel G de E^* , on a $F = (F^\perp)^\circ$ et $G = (G^\circ)^\perp$.

— Pour toute partie X de E , on a $(X^\perp)^\circ = \text{Vect}(X)$.

— Pour tous sous-espaces vectoriels F_1 et F_2 de E , on a :

$$(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp \text{ et } (F_1 \cap F_2)^\perp = F_1^\perp + F_2^\perp.$$

— Pour tous sous-espaces vectoriels G_1 et G_2 de E^* , on a :

$$(G_1 + G_2)^\circ = G_1^\circ \cap G_2^\circ \text{ et } (G_1 \cap G_2)^\circ = G_1^\circ + G_2^\circ.$$

18.7 Rang d'une famille de vecteurs ou d'une application linéaire

E et F sont des espaces vectoriels.

Définition

Le *rang* d'une famille $(x_i)_{i \in I}$ de vecteurs de E est la dimension de l'espace vectoriel engendré par ces vecteurs. On le note $\text{rg}(x_i)_{i \in I}$ et on dit que cette famille est de *rang fini* si $\text{rg}(x_i)_{i \in I}$ est fini.

Dans le cas où E est de dimension n , le rang d'une famille de vecteurs de E est au plus égal à n .

Théorème

Le rang d'une famille $(x_i)_{1 \leq i \leq p}$ de p vecteurs de E est au maximum égal à p et ce rang vaut p si, et seulement si, cette famille est libre.

Définition

On dit que $u \in \mathcal{L}(E, F)$ est *de rang fini*, si $\text{Im}(u)$ est de dimension finie. Dans ce cas, on dit que la dimension de $\text{Im}(u)$ est le *rang* de u , ce que l'on note $\text{rg}(u)$.

Théorème du rang

Pour E de dimension finie et $u \in \mathcal{L}(E, F)$, $\text{Im}(u)$ est isomorphe à un supplémentaire de $\text{Ker}(u)$ dans E et on a $\dim(E) = \dim(\text{Ker}(u)) + \text{rg}(u)$.

Pour E et F de dimension finie, on a $\text{rg}(u) \leq \min(\dim(E), \dim(F))$.

Avec les hypothèses du théorème précédent, l'espace E n'est pas somme directe de $\text{Ker}(u)$ et $\text{Im}(u)$, mais on a le cas particulier suivant.

Théorème

Pour E de dimension finie et $u \in \mathcal{L}(E)$, on a :

$$\left(\text{Im}(u) = \text{Im}(u^2) \right) \Leftrightarrow \left(E = \text{Ker}(u) \oplus \text{Im}(u) \right) \Leftrightarrow \left(\text{Ker}(u) = \text{Ker}(u^2) \right).$$

Théorème

Soit $u \in \mathcal{L}(E, F)$.

- Pour toute famille $X = (x_i)_{i \in I}$ dans E , on a $\text{rg}(u(X)) \leq \text{rg}(X)$.
- Pour E de dimension finie, on a $\text{rg}(u) \leq \dim(E)$ et l'égalité est réalisée si, et seulement si, u est injective.
- Pour F de dimension finie, on a $\text{rg}(u) \leq \dim(F)$ et l'égalité est réalisée si, et seulement si, u est surjective.
- u et $u^\top$ ont même rang.

Dans le cas où E et F sont de même dimension finie, on a :

$$(\text{rg}(u) = \dim(E)) \Leftrightarrow (u \text{ injective}) \Leftrightarrow (u \text{ surjective}) \Leftrightarrow (u \text{ bijective}).$$

Pour ce qui est de la composition des applications linéaires, on a les résultats suivants.

Théorème

Soient E, F, G des $\mathbb{K}$ -espaces vectoriels de dimension finie, $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$.

- $\text{rg}(v \circ u) \leq \min(\text{rg}(u), \text{rg}(v))$.
- Pour u surjectif, on a $\text{rg}(v \circ u) = \text{rg}(v)$;
pour u injectif, on a $\text{rg}(v \circ u) = \text{rg}(u)$.
- Pour $E = F = G$, on a $|\text{rg}(v) - \text{rg}(u)| \leq \text{rg}(u + v) \leq \text{rg}(u) + \text{rg}(v)$.

18.8 Applications linéaires et calcul matriciel

On se donne un espace vectoriel E de dimension n , une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ de E , un espace vectoriel F de dimension m , une base $\mathcal{B}' = (f_k)_{1 \leq k \leq m}$ de F et une application linéaire u de E dans F . Pour tout entier j compris entre 1 et n , il existe des scalaires a_{ij} tels que $u(e_j) = \sum_{i=1}^m a_{ij} f_i$.

En notant x_j , pour j compris entre 1 et n , les composantes de x dans la base $\mathcal{B}$ et y_i , pour i compris entre 1 et m , celles de $u(x)$ dans la base $\mathcal{B}'$, l'égalité $y = u(x)$ se traduit par le produit matriciel :

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = AX$$

qui s'effectue comme suit :

$$y_i = (a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^n a_{ij} x_j.$$

Définition

Avec les notations qui précèdent, on dit que

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

est la *matrice de u dans les bases $\mathcal{B}$ et $\mathcal{B}'$* .

Dans le cas où $E = F$ et $\mathcal{B} = \mathcal{B}'$, on dira simplement que c'est la *matrice de u dans la base $\mathcal{B}$* .

On note $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble de toutes les matrices à m lignes et n colonnes et à coefficients dans $\mathbb{K}$. Pour $n = m$, on note $\mathcal{M}_n(\mathbb{K})$ l'ensemble $\mathcal{M}_{n,n}(\mathbb{K})$ et on dit que c'est l'ensemble des *matrices carrées d'ordre n* à coefficients dans $\mathbb{K}$.

Le tableau
$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{K})$$
 est notée (a_{ij}) , le premier indice i étant le numéro de ligne et le deuxième j , le numéro de colonne.

Théorème

Si u et v sont deux applications linéaires de E dans F de matrices respectives $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B = (b_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ alors, pour tout scalaire λ , l'application linéaire $\lambda u + v$ a pour matrice dans ces bases la matrice $(\lambda a_{ij} + b_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$.

On définit donc naturellement dans $\mathcal{M}_{m,n}(\mathbb{K})$ la somme de deux matrices et le produit d'une matrice par un scalaire par : $A + B = (a_{ij} + b_{ij})$ et $\lambda A = (\lambda a_{ij})$.

Théorème

L'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ des matrices à m lignes et n colonnes est un espace vectoriel de dimension nm .

On note 0 (ou plus précisément $0_{m,n}$) la matrice nulle, c'est-à-dire l'élément de $\mathcal{M}_{m,n}(\mathbb{K})$ dont toutes les composantes sont nulles et pour $m = n$, I_n la matrice dont tous les coefficients a_{ij} pour $i \neq j$ sont nuls et tous les coefficients diagonaux a_{ii} valent 1.

Théorème

Étant donné une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ de E et une base $\mathcal{B}' = (f_k)_{1 \leq k \leq m}$ de F , l'application φ qui associe à toute application linéaire $u \in \mathcal{L}(E, F)$ sa matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est un isomorphisme d'espaces vectoriels de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{m,n}(\mathbb{K})$.

Théorème

Si G est un espace vectoriel de dimension r , $\mathcal{B}'' = (g_k)_{1 \leq k \leq r}$ une base de G , $u : E \rightarrow G$ une application linéaire de matrice $A = (a_{ij}) \in \mathcal{M}_{r,n}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}''$ et $v : F \rightarrow E$ une application linéaire de matrice $B = (b_{ij}) \in \mathcal{M}_{n,m}(\mathbb{K})$ dans les bases $\mathcal{B}'$ et $\mathcal{B}$, alors la matrice dans les bases $\mathcal{B}'$ et $\mathcal{B}''$ de l'application linéaire $u \circ v : F \rightarrow G$ est la matrice $C = (c_{ij}) \in \mathcal{M}_{r,m}(\mathbb{K})$ définie par $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ pour $1 \leq i \leq r$ et $1 \leq j \leq m$.

Le théorème précédent justifie la définition suivante.

Définition

Soient $A = (a_{ij}) \in \mathcal{M}_{r,n}(\mathbb{K})$ et $B = (b_{ij}) \in \mathcal{M}_{n,m}(\mathbb{K})$. Le *produit* AB de A par B est la matrice $C = (c_{ij}) \in \mathcal{M}_{r,m}(\mathbb{K})$ définie par $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ pour $1 \leq i \leq r$ et $1 \leq j \leq m$.

Théorème

L'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n est une algèbre non commutative de dimension n^2 .

Le choix d'une base de E permet de réaliser un isomorphisme d'anneaux de $\mathcal{L}(E)$ sur $\mathcal{M}_n(\mathbb{K})$.

On dit que deux matrices A et B dans $\mathcal{M}_n(\mathbb{K})$ *commutent* si $AB = BA$.

La matrice $I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$, appelée *matrice identité d'ordre n* , est

l'élément neutre pour le produit matriciel dans $\mathcal{M}_n(\mathbb{K})$.

L'associativité du produit matriciel permet de définir les puissances successives de $A \in \mathcal{M}_n(\mathbb{K})$ par la relation de récurrence $A^0 = I_n$ et $A^{p+1} = A^p A = A A^p$ pour tout $p \in \mathbb{N}$.

Pour A et B qui commutent dans $\mathcal{M}_n(\mathbb{K})$, on dispose de la formule du binôme de Newton $(A + B)^m = \sum_{k=0}^m \binom{m}{k} A^k B^{m-k}$.

Définition

On dit qu'une matrice A dans $\mathcal{M}_n(\mathbb{K})$ est *inversible* s'il existe une matrice A' dans $\mathcal{M}_n(\mathbb{K})$ telle que $AA' = A'A = I_n$. On dit alors que A' est l'*inverse* de A .

Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, son inverse est alors noté A^{-1} .

Théorème

Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, alors A^{-1} est aussi inversible et on a $(A^{-1})^{-1} = A$. Le produit de deux matrices inversibles A et B est inversible et on a $(AB)^{-1} = B^{-1}A^{-1}$.

L'ensemble des éléments inversibles de $\mathcal{M}_n(\mathbb{K})$ est noté $\text{GL}_n(\mathbb{K})$ et c'est un groupe multiplicatif isomorphe au groupe linéaire $\text{GL}(E)$ par le choix d'une base de E .

Théorème

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E . Un endomorphisme u de E est bijectif si, et seulement si, sa matrice A dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est inversible et dans ce cas A^{-1} est la matrice de u^{-1} dans les bases $\mathcal{B}'$ et $\mathcal{B}$.

Théorème

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si, et seulement si, l'unique solution du système linéaire $AX = 0$ est $X = 0$.

Corollaire

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors les assertions suivantes sont équivalentes :

- (i) A est inversible.
- (ii) il existe $A' \in \mathcal{M}_n(\mathbb{K})$ telle que $A'A = I_n$.
- (iii) il existe $A' \in \mathcal{M}_n(\mathbb{K})$ telle que $AA' = I_n$.

Définition

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. La *trace* de A est le scalaire $\text{tr}(A) = \sum_{k=1}^n a_{kk}$.

Soient E un $\mathbb{K}$ -espace vectoriel de dimension $n \in \mathbb{N}^*$ et $u \in \mathcal{L}(E)$. La *trace* de u , notée $\text{tr}(u)$, est la trace de la matrice de u dans une base de E .

La trace d'un endomorphisme de E ne dépend pas du choix d'une base de E .

Définition

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$. L'application linéaire *canoniquement associée* à A est l'application $u \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ dont la matrice dans les bases canoniques est A .

Définition

Soit $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$. En désignant, pour tout j compris entre 1 et n , par $C_j = (a_{ij})_{1 \leq i \leq m}$ le vecteur de $\mathbb{K}^m$ représentant la colonne numéro j de A , le *rang* de A est le rang de la famille $(C_1, \dots, C_n)$ de vecteurs de $\mathbb{K}^m$.

Théorème

Le rang d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est égal au rang de l'application linéaire $u \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ canoniquement associée à la matrice A .

Théorème

Si $u \in \mathcal{L}(E, F)$ a pour matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$, alors le rang de u est égal au rang de A .

En identifiant une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ à l'application linéaire $u \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ canoniquement associée, on note respectivement $\text{Ker}(A)$ et $\text{Im}(A)$ le noyau et l'image de u .

Théorème

Si $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est la matrice de $u \in \mathcal{L}(E, F)$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$, alors la matrice de $u^\top$ dans les bases duales $\mathcal{B}'^*$ et $\mathcal{B}^*$ est la matrice A' dont les colonnes sont les lignes de A .

Le théorème précédent justifie la définition suivante.

Définition

La *transposée* de $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est la matrice notée $A^\top$ (ou ${}^t A$) de $\mathcal{M}_{n,m}(\mathbb{K})$ dont les colonnes sont les lignes de A .

Pour $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$, on a donc $A^\top = (a'_{ij}) \in \mathcal{M}_{n,m}(\mathbb{K})$, où $a'_{ij} = a_{ji}$ pour tous i, j .

Théorème

L'application $A \mapsto A^T$ réalise un isomorphisme de l'espace vectoriel $\mathcal{M}_{m,n}(\mathbb{K})$ sur $\mathcal{M}_{n,m}(\mathbb{K})$.

Pour toutes $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,r}(\mathbb{K})$, on a $(A^T)^T = A$ et $(AB)^T = B^T A^T$.
Les matrices particulières suivantes sont souvent utiles.

Définition

On dit qu'une matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ est *triangulaire inférieure* [resp. *supérieure*] si $a_{ij} = 0$ pour $1 \leq i < j \leq n$ [resp. pour $1 \leq j < i \leq n$].

Une matrice *triangulaire inférieure* [resp. *triangulaire supérieure*] est de la forme

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad [\text{resp. de la forme} \quad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}].$$

Une matrice *diagonale* est une matrice triangulaire inférieure et supérieure, soit

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}, \quad \text{ce que l'on peut noter } \text{diag}(d_1, \dots, d_n).$$

Le produit de deux matrices triangulaires supérieures [resp. inférieures] est triangulaire supérieure [resp. inférieure] et le produit d'une matrice triangulaire supérieure par une matrice triangulaire inférieure est diagonale.

$$\text{Une matrice décomposée en blocs est de la forme} \quad \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix},$$

où les A_{ij} sont des matrices de tailles adéquates.

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est *symétrique* (ou autoadjointe) [resp. *antisymétrique*] si elle vérifie $A^T = A$ [resp. $A^T = -A$].

18.9 Formules de changement de base

On se donne un espace vectoriel E de dimension n et deux bases $\mathcal{B}_1 = (e_k)_{1 \leq k \leq n}$ et $\mathcal{B}_2 = (e'_k)_{1 \leq k \leq n}$ de E . Pour tout vecteur $x = \sum_{j=1}^n x_j e_j = \sum_{j=1}^n x'_j e'_j$ dans E , on note $X = (x_j)_{1 \leq j \leq n}$ et $X' = (x'_j)_{1 \leq j \leq n}$ les vecteurs de $\mathbb{K}^n$ respectivement formés des composantes de x dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$. Pour j compris entre 1 et n le vecteur e'_j s'écrit $e'_j = \sum_{i=1}^n p_{ij} e_i$, où les p_{ij} sont des scalaires uniquement déterminés et on a $x_i = \sum_{j=1}^n p_{ij} x'_j$ pour $1 \leq i \leq n$, ce qui peut se traduire par l'égalité matricielle

$$X = PX', \text{ où } P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \text{ ayant pour colonne } j \text{ le vecteur de}$$

$\mathbb{K}^n$ formé des composantes de e'_j dans $\mathcal{B}_1$.

On peut retenir cette formule sous la forme « $X_{\mathcal{B}_1} = P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2} X_{\mathcal{B}_2}$ » avec des notations évidentes.

Définition

Avec les notations qui précèdent, on dit que P est la *matrice de passage* de la base $\mathcal{B}_1$ à la base $\mathcal{B}_2$.

Théorème

Avec les notations qui précèdent, la matrice de passage P de $\mathcal{B}_1$ à $\mathcal{B}_2$ est inversible et son inverse est la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}_1$.

Théorème

Si u est un endomorphisme de E de matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ dans la base $\mathcal{B}_1$ et de matrice $A' = (a'_{ij}) \in \mathcal{M}_n(\mathbb{K})$ dans la base $\mathcal{B}_2$, on a alors $A' = P^{-1}AP$, où P est la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$.

La formule de changement de base s'écrit « $A_{\mathcal{B}_2} = P_{\mathcal{B}_2 \rightarrow \mathcal{B}_1} A_{\mathcal{B}_1} P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$ ».

19

Déterminants

$\mathbb{K}$ est un corps.

On se donne un entier $n \geq 1$ et $\mathcal{M}_n(\mathbb{K})$ est l'espace vectoriel des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$.

Pour i, j entiers compris entre 1 et n , on note E_{ij} la matrice dont tous les coefficients sont nuls sauf celui d'indice (i, j) (ligne i et colonne j) qui vaut 1. La famille $(E_{ij})_{1 \leq i, j \leq n}$ est une base de $\mathcal{M}_n(\mathbb{K})$ qui est donc de dimension n^2 .

Pour toute matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$, on note pour tout entier i compris entre 1 et n , $L_i = (a_{i1}, \dots, a_{in})$ sa ligne numéro i et pour tout entier j compris entre 1

et n , $C_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ sa colonne numéro j . On écrira $A = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}$ ou

$A = (C_1 \quad \cdots \quad C_n)$.

19.1 Opérations élémentaires. Matrices de dilatation et de transvection

On suppose que $n \geq 2$.

On appelle matrice déduite de A par opération élémentaire sur les lignes de A toute matrice $A_i(\lambda)$ avec $1 \leq i \leq n$ et $\lambda \in \mathbb{K}^*$ déduite de A en multipliant sa ligne numéro i par λ ou $A_{ij}(\lambda)$ avec $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$ déduite de A en ajoutant à la ligne numéro i la ligne numéro j multipliée par λ .

On appelle matrice déduite de A par opération élémentaire sur les colonnes de A toute matrice $A'_j(\lambda)$ avec $1 \leq j \leq n$ et $\lambda \in \mathbb{K}^*$ déduite de A en multipliant sa colonne numéro j par λ ou $A'_{ij}(\lambda)$ avec $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$ déduite de la matrice A en ajoutant à la colonne numéro j la colonne numéro i multipliée par λ .

Définition

Une *matrice de transvection* est une matrice de la forme $T_{ij}(\lambda) = I_n + \lambda E_{ij}$ avec $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$ et une *matrice de dilatation* est une matrice de la forme $D_i(\lambda) = I_n + (\lambda - 1) E_{ii}$ avec $1 \leq i \leq n$ et $\lambda \in \mathbb{K}^*$.

Théorème

Avec les notations précédentes, on a $A_i(\lambda) = D_i(\lambda)A$, $A'_j(\lambda) = AD_j(\lambda)$, $A_{ij}(\lambda) = T_{ij}(\lambda)A$ et $A'_{ij}(\lambda) = AT_{ij}(\lambda)$, ce qui signifie que :

- la multiplication à gauche [resp. à droite] par une matrice de dilatation $D_i(\lambda)$ [resp. $D_j(\lambda)$] a pour effet de multiplier la ligne i [resp. la colonne j] par λ ;
- la multiplication à gauche [resp. à droite] par une matrice de transvection $T_{ij}(\lambda)$ a pour effet de remplacer la ligne L_i [resp. la colonne C_j] par $L_i + \lambda L_j$ [resp. par $C_j + \lambda C_i$].

Définition

Une *matrice élémentaire* est une matrice de dilatation ou de transvection.

Lemme

Une matrice élémentaire est inversible avec $T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$ pour une transvection et $D_i(\lambda)^{-1} = D_i(\frac{1}{\lambda})$ pour une dilatation.

Théorème

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ (où $n \geq 2$) est inversible si, et seulement si, elle est produit de matrices élémentaires. Précisément si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, il existe alors des matrices de transvection $P_1, \dots, P_r$ et $Q_1, \dots, Q_s$ et une matrice de dilatation $D_n(\lambda)$ telles que $A = P_1 \dots P_r D_n(\lambda) Q_1 \dots Q_s$.

19.2 Déterminants des matrices carrées

Dans la définition suivante, $\mathcal{S}_n$ désigne le groupe symétrique d'ordre n et $\varepsilon(\sigma)$ la signature d'une permutation $\sigma \in \mathcal{S}_n$.

Définition

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. On appelle *déterminant* de A le scalaire noté

$$\det(A) \text{ ou } \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \text{ défini par } \det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i),i}.$$

Théorème

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. Alors, en notant A_{ij} la matrice carrée d'ordre $n-1$ déduite de A en supprimant la ligne i et la colonne j , on a :

- le déterminant d'un scalaire est le scalaire lui-même ;
- $\forall j \in \llbracket 1, n \rrbracket$, $\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$ (développement suivant la colonne j) ;
- $\forall i \in \llbracket 1, n \rrbracket$, $\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$ (développement suivant la ligne i).

Ce théorème permet de calculer tout déterminant par récursivité sans revenir à la définition. Avec les notations du théorème, on dit que $\det(A_{ij})$ est le *mineur* d'indice (i, j) de la matrice A et que $(-1)^{i+j} \det(A_{ij})$ est le *cofacteur* d'indice (i, j) de A .

Théorème

Soient A et B dans $\mathcal{M}_n(\mathbb{K})$. Alors

- $\det(A^T) = \det(A)$;
- $\det(AB) = \det(A) \det(B)$;
- A inversible si, et seulement si, $\det(A) \neq 0$ et dans ce cas, on a $\det(A^{-1}) = \frac{1}{\det(A)}$.

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors

- si A possède une ligne (ou une colonne) nulle, alors $\det(A) = 0$;
- on ne change pas la valeur de $\det(A)$ si on ajoute à une ligne (resp. colonne) une combinaison linéaire des autres lignes (resp. colonnes).

Dans le théorème suivant, en notant $C_1, \dots, C_n$ les colonnes d'une matrice carrée A , $\det(C_1, \dots, C_n)$ désigne le déterminant de A .

Théorème

Soit $\lambda \in \mathbb{K}$. Alors ¹

$$\begin{aligned} & \cdot \det(C_1, \dots, C_{i-1}, C_i + C'_i, C_{i+1}, \dots, C_n) \\ &= \det(C_1, \dots, C_{i-1}, C_i, C_{i+1}, \dots, C_n) + \det(C_1, \dots, C_{i-1}, C'_i, C_{i+1}, \dots, C_n); \\ & \cdot \det(C_1, \dots, C_{i-1}, \lambda C_i, C_{i+1}, \dots, C_n) = \lambda \det(C_1, \dots, C_{i-1}, C_i, C_{i+1}, \dots, C_n). \end{aligned}$$

Le théorème précédent se traduit en disant que le déterminant est linéaire par rapport à chaque colonne (ou chaque ligne).

Corollaire

Pour tous $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$, on a $\det(\lambda A) = \lambda^n \det(A)$.

Théorème

Si A' est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{K})$ en permutant deux colonnes (ou deux lignes), on a alors $\det(A') = -\det(A)$.

Le résultat précédent se traduit en disant que le déterminant est une forme alternée sur les colonnes (ou les lignes).

Théorème

Le déterminant d'une matrice triangulaire est égale au produit de ses termes diagonaux, soit $\det(A) = \prod_{i=1}^n a_{ii}$.

Théorème

Si $A = \begin{pmatrix} A_1 & B_{12} & \cdots & B_{1r} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & B_{r-1,r-1} \\ 0 & \cdots & 0 & A_r \end{pmatrix}$, où les A_k sont des matrices carrées, on a alors $\det(A) = \prod_{k=1}^r \det(A_k)$.

1. Comme $\det(A^T) = \det(A)$, le théorème est également vérifié sur des lignes.

Théorème

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, toute matrice de dilatation $D_i(\lambda)$ et toute matrice de transvection $T_{ij}(\lambda)$, on a $\det(D_i(\lambda)A) = \lambda \det(A)$ et $\det(T_{ij}(\lambda)A) = \det(A)$.

Définition

La *comatrice* de $A \in \mathcal{M}_n(\mathbb{K})$ est la matrice $\text{Com}(A) = ((-1)^{i+j} \det(A_{ij}))$ des cofacteurs de A .

Théorème

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, on a

$$A \cdot \text{Com}(A)^T = \text{Com}(A)^T \cdot A = \det(A)I_n$$

où $\text{Com}(A)^T$ est la transposée de $\text{Com}(A)$. Pour $\det(A) \neq 0$, la matrice A est inversible d'inverse $A^{-1} = \frac{1}{\det(A)} \text{Com}(A)^T$.

Soient $n \geq 2$ un entier, $\alpha_1, \dots, \alpha_n$ des scalaires et :

$$V(\alpha_1, \dots, \alpha_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

la *matrice de Vandermonde* associée. On note $\Delta(\alpha_1, \dots, \alpha_n)$, le déterminant de cette matrice.

Théorème (déterminant de Vandermonde)

On a $\Delta(\alpha_1, \dots, \alpha_n) = \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)$ et la matrice de Vandermonde $V(\alpha_1, \dots, \alpha_n)$ est inversible si, et seulement si, les α_i sont deux à deux distincts.

19.3 Déterminant d'une famille de vecteurs

Étant donné une famille $(x_j)_{1 \leq j \leq n}$ de n vecteurs de $\mathbb{K}^n$, on définit le déterminant de cette famille comme le déterminant de la matrice A dont les colonnes sont

formées de ces vecteurs. En notant, pour j compris entre 1 et n , $x_j = (x_{ij})_{1 \leq i \leq n}$ (vecteur colonne), on a donc $\det(x_1, \dots, x_n) = \det(x_{ij})$.

Théorème

Une famille $(x_j)_{1 \leq j \leq n}$ de n vecteurs de $\mathbb{K}^n$ est libre si, et seulement si, son déterminant est non nul.

Si E un espace vectoriel de dimension $n \geq 1$ et $\mathcal{B}$ une base E , en notant pour tout vecteur $x \in E$, X le vecteur colonne de $\mathbb{K}^n$ formé des composantes de x dans la base $\mathcal{B}$, on définit le déterminant d'une famille $(x_j)_{1 \leq j \leq n}$ de n vecteurs de E dans la base $\mathcal{B}$ par $\det_{\mathcal{B}}(x_1, \dots, x_n) = \det(X_1, \dots, X_n)$. Ce déterminant dépend du choix d'une base de E .

Théorème

Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E , alors pour tout n -uplet $(x_1, \dots, x_n)$ de vecteurs de E , on a $\det_{\mathcal{B}}(x_1, \dots, x_n) = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(x_1, \dots, x_n)$.

19.4 Déterminant d'un endomorphisme

E est un espace vectoriel de dimension $n \geq 1$.

Définition

Le *déterminant* d'un endomorphisme u de E est le déterminant de la matrice de u dans une base de E .

Ce déterminant ne dépend pas du choix d'une base de E .

Théorème

Si u, v sont deux endomorphismes de E , on a alors :

$$\det(u \circ v) = \det(v \circ u) = \det(u) \det(v).$$

Théorème

Un endomorphisme u de E est inversible si, et seulement si, son déterminant est non nul.

19.5 Systèmes linéaires

Formule de Cramer

On appelle *système de Cramer* tout système linéaire $Ax = b$ d'inconnue $x \in \mathbb{K}^n$, où $A \in \text{GL}_n(\mathbb{K})$ et $b \in \mathbb{K}^n$. L'unique solution d'un tel système est $x = A^{-1}b$.

En notant $x = (x_j)_{1 \leq j \leq n} \in \mathbb{K}^n$ cette solution, on a $b = \sum_{j=1}^n x_j C_j$, où les C_j sont les colonnes de A et en utilisant le caractère multilinéaire alterné du déterminant, on a pour tout entier k compris entre 1 et n :

$$\begin{aligned} \det(C_1, \dots, C_{k-1}, b, C_{k+1}, \dots, C_n) &= \sum_{j=1}^n x_j \det(C_1, \dots, C_{k-1}, C_j, C_{k+1}, \dots, C_n) \\ &= x_k \det(A) \end{aligned}$$

soit $x_k = \frac{\det(C_1, \dots, C_{k-1}, b, C_{k+1}, \dots, C_n)}{\det(A)}$ pour $1 \leq k \leq n$.

Ces dernières égalités sont appelées *formules de Cramer*. En calculant un déterminant par cette formule, cela nécessite $n!(n-1)$ multiplications et $(n!-1)$ additions, donc environ $nn!$ opérations élémentaires. Comme il y a $n+1$ déterminants à calculer puis n divisions à faire pour les formules de Cramer, on aura un total d'environ $n^2n!$ opérations élémentaires à effectuer, ce qui peut être beaucoup trop important pour de grandes valeurs de n .

Méthode du pivot de Gauss

La méthode du *pivot de Gauss* est plus efficace pour résoudre un système linéaire d'ordre $n \in \mathbb{N}^*$. En effectuant des opérations élémentaires sur les lignes du système linéaire $Ax = b$, on le transforme en un système triangulaire supérieur $Rx = c$.

On note ici L_i la ligne numéro i du système linéaire $Ax = b$.

Étape 0 — On se ramène à un système tel que a_{11} non nul.

Si pour tout $i = 1, \dots, n$, on a $a_{i1} = 0$, alors $\det(A) = 0$ et c'est fini. Sinon, il existe $i > 1$ tel que a_{i1} soit non nul, et en permutant les lignes 1 et i (si $i = 1$, on ne fait rien), on se ramène à un système $A^{(1)}x = b^{(1)}$, avec $a_{11}^{(1)}$ non nul. Le coefficient $a_{11}^{(1)}$ est le premier pivot. On a alors $\det(A) = \pm \det(A^{(1)})$, avec le signe moins si, et seulement si, il y a une permutation des lignes 1 et i avec $i > 1$.

Étape 1 — Élimination de x_1 dans les équations $2, \dots, n$.

On effectue pour cela les transformations élémentaires suivantes :

$$L_i^{(1)} \longleftarrow L_i^{(1)} - \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} L_1^{(1)} \quad (i = 2, \dots, n)$$

et après une éventuelle permutation des lignes 2 et $i > 3$, on obtient le système

$$A^{(2)}x = b^{(2)}, \text{ avec } a_{22}^{(2)} \text{ non nul où } A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix}.$$

Le coefficient $a_{22}^{(2)}$ est le deuxième pivot.

Étape k — Élimination de x_k dans les équations $k+1, \dots, n$.

À la fin de l'étape $k-1$, on a obtenu le système $A^{(k)}x = b^{(k)}$, avec :

$$A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}$$

et $a_{kk}^{(k)}$ non nul. Le coefficient $a_{kk}^{(k)}$ est le pivot numéro k .

On effectue alors les transformations élémentaires suivantes :

$$L_i^{(k)} \leftarrow L_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} L_k^{(k)} \quad (i = k+1, \dots, n)$$

puis une éventuelle permutation des lignes $k+1$ et $j > k+1$ pour se ramener à $a_{k+1,k+1}^{(k+1)}$ non nul.

Au bout de $n-1$ étapes, on est donc ramené à un système triangulaire supérieur $A^{(n)}x = b^{(n)}$.

De plus, on a $\det(A) = (-1)^p \det(A^{(n)})$, où p est le nombre de permutations qui ont été nécessaires pour avoir des pivots non nuls et $\det(A^{(n)})$ est le produit des pivots.

Définition

On appelle *matrice extraite* (ou *sous-matrice*) de $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$, toute matrice $A_{I,J} = (a_{ij})_{(i,j) \in I \times J}$, où $I = \{1 \leq i_1 < \dots < i_q \leq m\}$ et $J = \{1 \leq j_1 < \dots < j_p \leq n\}$. On appelle *déterminant extrait* de A (ou de $\det(A)$), le déterminant d'une matrice carrée extraite de A . Si δ_p est un déterminant extrait de A d'ordre p , on appelle *bordant* de δ tout déterminant δ_{p+1} extrait d'ordre $p+1$ de A tel que δ_p soit extrait de δ_{p+1} .

Théorème

Une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est de rang r si, et seulement si, il existe un déterminant extrait de A , δ_r d'ordre r et non nul tel que tous les bordants de δ_r sont nuls (si $r = \min(n, m)$, la deuxième condition n'est pas à prendre en compte).

Corollaire

Le rang d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est l'ordre du plus grand déterminant extrait de A qui est non nul.

Les formules de Cramer peuvent être utilisées pour déterminer l'ensemble des solutions d'un système linéaire de m équations à n inconnues $Ax = b$ d'inconnue $x \in \mathbb{K}^n$, où $A \in \mathcal{M}_{m,n}(\mathbb{K}) \setminus \{0\}$ est de rang $r \in \{1, \min(n, m)\}$ et $b \in \mathbb{K}^m$.

Il existe des matrices carrées d'ordre r extraites de A de déterminant non nul (un tel déterminant est appelé *déterminant principal* de A) et toute matrice carrée extraite d'ordre plus grand ou égal à $r + 1$ (s'il en existe) est de déterminant nul. En effectuant au besoin des permutations de lignes ou de colonnes du système linéaire, on se ramène à une matrice A tel que la matrice $A_r = (a_{ij})_{1 \leq i, j \leq r}$ soit inversible. Le système linéaire :

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad (1 \leq i \leq r)$$

est alors appelé *système d'équations principales* du système $Ax = b$.

Polynômes à coefficients dans un corps

20

$\mathbb{K}$ est un corps commutatif et pour tout couple d'entiers $(n, k) \in \mathbb{N}^2$, on note $\delta_{n,k}$ le *symbole de Kronecker* défini par $\delta_{n,k} = 1$ pour $k = n$ et $\delta_{n,k} = 0$ pour $k \neq n$.

20.1 L'algèbre $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$

On désigne par $\mathbb{K}^{\mathbb{N}}$ l'ensemble des suite $a = (a_k)_{k \in \mathbb{N}}$ d'éléments de $\mathbb{K}$. Muni des opérations d'addition et de multiplication externe définies par $a + b = (a_k + b_k)_{k \in \mathbb{N}}$ et $\lambda a = (\lambda a_k)_{k \in \mathbb{N}}$ pour a, b dans $\mathbb{K}^{\mathbb{N}}$ et λ dans $\mathbb{K}$, cet ensemble est un $\mathbb{K}$ -espace vectoriel. On définit une multiplication interne sur $\mathbb{K}^{\mathbb{N}}$ en posant pour $a = (a_k)_{k \in \mathbb{N}}$, $b = (b_k)_{k \in \mathbb{N}}$ dans $\mathbb{K}^{\mathbb{N}}$, $ab = (c_k)_{k \in \mathbb{N}}$, où les c_k sont définis par :

$$\forall k \in \mathbb{N}, c_k = \sum_{j=0}^k a_j b_{k-j} = \sum_{j=0}^k a_{k-j} b_j = \sum_{\substack{0 \leq i, j \leq k \\ i+j=k}} a_i b_j.$$

Muni de cette multiplication, $\mathbb{K}^{\mathbb{N}}$ est un anneau commutatif, le neutre pour le produit étant $P_0 = (\delta_{0,k})_{k \in \mathbb{N}} = (1, 0, 0, \dots, 0, \dots)$.

Définition

Le *support* d'une suite $a = (a_k)_{k \in \mathbb{N}}$ est $\text{Supp}(a) = \{k \in \mathbb{N}, a_k \neq 0\}$.

Définition

On appelle *polynôme* à coefficients dans $\mathbb{K}$ toute suite $P = (a_k)_{k \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ de support fini, c'est-à-dire une suite d'éléments de $\mathbb{K}$ dont tous les termes, sauf un nombre fini d'entre eux, sont nuls.

On note, temporairement, $\mathbb{K}^{(\mathbb{N})}$ l'ensemble de ces polynômes, 0 le polynôme nul, c'est-à-dire celui dont tous les termes sont nuls et 1 le polynôme $P_0 = (\delta_{0,k})_{k \in \mathbb{N}}$. Si $P = (a_k)_{k \in \mathbb{N}}$ est un polynôme non nul, il existe alors un nombre fini d'indices $k \in \mathbb{N}$ tels que $a_k \neq 0$ et on peut noter :

$$\deg(P) = \max\{k \in \mathbb{N}, a_k \neq 0\}, \quad \text{val}(P) = \min\{k \in \mathbb{N}, a_k \neq 0\}.$$

Définition

Avec les notations qui précèdent, on dit que $\deg(P)$ est le *degré* de P et $\text{val}(P)$ la *valuation* de P .

Par convention, on pose $\deg(0) = -\infty$ et $\text{val}(0) = +\infty$.

Un polynôme non nul de degré $n \geq 0$ peut être noté $P = (a_0, \dots, a_n, 0, \dots, 0, \dots)$ et on dit que $a_n \neq 0$ est le coefficient dominant de P .

Théorème

Pour tous P et Q dans $\mathbb{K}^{(\mathbb{N})}$, on a $\deg(P + Q) \leq \max(\deg(P), \deg(Q))$ avec égalité si $\deg(P) \neq \deg(Q)$, $\text{val}(P + Q) \geq \min(\text{val}(P), \text{val}(Q))$ avec égalité si $\text{val}(P) \neq \text{val}(Q)$, $\deg(PQ) = \deg(P) + \deg(Q)$ et $\text{val}(PQ) = \text{val}(P) + \text{val}(Q)$.

Théorème

$\mathbb{K}^{(\mathbb{N})}$ est une sous-algèbre de $\mathbb{K}^{\mathbb{N}}$ et la famille $(P_n)_{n \in \mathbb{N}}$ de polynômes définie par :
 $\forall n \in \mathbb{N}, P_n = (\delta_{n,k})_{k \in \mathbb{N}} = (0, \dots, 0, 1, 0, \dots, 0, \dots)$
 en est une base.

L'application $a_0 \in \mathbb{K} \mapsto (a_0, 0, \dots, 0, \dots) = a_0 P_0 \in \mathbb{K}^{(\mathbb{N})}$ réalise un morphisme d'anneaux injectif de $\mathbb{K}$ dans $\mathbb{K}^{(\mathbb{N})}$, ce qui permet d'identifier $\mathbb{K}$ au sous-anneau de $\mathbb{K}^{(\mathbb{N})}$ formé du polynôme nul et des polynômes non nuls de degré 0.

Un polynôme $a_0 P_0$, identifié à $a_0 \in \mathbb{K}$, est appelé *polynôme constant*.

En notant $X = P_1 = (0, 1, 0, \dots, 0, \dots)$, on a $X^n = P_n$ pour tout $n \in \mathbb{N}^*$. En définitive, un polynôme non nul de degré n s'écrit $P = \sum_{k=0}^n a_k X^k$, en convenant

que $X^0 = 1$. On notera aussi $P(X) = \sum_{k=0}^n a_k X^k$. On dit alors que X est une indéterminée et on note $\mathbb{K}[X]$ l'algèbre des polynômes à coefficients dans $\mathbb{K}$.

Un *monôme* est un polynôme de la forme λX^n où $\lambda \in \mathbb{K}$ et $n \in \mathbb{N}$ (pour $n = 0$, il s'agit d'un polynôme constant).

Pour tout entier naturel n , on note $\mathbb{K}_n[X]$ le sous-ensemble de $\mathbb{K}[X]$ formé des polynômes de degré au plus égal à n . On vérifie facilement que $\mathbb{K}_n[X]$ est un sous-

espace vectoriel de $\mathbb{K}[X]$. Ce sous-espace $\mathbb{K}_n[X]$ est de dimension $n + 1$, la famille $(X^k)_{0 \leq k \leq n}$ en étant une base (la base canonique de $\mathbb{K}_n[X]$).

Un polynôme $P(X) = \sum_{k=0}^n a_k X^k$ de degré $n \geq 0$ est dit *unitaire* si $a_n = 1$.

Théorème

Les éléments inversibles de l'anneau $\mathbb{K}[X]$ sont les polynômes constants non nuls.

Définition

Soient $P = \sum_{k=0}^n a_k X^k$ et Q deux polynômes $\mathbb{K}[X]$. La *composée* des polynômes P et Q est le polynôme $P \circ Q = \sum_{k=0}^n a_k Q^k$ avec la convention que $Q^0 = 1$ pour tout polynôme Q .

Théorème

Pour tous polynômes P et Q non nuls dans $\mathbb{K}[X]$, on a :

$$\deg(P \circ Q) = \deg(P) \deg(Q).$$

Définition

Une famille de polynômes non nuls $(P_k)_{k \in \mathbb{N}}$ est dite *étagée* en degrés [resp. en valuations] si on a $\deg(P_k) < \deg(P_{k+1})$ [resp. $\text{val}(P_k) < \text{val}(P_{k+1})$] pour tout $k \in \mathbb{N}$ et qu'elle est *échelonnée* en degrés [resp. en valuations] si on a $\deg(P_k) = k$ [resp. $\text{val}(P_k) = k$] pour tout $k \in \mathbb{N}$.

Théorème

Une famille $(P_k)_{k \in \mathbb{N}}$ de polynômes non nuls étagée en degrés [resp. en valuations] est libre dans l'espace vectoriel $\mathbb{K}[X]$.

Corollaire

Une famille $(P_k)_{k \in \mathbb{N}}$ [resp. $(P_k)_{0 \leq k \leq n}$] de polynômes non nuls échelonnée en degrés [resp. en valuations] est une base de $\mathbb{K}[X]$ [resp. de $\mathbb{K}_n[X]$].

20.2 Division euclidienne dans $\mathbb{K}[X]$

Théorème

Pour tout couple (A, B) de polynômes à coefficients dans $\mathbb{K}$ avec $B \neq 0$, il existe un unique couple (Q, R) de polynômes tel que $A = BQ + R$ et $\deg(R) < \deg(B)$.

Définition

On dit que deux polynômes A, B non nuls dans $\mathbb{K}[X]$ sont *associés*, s'il existe $\lambda \in \mathbb{K}^*$ tel que $B = \lambda A$.

Pour tout $P \in \mathbb{K}[X]$, l'ensemble $(P) = \{PQ ; Q \in \mathbb{K}[X]\}$ des multiples de P dans $\mathbb{K}[X]$ est l'*idéal engendré* par P et on le note aussi $P \cdot \mathbb{K}[X]$. Un tel idéal est dit *principal*. Plus généralement, pour toute famille $(P_k)_{1 \leq k \leq r}$ d'éléments de $\mathbb{K}[X]$, l'ensemble $(P_1, \dots, P_r) = \left\{ \sum_{k=1}^r Q_k P_k ; (Q_k)_{1 \leq k \leq r} \in \mathbb{K}[X]^r \right\}$ est un idéal. On dit que c'est l'idéal engendré par $P_1, \dots, P_r$. On le note aussi $\sum_{k=1}^r P_k \cdot \mathbb{K}[X]$.

Théorème

Pour tout idéal I de $\mathbb{K}[X]$ non réduit à $\{0\}$, il existe un unique polynôme unitaire A tel que $I = (A)$.

Le théorème précédent se traduit en disant que l'anneau $\mathbb{K}[X]$ est principal.

Théorème

Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r$ dans $\mathbb{K}[X]$ non tous nuls. Il existe un unique polynôme unitaire $\Delta \in \mathbb{K}[X]$ tel que $(A_1, \dots, A_r) = (\Delta)$. Il existe des polynômes $U_1, \dots, U_r$ tels que $\Delta = \sum_{k=1}^r U_k A_k$ et Δ est aussi caractérisé par la propriété suivante :

$$\begin{cases} \forall k \in \{1, \dots, r\}, \Delta \text{ divise } A_k \\ \text{tout diviseur commun à } A_1, \dots, A_r \text{ divise } \Delta \end{cases} \quad (20.1)$$

Définition

Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r$ dans $\mathbb{K}[X]$ non tous nuls. On appelle *plus grand commun diviseur* de $A_1, \dots, A_r$ le polynôme unitaire Δ qui vérifie $(A_1, \dots, A_r) = (\Delta)$ ou (20.1). On note $\text{pgcd}(A_1, \dots, A_r)$ ce polynôme ou encore $A_1 \wedge \dots \wedge A_r$.

La relation $\Delta = \sum_{k=1}^r U_k A_k$ est une *identité de Bézout*.

Théorème

Soient P, Q, R dans $\mathbb{K}[X]$ et $\lambda \in \mathbb{K}$, avec P et λ non nuls. En désignant par α le coefficient dominant de P , on a :

- $P \wedge 0 = \frac{1}{\alpha} P$; $P \wedge \lambda = 1$;
- $P \wedge Q = Q \wedge P$ (commutativité du pgcd) ;
- $P \wedge Q = \frac{1}{\alpha} P$ si, et seulement si, P divise Q ;
- $(PQ) \wedge (PR) = \frac{1}{\alpha} P (Q \wedge R)$;
- pour Q non nul, $P \wedge (Q \wedge R) = (P \wedge Q) \wedge R$ (associativité du pgcd).

Comme pour l'anneau $\mathbb{Z}$ des entiers relatifs, on dispose de l'algorithme d'Euclide pour obtenir le pgcd de deux polynômes A et B non nuls. Le principe de cet algorithme est le suivant pour A, B non nuls tels que $\deg(A) \geq \deg(B)$.

On définit la suite $(R_n)_{n \geq 0}$ de polynômes par :

- $R_0 = B$; R_1 est le reste dans la division euclidienne de A par B ; on a donc $R_1 = 0$ ou $0 \leq \deg(R_1) < \deg(R_0)$;
- pour $n \geq 2$, si $R_{n-1} = 0$ alors $R_n = 0$, sinon R_n est le reste dans la division euclidienne de R_{n-2} par R_{n-1} et on a $R_n = 0$ ou $0 \leq \deg(R_n) < \deg(R_{n-1})$.

Il existe alors un entier $p \geq 1$ tel que $R_p = 0$, $0 \leq \deg(R_{p-1}) < \dots < \deg(R_0)$ et $A \wedge B = R_0 \wedge R_1 = \dots = R_{p-1} \wedge R_p = R_{p-1}$ (à un facteur multiplicatif non nul près), c'est-à-dire que $A \wedge B$ est, à un facteur multiplicatif non nul près, le dernier reste non nul dans cette suite finie de divisions euclidiennes. On a en fait construit avec l'algorithme d'Euclide, dans le cas où B ne divise pas A , deux suites de polynômes $(R_n)_{0 \leq n \leq p}$ et $(Q_n)_{1 \leq n \leq p}$ de la manière suivante :

$$\left\{ \begin{array}{l} A = Q_1 R_0 + R_1 \quad (R_1 \neq 0 \text{ et } \deg(R_1) < \deg(R_0)) \\ R_0 = Q_2 R_1 + R_2 \quad (R_2 \neq 0 \text{ et } \deg(R_2) < \deg(R_1)) \\ \vdots \\ R_{p-3} = Q_{p-1} R_{p-2} + R_{p-1} \quad (R_{p-1} \neq 0 \text{ et } \deg(R_{p-1}) < \deg(R_{p-2})) \\ R_{p-2} = Q_p R_{p-1} + R_p \quad (R_p = 0) \end{array} \right.$$

On vérifie alors qu'il existe deux polynômes U_k et V_k tels que $R_k = AU_k + BV_k$. En particulier pour $k = p-1$ on a $A \wedge B = R_{p-1} = AU_{p-1} + BV_{p-1} = AU + BV$.

20.3 Polynômes premiers entre eux

Définition

Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r$ des polynômes dans $\mathbb{K}[X]$ tous non nuls. On dit que $A_1, \dots, A_r$ sont *premiers entre eux* dans leur ensemble si leur pgcd vaut 1. Pour $r = 2$, on dira simplement que A_1 et A_2 sont premiers entre eux.

Des polynômes deux à deux premiers entre eux sont premiers entre eux dans leur ensemble.

Théorème de Bézout

Soient $A_1, \dots, A_r$ des polynômes dans $\mathbb{K}[X]$ tous non nuls avec $r \geq 2$. Ces polynômes sont premiers entre eux dans leur ensemble si, et seulement si, il existe des polynômes $U_1, \dots, U_r$ tel que $\sum_{k=1}^r U_k A_k = 1$.

Corollaire

Soient $A_1, \dots, A_r$ des polynômes non nuls avec $r \geq 2$ et $\Delta = A_1 \wedge \dots \wedge A_r$. Il existe des polynômes, $B_1, \dots, B_r$ premiers entre eux dans leur ensemble, tels que $A_k = \Delta B_k$ pour tout k compris entre 1 et r .

Corollaire (Gauss)

Soient A, B, C trois polynômes non nuls. Si A divise BC et est premier avec B , alors A divise C .

Corollaire

Soient A, B, C trois polynômes non nuls. Si C est premier avec A alors $A \wedge B = A \wedge (BC)$ (le pgcd de deux polynômes est inchangé si on multiplie l'un d'eux par un polynôme premier avec l'autre).

Corollaire

Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r, C$ des polynômes non nuls. Si C est premier avec chacun des A_k , pour tout k compris entre 1 et r , il est alors premier avec leur produit $\prod_{k=1}^r A_k$.

20.4 Racines des polynômes

À tout polynôme $P(X) = \sum_{k=0}^n a_k X^k \in \mathbb{K}[X]$, on associe la *fonction polynomiale*

$\tilde{P} : \mathbb{K} \rightarrow \mathbb{K}$ définie par $\tilde{P}(x) = \sum_{k=0}^n a_k x^k$ pour tout $x \in \mathbb{K}$.

Pour $\alpha \in \mathbb{K}$, on note $P(\alpha) = \sum_{k=0}^n a_k \alpha^k$ l'évaluation de la fonction polynôme $\tilde{P}$ en α (ce qu'il faudrait noter en toute rigueur $\tilde{P}(\alpha)$). Cette évaluation en α est en fait la composition de P avec le polynôme constant α , qui est bien un polynôme constant. $P(0) = a_0$ est le coefficient constant de P .

Théorème

Pour tout $\alpha \in \mathbb{K}$, l'application $\varphi_\alpha : P \mapsto P(\alpha)$ réalise un morphisme d'algèbre de $\mathbb{K}[X]$ dans $\mathbb{K}$.

Définition

Soit $P \in \mathbb{K}[X]$. On dit que $\alpha \in \mathbb{K}$ est *racine* de P , si $P(\alpha) = 0$.

Définition

Soient $P \in \mathbb{K}[X] \setminus \mathbb{K}$ un polynôme non constant, $\alpha \in \mathbb{K}$ et $m \in \mathbb{N}^*$. On dit que α est *racine d'ordre* (ou de multiplicité) m de P si $(X - \alpha)^m$ divise P et $(X - \alpha)^{m+1}$ ne divise pas P .

Si α est racine de $P \in \mathbb{K}[X] \setminus \mathbb{K}$, sa multiplicité est alors comprise entre 1 et $\deg(P)$. Pour $m = 1$, on dit que α est une racine simple, pour $m = 2$, qu'elle est racine double et pour $m \geq 2$, qu'elle est racine multiple. Si α n'est pas racine de P , on peut dire que c'est une racine de multiplicité nulle.

Théorème

Soient $P \in \mathbb{K}[X] \setminus \mathbb{K}$, $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ deux à deux distincts et $m_1, \dots, m_r$ des entiers naturels non nuls. Les deux assertions suivantes sont équivalentes :

- pour tout k compris entre 1 et r , α_k est racine de P de multiplicité m_k ;
- il existe $Q \in \mathbb{K}[X]$ tel que $P(X) = Q(X) \prod_{k=1}^r (X - \alpha_k)^{m_k}$ et $Q(\alpha_k) \neq 0$ pour tout k compris entre 1 et r .

Corollaire

Si $P \in \mathbb{K}[X] \setminus \mathbb{K}$ admet $r \geq 1$ racines distinctes $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ de multiplicités respectives $m_1, \dots, m_r$, on a alors $\deg(P) \geq \sum_{k=1}^r m_k$.

Corollaire

Un polynôme $P \in \mathbb{K}[X]$ de degré $n \geq 1$ admet au plus n racines distinctes ou confondues dans $\mathbb{K}$.

Du corollaire précédent, on déduit que si P et Q dans $\mathbb{K}_n[X]$ coïncident en au moins $n + 1$ points distincts, ils sont alors égaux. Le corollaire précédent peut aussi s'exprimer en disant que si $P \in \mathbb{K}_n[X]$ admet $r \geq 1$ racines distinctes $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ de multiplicités respectives $m_1, \dots, m_r$ avec $\sum_{i=1}^r m_i > n$, c'est alors le polynôme nul.

Définition

On dit qu'un polynôme $P \in \mathbb{K}[X]$ est *scindé* sur $\mathbb{K}$, s'il est constant ou de degré $n \geq 1$ et admet $r \geq 1$ racines distinctes $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ de multiplicités respectives $m_1, \dots, m_r$ avec $\sum_{i=1}^r m_i = n$. Dans le cas où tous les m_i sont égaux à 1, on dit que le polynôme est *scindé à racines simples* ou encore qu'il est *simplement scindé*.

Un polynôme scindé non constant est donc de la forme $P(X) = \lambda \prod_{k=1}^r (X - \alpha_k)^{m_k}$, où λ est une constante non nulle, les α_k sont des scalaires deux à deux distincts et les m_k des entiers naturels non nuls.

Théorème de d'Alembert-Gauss ¹

Toute équation polynomiale à coefficients complexes de degré non nul n admet n racines complexes distinctes ou confondues.

Théorème (formules de Viète)

Soient $P(X) = \sum_{k=0}^n a_k X^k \in \mathbb{C}[X]$ de degré $n \in \mathbb{N}^*$ et $z_1, \dots, z_n$ les racines de P .

Soit, pour tout $k \in \llbracket 1, n \rrbracket$, $\sigma_k = \sum_{1 \leq i_1 < \dots < i_k \leq n} z_{i_1} \dots z_{i_k}$.

Alors, pour tout $k \in \llbracket 1, n \rrbracket$, $\sigma_k = (-1)^k \frac{a_{n-k}}{a_n}$.

En particulier, $\sigma_1 = z_1 + \dots + z_n = -\frac{a_{n-1}}{a_n}$ et $\sigma_n = z_1 \dots z_n = (-1)^n \frac{a_0}{a_n}$.

Par exemple pour $n = 3$, $\sigma_1 = z_1 + z_2 + z_3$, $\sigma_2 = z_1 z_2 + z_1 z_3 + z_2 z_3$ et $\sigma_3 = z_1 z_2 z_3$.

20.5 Théorème de Leibniz et formule de Taylor

Définition

Soit P un polynôme dans $\mathbb{K}[X]$. Le *polynôme dérivé* de P est le polynôme P' défini par :

$$P'(X) = \begin{cases} 0 & \text{si } P = a_0 \text{ est constant} \\ \sum_{k=1}^n k a_k X^{k-1} & \text{si } P(X) = \sum_{k=0}^n a_k X^k \text{ est de degré } n \geq 1 \end{cases}.$$

On a en particulier $(X^k)' = kX^{k-1}$ pour tout entier $k \geq 1$ et, pour tout polynôme P de degré $n \geq 1$, le polynôme P' est de degré $n - 1$.

1. également appelé *théorème fondamental de l'algèbre*.

Théorème

L'application $P \mapsto P'$ est linéaire de $\mathbb{K}[X]$ dans $\mathbb{K}[X]$, de $\mathbb{K}_p[X]$ dans $\mathbb{K}_{p-1}[X]$ pour $p \geq 1$ et pour tous polynômes P, Q dans $\mathbb{K}[X]$, on a :

$$(PQ)' = P'Q + PQ', \quad (P \circ Q)' = (P' \circ Q) Q'.$$

On définit les dérivées successives d'un polynôme P , en les notant $P^{(k)}$, par la relation de récurrence $P^{(0)} = P$ et $P^{(k)} = (P^{(k-1)})'$ pour tout $k \geq 1$. En particulier, on a $(X^k)^{(p)} = \frac{k!}{(k-p)!} X^{k-p}$ si $0 \leq p \leq k$ et $(X^k)^{(p)} = 0$ si $p \geq k+1$.

Théorème de Leibniz

Pour tous polynômes P, Q et tout entier naturel n , on a :

$$(PQ)^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n-k)} = \sum_{k=0}^n \binom{n}{k} P^{(n-k)} Q^{(k)}.$$

Théorème (formule de Taylor)

Pour tout polynôme $P(X) = \sum_{k=0}^n a_k X^k$ et tout scalaire $a \in \mathbb{K}$, on a :

$$P(X) = \sum_{k=0}^n \frac{P^{(k)}(a)}{k!} (X-a)^k.$$

Théorème

Soient $P \in \mathbb{K}[X] \setminus \{0\}$, $\alpha \in \mathbb{K}$ et $m \in \mathbb{N}^*$. Les assertions suivantes sont équivalentes :

- α est racine d'ordre $m \geq 1$ de P ;
- il existe un polynôme $Q \in \mathbb{K}[X]$ tel que $P(X) = Q(X) (X - \alpha)^m$ et $Q(\alpha) \neq 0$;
- $P^{(k)}(\alpha) = 0$ pour k compris entre 0 et $m-1$ et $P^{(m)}(\alpha) \neq 0$.

20.6 Polynômes irréductibles

Définition

Un polynôme $P \in \mathbb{K}[X] \setminus \{0\}$ est dit *irréductible* s'il est non constant et n'est divisible que par les constantes non nulles ou les polynômes λP avec $\lambda \in \mathbb{K}^*$.

Théorème d'Euclide

Soient $P \in \mathbb{K}[X]$ un polynôme irréductible. Si P divise un produit $\prod_{k=1}^r A_k$ de $r \geq 2$ polynômes non nuls, il divise alors l'un des A_k .

Théorème

Tout polynôme non constant $P \in \mathbb{K}[X]$ est produit de polynômes irréductibles et une telle décomposition est unique² à l'ordre près des facteurs, ce qui signifie qu'il existe une constante $\lambda \in \mathbb{K}^*$, un entier $r \geq 1$, des polynômes unitaires $P_1, \dots, P_r$ deux à deux distincts et irréductibles et des entiers naturels non nuls $m_1, \dots, m_r$ tels que $P = \lambda \prod_{k=1}^r P_k^{m_k}$.

Corollaire

Tout polynôme non constant $P \in \mathbb{K}[X]$ admet au moins un diviseur irréductible.

Pour la suite de ce paragraphe, on s'intéresse aux polynômes irréductibles dans $\mathbb{R}[X]$. On sait déjà que les polynômes de degré 1 sont irréductibles.

Théorème

Les polynômes réels irréductibles sont les polynômes de degré 1 et les polynômes de degré 2, $P(X) = aX^2 + bX + c$ tels que $b^2 - 4ac < 0$.

2. L'unicité signifie que si on a une autre décomposition du même type $P = \mu \prod_{k=1}^s Q_k^{n_k}$, on a alors $\mu = \lambda$, $s = r$ et il existe une permutation $\sigma \in \mathcal{S}_r$ telle que $Q_k = P_{\sigma(k)}$ et $n_k = m_{\sigma(k)}$ pour tout k compris entre 1 et r .

Théorème

Tout polynôme non constant $P \in \mathbb{R}[X]$ est produit de polynômes irréductibles de degré 1 ou 2, c'est-à-dire qu'il existe une constante $\lambda \in \mathbb{R}^*$, deux entiers naturels r et s , des réels $a_1, \dots, a_r$ deux à deux distincts, des entiers naturels non nuls $n_1, \dots, n_r$, des couples de réels $(b_1, c_1), \dots, (b_s, c_s)$ deux à deux distincts tels que $b_k^2 - 4c_k < 0$ pour tout k et des entiers naturels non nuls $m_1, \dots, m_s$ tels que
$$P = \lambda \prod_{k=1}^r (X - a_k)^{n_k} \prod_{k=1}^s (X^2 + b_k X + c_k)^{m_k}.$$
 Une telle décomposition est unique à l'ordre près des facteurs.

Lemme

Soient P, Q deux polynômes unitaires non constants $\mathbb{K}[X]$ et $P = \prod_{k=1}^r P_k^{m_k}$, $Q = \prod_{k=1}^r P_k^{n_k}$ leurs décompositions en produit de polynômes irréductibles avec les P_k unitaires irréductibles deux à deux distincts et les n_k, m_k entiers naturels (certains de ces entiers pouvant être nuls). On a alors :

$$(P \text{ divise } Q) \Leftrightarrow (\forall k \in \{1, \dots, r\}, m_k \leq n_k).$$

Théorème

Avec les notations du lemme précédent, on a $P \wedge Q = \prod_{k=1}^r P_k^{\min(m_k, n_k)}$.

De manière plus générale si $P_1, \dots, P_n$ sont des polynômes unitaires non constants et $P_j = \prod_{k=1}^r P_k^{m_{jk}}$ la décomposition en produit de polynômes irréductibles de chacun de ces polynômes, on a $P_1 \wedge \dots \wedge P_n = \prod_{k=1}^r P_k^{\min(m_{1,k}, \dots, m_{n,k})}$.

20.7 Décomposition des fractions rationnelles en éléments simples

Une *fonction rationnelle* est une fonction $x \mapsto f(x) = \frac{P(x)}{Q(x)}$, où P, Q sont deux fonctions polynomiales à coefficients complexes avec $Q \neq 0$ premier avec P . Elle est définie sur $\mathbb{R}$ privé des racines réelles de Q .

Une *fraction rationnelle* est le quotient $\frac{P}{Q}$ de deux polynômes à coefficients complexes avec $Q \neq 0$ premier avec P .

Définition

$\alpha \in \mathbb{C}$ est un *pôle* d'ordre m de $\frac{P}{Q}$, s'il est racine d'ordre m de Q .

Théorème

Soient $f = \frac{P}{Q}$ une fraction rationnelle (P et Q étant premiers entre eux) et $\alpha_1, \dots, \alpha_p$ les pôles de f d'ordres respectifs $m_1, \dots, m_p$. Il existe des nombres complexes $(\lambda_{jk})_{\substack{1 \leq j \leq p \\ 1 \leq k \leq m_j}}$, avec $\lambda_{j,m_j} \neq 0$ pour tout j compris entre

1 et p , et une fonction polynomiale R , uniquement déterminés, tels que :

$$f(X) = \sum_{j=1}^p \sum_{k=1}^{m_j} \frac{\lambda_{jk}}{(X - \alpha_j)^k} + R(X) \quad (20.2)$$

Définition

Avec les notations du théorème précédent, on dit que la décomposition (20.2) est la *décomposition en éléments simples* de f et, pour tout j compris entre 1 et p , on dit que la fraction rationnelle $\sum_{k=1}^{m_j} \frac{\lambda_{jk}}{(X - \alpha_j)^k}$ est la *partie polaire* de f associée au pôle α_j .

Dans une telle décomposition en éléments simples, le polynôme R est le quotient dans la division euclidienne de P par Q .

Théorème

Soit $f = \frac{P}{Q}$ une fraction rationnelle réelle (P et Q étant premiers entre eux) avec $Q(X) = \prod_{k=1}^r (X - x_k)^{m_k} \prod_{k=1}^s (X^2 + \beta_k X + \gamma_k)^{n_k}$, où $x_1, \dots, x_r$ sont des réels deux à deux distincts, $(\beta_1, \gamma_1), \dots, (\beta_s, \gamma_s)$ des couples de réels deux à deux distincts tels que $\beta_k^2 - 4\gamma_k < 0$ pour tout k et $m_1, \dots, m_r, n_1, \dots, n_s$ des entiers naturels non nuls. Il existe des réels $(\lambda_{jk})_{\substack{1 \leq j \leq r \\ 1 \leq k \leq m_j}}, (\mu_{jk})_{\substack{1 \leq j \leq s \\ 1 \leq k \leq n_j}}, (\nu_{jk})_{\substack{1 \leq j \leq s \\ 1 \leq k \leq n_j}}$ et une fonction polynomiale réelle R , uniquement déterminés, tels que :

$$f(X) = \sum_{j=1}^r \sum_{k=1}^{m_j} \frac{\lambda_{jk}}{(X - \alpha_j)^k} + \sum_{j=1}^s \sum_{k=1}^{n_j} \frac{\mu_{jk}X + \nu_{jk}}{(X^2 + \beta_j X + \gamma_j)^k} + R(X).$$

20.8 Polynômes d'interpolation de Lagrange

Pour tout entier naturel n , on se donne une famille $(x_i)_{0 \leq i \leq n}$ de $n+1$ éléments de $\mathbb{K}$ deux à deux distincts et à cette famille on associe la famille $(L_i)_{0 \leq i \leq n}$ de polynômes définie par $L_i(X) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{X - x_j}{x_i - x_j}$ pour $0 \leq i \leq n$. On peut remarquer que $L_i(x_j) = \delta_{ij}$ pour $0 \leq i, j \leq n$.

Théorème

Avec les notations qui précèdent, la famille de polynômes $(L_i)_{0 \leq i \leq n}$ est une base de $\mathbb{K}_n[X]$.

Tout polynôme $P \in \mathbb{K}_n[X]$ s'écrit donc, $P = \sum_{i=1}^n \lambda_i L_i$, l'évaluation en x_j nous donnant $\lambda_j = P(x_j)$, pour tout j compris entre 0 et n .

Théorème

Avec les notations qui précèdent, pour tous scalaires $y_0, y_1, \dots, y_n$, il existe un unique polynôme $P \in \mathbb{K}_n[X]$ tel que $P(x_j) = y_j$ pour tout j compris entre 0 et n . Ce polynôme est donné par $P = \sum_{i=1}^n y_i L_i$.

Avec les notations du théorème précédent, on dit que P est le *polynôme d'interpolation de Lagrange* associé à $x = (x_i)_{0 \leq i \leq n}$ et $y = (y_i)_{0 \leq i \leq n}$.

Théorème

Soient $a < b$ deux réels, $n \in \mathbb{N}^*$ et $(x_k)_{0 \leq k \leq n}$ une suite de points deux à deux distincts dans $[a, b]$. Il existe une unique suite de réels $(\mu_k)_{0 \leq k \leq n}$ telle que pour tout polynôme P dans $\mathbb{R}_n[X]$ on a $\int_a^b P(x) dx = \sum_{k=0}^n \mu_k P(x_k)$.

Réduction des endomorphismes

21

$\mathbb{K}$ est un corps, E un $\mathbb{K}$ -espace vectoriel de dimension finie $n \geq 1$ et on se donne $u \in \mathcal{L}(E)$.

21.1 Valeurs et vecteurs propres

u est un endomorphisme de E .

Définition

On dit que $\lambda \in \mathbb{K}$ est *valeur propre* de u , si on a, $\text{Ker}(u - \lambda \text{Id}) \neq \{0\}$.

La définition précédente est en fait valable pour E de dimension infinie.

Dire que $\lambda \in \mathbb{K}$ est valeur propre de u équivaut à dire qu'il existe $x \in E \setminus \{0\}$ tel que $u(x) = \lambda x$, ce qui revient aussi à dire que $u - \lambda \text{Id}$ est non inversible puisque E est de dimension finie. On dit alors que x est un *vecteur propre* de u associé à la valeur propre λ et que le sous-espace vectoriel $E_\lambda = \text{Ker}(u - \lambda \text{Id})$ de E est le *sous-espace propre* associé à λ .

L'ensemble des valeurs propres de u est appelé le *spectre* de u et noté $\text{Sp}(u)$.

Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de $u \in \mathcal{L}(E)$ si, et seulement si, le *polynôme caractéristique* de u défini par $\chi_u(X) = \det(X \text{Id} - u)$ est nul pour $X = \lambda$. On a donc $\text{Sp}(u) = \chi_u^{-1}(\{0\}) = \{\lambda \in \mathbb{K}, \det(\lambda \text{Id} - u) = 0\}$ et c'est une partie finie de $\mathbb{K}$ ayant au plus n éléments. Pour $\mathbb{K} = \mathbb{R}$, ce spectre peut être vide.

Définition

On dit que λ dans $\mathbb{K}$ est *valeur propre* de $A \in \mathcal{M}_n(\mathbb{K})$ s'il existe un vecteur non nul X dans $\mathbb{K}^n$ tel que $AX = \lambda X$. On dit alors que X est un *vecteur propre* de A associé à la valeur propre λ et que le sous-espace vectoriel de $\mathbb{K}^n$, $E_\lambda = \{X \in \mathbb{K}^n, AX = \lambda X\}$ (que l'on peut noter $\text{Ker}(A - \lambda I_n)$) est le *sous-espace propre* associé à λ .

L'ensemble $\text{Sp}(A)$ des valeurs propres de A est appelé le *spectre* de A et le *polynôme caractéristique* de $A \in \mathcal{M}_n(\mathbb{K})$ est défini par $\chi_A(X) = \det(XI_n - A)$.

Définition

Un sous-espace vectoriel F de E est dit *stable* par u , s'il vérifie $u(F) \subset F$.

Lemme

Si F est un sous-espace vectoriel de E stable par u , le polynôme caractéristique de la restriction de u à F divise alors celui de u .

Théorème

Si $\lambda \in \mathbb{K}$ est une valeur propre de u de multiplicité α en tant que racine du polynôme caractéristique de u , on a alors $1 \leq \dim(\text{Ker}(u - \lambda \text{Id})) \leq \alpha$.

Théorème

Si $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$, les sous-espaces propres $\text{Ker}(u - \lambda_k \text{Id})$, pour k compris entre 1 et p , sont alors en somme directe.

21.2 Diagonalisation

Définition

Soient A et A' deux matrices carrées d'ordre $n \in \mathbb{N}^*$. On dit que A et A' sont *semblables* (ou que A est *semblable* à A') s'il existe une matrice carrée inversible P d'ordre n telle que $A' = P^{-1}AP$.

Théorème

Deux matrices semblables ont même trace, même déterminant, même rang et même polynôme caractéristique.

Définition

On dit que u est *diagonalisable*, s'il existe une base de E dans laquelle la matrice de u est diagonale.

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est *diagonalisable*, si elle est semblable à une matrice diagonale.

La matrice A est donc diagonalisable si, et seulement si, il existe une matrice diagonale D et une matrice inversible P telles que $D = P^{-1}AP$.

Théorème

Si u a n valeurs propres distinctes dans $\mathbb{K}$, il est alors diagonalisable.

Théorème

Les conditions suivantes sont équivalentes :

- l'endomorphisme u est diagonalisable ;
- en notant $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$, on a $E = \bigoplus_{k=1}^p \text{Ker}(u - \lambda_k \text{Id})$;
- en notant $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$, on a $\sum_{k=1}^p \dim(\text{Ker}(u - \lambda_k \text{Id})) = n$;
- le polynôme caractéristique de u est scindé sur $\mathbb{K}$ de racines deux à deux distinctes $\lambda_1, \dots, \lambda_p$ dans $\mathbb{K}$, chaque λ_k ($1 \leq k \leq p$) étant de multiplicité $\alpha_k = \dim(\text{Ker}(u - \lambda_k \text{Id}))$.

21.3 Endomorphismes trigonalisables

Définition

On dit que u est *trigonalisable*, s'il existe une base de E dans laquelle la matrice de u est triangulaire.

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est *trigonalisable*, si elle est semblable à une matrice triangulaire.

Théorème

L'endomorphisme u est trigonalisable sur $\mathbb{K}$ si, et seulement si, son polynôme caractéristique est scindé sur $\mathbb{K}$.

Corollaire

Sur $\mathbb{C}$, tout endomorphisme $u \in \mathcal{L}(E)$ est trigonalisable.

Corollaire

Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est trigonalisable.

Corollaire

Si $u \in \mathcal{L}(E)$ est trigonalisable et si F est un sous-espace vectoriel de E stable par u , alors la restriction de u à F est aussi trigonalisable.

Corollaire

Si $u \in \mathcal{L}(E)$ [resp. $A \in \mathcal{M}_n(\mathbb{K})$] est trigonalisable, alors la trace de u [resp. de A] est égale à la somme des valeurs propres de u [resp. de A] et le déterminant de u [resp. de A] est égal au produit des valeurs propres de u [resp. de A].

21.4 Polynômes d'endomorphismes, polynôme minimal

On définit les puissances successives de u par $u^0 = \text{Id}$ et $u^{k+1} = u \circ u^k$ pour tout $k \in \mathbb{N}$, ce qui permet de définir, pour tout polynôme $P(X) = \sum_{k=0}^p a_k X^k$

dans $\mathbb{K}[X]$, l'endomorphisme $P(u)$, par $P(u) = \sum_{k=0}^p a_k u^k$. On note naturellement

$\mathbb{K}[u] = \{P(u) ; P \in \mathbb{K}[X]\}$ et cet ensemble est un sous-espace vectoriel et un sous-anneau commutatif de $\mathcal{L}(E)$. Précisément, $\mathbb{K}[u]$ est une sous-algèbre de $\mathcal{L}(E)$ et

pour ce qui est de la commutativité, on a pour tous P, Q dans $\mathbb{K}[X]$:

$$(PQ)(u) = P(u) \circ Q(u) = Q(u) \circ P(u) = (QP)(u).$$

On définit de manière analogue la sous-algèbre $\mathbb{K}[A]$ de $\mathcal{M}_n(\mathbb{K})$ engendrée par une matrice $A \in \mathcal{M}_n(\mathbb{K})$.

Si $A \in \mathcal{M}_n(\mathbb{K})$ est la matrice de $u \in \mathcal{L}(E)$ dans une base $\mathcal{B}$, alors $P(A)$ est la matrice de $P(u)$ dans $\mathcal{B}$.

Définition

Soient $P \in \mathbb{K}[X]$ non nul et $u \in \mathcal{L}(E)$. On dit que P est un *polynôme annulateur* de u (ou que P annule u) si $P(u) = 0$.

Théorème

Soit $P \in \mathbb{K}[X]$. Pour toute valeur propre $\lambda \in \text{Sp}(u)$, $P(\lambda)$ est valeur propre de $P(u)$. Pour $\mathbb{K} = \mathbb{C}$, on a $\text{Sp}(P(u)) = \{P(\lambda) ; \lambda \in \text{Sp}(u)\}$.

Corollaire

Soient $P \in \mathbb{K}[X]$ non nul et $u \in \mathcal{L}(E)$. Si P annule u , toute valeur propre de u est racine de P .

L'espace vectoriel $\mathcal{L}(E)$ étant de dimension n^2 , la famille $(u^k)_{0 \leq k \leq n^2}$ est liée, ce qui se traduit en disant qu'il existe un polynôme P non nul dans $\mathbb{K}[X]$ tel que $P(u) = 0$. Il en résulte que l'ensemble $I_u = \{P \in \mathbb{K}[X], P(u) = 0\}$ n'est pas réduit au polynôme nul. Cet ensemble qui est le noyau du morphisme d'anneaux $P \mapsto P(u)$ est un idéal de $\mathbb{K}[X]$. L'anneau $\mathbb{K}[X]$ étant principal on peut donner la définition suivante.

Définition

I_u est appelé *idéal annulateur* de u . On appelle *polynôme minimal* de u l'unique polynôme π_u tel que $I_u = \pi_u \cdot \mathbb{K}[X]$.

On définit de manière analogue l'idéal annulateur et le polynôme minimal d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$.

Si u a pour matrice A dans une base $\mathcal{B}$ de E , alors A et u ont même idéal annulateur et même polynôme minimal.

Deux matrices semblables ont même idéal annulateur et même polynôme minimal.

Théorème

Les valeurs propres de u sont les racines de son polynôme minimal.

Théorème

Dans le cas où le polynôme minimal π_u est scindé sur $\mathbb{K}$, le déterminant de u est le produit des valeurs propres et la trace de u est la somme des valeurs propres, ces dernières étant comptées avec leurs multiplicités.

Lemme de décomposition des noyaux

Soient p un entier supérieur ou égal à 2, $P_1, \dots, P_p$ des polynômes non nuls dans $\mathbb{K}[X]$ deux à deux premiers entre eux et $P = \prod_{k=1}^p P_k$.

On a $\text{Ker}(P(u)) = \bigoplus_{k=1}^p \text{Ker}(P_k(u))$ et les projecteurs associés à cette décomposition sont des éléments de $\mathbb{K}[u]$.

En notant $\pi_u = \prod_{k=1}^p P_k^{\beta_k}$ la décomposition du polynôme minimal π_u en facteurs irréductibles dans $\mathbb{K}[X]$, on déduit du lemme de décomposition des noyaux que $E = \bigoplus_{k=1}^p \text{Ker}(P_k^{\beta_k}(u))$, chaque sous-espace vectoriel $E_k = \text{Ker}(P_k^{\beta_k}(u))$ étant stable par u .

Théorème de Cayley-Hamilton

Si χ_u est le polynôme caractéristique de u , on a alors $\chi_u(u) = 0$.

Corollaire

Le polynôme minimal π_u divise le polynôme caractéristique χ_u . On a donc $\dim(\mathbb{K}[u]) = \deg(\pi_u) \leq n$.

Pour $\mathbb{K} = \mathbb{C}$, on a $\chi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$ avec $\alpha_k \in \mathbb{N}^*$ et les λ_k deux à deux distincts et $\pi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\beta_k}$ avec $1 \leq \beta_k \leq \alpha_k$.

Dans le cas où l'endomorphisme u est inversible, le théorème de Cayley-Hamilton nous donne un moyen de calculer l'inverse de u . En notant $\chi_u(X) = \sum_{k=0}^n a_k X^k$, on

$$a u^{-1} = -\frac{1}{a_0} \sum_{k=1}^n a_k u^{k-1}.$$

Le théorème de Cayley-Hamilton permet également de calculer u^p pour tout entier $p \geq n$ en fonction de $\text{Id}, u, \dots, u^{n-1}$. Pour $p = n$, on a $u^n = -\sum_{k=0}^{n-1} a_k u^k$ et pour $p > n$ la division euclidienne de X^p par χ_u , $X^p = Q\chi_u + R$ avec $\deg(R) < n$, donne $u^p = R(u)$.

Pour la réduction des endomorphismes, on dispose des théorèmes suivants.

Théorème

Les conditions suivantes sont équivalentes :

- l'endomorphisme u est diagonalisable ;
- il existe un polynôme annulateur de u scindé à racines simples dans $\mathbb{K}$;
- le polynôme minimal π_u est scindé à racines simples dans $\mathbb{K}$.

Théorème

Les conditions suivantes sont équivalentes :

- l'endomorphisme u est trigonalisable ;
- il existe un polynôme annulateur de u qui est scindé sur $\mathbb{K}$;
- le polynôme minimal π_u est scindé sur $\mathbb{K}$.

Définition

On dit que u est *nilpotent* s'il existe un entier $r \geq 1$ tel que $u^{r-1} \neq 0$ et $u^r = 0$. On dit que r est l'*indice* (ou l'*ordre*) de *nilpotence* de u .

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est *nilpotente* s'il existe un entier $r \geq 1$ tel que $A^{r-1} \neq 0$ et $A^r = 0$. On dit que r est l'*indice* (ou l'*ordre*) de *nilpotence* de A .

Il est facile de vérifier que 0 est la seule valeur propre d'un endomorphisme [resp. d'une matrice] nilpotent [resp. nilpotente].

L'indice de nilpotence d'un endomorphisme de E est au plus égal à $\dim(E)$.

Si u et v sont deux endomorphismes nilpotents qui commutent, leur somme $u + v$ est alors un endomorphisme nilpotent.

Théorème

L'endomorphisme u est nilpotent si, et seulement si, son polynôme caractéristique est égal à X^n . Il est nilpotent d'indice $r \geq 1$ si, et seulement si, son polynôme minimal est égal à X^r .

Du théorème précédent, on déduit qu'un endomorphisme [resp. une matrice] nilpotent [resp. nilpotente] est trigonalisable. Réciproquement, un endomorphisme trigonalisable ayant 0 pour unique valeur propre est trigonalisable.

Théorème

Si u est nilpotent, on a alors $\text{Tr}(u^k) = 0$ pour tout entier $k \geq 1$. Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$, l'endomorphisme u est nilpotent si, et seulement si, il vérifie $\text{Tr}(u^k) = 0$ pour tout entier k compris entre 1 et n .

21.5 Sous-espaces caractéristiques

Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique s'écrit $\chi_u(X) = \prod_{k=1}^r (X - \lambda_k)^{\alpha_k}$ avec $\alpha_k \in \mathbb{N}^*$ et les $\lambda_k \in \mathbb{K}$ deux à deux distincts. Le polynôme minimal de u s'écrit alors $\pi_u(X) = \prod_{k=1}^r (X - \lambda_k)^{\beta_k}$ avec $1 \leq \beta_k \leq \alpha_k$.

Définition

Avec les notations qui précèdent, on appelle *sous-espaces caractéristiques* de u les sous espaces vectoriels $N_k = \text{Ker}(u - \lambda_k \text{Id})^{\alpha_k}$ où $k \in \{1, \dots, r\}$.

Théorème

Avec les notations qui précèdent, on a $E = \bigoplus_{k=1}^r N_k$ et pour tout k compris entre 1 et r , on a :

1. $N_k = \text{Ker}(u - \lambda_k \text{Id})^{\beta_k}$;
2. λ_k est la seule valeur propre de la restriction de u à N_k ;
3. $\dim(N_k) = \alpha_k$;
4. la restriction de $u - \lambda_k \text{Id}$ à N_k est nilpotente d'indice β_k .

Lemme

Avec les notations qui précèdent, pour tout k compris entre 1 et r , le projecteur π_k de E sur N_k parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq k}}^r N_j$ est un polynôme en u .

Les projecteurs π_k de E sur N_k parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq k}}^r N_j$ sont les *projecteurs spectraux* de u .

Théorème de Dunford

Soit u un endomorphisme de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$. Il existe un unique couple (d, v) d'endomorphismes de E tel que d soit diagonalisable, v soit nilpotent, d et v commutent et $u = d + v$.

Pratiquement la décomposition de Dunford d'un endomorphisme de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$ passe par le calcul des projecteurs spectraux π_k . Pour ce faire il suffit de disposer d'un polynôme annulateur de u de la forme $P(X) = \prod_{k=1}^r (X - \lambda_k)^{m_k}$ avec $m_k \geq \beta_k$ pour tout k compris entre 1 et

p où les λ_k sont les valeurs propres de u et β_k est la multiplicité de λ_k comme racine du polynôme minimal (on peut prendre pour polynôme P le polynôme caractéristique, ou mieux le polynôme minimal, de u). À partir de la décomposition en éléments simples $\frac{1}{P(X)} = \sum_{k=1}^p \sum_{i=1}^{m_k} \frac{\alpha_{ik}}{(X - \lambda_k)^i}$, en posant pour k compris

entre 1 et r , $Q_k(X) = \sum_{i=1}^{m_k} \alpha_{ik} (X - \lambda_k)^{m_k - i}$, $P_k(X) = \prod_{\substack{j=1 \\ j \neq k}}^r (X - \lambda_j)^{m_j}$, on obtient

$\frac{1}{P(X)} = \sum_{k=1}^r \frac{Q_k(X)}{(X - \lambda_k)^{m_k}}$ et la décomposition de Bézout $1 = \sum_{k=1}^r Q_k P_k$ qui permet d'obtenir les projecteurs spectraux $\pi_k = (Q_k P_k)(u)$. On a alors $u = d + v$ avec $d = \sum_{k=1}^r \lambda_k \pi_k$ et $v = u - d$.

La décomposition de Dunford d'un endomorphisme u de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$ permet le calcul de ses puissances successives. En effet comme d et v commutent, on peut utiliser la formule du binôme de Newton pour écrire :

$$\forall p \geq 1, u^p = (d + v)^p = \sum_{k=0}^p \binom{p}{k} d^k \circ v^{p-k}.$$

Le calcul des puissances successives de l'endomorphisme d peut se faire dans une base de diagonalisation ou en utilisant les propriétés des projecteurs pour écrire

que :

$$\forall p \geq 1, d^p = \left(\sum_{k=1}^r \lambda_k \pi_k \right)^p = \sum_{k=1}^r \lambda_k^p \pi_k$$

et le calcul des puissances successives de l'endomorphisme nilpotent v s'arrête à v^{q-1} où q est son indice de nilpotence.

On peut aussi calculer v^p avec :

$$\forall p \geq 1, v^p = \left(\sum_{k=1}^r (u - \lambda_k \text{Id}) \circ \pi_k \right)^p = \sum_{k=1}^r (u - \lambda_k \text{Id})^p \circ \pi_k.$$

Formes bilinéaires et quadratiques

22

On notera $\mathbb{K}$ le corps de réels ou des complexes, en précisant quand cela sera nécessaire s'il s'agit de $\mathbb{R}$ ou $\mathbb{C}$. Par scalaire on entend réel ou complexe.

On désigne par E un espace vectoriel réel ou complexe de dimension finie ou non et non réduit à $\{0\}$.

22.1 Formes bilinéaires

Définition

Une *forme bilinéaire* sur E est une application :

$$\begin{aligned}\varphi : E \times E &\rightarrow \mathbb{K} \\ (x, y) &\mapsto \varphi(x, y)\end{aligned}$$

telle que pour tout x dans E l'application $y \mapsto \varphi(x, y)$ est linéaire et pour tout y dans E l'application $x \mapsto \varphi(x, y)$ est linéaire.

Définition

On dit qu'une forme bilinéaire φ sur E est *symétrique* si $\varphi(y, x) = \varphi(x, y)$ pour tous x, y dans E .

Définition

On dit qu'une forme bilinéaire φ sur E est *antisymétrique* (ou *alternée*) si $\varphi(y, x) = -\varphi(x, y)$ pour tous x, y dans E .

L'ensemble $\text{Bil}(E)$ de toutes les formes bilinéaires sur E est un espace vectoriel.

22.2 Expression matricielle des formes bilinéaires.

On désigne par $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E . Tout vecteur $x \in E$ s'écrit de manière unique sous la forme $x = \sum_{j=1}^n x_j e_j$. On associe à un tel x le vecteur colonne $X = (x_i) \in \mathbb{K}^n$.

Définition

La *matrice d'une forme bilinéaire* $\varphi \in \text{Bil}(E)$ dans une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E est la matrice $A = (\varphi(e_i, e_j)) \in \mathcal{M}_n(\mathbb{K})$.

Théorème

Soit φ une forme bilinéaire sur E et A la matrice de φ dans la base $\mathcal{B}$. Pour tous vecteurs x, y dans E , on a $\varphi(x, y) = X^T A Y$.

Théorème

Une application φ de $E \times E$ dans $\mathbb{K}$ est une forme bilinéaire sur E si, et seulement si, il existe une matrice $A = (a_{ij})$ dans $\mathcal{M}_n(\mathbb{K})$ et des formes linéaires $\ell_1, \dots, \ell_n$ linéairement indépendantes telles que :

$$\forall (x, y) \in E \times E, \varphi(x, y) = \sum_{1 \leq i, j \leq n} a_{ij} \ell_i(x) \ell_j(y).$$

L'application qui associe à une forme bilinéaire φ sur un espace vectoriel E de dimension n sa matrice dans une base donnée de E réalise un isomorphisme de $\text{Bil}(E)$ sur l'espace $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n , donc $\dim_{\mathbb{K}}(\text{Bil}(E)) = n^2$.

Théorème

Une forme bilinéaire φ sur E est symétrique [resp. alternée] si, et seulement si, sa matrice A dans une quelconque base $\mathcal{B}$ de E est symétrique [resp. alternée].

Théorème

Soient $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E et P la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$. Si A_1 et A_2 sont les matrices d'une forme bilinéaire φ sur E dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ respectivement, on a alors $A_2 = P^T A_1 P$.

Définition

Le *discriminant* dans une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E d'une forme bilinéaire φ est le déterminant de la matrice $A = (\varphi(e_i, e_j)) \in \mathcal{M}_n(\mathbb{K})$ de φ dans cette base. On le note $\Delta_{\mathcal{B}}(\varphi)$.

Si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de E et P la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$, on a alors pour toute forme bilinéaire φ sur E , $\Delta_{\mathcal{B}_2}(\varphi) = (\det(P))^2 \Delta_{\mathcal{B}_1}(\varphi)$.

22.3 Formes quadratiques

Définition

On appelle *forme quadratique* sur E une application q définie de E dans $\mathbb{K}$ par $q(x) = \varphi(x, x)$ pour tout $x \in E$, où φ est une forme bilinéaire.

L'ensemble $Q(E)$ des formes quadratiques sur E est un espace vectoriel.

Théorème

Si q est une forme quadratique sur E , il existe alors une unique forme bilinéaire symétrique φ telle que $q(x) = \varphi(x, x)$ pour tout $x \in E$.

Définition

Avec les notations du théorème qui précède, on dit que φ est la *forme polaire* de la forme quadratique q .

On a les expressions suivantes de cette forme polaire :

$$\varphi(x, y) = \frac{q(x+y) - q(x) - q(y)}{2} = \frac{q(x+y) - q(x-y)}{4}$$

Pour tout scalaire λ et tout vecteur x , on a :

$$q(\lambda x) = \varphi(\lambda x, \lambda x) = \lambda^2 \varphi(x, x) = \lambda^2 q(x)$$

ce qui se traduit en disant que q est une fonction homogène de degré 2.

L'application qui associe à une forme quadratique q sa forme polaire φ réalise un isomorphisme d'espaces vectoriels de $Q(E)$ sur l'espace $\text{Bil}_s(E)$ des formes bilinéaires symétriques sur E . Donc $Q(E)$ est de dimension $\frac{n(n+1)}{2}$.

Dans le cas des espaces vectoriels de dimension finie on peut utiliser les matrices pour définir les formes quadratiques.

Définition

Soit E un espace vectoriel de dimension n et $\mathcal{B}$ une base de E . Si q est une forme quadratique sur E de forme polaire φ , on dit alors que la matrice de φ dans la base $\mathcal{B}$ est la matrice de q dans cette base.

22.4 Théorème de réduction de Gauss

On note $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E et pour tout vecteur x de E , $x_1, \dots, x_n$ désignent les coordonnées de x dans cette base, soit $x = \sum_{i=1}^n x_i e_i$. On associe toujours à ce vecteur x de E le vecteur colonne $X = (x_i)_{1 \leq i \leq n}$ dans $\mathbb{K}^n$.

Théorème de réduction de Gauss

Pour toute forme quadratique $q \neq 0$ sur E , il existe un entier p compris entre 1 et n , des scalaires non nuls $\lambda_1, \dots, \lambda_p$ et des formes linéaires $\ell_1, \dots, \ell_p$ indépendantes dans E^* tels que $q(x) = \sum_{j=1}^p \lambda_j \ell_j^2(x)$ pour tout $x \in E$.

Pour la suite de ce paragraphe, q désigne une forme quadratique non nulle sur E et $q = \sum_{j=1}^p \lambda_j \ell_j^2$ une réduction de Gauss de cette forme quadratique où p est un entier compris entre 1 et n , $\lambda_1, \dots, \lambda_p$ sont des scalaires non nuls et $\ell_1, \dots, \ell_p$ des formes linéaires indépendantes. On notera φ la forme polaire de q .

Théorème

Avec les notations qui précèdent, la forme polaire φ de q est définie par :

$$\forall (x, y) \in E \times E, \varphi(x, y) = \sum_{j=1}^p \lambda_j \ell_j(x) \ell_j(y).$$

Théorème

Étant donné une base $(\ell_i)_{1 \leq i \leq n}$ de l'espace vectoriel E^* des formes linéaires sur E , il existe une base $(f_i)_{1 \leq i \leq n}$ de E telle que $\ell_i(f_j) = \delta_{ij}$ pour tous i, j compris entre 1 et n .

Dans la situation du théorème précédent, on dit que $(\ell_i)_{1 \leq i \leq n}$ est la *base duale* de $(f_i)_{1 \leq i \leq n}$ ou que $(f_i)_{1 \leq i \leq n}$ est la *base anté-duale* de $(\ell_i)_{1 \leq i \leq n}$.

Théorème

Avec les notations qui précèdent, il existe une base $(f_i)_{1 \leq i \leq n}$ de E dans

laquelle la matrice de q est diagonale de la forme
$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

(les p premiers λ_i sont non nuls et les suivants sont nuls).

Une telle base $(f_i)_{1 \leq i \leq n}$ est dite *orthogonale* pour la forme quadratique q .

Corollaire

Si A est une matrice symétrique d'ordre n à coefficients dans $\mathbb{K}$, il existe alors une matrice inversible P telle que la matrice $P^T A P$ soit diagonale.

22.5 Orthogonalité, noyau et rang

Pour ce paragraphe, φ est une forme bilinéaire symétrique sur E et q la forme quadratique associée.

Définition

On dit que deux vecteurs x, y de E sont *orthogonaux* relativement à φ si $\varphi(x, y) = 0$.

Définition

Si X est une partie non vide E , l'*orthogonal* de X relativement à φ est le sous-ensemble de E formé des vecteurs orthogonaux à tous les vecteurs de X .

L'orthogonal d'une partie non vide X de E est notée $X^\perp$ et on a :

$$X^\perp = \{y \in E, \forall x \in X, \varphi(x, y) = 0\}.$$

Pour $X = \{0\}$, on a $X^\perp = E$.

Théorème

Soient X et Y deux parties non vide de E .

- $X^\perp$ est un sous-espace vectoriel de E , $X \subset (X^\perp)^\perp$.
- Si $X \subset Y$, alors $Y^\perp \subset X^\perp$.

Définition

On dit qu'un vecteur x de E est *isotrope* relativement à φ s'il est orthogonal à lui même.

Définition

L'ensemble des vecteurs isotropes de E , relativement à φ , est le *cône isotrope* de φ .

Définition

Le *noyau* de φ est l'orthogonal de E .

En notant $\text{Ker}(\varphi)$ le noyau de φ , on a $\text{Ker}(\varphi) = E^\perp$ et ce noyau est un sous-espace vectoriel de E . On dit aussi que $\text{Ker}(\varphi)$ est le noyau de la forme quadratique q et on le note alors $\text{Ker}(q)$.

Lemme

Le noyau de φ est contenu dans son cône isotrope.

Théorème

Soient $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E , A la matrice de la forme bilinéaire φ dans la base $\mathcal{B}$ et u l'endomorphisme de E de matrice A dans la base $\mathcal{B}$. On a $\text{Ker}(\varphi) = \text{Ker}(u)$.

On retiendra qu'en dimension finie, le noyau de φ se calcule en résolvant le système $AX = 0$.

Définition

On dit que la forme bilinéaire symétrique φ (ou de manière équivalente la forme quadratique q) est *non dégénérée* si son noyau est réduit à $\{0\}$.

Une forme bilinéaire symétrique est non dégénérée si, et seulement si, sa matrice dans une quelconque base de E est inversible, ce qui équivaut à dire que son déterminant est non nul.

Définition

Le *rang* de φ (ou de q) est l'entier $\text{rg}(q) = n - \dim(\text{Ker}(q))$.

Le rang d'une forme quadratique est égal à celui de sa matrice dans une quelconque base.

Pour la suite de ce paragraphe, q désigne une forme quadratique non nulle sur un espace vectoriel E de dimension n et $q = \sum_{j=1}^p \lambda_j \ell_j^2$ la réduction de Gauss de cette forme quadratique où p est un entier compris entre 1 et n , $\lambda_1, \dots, \lambda_p$ sont des scalaires non nuls et $\ell_1, \dots, \ell_p$ des formes linéaires indépendantes. La forme polaire de q est définie par $\varphi(x, y) = \sum_{j=1}^p \lambda_j \ell_j(x) \ell_j(y)$.

Théorème

Avec les notations qui précèdent, on a $\text{rg}(q) = p$ et :

$$\text{Ker}(q) = \{x \in E, \ell_1(x) = \dots = \ell_p(x) = 0\}.$$

Dans le cas où $p = n$, la forme q est non dégénérée et une telle base q -orthogonale se calcule en résolvant les n systèmes linéaires :

$$\ell_i(f_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (1 \leq i, j \leq n)$$

ce qui revient à inverser la matrice $Q = (\alpha_{ij})_{1 \leq i, j \leq n}$, où les α_{ij} sont définis par $\ell_i(x) = \alpha_{i1}x_1 + \dots + \alpha_{in}x_n$ (les ℓ_i étant exprimés dans une base de E).

Dans le cas où $1 \leq p \leq n-1$, on détermine tout d'abord une base $(f_{p+1}, \dots, f_n)$ du noyau de q en résolvant le système linéaire de p équations à n inconnues :

$$\ell_i(x) = 0 \quad (1 \leq i \leq p).$$

Ces vecteurs sont deux à deux orthogonaux puisque orthogonaux à tout vecteur de E . Il suffit ensuite de résoudre les p systèmes linéaires :

$$\ell_i(f_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (1 \leq i, j \leq p)$$

ce qui fournit une famille q -orthogonale $(f_1, \dots, f_p)$ formée de vecteurs non nuls. La famille $(f_i)_{1 \leq i \leq n}$ est alors une base q -orthogonale de E . Dans la pratique, on résout d'abord le système :

$$\ell_1(x) = b_1, \dots, \ell_p(x) = b_p$$

où $b = (b_1, \dots, b_p)$ est un élément quelconque de $\mathbb{K}^p$. La valeur $b = 0$ nous donne une base du noyau de q , puis les valeurs successives $b = (1, 0, \dots, 0)$, $b = (0, 1, 0, \dots, 0), \dots, b = (0, \dots, 0, 1)$ nous permettent de déterminer des vecteurs $f_1, \dots, f_p$.

22.6 Signature d'une forme quadratique réelle en dimension finie

q est une forme quadratique non nulle et on note φ sa forme polaire.

Théorème

Il existe un unique couple (s, t) d'entiers naturels tel que pour toute base $(e_i)_{1 \leq i \leq n}$ de E qui est orthogonale relativement à q , le nombre de vecteurs e_i tels que $q(e_i) > 0$ est égal à s et le nombre de vecteurs e_i tels que $q(e_i) < 0$ est égal à t . De plus, on a $s + t = \text{rg}(q)$.

Définition

Le couple (s, t) d'entiers naturels défini par le théorème précédent est appelé *signature* de q et on le note $\text{sgn}(q)$.

Une forme quadratique q est de signature (s, t) si, et seulement si, elle admet une réduction de Gauss de la forme $q = \sum_{j=1}^s \lambda_j \ell_j^2 - \sum_{j=s+1}^{s+t} \lambda_j \ell_j^2$, où les λ_j sont tous strictement positifs (pour $s = 0$ la première somme n'existe pas et $s = n$ c'est la deuxième qui n'existe pas).

En définissant les formes linéaires L_j par $L_j(x) = \ell_j(\sqrt{\lambda_j}x)$, on a la décomposition

$q = \sum_{j=1}^s L_j^2 - \sum_{j=s+1}^{s+t} L_j^2$ et à cette décomposition est associée une base q -orthogonale

de E dans laquelle la matrice de q est $D = \begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}$, où I_r est la matrice

identité d'ordre r . Les blocs diagonaux $I_s, -I_t$ ou 0 n'existent pas si $s = 0$, $s = n$ ou $s + t = n$.

Définition

On dit que la forme bilinéaire symétrique φ (ou de manière équivalente la forme quadratique q) est *positive* [resp. *définie positive*] si $q(x) \geq 0$ [resp. $q(x) > 0$] pour tout x dans E [resp. dans $E \setminus \{0\}$].

Une forme quadratique non nulle est positive [resp. définie positive] si, et seulement si, sa signature est $(s, 0)$ [resp. $(n, 0)$] où s est compris entre 1 et n .

On définit de manière analogue les formes quadratiques négative [resp. définie négative] et une forme quadratique non nulle est négative [resp. définie négative] si, et seulement si, sa signature est $(0, t)$ [resp. $(0, n)$] avec t compris entre 1 et n .

À partir d'une réduction de Gauss, $q = \sum_{j=1}^p \lambda_j \ell_j^2$, on déduit que q est positive [resp. définie positive] si, et seulement si, tous les λ_j sont strictement positifs [resp. $p = n$ et tous les λ_j sont strictement positifs].

Lemme

Soit F un sous-espace vectoriel de E . La restriction de q à F est non dégénérée si, et seulement si, on a $F \cap F^\perp = \{0\}$.

Lemme

Soit F un sous-espace vectoriel de E . Si la restriction de q à F est non dégénérée, on a alors $E = F \oplus F^\perp$.

Théorème

En désignant par $\mathcal{P}$ [resp. $\mathcal{N}$] l'ensemble de tous les sous-espaces vectoriels F de E tels que la restriction de q à F soit définie positive [resp. définie négative] ($\mathcal{P}$ ou $\mathcal{N}$ peut être vide), la signature (s, t) de q est donnée par :

$$s = \begin{cases} 0 & \text{si } \mathcal{P} = \emptyset \\ \max_{F \in \mathcal{P}} \dim(F) & \text{si } \mathcal{P} \neq \emptyset \end{cases} \quad \text{et} \quad t = \begin{cases} 0 & \text{si } \mathcal{N} = \emptyset \\ \max_{F \in \mathcal{N}} \dim(F) & \text{si } \mathcal{N} \neq \emptyset \end{cases}.$$

Théorème

Soit q une forme quadratique non nulle sur un espace vectoriel réel E de dimension n de matrice $A = (a_{ij})$ dans une base $(e_i)_{1 \leq i \leq n}$. La forme q est définie positive si, et seulement si, tous les mineurs principaux de A sont strictement positifs.

22.7 Quadriques dans $\mathbb{R}^n$ ou $\mathbb{C}^n$

Pour $n \geq 2$, on munit $\mathbb{K}^n$ de sa base canonique et les coordonnées d'un vecteur X de $\mathbb{K}^n$ sont notées $x_1, \dots, x_n$. Pour $n = 2$ [resp. $n = 3$] et $\mathbb{K} = \mathbb{R}$ les coordonnées seront notées x, y [resp. x, y, z].

Définition

On appelle *quadrique* dans $\mathbb{K}^n$ toute partie $\mathcal{C}$ de $\mathbb{K}^n$ définie par :

$$\mathcal{C} = \left\{ X \in \mathbb{K}^n, \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j + \sum_{i=1}^n b_i x_i + c = 0 \right\}$$

où c , les a_{ij} et les b_i sont des scalaires avec $(a_{11}, \dots, a_{nn}) \neq 0$.

On dit aussi que $\mathcal{C}$ est la courbe d'équation :

$$P(X) = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j + \sum_{i=1}^n b_i x_i + c = 0$$

dans la base canonique. On notera aussi $\mathcal{C} = P^{-1}\{0\}$ où P est une fonction polynomiale de degré 2 sur $\mathbb{K}^n$. Une telle courbe peut être vide comme le montre l'exemple de $P(X) = x^2 + y^2 + 1$ dans $\mathbb{R}^2$. Le polynôme P s'écrit $P = q + \ell + c$, où q est une forme quadratique, ℓ une forme linéaire et c une constante.

Définition

On dit que la quadrique $\mathcal{C} = P^{-1}\{0\}$ est à *centre* s'il existe un unique élément X_0 dans $\mathbb{K}^n$ tel que $P(X + X_0) = q(X) + d$ pour tout X dans $\mathbb{K}^n$, où d est une constante.

Théorème

La quadrique $\mathcal{C} = P^{-1}\{0\}$ est à centre si, et seulement si, la forme quadratique q est non dégénérée.

Le système linéaire permettant de déterminer le centre, quand il est unique, est donné par : $2 \sum_{j=1}^n a_{ij}x_j + b_i = 0 \quad (1 \leq i \leq n)$.

Dans $\mathbb{R}^2$, une quadrique a une équation de la forme $ax^2 + 2bxy + cy^2 + dx + ey + f = 0$ et dans une base adaptée, une quadrique de $\mathbb{R}^2$ a une équation de l'une des formes suivantes : $x^2 + y^2 = \beta$; $x^2 - y^2 = \alpha$; $dx + ey + f = 0$; $x^2 = \alpha$; $ax^2 + y + f = 0$.

On désigne par E un espace vectoriel réel non réduit à $\{0\}$.

23.1 Produit scalaire

Définition

On dit qu'une forme bilinéaire symétrique φ sur E est :

- *positive* si $\varphi(x, x) \geq 0$ pour tout x dans E ;
- *définie* si pour x dans E l'égalité $\varphi(x, x) = 0$ équivaut à $x = 0$.

Définition

On appelle *produit scalaire* sur E toute forme bilinéaire symétrique définie positive. Un *espace préhilbertien* est un espace vectoriel réel muni d'un produit scalaire. Un espace préhilbertien de dimension finie est dit *euclidien*.

Dans le cas où E est un espace euclidien, on peut aussi dire qu'un produit scalaire sur E est la forme polaire d'une forme quadratique de signature $(n, 0)$. On notera $(x, y) \mapsto \langle x | y \rangle$ quand il n'y a pas d'ambiguïté un tel produit scalaire et pour $y = x$, on note $\|x\| = \sqrt{\langle x | x \rangle}$.

Théorème

Pour tous x, y dans E on a :

$$\langle x | y \rangle = \frac{\|x + y\|^2 - \|x\|^2 - \|y\|^2}{2} = \frac{\|x + y\|^2 - \|x - y\|^2}{4}$$

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

La deuxième identité est l'*égalité du parallélogramme*. Elle est caractéristique des produits scalaires dans le sens où une norme est déduite d'un produit scalaire si, et seulement si, elle vérifie l'identité du parallélogramme.

Dans tout ce qui suit E désigne un espace préhilbertien.

Théorème de Cauchy-Schwarz

Pour tous x, y dans E on a $|\langle x | y \rangle| \leq \|x\| \|y\|$, l'égalité étant réalisée si, et seulement si, x et y sont liés.

Sur $\mathbb{R}^n$ muni du produit scalaire canonique, l'inégalité de Cauchy-Schwarz prend la forme suivante :

$$\left| \sum_{k=1}^n x_k y_k \right|^2 \leq \left(\sum_{k=1}^n x_k^2 \right) \left(\sum_{k=1}^n y_k^2 \right).$$

Dans l'espace $\mathcal{C}^0([a, b], \mathbb{R})$ des fonctions continues de $[a, b]$ dans $\mathbb{R}$ muni du produit scalaire $(f, g) \mapsto \int_a^b f(t)g(t) dt$, l'inégalité de Cauchy-Schwarz s'écrit :

$$\left(\int_a^b f(t)g(t) dt \right)^2 \leq \int_a^b f^2(t) dt \int_a^b g^2(t) dt.$$

Théorème de Minkowski

Pour tous x, y dans E on a $\|x + y\| \leq \|x\| + \|y\|$, l'égalité étant réalisée si, et seulement si, $x = 0$ ou $x \neq 0$ et $y = \lambda x$ avec $\lambda \geq 0$ (on dit que x et y sont *positivement liés*).

L'inégalité de Minkowski ajoutée aux propriétés de positivité ($\|x\| > 0$ pour tout $x \neq 0$) et d'homogénéité ($\|\lambda x\| = |\lambda| \|x\|$ pour tout réel λ et tout vecteur x) se traduit en disant que l'application $x \mapsto \|x\| = \sqrt{\langle x | x \rangle}$ définit une norme sur E . De cette inégalité, on déduit l'inégalité suivante :

$$\forall (x, y) \in E^2, \quad \left| \|x\| - \|y\| \right| \leq \|x + y\|.$$

Par récurrence, on montre facilement que pour tous vecteurs $x_1, \dots, x_p$, on a :

$$\|x_1 + \dots + x_p\| \leq \|x_1\| + \dots + \|x_p\|.$$

Sur $\mathbb{R}^n$ muni du produit scalaire canonique, l'inégalité de Minkowski prend la forme suivante :

$$\sqrt{\sum_{k=1}^n (x_k + y_k)^2} \leq \sqrt{\sum_{k=1}^n x_k^2} + \sqrt{\sum_{k=1}^n y_k^2}.$$

Dans l'espace $\mathcal{C}^0([a, b], \mathbb{R})$ des fonctions continues de $[a, b]$ dans $\mathbb{R}$ muni du produit scalaire $(f, g) \mapsto \int_a^b f(t)g(t) dt$, l'inégalité de Minkowski s'écrit :

$$\sqrt{\int_a^b (f(t) + g(t))^2 dt} \leq \sqrt{\int_a^b f^2(t) dt} + \sqrt{\int_a^b g^2(t) dt}.$$

L'inégalité de Cauchy-Schwarz nous dit que pour tous vecteurs u et v non nuls dans E , on a $-1 \leq \frac{\langle u | v \rangle}{\|u\| \|v\|} \leq 1$, ce qui implique qu'il existe un unique réel $\theta \in [0, \pi]$ tel que $\langle u | v \rangle = \cos(\theta) \|u\| \|v\|$.

Définition

Le réel θ est la mesure dans $[0, \pi]$ de l'angle géométrique (ou angle non orienté) que font les vecteurs u et v dans $E \setminus \{0\}$.

On note $\widehat{(u, v)}$ cette mesure et on a $\widehat{(u, v)} = \arccos\left(\frac{\langle u | v \rangle}{\|u\| \|v\|}\right) \in [0, \pi]$.

Pour $\theta \in \{0, \pi\}$, les vecteurs u et v sont liés et pour $\theta = \frac{\pi}{2}$, ils sont orthogonaux. De manière générale, on a $\|u + v\|^2 = \|u\|^2 + 2 \cos(\theta) \|u\| \|v\| + \|v\|^2$, où θ est la mesure dans $[0, \pi]$ de l'angle que font les vecteurs non nuls u et v .

Si λ, μ sont deux réels strictement positifs, on a alors $\widehat{(\lambda u, \mu v)} = \widehat{(u, v)}$, ce qui permet de définir la mesure dans $[0, \pi]$ de l'angle géométrique de deux demi-droites $\Delta_1 = \mathbb{R}^+ u_1$ et $\Delta_2 = \mathbb{R}^+ u_2$ par $\widehat{(\Delta_1, \Delta_2)} = \widehat{(u_1, u_2)}$, où u_1 est un vecteur directeur de Δ_1 et u_2 un vecteur directeur de Δ_2 .

On a :

- $\widehat{(\Delta_1, \Delta_2)} = 0$ si, et seulement si, $\Delta_1 = \Delta_2$;
- $\widehat{(\Delta_1, \Delta_2)} = \pi$ si, et seulement si, $\Delta_1 = -\Delta_2$ (Δ_1 et Δ_2 sont opposées) ;
- $\widehat{(\Delta_1, \Delta_2)} = \frac{\pi}{2}$ si, et seulement si, Δ_1 et Δ_2 sont orthogonales (voir le paragraphe suivant).

23.2 Orthogonalité

Définition

On dit que deux vecteurs x et y appartenant à E sont *orthogonaux* s'ils vérifient $\langle x | y \rangle = 0$.

Théorème de Pythagore

Les vecteurs x et y sont orthogonaux dans E si, et seulement si, on a :

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

On déduit par récurrence sur $p \geq 2$, que si $x_1, \dots, x_p$ sont deux à deux orthogonaux, on a alors $\left\| \sum_{k=1}^p x_k \right\|^2 = \sum_{k=1}^p \|x_k\|^2$.

Définition

On appelle *famille orthogonale* dans E toute famille $(e_i)_{i \in I}$ de vecteurs de E telle que $\langle e_i | e_j \rangle = 0$ pour tous $i \neq j$ dans I . Si de plus $\|e_i\| = 1$ pour tout $i \in I$, on dit alors que cette famille est *orthonormée* ou *orthonormale*.

Définition

L' orthogonal d'une partie non vide X de E est l'ensemble :

$$X^\perp = \{y \in E, \forall x \in X, \langle x | y \rangle = 0\}.$$

$X^\perp$ est un sous-espace vectoriel de E .

Théorème

Une famille orthogonale de vecteurs non nuls de E est libre.

La famille $\{\cos(nt), \sin(mt) ; (n, m) \in \mathbb{N} \times \mathbb{N}^*\}$ est orthogonale dans l'espace vectoriel $\mathcal{F}$ des fonctions continues et 2π -périodiques sur $\mathbb{R}$ muni du produit scalaire $(f, g) \mapsto \langle f | g \rangle = \int_{-\pi}^{\pi} f(t)g(t) dt$. On en déduit que la famille de fonctions :

$$\left\{ \frac{1}{\sqrt{2\pi}} \right\} \cup \left\{ \frac{1}{\sqrt{\pi}} \cos(nt) ; n \in \mathbb{N}^* \right\} \cup \left\{ \frac{1}{\sqrt{\pi}} \sin(mt) ; m \in \mathbb{N}^* \right\}$$

est orthonormée dans $\mathcal{F}$. Pour tout $n \in \mathbb{N}^*$, la famille :

$$\mathcal{T}_n = \left\{ \frac{1}{\sqrt{2\pi}} \right\} \cup \left\{ \frac{\cos(jt)}{\sqrt{\pi}} ; 1 \leq j \leq n \right\} \cup \left\{ \frac{\sin(kt)}{\sqrt{\pi}} ; 1 \leq k \leq n \right\}$$

est une base orthonormée de l'espace $\mathcal{P}_n$ des polynômes trigonométriques de degré inférieur ou égal à n .

23.3 Orthonormalisation de Gram-Schmidt

Théorème d'orthonormalisation de Gram-Schmidt

Pour toute famille libre $(x_i)_{1 \leq i \leq p}$ dans E , il existe une unique famille orthonormée $(e_i)_{1 \leq i \leq p}$ dans E telle que :

$$\forall k \in \{1, \dots, p\}, \quad \begin{cases} \text{Vect}\{e_1, \dots, e_k\} = \text{Vect}\{x_1, \dots, x_k\}; \\ \langle x_k | e_k \rangle > 0. \end{cases}$$

La construction d'une famille orthonormée $(e_i)_{1 \leq i \leq p}$ peut se faire en utilisant l'al-

gorithme suivant :
$$\begin{cases} y_1 = x_1, & e_1 = \frac{1}{\|y_1\|} y_1; \\ y_k = x_k - \sum_{j=1}^{k-1} \langle x_k | e_j \rangle e_j, & e_k = \frac{1}{\|y_k\|} y_k, \quad (k = 2, \dots, p). \end{cases}$$

Le calcul de $\|y_k\|$ peut être simplifié en écrivant que :

$$\|y_k\|^2 = \|x_k\|^2 - \sum_{j=1}^{k-1} \langle x_k | e_j \rangle^2.$$

Les $\langle x_k | e_j \rangle$ étant déjà calculés (pour obtenir y_k), il suffit donc de calculer $\|x_k\|^2$.

Corollaire

Si F est un sous-espace vectoriel de dimension finie ou infinie dénombrable de E , il existe alors une base orthonormée pour F .

Si E est un espace euclidien de dimension finie et $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base orthonormée de E , alors tout vecteur $x \in E$ s'écrit $x = \sum_{k=1}^n \langle x | e_k \rangle e_k = \sum_{k=1}^n x_k e_k$ et on a pour tous vecteurs x, y dans E , en notant X la matrice de x dans la base $\mathcal{B}$:

$$\langle x | y \rangle = \sum_{k=1}^n \langle x | e_k \rangle \langle y | e_k \rangle = \sum_{k=1}^n x_k y_k = X^T Y$$

$$\text{et } \|x\|^2 = \sum_{k=1}^n \langle x | e_k \rangle^2 = \sum_{k=1}^n x_k^2.$$

Théorème

Soient E un espace euclidien de dimension $n \geq 1$, $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ deux bases orthonormées de E . La matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est telle que $P^{-1} = P^T$. En particulier, on a $\det(P) = \pm 1$.

Définition

On appelle *matrice orthogonale* toute matrice réelle d'ordre n inversible telle que $P^{-1} = P^T$.

La matrice de passage d'une base orthonormée $\mathcal{B}$ de E à une autre base orthonormée $\mathcal{B}'$ est une matrice orthogonale et réciproquement une telle matrice est la matrice de passage d'une base orthonormée de E à une autre.

23.4 Adjoint d'un endomorphisme

$(E, \langle \cdot | \cdot \rangle)$ est un espace euclidien de dimension n .

Lemme

Pour toute forme linéaire ℓ sur E , il existe un unique vecteur $a \in E$ tel que $\ell(x) = \langle x | a \rangle$ pour tout $x \in E$.

Théorème

Pour tout endomorphisme $u \in \mathcal{L}(E)$, il existe un unique endomorphisme $u^* \in \mathcal{L}(E)$ tel que $\langle u(x) | y \rangle = \langle x | u^*(y) \rangle$ pour tous x, y dans E .

Définition

Avec les notations du théorème précédent, on dit que u^* est l'*adjoint* de u .

Théorème

Soient $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E et la matrice $G = (\langle e_i | e_j \rangle)$. Si $u \in \mathcal{L}(E)$ a pour matrice A dans la base $\mathcal{B}$, alors la matrice de u^* dans $\mathcal{B}$ est $B = G^{-1}A^T G$. Dans le cas où la base $\mathcal{B}$ est orthonormée, on a $B = A^T$.

Sachant qu'une matrice et sa transposée ont même déterminant, on déduit du théorème précédent que pour tout $u \in \mathcal{L}(E)$, on a $\det(u^*) = \det(u)$.

Théorème

Pour tous endomorphismes u, v dans $\mathcal{L}(E)$, on a :

- $(\lambda u + v)^* = \lambda u^* + v^*$; $(u^*)^* = u$; $(u \circ v)^* = v^* \circ u^*$;
- si $u \in \text{GL}(E)$, alors $u^* \in \text{GL}(E)$ et $(u^*)^{-1} = (u^{-1})^*$;
- $\text{Ker}(u^*) = (\text{Im}(u))^\perp$ et $\text{Im}(u^*) = (\text{Ker}(u))^\perp$; $\text{rg}(u^*) = \text{rg}(u)$;
- si F est un sous-espace vectoriel de E stable par u , alors $F^\perp$ est stable par u^* .

23.5 Projection orthogonale sur un sous-espace de dimension finie

$(E, \langle \cdot | \cdot \rangle)$ est toujours un espace préhilbertien de dimension finie ou non et pour toute partie non vide F de E , on note $d(x, F) = \inf_{z \in F} \|x - z\|$ la distance de x à F .

Théorème de projection orthogonale

Soit F un sous-espace vectoriel de dimension finie de E non réduit à $\{0\}$. Pour tout vecteur $x \in E$, il existe un unique vecteur y dans F tel que $\|x - y\| = d(x, F)$. Ce vecteur est également l'unique vecteur appartenant à F tel que $x - y \in F^\perp$. Son expression dans une base orthonormée $(e_i)_{1 \leq i \leq n}$ de F est donnée par $y = \sum_{k=1}^n \langle x | e_k \rangle e_k$ et on a :

$$\|x - y\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum_{k=1}^n \langle x | e_k \rangle^2 \quad (23.1)$$

Si x est un vecteur de E , alors le vecteur y de F qui lui est associé dans le théorème précédent est la meilleure approximation de x dans F . On dit aussi que y est la *projection orthogonale* de x sur F . On note $y = p_F(x)$ et on a :

$$(y = p_F(x)) \Leftrightarrow (y \in F \text{ et } x - y \in F^\perp) \Leftrightarrow (y \in F \text{ et } \|x - y\| = d(x, F)).$$

Dans une base orthonormée de F , une expression de p_F est :

$$\forall x \in E, p_F(x) = \sum_{k=1}^n \langle x | e_k \rangle e_k.$$

Définition

Avec les notations et hypothèses précédentes, l'application p_F est la *projection orthogonale* de E sur F .

p_F est un projecteur.

Si $D = \mathbb{R}a$ est une droite vectorielle, on a alors $p_D(x) = \frac{\langle x | a \rangle}{\|a\|^2} a$ pour tout $x \in E$. Pour E euclidien et H hyperplan orthogonal à la droite $D = \mathbb{R}a$, on a, pour tout $x \in E$, $p_H(x) = x - \frac{\langle x | a \rangle}{\|a\|^2} a$.

De l'inégalité (23.1), on déduit que pour tout vecteur $x \in E$, on a :

$$\|p_F(x)\|^2 = \sum_{k=1}^n \langle x | e_k \rangle^2 \leq \|x\|^2.$$

Cette inégalité est l'*inégalité de Bessel*.

Si $(e_i)_{1 \leq i \leq n}$ est une base (non nécessairement orthonormée) de F , alors la projection orthogonale d'un vecteur x de E sur F est le vecteur $y = \sum_{j=1}^n y_j e_j$, où les composantes y_j , pour j compris entre 1 et n , sont solutions du système linéaire :

$$\langle x - y \mid e_i \rangle = 0 \quad (1 \leq i \leq n)$$

soit :

$$\sum_{j=1}^n \langle e_i \mid e_j \rangle y_j = \langle x \mid e_i \rangle \quad (1 \leq i \leq n).$$

Ce système est appelé système d'équations normales.

Sur l'espace vectoriel $\mathcal{F}$ des fonctions continues et 2π -périodiques muni du produit scalaire $(f, g) \mapsto \langle f \mid g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$, la meilleure approximation, pour la norme déduite de ce produit scalaire, d'une fonction $f \in \mathcal{F}$ par un polynôme trigonométrique de degré inférieur ou égal à n est donnée par :

$$\begin{aligned} S_n(f) &= \left\langle f \mid \frac{c_0}{\sqrt{2\pi}} \right\rangle \frac{c_0}{\sqrt{2\pi}} + \sum_{k=1}^n \left\langle f \mid \frac{c_k}{\sqrt{\pi}} \right\rangle \frac{c_k}{\sqrt{\pi}} + \sum_{k=1}^n \left\langle f \mid \frac{s_k}{\sqrt{\pi}} \right\rangle \frac{s_k}{\sqrt{\pi}} \\ &= \frac{1}{2\pi} \langle f \mid c_0 \rangle c_0 + \frac{1}{\pi} \sum_{k=1}^n \langle f \mid c_k \rangle c_k + \frac{1}{\pi} \sum_{k=1}^n \langle f \mid s_k \rangle s_k \end{aligned}$$

où $c_k : x \mapsto \cos(kx)$ pour $k \geq 0$ et $s_k : x \mapsto \sin(kx)$ pour $k \geq 1$. Soit :

$$S_n(f)(x) = \frac{a_0(f)}{2} + \sum_{k=1}^n a_k(f) \cos(kx) + \sum_{k=1}^n b_k(f) \sin(kx)$$

où les $a_k(f)$ et $b_k(f)$ sont les coefficients de Fourier trigonométriques de f définis par :

$$a_k(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) dt \text{ et } b_k(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) dt.$$

L'opérateur S_n de projection orthogonale de $\mathcal{F}$ sur l'espace $\mathcal{P}_n$ des polynômes trigonométriques de degré inférieur ou égal à n est l'opérateur de Fourier.

La série $a_0(f) + \sum (a_n(f) \cos(nx) + b_n(f) \sin(nx))$ est la série de Fourier de f .

L'inégalité de Bessel s'écrit :

$$\left\langle f \mid \frac{c_0}{\sqrt{2\pi}} \right\rangle^2 + \sum_{k=1}^n \left\langle f \mid \frac{c_k}{\sqrt{\pi}} \right\rangle^2 + \sum_{k=1}^n \left\langle f \mid \frac{s_k}{\sqrt{\pi}} \right\rangle^2 \leq \|f\|^2$$

ou encore :

$$\frac{a_0^2(f)}{2} + \sum_{k=1}^n (a_k^2(f) + b_k^2(f)) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(t) dt.$$

La série numérique $a_0^2(f) + \frac{1}{2} \sum (a_n^2(f) + b_n^2(f))$ converge avec :

$$\frac{a_0^2(f)}{2} + \sum_{n=1}^{+\infty} (a_n^2(f) + b_n^2(f)) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(t) dt \quad (\text{inégalité de Bessel}).$$

On peut monter qu'on a fait l'égalité (théorème de Parseval).

De l'inégalité de Bessel, on déduit que $\lim_{n \rightarrow +\infty} a_n(f) = \lim_{n \rightarrow +\infty} b_n(f) = 0$ (théorème de Riemann-Lebesgue).

Corollaire

Pour tout sous-espace vectoriel F de dimension finie de E on a $E = F \oplus F^\perp$ et $(F^\perp)^\perp = F$.

Pour E de dimension finie, on a $\dim(E) = \dim(F) + \dim(F^\perp)$.

Théorème

Soit F un sous-espace vectoriel de dimension finie de E .

- Pour $x \in E$, on a $x \in F$ si, et seulement si, $p_F(x) = x$.
- $p_F \circ p_F = p_F$.
- La projection orthogonale p_F de E sur F est une application linéaire surjective de E sur F .
- Le noyau de p_F est $F^\perp$.
- Pour tous x, y dans E , on a :

$$\langle p_F(x) \mid y \rangle = \langle x \mid p_F(y) \rangle = \langle p_F(x) \mid p_F(y) \rangle.$$

(on dit que p_F est *auto-adjoint*).

- Pour E de dimension finie, on a $p_F + p_{F^\perp} = \text{Id}$.
- Pour E de dimension finie, on a $p_F \circ p_{F^\perp} = p_{F^\perp} \circ p_F = 0$.

Théorème

Soit p un projecteur d'un espace euclidien E . Les propriétés suivantes sont équivalentes :

- p est un projecteur orthogonal ;
- pour tous x, y dans E , on a : $\langle p(x) \mid y \rangle = \langle x \mid p(y) \rangle$;
- pour tout $x \in E$, on a $\|p(x)\| \leq \|x\|$.

23.6 Symétries orthogonales dans les espaces euclidiens

E est un espace euclidien de dimension n .

Définition

Si F est un sous-espace vectoriel de E , la *symétrie orthogonale* par rapport à F est l'application définie sur E par :

$$\forall x \in E, s_F(x) = p_F(x) - p_{F^\perp}(x).$$

Comme p_F et $p_{F^\perp}$, l'application s_F est linéaire. Elle est aussi définie par :

$$\forall x \in E, s_F(x) = 2p_F(x) - x = x - 2p_{F^\perp}(x).$$

Si $D = \mathbb{R}a$ est une droite vectorielle, on a alors :

$$s_D(x) = 2p_D(x) - x = 2 \frac{\langle x | a \rangle}{\|a\|^2} a - x.$$

Si $H = (\mathbb{R}a)^\perp$ est un hyperplan, on a alors :

$$s_H(x) = 2p_H(x) - x = x - 2 \frac{\langle x | a \rangle}{\|a\|^2} a.$$

Définition

On appelle *réflexion* une symétrie orthogonale par rapport à un hyperplan et *demi-tour* ou *retournement* une symétrie orthogonale par rapport à une droite.

Théorème

Soit F un sous espace vectoriel de E .

- Pour $x \in E$, on a $x \in F$ si, et seulement si, $s_F(x) = x$ et $x \in F^\perp$ si, et seulement si, $s_F(x) = -x$.
- $s_F \circ s_F = \text{Id}$ (on dit que s_F est une involution). Une symétrie orthogonale est donc un automorphisme de E avec $s_F^{-1} = s_F$.
- Pour tous x, y dans E , on a $\langle s_F(x) | y \rangle = \langle x | s_F(y) \rangle$ (s_F est auto-adjoint).
- Pour tous x, y dans E , on a $\langle s_F(x) | s_F(y) \rangle = \langle x | y \rangle$ (s_F est une isométrie).
- On a $s_F + s_{F^\perp} = 0$ et $s_F \circ s_{F^\perp} = s_{F^\perp} \circ s_F = -\text{Id}$.
- Si F est de dimension $p \in \{1, \dots, n-1\}$, il existe alors une base orthonormée de E dans laquelle la matrice de s_F est $\begin{pmatrix} I_p & 0 \\ 0 & -I_{n-p} \end{pmatrix}$ et on a $\det(s_F) = (-1)^{n-p}$.

23.7 Les endomorphismes symétriques réels

Théorème

Les valeurs propres d'une matrice symétrique réelle A sont toutes réelles.

Pour ce paragraphe, $(E, \langle \cdot | \cdot \rangle)$ est un espace réel euclidien de dimension $n \geq 1$. Un endomorphisme $u \in \mathcal{L}(E)$ est dit *symétrique* (ou *autoadjoint*) si :

$$\forall (x, y) \in E \times E, \langle u(x) | y \rangle = \langle x | u(y) \rangle.$$

Un endomorphisme $u \in \mathcal{L}(E)$ est symétrique si, et seulement si, pour toute base orthonormée $\mathcal{B}$ de E la matrice A de u dans $\mathcal{B}$ est symétrique.

On note $\mathcal{S}(E)$ l'ensemble des endomorphismes symétriques de E et $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques réelles.

Théorème spectral

Un endomorphisme symétrique réel $u \in \mathcal{S}(E)$ a n valeurs propres réelles distinctes ou confondues et se diagonalise dans une base orthonormée.

Définition

Une *matrice orthogonale* est une matrice réelle $A \in \mathcal{M}_n(\mathbb{R})$ qui vérifie $A^T A = A A^T = I_n$.

Il revient au même de dire qu'une matrice orthogonale est une matrice inversible A d'inverse A^T ou que ses colonnes forment une base orthonormée de $\mathbb{R}^n$ muni de sa structure euclidienne canonique.

On note $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales.

Théorème

Pour toute matrice A dans $\mathcal{O}_n(\mathbb{R})$, on a $\det(A) = \pm 1$ et $\mathcal{O}_n(\mathbb{R})$ est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$.

Corollaire

Toute matrice symétrique réelle $A \in \mathcal{S}_n(\mathbb{R})$ se diagonalise dans une base orthonormée, c'est-à-dire qu'il existe une matrice orthogonale $P \in \mathcal{O}_n(\mathbb{R})$ et une matrice diagonale D telles que $P^T A P$ soit diagonale.

Définition

Un endomorphisme $u \in \mathcal{L}(E)$ est dit *symétrique positif* [resp. *défini positif*] s'il est symétrique avec $\langle x | u(x) \rangle \geq 0$ pour tout $x \in E$ [resp. $\langle x | u(x) \rangle > 0$ pour tout $x \in E \setminus \{0\}$].

On note $\mathcal{S}^+(E)$ [resp. $\mathcal{S}^{++}(E)$] l'ensemble des endomorphismes symétriques positifs [resp. définis positifs] de E . Si $u \in \mathcal{S}^+(E)$ [resp. $u \in \mathcal{S}^{++}(E)$] a pour matrice A dans une base orthonormée, en désignant par X le vecteur colonne constitué des coordonnées de $x \in E$ dans cette base, on a pour tout $X \in \mathbb{R}^n$, $\langle X | AX \rangle = \langle x | u(x) \rangle \geq 0$ [resp. $\langle X | AX \rangle = \langle x | u(x) \rangle > 0$], ce qui nous conduit à la définition suivante.

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est dite *symétrique positive* [resp. *définie positive*] si elle est symétrique avec $\langle x | Ax \rangle \geq 0$ pour tout $x \in \mathbb{R}^n$ [resp. $\langle x | Ax \rangle > 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$].

Une matrice symétrique positive [resp. définie positive] définit donc dans la base canonique un endomorphisme symétrique positif [resp. défini positif].

On note $\mathcal{S}_n^+(\mathbb{R})$ [resp. $\mathcal{S}_n^{++}(\mathbb{R})$] l'ensemble des matrices symétriques positives [resp. définies positives].

Le théorème spectral nous permet de caractériser les endomorphismes symétriques positifs [resp. définis positifs].

Théorème

Soit $u \in \mathcal{S}(E)$. On a $u \in \mathcal{S}^+(E)$ [resp. $u \in \mathcal{S}^{++}(E)$] si, et seulement si, toutes ses valeurs propres sont positives [resp. strictement positives].

Si $u \in \mathcal{S}^+(E)$ [resp. $u \in \mathcal{S}^{++}(E)$], on a $\det(u) = \prod_{k=1}^n \lambda_k \geq 0$ [resp. $\det(u) > 0$], où les λ_k sont les valeurs propres distinctes ou confondues de u . De même, pour tout $A \in \mathcal{S}_n^+(\mathbb{R})$ [resp. $A \in \mathcal{S}_n^{++}(\mathbb{R})$], on a $\det(A) \geq 0$ [resp. $\det(A) > 0$].

Corollaire

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

Alors A est dans $\mathcal{S}_n^+(\mathbb{R})$ si, et seulement si, il existe $B \in \mathcal{M}_n(\mathbb{R})$ telle que $A = B^T B$.

L'utilisation des mineurs principaux nous permet de savoir si un endomorphisme ou une matrice symétrique est défini [resp. définie] positif [resp. positive] ou non.

On rappelle que si $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{R})$, ses mineurs principaux sont les déterminants des matrices extraites $A_k = (a_{ij})_{1 \leq i, j \leq k}$ où k est un entier compris entre 1 et n .

Théorème

Une matrice symétrique est définie positive si, et seulement si, tous ses mineurs principaux sont strictement positifs.

Théorème

Soit $u \in \mathcal{S}(E)$ de matrice $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{S}_n(\mathbb{R})$ dans une base orthonormée $\mathcal{B} = (e_i)_{1 \leq i \leq n}$. L'endomorphisme symétrique u est défini positif si, et seulement si, tous les mineurs principaux de A sont strictement positifs.

23.8 Isométries vectorielles

E est un espace préhilbertien.

Définition

Une *isométrie* (ou application orthogonale) est une application $u : E \rightarrow E$ qui conserve le produit scalaire, c'est-à-dire telle que :

$$\forall (x, y) \in E \times E, \langle u(x) \mid u(y) \rangle = \langle x \mid y \rangle.$$

On note $\mathcal{O}(E)$ l'ensemble des isométries de E .

Une isométrie conserve l'orthogonalité, c'est-à-dire que pour x, y dans E , on a :

$$\langle x \mid y \rangle = 0 \Rightarrow \langle u(x) \mid u(y) \rangle = 0$$

mais une application qui conserve l'orthogonalité n'est pas nécessairement une isométrie.

Théorème

Une application $u : E \rightarrow E$ est une isométrie si, et seulement si, elle est linéaire et conserve la norme, c'est-à-dire que :

$$\forall x \in E, \|u(x)\| = \|x\|.$$

Théorème

Si E est un espace euclidien (donc de dimension finie), alors une isométrie est un automorphisme de E et $\mathcal{O}(E)$ est un sous-groupe de $\text{GL}(E)$.

On dit, pour E de dimension finie, que $\mathcal{O}(E)$ est le groupe orthogonal de E .

Théorème

Soit u une isométrie de l'espace préhilbertien E . Si F est un sous-espace vectoriel de E de dimension finie stable par u , alors son orthogonal $F^\perp$ est aussi stable par u .

Théorème

Soient E un espace euclidien, $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base orthonormée de E et u une application linéaire de E dans E . L'application u est une isométrie si, et seulement si, elle transforme $\mathcal{B}$ en une base orthonormée de E .

Théorème

Soient E un espace euclidien, $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base orthonormée de E et u une application linéaire de E dans E de matrice A dans $\mathcal{B}$. L'application u est une isométrie si, et seulement si, la matrice A est orthogonale.

Corollaire

Si u est une isométrie d'un espace euclidien E , on a alors $\det(u) = \pm 1$.

On note :

$$\mathcal{O}^+(E) = \{u \in \mathcal{O}(E), \det(u) = 1\}, \quad \mathcal{O}_n^+(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}), \det(A) = 1\}$$

$$\mathcal{O}^-(E) = \{u \in \mathcal{O}(E), \det(u) = -1\}, \quad \mathcal{O}_n^-(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}), \det(A) = -1\}$$

et on dit que les éléments de $\mathcal{O}^+(E)$ [resp. $\mathcal{O}_n^+(\mathbb{R})$] sont des *automorphismes orthogonaux positifs* ou des *isométries directes* ou des *rotations vectorielles* [resp. des *matrices orthogonales positives*] et les éléments de $\mathcal{O}^-(E)$ [resp. $\mathcal{O}_n^-(\mathbb{R})$] sont des *automorphismes orthogonaux négatifs* [resp. les *matrices orthogonales négatives*]. L'ensemble $\mathcal{O}^+(E)$ [resp. $\mathcal{O}_n^+(\mathbb{R})$] est un sous-groupe de $\mathcal{O}(E)$ [resp. de $\mathcal{O}_n(\mathbb{R})$] appelé *groupe spécial orthogonal*. On le notera par la suite $\mathcal{SO}(E)$ [resp. $\mathcal{SO}_n(\mathbb{R})$].

23.9 Orientation d'un espace euclidien

E est un espace euclidien de dimension $n \geq 2$.

Théorème

Si $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ sont deux bases orthonormées de E , alors la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est une matrice orthogonale.

Avec les notations du théorème, on a $\det(P) = \pm 1$.

On définit une relation sur l'ensemble des bases orthonormées de E en disant qu'une base orthonormée $\mathcal{B}$ est en relation avec une base orthonormée $\mathcal{B}'$ si, et seulement si, la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est dans $\mathcal{SO}_n(\mathbb{R})$.

On notera $\sim$ cette relation.

Théorème

La relation $\sim$ ainsi définie est une relation d'équivalence et il y a exactement deux classes d'équivalence pour cette relation.

Orienter l'espace euclidien E revient à choisir une base orthonormée E .

Ce théorème nous dit qu'il n'y a que deux orientations possibles pour E .

Définition

Si l'espace E est orienté par le choix d'une base orthonormée $\mathcal{B}$, on dit qu'une base orthonormée $\mathcal{B}'$ est *directe* (ou qu'elle définit la même orientation que $\mathcal{B}$) si $\mathcal{B}'$ est dans la classe d'équivalence de $\mathcal{B}$ et on dit que cette base $\mathcal{B}'$ est *indirecte* dans le cas contraire.

23.10 Produit vectoriel dans un espace euclidien

On désigne par E un espace euclidien de dimension $n \geq 3$ orienté par le choix d'une base orthonormée $\mathcal{B}_0 = (e_i)_{1 \leq i \leq n}$.

Si $\mathcal{B}$ est une autre base de E , alors pour tout n -uplet $(x_1, \dots, x_n)$ de vecteurs de E , on a $\det_{\mathcal{B}_0}(x_1, \dots, x_n) = \det_{\mathcal{B}_0}(\mathcal{B}) \det_{\mathcal{B}}(x_1, \dots, x_n)$. Il en résulte que la quantité $\det_{\mathcal{B}}(x_1, \dots, x_n)$ est indépendante du choix d'une base orthonormée directe $\mathcal{B}$ de E . On la note $\det(x_1, \dots, x_n)$ (ce qui suppose le choix d'une orientation de E) et on dit que c'est le *produit mixte* des vecteurs ordonnés $x_1, \dots, x_n$. On le note parfois $[x_1, \dots, x_n]$.

Lemme

Pour toute forme linéaire ℓ sur E , il existe un unique vecteur $a \in E$ tel que :

$$\forall x \in E, \ell(x) = \langle a | x \rangle.$$

Pour tout $(x_i)_{1 \leq i \leq n-1} \in E^{n-1}$, l'application $x \mapsto \det(x_1, \dots, x_{n-1}, x)$ est une forme linéaire, donc il existe un unique vecteur $a \in E$ tel que :

$$\forall x \in E, \det(x_1, \dots, x_{n-1}, x) = \langle a | x \rangle \quad (23.2)$$

ce vecteur a étant fonction des vecteurs $x_1, \dots, x_{n-1}$.

Définition

Le *produit vectoriel* des $n - 1$ vecteurs $x_1, \dots, x_{n-1}$ de E est le vecteur a défini par (23.2). On le note $x_1 \wedge \dots \wedge x_{n-1}$.

Dans la base orthonormée $\mathcal{B}_0$, en notant $x_i = \sum_{j=1}^n x_{ij} e_j$ pour tout i compris entre 1 et n , les réels $\det(x_1, \dots, x_{n-1}, e_i) = \langle x_1 \wedge \dots \wedge x_{n-1} | e_i \rangle$ sont les composantes du vecteur $x_1 \wedge \dots \wedge x_{n-1}$ dans la base $\mathcal{B}_0$. On a donc

$$x_1 \wedge \dots \wedge x_{n-1} = \sum_{i=1}^n (-1)^{i+n} \delta_i e_i$$

où δ_i est le déterminant de la matrice d'ordre $n - 1$ déduite de $(X_1, \dots, X_{n-1})$ en supprimant de cette matrice la ligne numéro i (X_i étant le vecteur de $\mathbb{R}^n$ formé des composantes de x_i dans la base $\mathcal{B}_0$).

Par exemple dans l'espace euclidien $E = \mathbb{R}^3$ muni de sa base canonique, le produit vectoriel de $x = (x_1, x_2, x_3)$ et $y = (y_1, y_2, y_3)$ est le vecteur $z = (z_1, z_2, z_3)$ défini par $z_1 = x_2 y_3 - x_3 y_2$, $z_2 = x_3 y_1 - x_1 y_3$ et $z_3 = x_1 y_2 - x_2 y_1$.

Théorème

Soit $(x_i)_{1 \leq i \leq n-1} \in E^{n-1}$.

- Le vecteur $x_1 \wedge \dots \wedge x_{n-1}$ est orthogonal à tous les vecteurs x_i ;
- $x_1 \wedge \dots \wedge x_{n-1} = 0$ si, et seulement si, la famille $(x_1, \dots, x_{n-1})$ est liée ;
- si la famille $(x_1, \dots, x_{n-1})$ est libre, alors la famille $(x_1, \dots, x_{n-1}, x_1 \wedge \dots \wedge x_{n-1})$ est une base de E .

Théorème

Si H est un hyperplan de E et $(x_1, \dots, x_{n-1})$ une base de H , alors la droite $D = H^\perp$ est dirigée par le vecteur $x_1 \wedge \dots \wedge x_{n-1}$ et pour tout vecteur x de E , la projection orthogonale de x sur H est :

$$p_H(x) = x - \frac{\langle x_1 \wedge \dots \wedge x_{n-1} \mid x \rangle}{\|x_1 \wedge \dots \wedge x_{n-1}\|^2} (x_1 \wedge \dots \wedge x_{n-1})$$

et la distance de x à H est donnée par $d(x, H) = \frac{|\langle x_1 \wedge \dots \wedge x_{n-1} \mid x \rangle|}{\|x_1 \wedge \dots \wedge x_{n-1}\|}$.

Le théorème précédent nous dit aussi qu'une équation de l'hyperplan H de base $(x_1, \dots, x_{n-1})$ est donnée par $\langle x_1 \wedge \dots \wedge x_{n-1} \mid x \rangle = 0$.

23.11 Isométries vectorielles en dimension 2

E est un espace euclidien de dimension 2, soit un *plan euclidien*, orienté par le choix d'une base orthonormée $\mathcal{B}_0 = (e_1, e_2)$.

Théorème

Un endomorphisme u de E est une isométrie positive [resp. négative] si, et seulement si, il existe un réel θ tel que la matrice de u dans la base orthonormée $\mathcal{B}_0$ soit de la forme $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ [resp. de la forme $S_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$].

Le réel θ qui intervient dans le théorème précédent est unique si on le prend dans $[-\pi, \pi[$.

Corollaire

Les groupes $\mathcal{SO}(E)$ et $\mathcal{SO}_2(\mathbb{R})$ sont commutatifs.

Théorème

Soit E un plan vectoriel euclidien orienté. Si u est une isométrie directe, il existe alors un réel θ tel que la matrice de u dans toute base orthonormée directe de E est $R_\theta = \begin{pmatrix} \cos(\theta) & -\varepsilon \sin(\theta) \\ \sin(\theta) & \varepsilon \cos(\theta) \end{pmatrix}$.

Une isométrie $u \in \mathcal{SO}(E)$ de matrice R_θ dans une base orthonormée directe $\mathcal{B}$, est une *rotation* et θ est une mesure de l'angle de cette rotation. Si $\theta \in [-\pi, \pi[$, on dit que c'est la *mesure principale de la rotation*. Dans une base indirecte, cette mesure principale est $-\theta$. On dit de manière plus précise que $\bar{\theta} = \{\theta + 2k\pi \mid k \in \mathbb{Z}\}$ est l'angle de la rotation dans l'espace orienté E . Par abus de langage, on dit parfois que θ est l'angle de la rotation, étant entendu que le réel θ est définie modulo 2π . L'inverse de la rotation d'angle $\bar{\theta}$ est la rotation d'angle $-\bar{\theta}$ et la composée des rotations ρ d'angle $\bar{\theta}$ et ρ' d'angle $\bar{\theta}'$ est la rotation $\rho \circ \rho' = \rho' \circ \rho$ d'angle $\bar{\theta} + \bar{\theta}'$.

Théorème

Si x, y sont deux vecteurs non nuls dans E , il existe alors une unique rotation $\rho \in \mathcal{SO}(E)$ telle que $\frac{1}{\|y\|}y = \rho\left(\frac{1}{\|x\|}x\right)$.

Si, avec les notations du théorème qui précède, ρ est la rotation d'angle $\bar{\theta}$, on dit alors que $\bar{\theta}$ est l'angle orienté des vecteurs x et y et on note $\widehat{(x, y)} = \bar{\theta}$. Un réel θ dans la classe d'équivalence $\bar{\theta}$ est une mesure de l'angle orienté $\widehat{(x, y)}$, le représentant $\theta \in [-\pi, \pi[$ est la mesure principale de $\widehat{(x, y)}$.

Théorème

Les isométries indirectes d'un plan euclidien sont les *réflexions*.

23.12 Isométries vectorielles en dimension 3

E est un espace euclidien de dimension 3 orienté par le choix d'une base orthonormée $\mathcal{B}_0 = (e_1, e_2, e_3)$.

Théorème

Soit $u \in \mathcal{SO}(E) \setminus \{\text{Id}\}$. L'ensemble des points fixes de u est une droite D . Si (f_1, f_2) est une base orthonormée du plan $D^\perp$, alors $(f_1, f_2, f_1 \wedge f_2)$ est une base orthonormée directe de E et la matrice de u dans cette base est

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Avec les notations du théorème, on dit que u est la *rotation* d'axe D orienté par $f_1 \wedge f_2$ et d'angle θ (défini modulo 2π).

Si $A \in \mathcal{SO}_3(\mathbb{R}) \setminus \{I_3\}$ est la matrice de $u \in \mathcal{SO}(E) \setminus \{\text{Id}\}$ dans la base orthonormée $\mathcal{B}_0$, alors l'axe de la rotation u est obtenu en déterminant le noyau de $u - \text{Id}$, ce

qui revient à résoudre un système linéaire $(A - I_3)X = 0$, où la matrice $A - I_3$ est de rang 2.

Pour ce qui est de la mesure principale $\theta \in [-\pi, \pi[\setminus \{0\}$ de l'angle de cette rotation, avec $\text{Tr}(u) = \text{Tr}(A) = \text{Tr}(R_\theta) = 2\cos(\theta) + 1$, on en déduit la valeur de $\cos(\theta)$ et celle de θ au signe près.

Si $\text{Tr}(u) = -1$, on a alors $\cos(\theta) = -1$, donc $\sin(\theta) = 0$ et $\theta = -\pi$.

Dans le cas où $\theta \in]-\pi, \pi[\setminus \{0\}$, on peut déterminer le signe de $\sin(\theta)$, et donc de θ , comme suit.

Avec $u(f_1) = \cos(\theta)f_1 + \sin(\theta)f_2$, on déduit que $\sin(\theta) = \langle u(f_1) | f_2 \rangle$. De plus, en notant $f_3 = f_1 \wedge f_2$, on a $f_3 \wedge f_1 = f_2$ et :

$$\sin(\theta) = \langle u(f_1) | f_2 \rangle = \langle u(f_1) | f_3 \wedge f_1 \rangle = \det(f_3, f_1, u(f_1)) = \det(f_1, u(f_1), f_3)$$

ce qui permet de déterminer $\sin(\theta)$ et θ .

Comme seul le signe de θ nous importe, on choisit un vecteur non nul x dans $D^\perp$, on pose $f_1 = \frac{1}{\|x\|}x$, on complète ce vecteur en une base orthonormée (f_1, f_2, f_3)

de E et $\sin(\theta) = \frac{\det(x, u(x), f_3)}{\|x\|^2}$ est du signe de $\det(x, u(x), f_3)$, ce qui permet de déterminer θ .

23.13 Espaces vectoriels hermitiens

Pour ce paragraphe, E est un espace vectoriel sur $\mathbb{C}$ de dimension $n \geq 1$.

Définition

On appelle *produit scalaire hermitien* sur E toute application

$$\begin{aligned} \varphi : E \times E &\rightarrow \mathbb{C} \\ (x, y) &\mapsto \varphi(x, y) \end{aligned}$$

qui vérifie les propriétés suivantes :

- pour tout y dans E l'application $x \mapsto \varphi(x, y)$ est linéaire ;
- $\varphi(y, x) = \overline{\varphi(x, y)}$ pour tous x, y dans E ;
- $\varphi(x, x) \geq 0$ pour tout x dans E ;
- pour x dans E , $\varphi(x, x) = 0$ équivaut à $x = 0$.

On note en général $(x, y) \mapsto \langle x | y \rangle$ un tel produit scalaire.

Définition

Un *espace hermitien* est un $\mathbb{C}$ -espace vectoriel de dimension finie muni d'un produit scalaire hermitien.

Dans la suite de ce paragraphe E désigne un espace hermitien et on note pour tout x dans E , $\|x\| = \sqrt{\langle x|x \rangle}$.

Théorème de Cauchy-Schwarz

Pour tous x, y dans E on a $|\langle x | y \rangle| \leq \|x\| \|y\|$, l'égalité étant réalisée si, et seulement si, x et y sont liés.

Théorème de Minkowski

Pour tous x, y dans E on a $\|x + y\| \leq \|x\| + \|y\|$, l'égalité étant réalisée si, et seulement si, $x = \lambda y$ avec $\lambda \geq 0$ (on dit que x et y sont *positivement liés*).

Comme dans le cas euclidien, on dit que deux vecteurs x et y dans E sont *orthogonaux* si $\langle x | y \rangle = 0$.

Le théorème de Pythagore est encore vrai sous la forme : si les vecteurs x et y sont orthogonaux dans E , on a alors $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. La réciproque étant fautive dans le cas hermitien (l'égalité $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ entraîne seulement $\operatorname{Re} \langle x | y \rangle = 0$).

On définit, comme dans le cas euclidien, les notions de famille orthogonale et orthonormale. Une famille orthogonale de vecteurs non nuls de E est libre.

Définition

L'orthogonal de $X \neq \emptyset$ dans E est $X^\perp = \{y \in E, \forall x \in X, \langle x | y \rangle = 0\}$.

$X^\perp$ est un sous-espace vectoriel de E .

Théorème d'orthonormalisation de Gram-Schmidt

Pour toute famille libre $(x_k)_{1 \leq k \leq p}$ dans E , il existe une unique famille orthonormée $(e_k)_{1 \leq k \leq p}$ dans E telle que :

$$\forall k \in \{1, \dots, p\}, \quad \begin{cases} \operatorname{Vect}\{e_1, \dots, e_k\} = \operatorname{Vect}\{x_1, \dots, x_k\}; \\ \langle x_k | e_k \rangle > 0. \end{cases}$$

En notant $\mathcal{E} = (e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$ muni de sa structure hermitienne canonique avec le produit scalaire $(x, y) \mapsto \langle x | y \rangle = \sum_{k=1}^n x_k \overline{y_k}$, tout système orthonormé $\mathcal{F} = (f_k)_{1 \leq k \leq n}$ est une base de $\mathbb{C}^n$ et la matrice de passage U de la base canonique $\mathcal{E}$ à la base $\mathcal{F}$ est telle que $U^* U = I_n$, où $U^* = (u_{ij}^*)$ désigne la

matrice adjointe de la matrice complexe $U = (u_{ij})$, définie par $u_{ij}^* = \overline{u_{ji}}$ pour $1 \leq i, j \leq n$.

Définition

On dit qu'une matrice carrée A à coefficients complexes d'ordre n est *unitaire* si $U^*U = I_n$. On note $\mathcal{U}_n(\mathbb{C})$ l'ensemble des matrices unitaires.

On peut également dire qu'une matrice unitaire est la matrice de passage de la base canonique de $\mathbb{C}^n$ à une base orthonormée où $\mathbb{C}^n$ est muni de sa structure hermitienne canonique.

23.14 Réduction des matrices normales

On note $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$ et le produit scalaire hermitien canonique de $\mathbb{C}^n$ est noté :

$$(x, y) = \left(\sum_{i=1}^n x_i e_i, \sum_{i=1}^n y_i e_i \right) \mapsto \langle x | y \rangle = \sum_{i=1}^n x_i \overline{y_i}.$$

La norme associée est la norme hermitienne notée $x \mapsto \|x\|$.

Définition

Une matrice complexe A est dite *hermitienne* si $A^* = A$. On dit qu'elle est *hermitienne positive*, si on a de plus $\langle Ax | x \rangle \geq 0$ pour tout $x \in E$. On dit qu'elle est *hermitienne définie positive*, si on a de plus $\langle Ax | x \rangle > 0$ pour tout $x \in E \setminus \{0\}$.

Définition

Une matrice complexe A est dite *normale* si $A^*A = AA^*$.

Les matrices hermitiennes et unitaires sont des cas particuliers de matrices normales.

Lemme

Les valeurs propres d'une matrice hermitienne [resp. hermitienne positive] sont réelles [resp. réelles et positives]. Les valeurs propres d'une matrice unitaire sont de module égal à 1.

Théorème

Si A une matrice complexe normale, elle se diagonalise alors dans une base orthonormée, c'est-à-dire qu'il existe une matrice unitaire U et une matrice diagonale D telles que $A = UDU^*$, les coefficients diagonaux de D étant les valeurs propres dans $\mathbb{C}$ de la matrice A .

Corollaire

Une matrice hermitienne [resp. unitaire] a ses valeurs propres réelles [resp. de module 1] et se diagonalise dans une base orthonormée.

24.1 Espace affine associé à un espace vectoriel

E est un $\mathbb{R}$ -espace vectoriel.

Définition

Un *espace affine* de direction E est un ensemble non vide $\mathcal{E}$ tel que l'on puisse définir une loi externe $(x, u) \in \mathcal{E} \times E \mapsto x + u$ vérifiant les propriétés suivantes :

- pour tout $x \in \mathcal{E}$ l'application $u \mapsto x + u$ est bijective ;
- pour tous $x \in \mathcal{E}$ et u, u' dans E , on a $(x + u) + u' = x + (u + u')$.

E est un espace affine sur lui même grâce à l'application $(x, y) \in E^2 \mapsto x + y$. Les éléments de $\mathcal{E}$ sont appelés points et ceux de E , vecteurs. Un vecteur $u \in E$ sera parfois noté $\overrightarrow{u}$. La direction de l'espace affine $\mathcal{E}$ est parfois notée $E = \overrightarrow{\mathcal{E}}$. La dimension de l'espace affine $\mathcal{E}$ est la dimension de l'espace vectoriel E . Les espaces affines de dimension 1 [resp. de dimension 2] sont appelés *droites affines* [resp. *plans affines*].

Pour tout couple (x, y) de points de $\mathcal{E}$, le vecteur $u \in E$ tel que $y = x + u$ est noté $\overrightarrow{xy}$ et il est caractérisé par $x + \overrightarrow{xy} = y$. Il sera parfois commode de noter $\overrightarrow{xy} = y - x$. Avec la première notation, on a pour tous points x, y, z de $\mathcal{E}$:

$$\overrightarrow{xy} + \overrightarrow{yz} = \overrightarrow{xz}, \quad \overrightarrow{xx} = \vec{0}, \quad \overrightarrow{yx} = -\overrightarrow{xy}.$$

La première des identités précédentes est la *relation de Chasles*.

Un segment dans $\mathcal{E}$ d'extrémités a, b est l'ensemble $[a, b] = \{a + \lambda \overrightarrow{ab} ; \lambda \in [0, 1]\}$. Pour tout vecteur $u \in E$, l'application $t_u : x \in \mathcal{E} \mapsto x + u \in \mathcal{E}$ est une bijection de $\mathcal{E}$ sur $\mathcal{E}$ d'inverse $t_{-u} : x \mapsto x + (-u)$ (que l'on notera $x - u$) et l'application

$t : u \mapsto t_u$ est un morphisme de groupes injectif de $(E, +)$ dans $(\mathcal{S}(\mathcal{E}), \circ)$. Une telle application t_u est appelée *translation* de vecteur u .

Une translation t_u de vecteur u non nul n'a pas de point fixe (pour $u = 0$, on a $t_u = \text{Id}_{\mathcal{E}}$ et tous les points de $\mathcal{E}$ sont fixes). L'ensemble $\mathcal{T}(\mathcal{E})$ de ces translations est un sous-groupe de $(\mathcal{S}(\mathcal{E}), \circ)$ isomorphe au groupe $(E, +)$.

Pour tout point ω de $\mathcal{E}$, l'application $x \in \mathcal{E} \mapsto \overrightarrow{\omega x} \in E$ est bijective d'inverse $u \mapsto \omega + u$, ce qui permet d'identifier l'espace affine $\mathcal{E}$ à l'espace vectoriel E , cette identification dépendant du choix d'une origine ω . On note $\mathcal{E} = \omega + E$ ou $\mathcal{E} = \omega + \overrightarrow{\mathcal{E}}$. Une droite affine est de la forme $\mathcal{D} = \omega + \mathbb{R}\overrightarrow{v}$ où $\overrightarrow{v} \in \overrightarrow{\mathcal{D}} \setminus \{\overrightarrow{0}\}$ est un *vecteur directeur* de $\overrightarrow{\mathcal{D}}$. Un plan affine est de la forme $\mathcal{P} = \omega + \mathbb{R}\overrightarrow{u} \oplus \mathbb{R}\overrightarrow{v}$ où $\overrightarrow{u}$ et $\overrightarrow{v}$ sont linéairement indépendants dans $\overrightarrow{\mathcal{P}} \setminus \{\overrightarrow{0}\}$.

Pour E de dimension n , le choix d'une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ de E permet d'identifier l'espace vectoriel E à $\mathbb{R}^n$ et le choix d'une origine $O \in \mathcal{E}$ permet d'identifier l'espace affine $\mathcal{E}$ à $\mathbb{R}^n$: pour tout x de $\mathcal{E}$, il existe un unique n -uplet $(x_k)_{1 \leq k \leq n} \in \mathbb{R}^n$ tel

que $x = O + \sum_{k=1}^n x_k \overrightarrow{e_k}$. On dit que $\mathcal{R} = (O, \mathcal{B})$ est un *repère affine* de $\mathcal{E}$ et que les x_k , pour k compris entre 1 et n , sont les coordonnées de x dans le repère $\mathcal{R}$.

Si $\mathcal{R}' = (O', \mathcal{B}')$ est un autre repère affine de $\mathcal{E}$, en utilisant la relation de Chasles $\overrightarrow{Ox} = \overrightarrow{OO'} + \overrightarrow{O'x}$, on a les formules de changement de repère :

$$x_i = O'_i + \sum_{j=1}^n p_{ij} x_j \quad (1 \leq i \leq n)$$

où les O'_i sont les coordonnées de O' dans le repère $\mathcal{R}$, ce qui peut aussi s'écrire sous forme matricielle $X = \Omega' + PX$ en notant $P = (p_{ij})$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

Définition

Une partie $\mathcal{F}$ d'un espace affine $\mathcal{E}$ est un *sous-espace affine* s'il existe un point ω de $\mathcal{F}$ tel que l'ensemble $F = \{\overrightarrow{\omega x} ; x \in \mathcal{F}\}$ soit un sous-espace vectoriel de E .

Si $(\mathcal{F}_i)_{i \in I}$ est une famille de sous-espaces affines de $\mathcal{E}$, leur intersection $\bigcap_{i \in I} \mathcal{F}_i$ est alors un sous-espace affine de $\mathcal{E}$.

Si $\mathcal{F}$ est un sous-espace affine de $\mathcal{E}$, on a alors $F = \{\overrightarrow{x y} ; y \in \mathcal{F}\}$ pour tout point x de $\mathcal{F}$ et $\mathcal{F}$ est un espace affine de direction F . Par tout point ω de $\mathcal{E}$ passe un unique sous-espace affine $\mathcal{F}_\omega$ de $\mathcal{E}$ de direction F . Pour ω, ω' dans $\mathcal{E}$, on a :

$$\mathcal{F}_{\omega'} = \{\omega' + u ; u \in F\} = \left\{ \omega + \overrightarrow{\omega \omega'} + u ; u \in F \right\} = \overrightarrow{t_{\omega \omega'}}(\mathcal{F}_\omega).$$

Deux sous-espaces affines de $\mathcal{E}$ de même direction sont dits *parallèles*. La relation de *parallélisme* sur les sous-espaces affines de $\mathcal{E}$ est une relation d'équivalence.

Un hyperplan de $\mathcal{E}$ est un sous-espace affine dont la direction est un hyperplan de E . Pour $\mathcal{E}$ de dimension $n \geq 2$, les hyperplans sont ses sous-espaces affines de dimension $n - 1$.

Si $\mathcal{H}$ est hyperplan de $\mathcal{E}$ de direction $H = \text{Ker}(\ell)$, où ℓ est une forme linéaire non nulle, on a alors $\mathcal{H} = \{\omega + u ; u \in E, \ell(u) = 0\}$. En dimension finie, dans un repère affine $\mathcal{R} = (\omega, \mathcal{B})$, l'hyperplan $\mathcal{H}$ est l'ensemble des points de $\mathcal{E}$ d'équation cartésienne $\sum_{k=1}^n \alpha_k x_k = \lambda$, où les réels α_k ne sont pas tous nuls et λ est un réel.

En dimension 2, l'équation d'une droite est de la forme $ax + by = c$ et en dimension 3, l'équation d'un plan est de la forme $ax + by + cz = d$.

Définition

Soient $\mathcal{E}, \mathcal{E}'$ deux espaces affines de directions respectives E, E' . On dit qu'une application $f : \mathcal{E} \rightarrow \mathcal{E}'$ est *affine* s'il existe une application linéaire $\varphi : E \rightarrow E'$ telle que $f(x + u) = f(x) + \varphi(u)$ pour tout $(x, u) \in \mathcal{E} \times E$, ce qui revient à dire que :

$$\forall u \in E, f \circ t_u = t_{\varphi(u)} \circ f.$$

La condition $f(x + u) = f(x) + \varphi(u)$ pour tout $(x, u) \in \mathcal{E} \times E$ est équivalente à $\overrightarrow{f(x)f(y)} = \varphi(\overrightarrow{xy})$ pour tout $(x, y) \in \mathcal{E}^2$. L'application linéaire φ associée à une application affine f est uniquement déterminée et f est injective [resp. surjective] si, et seulement si, φ est injective [resp. surjective]. Il sera parfois commode de noter $\varphi = \overrightarrow{f}$.

Si une application affine $f : \mathcal{E} \rightarrow \mathcal{E}'$ est bijective, son inverse est alors affine d'application linéaire associée $(\overrightarrow{f})^{-1}$.

La composée $f \circ g$ de deux applications affines f et g est affine d'application linéaire associée $\overrightarrow{f} \circ \overrightarrow{g}$.

Dans le cas où $\mathcal{E}$ et $\mathcal{E}'$ sont de dimensions respectives m et n , si $\mathcal{R} = (O, \mathcal{B})$ est un repère affine de $\mathcal{E}$ et $\mathcal{R}' = (O', \mathcal{B}')$ est un repère affine de $\mathcal{E}'$, en désignant par X le vecteur de $\mathbb{R}^m$ formé des coordonnées de $x \in \mathcal{E}$ dans le repère $\mathcal{R}$, par Y le vecteur de $\mathbb{R}^n$ formé des coordonnées de $y = f(x)$ dans le repère $\mathcal{R}'$, par B le vecteur de $\mathbb{R}^n$ formé des coordonnées de $f(O)$ dans le repère $\mathcal{R}'$ et par A la matrice de $\overrightarrow{f}$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$, l'égalité $\overrightarrow{f(O)f(x)} = \varphi(\overrightarrow{Ox})$ se traduit par $Y = AX + B$.

Pour $m = n$, l'application f est bijective si, et seulement si, la matrice A est inversible.

Une translation t_u est une bijection affine d'application linéaire associée Id_E .

Théorème

Soient $\mathcal{E}, \mathcal{E}'$ deux espaces affines de directions respectives E, E' . Pour toute application linéaire $\varphi : E \rightarrow E'$, tout point a de $\mathcal{E}$ et tout point a' de $\mathcal{E}'$, il existe une unique application affine $f : \mathcal{E} \rightarrow \mathcal{E}'$ telle que $\overrightarrow{f} = \varphi$ et $f(a) = a'$, c'est l'application $f : x \mapsto a' + \varphi(\overrightarrow{ax})$.

Du théorème précédent, on déduit que deux applications affines qui ont même application linéaire sont égales si, et seulement si, elles coïncident en un point.

Définition

Soient ω un point de $\mathcal{E}$ et λ un réel non nul. L'*homothétie* de centre ω et de rapport λ est l'application affine $f : \mathcal{E} \rightarrow \mathcal{E}$ d'application linéaire associée $\overrightarrow{f} = \lambda \text{Id}_E$ telle que $f(\omega) = \omega$.

On note $h(\omega, \lambda)$ une telle homothétie. Elle est aussi définie par $\overrightarrow{\omega f(x)} = \lambda \overrightarrow{\omega x}$ pour tout $x \in \mathcal{E}$. Pour $\lambda = 1$, on a $f = \text{Id}_{\mathcal{E}}$ et tous les points de $\mathcal{E}$ sont fixes par f . Pour $\lambda = -1$, on a $\overrightarrow{\omega f(x)} = -\overrightarrow{\omega x}$ pour tout $x \in \mathcal{E}$ et f est la symétrie de centre ω . Pour $\lambda \neq 1$, ω est l'unique point fixe de f .

Théorème

Soient $\mathcal{E}, \mathcal{E}'$ deux espaces affines et f, g deux applications affines de $\mathcal{E}$ dans $\mathcal{E}'$. L'ensemble $\mathcal{F} = \{x \in \mathcal{E}, f(x) = g(x)\}$ est soit vide, soit un sous-espace affine de $\mathcal{E}$ de direction $\text{Ker}(\overrightarrow{f} - \overrightarrow{g})$.

Pour $\mathcal{E} = \mathcal{E}'$ de dimension finie et $g = \text{Id}_{\mathcal{E}}$, l'ensemble des points fixes d'une application affine $f : \mathcal{E} \rightarrow \mathcal{E}$ est soit vide, soit un sous-espace affine de $\mathcal{E}$ de direction $\text{Ker}(\overrightarrow{f} - \text{Id}_E)$.

Théorème

Soient $\mathcal{F}$ un sous-espace affine d'un espace affine $\mathcal{E}$ et G un supplémentaire dans E de la direction F de $\mathcal{F}$. Pour tout $x \in \mathcal{E}$, l'intersection $\mathcal{F} \cap (x + G)$ est réduite à un point $p(x)$. L'application $p : \mathcal{E} \rightarrow \mathcal{F}$ est affine d'application linéaire associée la projection vectorielle de E sur F parallèlement à G .

Définition

Avec les notations et hypothèses du théorème précédent, on dit que p est la *projection affine* de $\mathcal{E}$ sur $\mathcal{F}$ parallèlement à G .

Théorème

Une application affine $p : \mathcal{E} \rightarrow \mathcal{F}$ est une projection si, et seulement si, elle vérifie $p \circ p = p$.

Théorème de Thalès

Soient $\mathcal{H}$ un hyperplan de direction H d'un espace affine $\mathcal{E}$ de direction E et $\mathcal{D}, \mathcal{D}'$, deux droites dans $\mathcal{E}$ de directions D, D' supplémentaires de H dans E . Étant donné trois hyperplans deux à deux distincts $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ de même direction H rencontrant $\mathcal{D}$ et $\mathcal{D}'$ en x_1, x_2, x_3 et x'_1, x'_2, x'_3 respectivement, on a $\frac{\overrightarrow{x_1x_3}}{\overrightarrow{x_1x_2}} = \frac{\overrightarrow{x'_1x'_3}}{\overrightarrow{x'_1x'_2}}$.

Théorème

Soient $\mathcal{F}$ un sous-espace affine d'un espace affine $\mathcal{E}$, G un supplémentaire dans E de la direction F de $\mathcal{F}$ et p la projection de $\mathcal{E}$ sur $\mathcal{F}$ parallèlement à G . L'application $s : \mathcal{E} \rightarrow \mathcal{E}$ définie par $s(x) = x + 2\overrightarrow{xp(x)}$ est affine d'application linéaire associée la symétrie vectorielle par rapport à F et parallèlement à G .

Définition

Avec les notations et hypothèses du théorème précédent, on dit que s est la *symétrie affine* par rapport à $\mathcal{F}$ et parallèlement à G .

Théorème

Une application affine $s : \mathcal{E} \rightarrow \mathcal{E}$ est une symétrie si, et seulement si, elle vérifie $s \circ s = \text{Id}$.

24.2 Le groupe affine $\text{GA}(\mathcal{E})$ en dimension finie

$\mathcal{E}$ est un espace affine de dimension $n \geq 1$ et de direction E . On note $\text{GA}(\mathcal{E})$ l'ensemble des bijections affines de $\mathcal{E}$ sur $\mathcal{E}$, c'est un groupe pour la composition

des applications et l'application $f \in \text{GA}(\mathcal{E}) \mapsto \overrightarrow{f} \in \text{GL}(E)$ est un morphisme surjectif de groupes.

Pour tout $\omega \in \mathcal{E}$, on note $\text{GA}(\mathcal{E})_\omega = \{f \in \text{GA}(\mathcal{E}), f(\omega) = \omega\}$ l'ensemble des bijections affines qui laissent fixe le point ω . C'est un sous-groupe de $\text{GA}(\mathcal{E})$.

Théorème

- Pour tout point ω de $\mathcal{E}$, l'application $f \mapsto \overrightarrow{f}$ réalise un isomorphisme de groupes de $\text{GA}(\mathcal{E})_\omega$ sur $\text{GL}(E)$.
- Soit $\omega \in \mathcal{E}$. Pour tout $f \in \text{GA}(\mathcal{E})$, il existe deux couples (u, g) et (v, h) uniquement déterminés dans $E \times \text{GA}(\mathcal{E})_\omega$ tels que $f = t_u \circ g = h \circ t_v$.

Théorème

- L'ensemble $\mathcal{T}(\mathcal{E})$ des translations de $\mathcal{E}$ est un sous-groupe commutatif de $\text{GA}(\mathcal{E})$.
- Une application f de $\mathcal{E}$ dans $\mathcal{E}$ est une translation si, et seulement si, elle commute à toute translation.
- Pour toute bijection affine f et tout $u \in E$, on a $f \circ t_u \circ f^{-1} = t_{\overrightarrow{f}(u)}$.
- L'ensemble $\mathcal{H}(\mathcal{E}) = \{f \in \text{GA}(\mathcal{E}), \overrightarrow{f} \in \mathbb{R}^* \cdot \text{Id}_E\}$ est un sous-groupe de $\text{GA}(\mathcal{E})$.

Le groupe $\mathcal{H}(\mathcal{E})$ est le groupe des *homothéties ou translations*.

Théorème

- Pour tout point ω de $\mathcal{E}$, l'ensemble $\mathcal{H}_\omega(\mathcal{E})$ des homothéties de centre ω est un sous-groupe de $\mathcal{H}(\mathcal{E})$.
- On a $\mathcal{H}(\mathcal{E}) = \mathcal{T}(\mathcal{E}) \cup \left(\bigcup_{\omega \in \mathcal{E}} \mathcal{H}_\omega(\mathcal{E}) \right)$.
- Pour toute homothétie $h(\omega, \lambda)$ et toute bijection affine $f \in \text{GA}(\mathcal{E})$, on a $f \circ h(\omega, \lambda) \circ f^{-1} = h(f(\omega), \lambda)$.

Théorème

Soient $h(\omega, \lambda)$, $h(\omega', \lambda')$ deux homothéties et $h = h(\omega, \lambda) \circ h(\omega', \lambda')$.

- Pour $\omega = \omega'$, h est une homothétie de rapport $\lambda\lambda'$ et de centre ω (pour $\lambda\lambda' = 1$, on a $h = \text{Id}_{\mathcal{E}}$).
- Pour $\omega \neq \omega'$ et $\lambda\lambda' \neq 1$, h est l'homothétie de rapport $\lambda\lambda'$ et de centre $\omega'' = \omega' + \frac{\lambda-1}{1-\lambda\lambda'}\overrightarrow{\omega\omega'} \in (\omega\omega')$ (droite passant par les points ω et ω').
- Pour $\omega \neq \omega'$ et $\lambda\lambda' = 1$, on a $h = \text{Id}_{\mathcal{E}}$ si $\lambda = 1$ ou $h = t_{(\lambda-1)\overrightarrow{\omega\omega'}}$ si $\lambda \neq 1$.

Le théorème précédent nous dit que le groupe $\mathcal{H}(\mathcal{E})$ n'est pas commutatif. Plus précisément, deux translations commutent, une homothétie de rapport différent de 1 et une translation de vecteur non nul ne commutent pas et deux homothéties commutent si, et seulement si, elles ont même centre.

24.3 Espaces affines euclidiens

Définition

Un *espace affine euclidien* est un espace affine de direction un espace euclidien.

Pour ce qui suit, $\mathcal{E}$ est un espace affine euclidien de direction $(E, \langle \cdot | \cdot \rangle)$. On note $\|\cdot\|$ la norme associée au produit scalaire $\langle \cdot | \cdot \rangle$ sur E .

L'application $d : (x, y) \rightarrow \|\overrightarrow{xy}\|$ est une distance sur $\mathcal{E}$ dite euclidienne.

Un repère $\mathcal{R} = (O, \mathcal{B})$ de $\mathcal{E}$ est dit *orthonormé* si $\mathcal{B}$ est une base orthonormée de

E . Dans un tel repère, on a $d(x, y) = \sqrt{\sum_{k=1}^n (y_k - x_k)^2}$ avec des notations évidentes.

Deux droites affines dans $\mathcal{E}$ sont dites *perpendiculaires* si leurs vecteurs directeurs sont orthogonaux.

Définition

On dit qu'une partie S de $\mathcal{E}$ est une *sphère* s'il existe un point ω dans $\mathcal{E}$ et un réel R positif ou nul tels que $S = \{x \in E, d(\omega, x) = R\}$. On dit alors que ω est un centre et R un rayon de cette sphère.

On note $S(\omega, R)$ une telle sphère et on vérifie que le centre et le rayon sont uniquement déterminés.

En dimension 2, une sphère est appelée cercle.

Définition

Soient $\omega \in \mathcal{E}$ et R un réel positif ou nul. La *boule fermée* [resp. ouverte] de centre ω et de rayon R est l'ensemble $B(\omega, R) = \{x \in E, d(\omega, x) \leq R\}$ [resp. $\overset{\circ}{B}(\omega, r) = \{x \in E, d(\omega, x) < R\}$].

Définition

Si $S(\omega, R)$ une sphère de centre ω et de rayon R , on appelle *diamètre* de cette sphère tout segment $[a, b]$ où a et b sont deux points de $S(\omega, R)$ tels que $d(a, b) = 2R$.

Théorème

Soient a, b deux points de $\mathcal{E}$. L'ensemble $S = \{x \in \mathcal{E}, \langle \overrightarrow{ax} | \overrightarrow{bx} \rangle = 0\}$ est une sphère de diamètre $[a, b]$.

Si $\mathcal{R} = (O, \mathcal{B})$ est un repère orthonormé de $\mathcal{E}$ où $\mathcal{B} = (e_i)_{1 \leq i \leq n}$, on a alors en notant $x_1, \dots, x_n$ les coordonnées d'un point x de $\mathcal{E}$ dans ce repère :

$$(x \in S(\omega, R)) \Leftrightarrow \left(\sum_{k=1}^n x_k^2 - 2 \sum_{k=1}^n \omega_k x_k + \sum_{k=1}^n \omega_k^2 - R^2 = 0 \right).$$

Réciproquement si $\omega_1, \dots, \omega_n$ et c sont des réels, alors en notant $\omega = O + \sum_{k=1}^n \omega_k e_k$,

l'ensemble des points $x = O + \sum_{k=1}^n x_k e_k$ de $\mathcal{E}$ tels que $\sum_{k=1}^n x_k^2 - 2 \sum_{k=1}^n \omega_k x_k + c = 0$ est

: l'ensemble vide si $c > \sum_{k=1}^n \omega_k^2$; réduit à $\{\omega\}$ si $c = \sum_{k=1}^n \omega_k^2$; la sphère de centre ω

et de rayon $R = \sqrt{\sum_{k=1}^n \omega_k^2 - c}$ si $c < \sum_{k=1}^n \omega_k^2$.

Dans le cas où $\mathcal{E}$ est un plan affine euclidien, tout point x du cercle $S(\omega, R)$ s'écrit de manière unique $x = \omega + R \cos(\theta) e_1 + R \sin(\theta) e_2$ où $\theta \in]-\pi, \pi]$.

Dans le cas où $n = 3$, tout point x de la sphère $S(\omega, R)$ s'écrit de manière unique $x = \omega + R \cos(\theta_2) \sin(\theta_3) e_1 + R \sin(\theta_2) \sin(\theta_3) e_2 + R \cos(\theta_3) e_3$ où $\theta_3 \in [0, \pi]$ et $\theta_2 \in]-\pi, \pi]$.

Pour $n \geq 3$, une paramétrisation de la sphère $S(\omega, R)$ est donnée par

$$\begin{cases} x_1 = \omega_1 + R \cos(\theta_2) \sin(\theta_3) \cdots \sin(\theta_n) \\ x_2 = \omega_2 + R \sin(\theta_2) \sin(\theta_3) \cdots \sin(\theta_n) \\ x_3 = \omega_3 + R \cos(\theta_3) \sin(\theta_4) \cdots \sin(\theta_n) \\ \vdots \\ x_{n-2} = \omega_{n-2} + R \cos(\theta_{n-2}) \sin(\theta_{n-1}) \sin(\theta_n) \\ x_{n-1} = \omega_{n-1} + R \cos(\theta_{n-1}) \sin(\theta_n) \\ x_n = \omega_n + R \cos(\theta_n) \end{cases}$$

avec $\theta_2 \in]-\pi, \pi]$ et $\theta_3, \dots, \theta_n$ dans $[0, \pi]$.

24.4 Hyperplans dans un espace euclidien

$\mathcal{E}$ est un espace affine euclidien de dimension $n \geq 2$ et $\mathcal{R} = (O, \mathcal{B})$ est un repère orthonormé de $\mathcal{E}$ où $\mathcal{B} = (e_i)_{1 \leq i \leq n}$.

Un hyperplan affine est un sous-espace affine de $\mathcal{E}$ de la forme :

$$\mathcal{H} = \{\omega + u ; u \in E, \ell(u) = 0\}$$

où ℓ est une forme linéaire non nulle. C'est l'hyperplan affine passant par ω et dirigé par $\text{Ker}(\ell)$.

Sachant que pour toute forme linéaire ℓ sur E il existe un unique vecteur $a \in E$ tel que $\ell(x) = \langle a | x \rangle$ pour tout $x \in E$, on déduit que $\mathcal{H} = \omega + \{a\}^\perp$ où a est un vecteur non nul (hyperplan passant par ω et orthogonal au vecteur a). Pour tout $x \in \mathcal{H}$, on a $\langle \overrightarrow{\omega x} | \overrightarrow{a} \rangle = 0$, soit $\langle \overrightarrow{Ox} | \overrightarrow{a} \rangle = \lambda$ où $\lambda = \langle \overrightarrow{\omega O} | \overrightarrow{a} \rangle$. Réciproquement, pour tout vecteur $\overrightarrow{a} \in E \setminus \{\overrightarrow{0}\}$ et tout réel λ , l'ensemble $\mathcal{H} = \{x \in \mathcal{E}, \langle \overrightarrow{Ox} | \overrightarrow{a} \rangle = \lambda\}$ est un hyperplan affine de direction $\{\overrightarrow{a}\}^\perp$. Dans le repère $\mathcal{R}$ une équation de cet hyperplan est $\sum_{k=1}^n a_k x_k = \lambda$, les réels a_k n'étant pas tous nuls.

Théorème

Soit $\mathcal{H}$ un hyperplan de $\mathcal{E}$ d'équation $\sum_{k=1}^n a_k x_k = \lambda$ dans le repère $\mathcal{R}$. Pour tout $x \in \mathcal{E}$ de coordonnées $x_1, \dots, x_n$ dans le repère $\mathcal{R}$, la distance de x à

$$\mathcal{H} \text{ est } d(x, H) = \frac{\left| \sum_{k=1}^n a_k x_k - \lambda \right|}{\sqrt{\sum_{k=1}^n a_k^2}}.$$

Théorème

Soient S une sphère de centre ω et de rayon $R > 0$ et $\mathcal{H}$ un hyperplan affine de $\mathcal{E}$. On note $d = d(\omega, \mathcal{H})$ la distance de ω à $\mathcal{H}$.

- Si $d > R$, alors $\mathcal{H} \cap S = \emptyset$.
- Si $d = R$, alors $\mathcal{H} \cap S = \{p_{\mathcal{H}}(\omega)\}$, où $p_{\mathcal{H}}(\omega)$ est la projection orthogonale de ω sur $\mathcal{H}$ (voir la définition 24.5).
- Si $d < R$, alors $\mathcal{H} \cap S$ est une sphère de $\mathcal{H}$ de centre $p_{\mathcal{H}}(\omega)$ et de rayon $\sqrt{R^2 - d^2}$.

24.5 Isométries affines d'un espace euclidien

$\mathcal{E}$ est un espace affine euclidien.

Définition

On appelle *isométrie affine* toute application $f : \mathcal{E} \rightarrow \mathcal{E}$ telle que :

$$\forall (x, y) \in \mathcal{E}^2, d(f(x), f(y)) = d(x, y).$$

On note $\text{Is}(\mathcal{E})$ l'ensemble de toutes les isométries affines de $\mathcal{E}$.

Définition

Une *projection affine orthogonale* sur $\mathcal{E}$ est une projection sur un sous-espace affine $\mathcal{F}$ parallèlement à $\vec{\mathcal{F}}^\perp$. Une *symétrie affine orthogonale* sur $\mathcal{E}$ est une symétrie par rapport à un sous-espace affine $\mathcal{F}$ parallèlement à $\vec{\mathcal{F}}^\perp$. Une *réflexion affine* est une symétrie orthogonale par rapport à un hyperplan affine.

Théorème

On a $\text{Is}(\mathcal{E}) = \{f \in \text{GA}(\mathcal{E}), \vec{f} \in \mathcal{O}(E)\}$ et $\text{Is}(\mathcal{E})$ est un sous-groupe de $\text{GA}(\mathcal{E})$ qui contient le groupe $\mathcal{T}(\mathcal{E})$ des translations et l'ensemble des symétries orthogonales.

L'application $f \in \text{Is}(\mathcal{E}) \mapsto \det(\vec{f}) \in \{-1, 1\}$ est un morphisme de groupes surjectif. Son noyau est noté $\text{Is}^+(\mathcal{E})$ et ses éléments sont appelés *déplacements* ou isométries directes de $\mathcal{E}$.

On note $\text{Is}^-(\mathcal{E}) = \text{Is}(\mathcal{E}) \setminus \text{Is}^+(\mathcal{E})$ et les éléments de cet ensemble sont appelés *anti-déplacements* ou isométries indirectes.

Théorème

Soit $f \in \text{Is}(\mathcal{E})$. Il existe un unique couple (a, g) formé d'un vecteur $a \in E$ et d'une isométrie g admettant au moins un point fixe tel que $f = g \circ t_a = t_a \circ g$. De plus, on a $\overrightarrow{f}(a) = \overrightarrow{g}(a) = a$.

Corollaire

Soit $f \neq \text{Id}$ un déplacement d'un plan affine euclidien orienté $\mathcal{P}$.

- Si f n'admet pas de point fixe, c'est alors une translation.
- Si f admet au moins un point fixe, l'application linéaire associée $\overrightarrow{f}$ est alors une rotation d'angle θ et f admet un unique point fixe ω .

Dans le deuxième cas du corollaire précédent, on dit que f est une *rotation* de centre ω et d'angle θ .

Corollaire

Soit $f \neq \text{Id}$ un anti-déplacement d'un plan affine euclidien orienté $\mathcal{P}$.

- Si f admet au moins un point fixe, c'est alors une réflexion.
- Si f n'admet pas de point fixe, c'est alors la composée commutative d'une translation t_a et d'une réflexion, le vecteur a dirigeant l'axe $\mathcal{D}$ de la réflexion (*symétrie glissée* d'axe $\mathcal{D}$ et de vecteur a).

24.6 Barycentres

$\mathcal{E}$ est un espace affine euclidien de dimension n et E l'espace vectoriel euclidien associé.

Lemme

Soit $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$ un système de points pondérés dans $\mathcal{E}$, où les α_k sont des réels tels que $\sum_{k=1}^p \alpha_k \neq 0$. Il existe un unique point $G \in \mathcal{E}$ qui vérifie l'une des conditions équivalentes suivantes :

- $\sum_{k=1}^p \alpha_k \overrightarrow{GM_k} = \vec{0}$;
- il existe $\Omega \in \mathcal{E}$ tel que $\left(\sum_{k=1}^p \alpha_k\right) \overrightarrow{\Omega G} = \sum_{k=1}^p \alpha_k \overrightarrow{\Omega M_k}$;
- pour tout $\Omega \in \mathcal{E}$, on a $\left(\sum_{k=1}^p \alpha_k\right) \overrightarrow{\Omega G} = \sum_{k=1}^p \alpha_k \overrightarrow{\Omega M_k}$.

Définition

Avec les notations du lemme précédent, on dit que G est le *barycentre* de la famille $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$. Dans le cas où tous les α_k sont égaux, on dit que G est l'*isobarycentre* de la famille $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$.

On note $G = \text{Bar} \left\{ \{(M_k, \alpha_k)\}_{1 \leq k \leq p} \right\}$.

Le barycentre étant inchangé si on multiplie les poids α_k par un même réel non nul, on peut supposer, dans la définition du barycentre, que $\sum_{k=1}^p \alpha_k = 1$ et dans ce cas, on peut noter $G = \sum_{k=1}^p \alpha_k M_k$ (on a « vectorialisé » $\mathcal{E}$ en se fixant une origine). Si A, B sont deux points distincts de $\mathcal{E}$, alors :

- la droite (AB) est l'ensemble des barycentres de $\{(A, \alpha), (B, 1 - \alpha)\}$ où α décrit $\mathbb{R}$;
- le segment $[A, B]$ est l'ensemble des barycentres de $\{(A, \alpha), (B, 1 - \alpha)\}$ où α décrit $[0, 1]$;
- le milieu I du segment $[A, B]$ est par définition l'isobarycentre de $\{A, B\}$, il est caractérisé par $\overrightarrow{AI} = \frac{1}{2} \overrightarrow{AB}$ ou par $\overrightarrow{AI} = \overrightarrow{IB}$.
- Le sous-espace affine de $\mathcal{E}$ engendré par $(a_k)_{1 \leq k \leq p} \in \mathcal{E}^p$ est l'ensemble des barycentres de $\{(A_k, \alpha_k)\}_{1 \leq k \leq p}$ où $\sum_{k=1}^p \alpha_k = 1$.

Théorème

Soient $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$ un système de points pondérés dans l'espace affine $\mathcal{E}$ avec $\sum_{k=1}^p \alpha_k \neq 0$ et $\{1, \dots, p\} = \bigcup_{j=1}^r I_j$ une partition de $\{1, \dots, p\}$ telle que $\sigma_j = \sum_{k \in I_j} \alpha_k \neq 0$ pour tout j compris entre 1 et r . En désignant, pour tout j compris entre 1 et r , par G_j le barycentre de $\{(M_k, \alpha_k)\}_{k \in I_j}$, on a :

$$\text{Bar} \left\{ \{(M_k, \alpha_k)\}_{1 \leq k \leq p} \right\} = \text{Bar} \left\{ \{(G_j, \sigma_j)\}_{1 \leq j \leq r} \right\}$$

Théorème

Une partie non vide $\mathcal{F}$ de $\mathcal{E}$ est un sous-espace affine si, et seulement si, pour toute famille $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$ de points pondérés de $\mathcal{F}$, le barycentre $\text{Bar} \left\{ \{(M_k, \alpha_k)\}_{1 \leq k \leq p} \right\}$ est dans $\mathcal{F}$.

Théorème

Soient $\mathcal{E}'$ un espace affine de dimension finie et f une application de $\mathcal{E}$ dans $\mathcal{E}'$. L'application f est affine si, et seulement si, pour toute famille $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$ de points pondérés de $\mathcal{E}$, le barycentre de $\{(f(M_k), \alpha_k)\}_{1 \leq k \leq p}$ est $f(G)$, où G est le barycentre de $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$.

On se donne un repère affine $\mathcal{R} = (A_k)_{0 \leq k \leq n}$ de $\mathcal{E}$ c'est-à-dire $(\overrightarrow{A_0 A_k})_{1 \leq k \leq n}$ est une base de E .

Théorème

Pour tout $M \in \mathcal{E}$, il existe un unique $(n+1)$ -uplet $(\lambda_0, \dots, \lambda_n)$ dans $\mathbb{R}^{n+1}$ tel que $M = \text{Bar} \left\{ (A_k, \lambda_k) \right\}_{0 \leq k \leq n}$ et $\sum_{k=0}^n \lambda_k = 1$.

Théorème

Avec les notations du théorème précédent, on appelle *coordonnées barycentriques* de M dans le repère $\mathcal{R}$, tout n -uplet $(\gamma \lambda_1, \dots, \gamma \lambda_n)$, où γ est un réel non nul.

L'application $M \mapsto (\lambda_0, \dots, \lambda_n)$ est une bijection de $\mathcal{E}$ sur $\mathbb{R}^{n+1}$.

Soient ABC trois points non alignés du plan affine euclidien tels que le triangle $\mathcal{T} = ABC$ soit non rectangle. En notant $a = BC$, $b = AC$, $c = AB$, $\widehat{A} = \widehat{BAC}$, $\widehat{B} = \widehat{ABC}$ et $\widehat{C} = \widehat{ACB}$, les coordonnées barycentriques dans le repère affine (A, B, C) du centre de gravité G , de l'orthocentre H , du centre I du cercle inscrit, du centre O du cercle circonscrit et des centres I_A, I_B, I_C des cercles exinscrits sont respectivement :

- $G(1, 1, 1)$; $H \left(\tan(\widehat{A}), \tan(\widehat{B}), \tan(\widehat{C}) \right)$; $I(a, b, c)$;
- $O \left(\tan(\widehat{B}) + \tan(\widehat{C}), \tan(\widehat{A}) + \tan(\widehat{C}), \tan(\widehat{A}) + \tan(\widehat{B}) \right)$;
- $I_A(-a, b, c)$, $I_B(a, -b, c)$, $I_C(a, b, -c)$.

Définition

On dit qu'une partie non vide C de $\mathcal{E}$ est *convexe*, si pour tout couple (A, B) de points de $\mathcal{E}$, le segment $[A, B]$ est contenu dans $\mathcal{E}$.

Théorème

Soit C une partie non vide de $\mathcal{E}$. Cet ensemble C est convexe si, et seulement si, pour toute famille $\{(M_k, \alpha_k)\}_{1 \leq k \leq p}$ de points pondérés de $\mathcal{E}$, où les α sont des réels positifs ou nuls tels que $\sum_{k=1}^p \alpha_k = 1$, le barycentre $\text{Bar} \left\{ \{(A_k, \alpha_k)\}_{0 \leq k \leq n} \right\}$ est dans C .

Une intersection de parties convexes est convexe. L'adhérence et l'intérieur d'un convexe sont convexes.

Théorème

Soient $\mathcal{E}'$ un espace affine de dimension finie et $f : \mathcal{E} \rightarrow \mathcal{E}'$ affine. Pour tout convexe C dans $\mathcal{E}$ [resp. C' dans $\mathcal{E}'$] $f(C)$ [resp. $f^{-1}(C')$] est un convexe de $\mathcal{E}'$ [resp. de $\mathcal{E}$].

PROBABILITÉS

par Jean-Étienne ROMBALDI

25.1 Dénombrément

Ω est un ensemble non vide et $\mathcal{P}(\Omega)$ est l'ensemble de toutes les parties de Ω . Pour tout ensemble non vide F , on note F^Ω l'ensemble des applications de Ω dans F . Le *cardinal* d'un ensemble fini E est le nombre de ses éléments. Il est noté $\text{Card}(E)$. Deux ensembles finis ont le même cardinal si, et seulement si, ils sont en bijection.

Définition

Deux ensembles E et F sont *équipotents* s'il existe une bijection de E sur F .

Par exemple l'ensemble des entiers naturels pairs est équipotent à $\mathbb{N}$ de même que $\mathbb{N}^2$.

La relation d'équipotence est une relation d'équivalence.

Définition

Un ensemble est *dénombrable* s'il est équipotent à $\mathbb{N}$.

Un ensemble est *au plus dénombrable* s'il est fini ou dénombrable.

Par exemple, $\mathbb{Z}$ et $\mathbb{Q}$ sont dénombrables. $\mathbb{R}$ et l'ensemble des parties de $\mathbb{N}$ ne sont pas dénombrables.

Les termes « énumérable » ou « numérotable » seraient peut-être plus compréhensibles que « dénombrable » pour les néophytes car ce dernier pourrait laisser entendre qu'un ensemble dénombrable est de cardinal fini.

Un sous-ensemble d'un ensemble dénombrable, une réunion finie ou infinie dénombrable d'ensembles dénombrables sont dénombrables.

Théorème

Soient $E, F, E_1, \dots, E_n$ des ensembles finis avec $n \geq 2$.

- Si les E_k sont 2 à 2 disjoints, on a $\text{Card}\left(\bigcup_{k=1}^n E_k\right) = \sum_{k=1}^n \text{Card}(E_k)$.
- Pour tout $A \in \mathcal{P}(E)$, on a $\text{Card}(E \setminus A) = \text{Card}(E) - \text{Card}(A)$ et $\text{Card}(A) \leq \text{Card}(E)$, l'égalité étant réalisée si, et seulement si, $A = E$.
- On a $\text{Card}(E \cup F) = \text{Card}(E) + \text{Card}(F) - \text{Card}(E \cap F)$.
- On a $\text{Card}\left(\prod_{k=1}^n E_k\right) = \prod_{k=1}^n \text{Card}(E_k)$.

Théorème

Soient E, F deux ensembles finis non vides de cardinaux respectifs p, n et φ dans F^E .

- On a $\text{Card}(F^E) = (\text{Card}(F))^{\text{Card}(E)}$.
- Si φ est injective, on a alors $\text{Card}(E) \leq \text{Card}(F)$.
- Pour $n \geq p$, le nombre d'injections de E dans F est $\frac{n!}{(n-p)!}$.
- Si φ est surjective, on a alors $\text{Card}(E) \geq \text{Card}(F)$.
- Si φ est bijective, on a alors $\text{Card}(E) = \text{Card}(F)$.
- Pour $n = p$, le nombre d'applications bijectives de E dans F est $n!$.
- On a $\text{Card}(\varphi(E)) \leq \min(\text{Card}(E), \text{Card}(F))$ et φ est injective si, et seulement si, on a $\text{Card}(\varphi(E)) = \text{Card}(E)$, φ est surjective si, et seulement si, on a $\text{Card}(\varphi(E)) = \text{Card}(F)$.
- Pour $n = p$, on a :

$$(\varphi \text{ injective}) \Leftrightarrow (\varphi \text{ surjective}) \Leftrightarrow (\varphi \text{ bijective}).$$
- S'il existe $r \in \mathbb{N}^*$ tel que $\text{Card}(\varphi^{-1}\{y\}) = r$ pour tout $y \in F$, φ est alors surjective et on a $\text{Card}(E) = r \text{Card}(F)$ (principe des bergers).

Théorème de Poincaré

Si $(A_k)_{1 \leq k \leq n}$ est une suite de parties finies d'un ensemble Ω , on a alors :

$$\text{Card}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n (-1)^{k-1} p_{k,n}$$

où $p_{k,n} = \sum_{1 \leq i_1 < \dots < i_k \leq n} \text{Card}(A_{i_1} \cap \dots \cap A_{i_k})$ pour $1 \leq k \leq n$.

En probabilités élémentaires, on peut être conduit à considérer des tirages de p objets pris dans un ensemble F à n éléments. Il y a trois façons de procéder.

- On peut effectuer des tirages avec remise et avec ordre, ce qui signifie que l'on extrait les objets un à un en remettant, après chaque tirage, l'objet dans l'ensemble F . Un tel tirage est une p -liste (ou un p -uplet) et revient à se donner une application de $E = \{1, \dots, p\}$ dans F , ce qui donne n^p possibilités.
- On peut effectuer des tirages sans remise et avec ordre, ce qui signifie que l'on extrait les objets un à un sans les remettre dans l'ensemble F . Un tel tirage est une p -liste sans répétitions et revient à se donner une application injective de $E = \{1, \dots, p\}$ dans F , ce qui donne 0 possibilités si $p > n$ et $\frac{n!}{(n-p)!}$ possibilités si $p \leq n$.
- On peut extraire les p objets simultanément.

Un tel tirage est une p -combinaison c'est-à-dire une partie de F à p éléments. Ce nombre de p -combinaisons vaut $\binom{n}{p} = 0$ si $p > n$ ou $\binom{n}{p} = \frac{n!}{p!(n-p)!}$ si $p \leq n$. Cette quantité est aussi le nombre de parties à p éléments d'un ensemble à n éléments.

Les relations suivantes sur les coefficients binomiaux $\binom{n}{p}$ sont à connaître :

- ▷ $\binom{n}{p} = \binom{n}{n-p}$ pour $0 \leq p \leq n$; $\binom{n}{n} = \binom{n}{0} = 1$ pour $n \geq 0$;
- ▷ $\binom{n}{p} = \frac{n}{p} \binom{n-1}{p-1}$ pour $1 \leq p \leq n-1$;
- ▷ $\binom{n}{p} = \binom{n-1}{p} + \binom{n-1}{p-1}$ pour $1 \leq p \leq n-1$ (triangle de Pascal) ;
- ▷ $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$ pour $n \in \mathbb{N}$ et a et b dans un anneau commutatif $\mathbb{A}$ (formule du binôme de Newton) ;
- ▷ $\sum_{k=0}^n \binom{n}{k} = 2^n$ et $\sum_{k=0}^n \binom{n}{k} (-1)^k = 0$, ces égalités entraînant que, pour $n \in \mathbb{N}^*$
on a $\sum_{0 \leq 2j \leq n} \binom{n}{2j} = \sum_{0 \leq 2j+1 \leq n} \binom{n}{2j+1} = 2^{n-1}$.

25.2 Fonctions indicatrices d'ensembles

À toute partie A de Ω , on associe sa *fonction indicatrice* $\mathbf{1}_A : \Omega \rightarrow \{0, 1\}$ définie par $\mathbf{1}_A(x) = 1$ si $x \in A$ et $\mathbf{1}_A(x) = 0$ sinon.

Les fonctions indicatrices permettent de transformer des opérations ensemblistes ou probabilistes en opérations algébriques sur des fonctions.

Théorème

L'application $\mathbf{1} : A \in \mathcal{P}(\Omega) \rightarrow \mathbf{1}_A \in \{0, 1\}^\Omega$ est une bijection d'inverse $\mathbf{1}^{-1} : \gamma \in \{0, 1\}^\Omega \mapsto \gamma^{-1}\{1\} \in \mathcal{P}(\Omega)$.

Corollaire

Pour Ω fini, on a $\text{Card}(\mathcal{P}(\Omega)) = 2^{\text{Card}(\Omega)}$.

Théorème de Cantor

Il n'existe pas de bijection de Ω sur $\mathcal{P}(\Omega)$.

On déduit en particulier du théorème précédent que $\mathcal{P}(\mathbb{N})$ et $\{0, 1\}^{\mathbb{N}}$ ne sont pas dénombrables et en utilisant l'écriture dyadique propre (en base 2) des nombres réels, il en résulte que $[0, 1]$ n'est pas dénombrable.

25.3 Expérience aléatoire

De façon intuitive, une *expérience aléatoire* est une expérience dont le résultat ne peut être annoncé avec certitude *a priori*, c'est-à-dire avant la réalisation de cette expérience. Un *événement* est une proposition logique susceptible de se produire ou non à l'issue de la réalisation d'une expérience aléatoire. Enfin à chaque événement on essaye d'associer un nombre $p \in [0, 1]$ qui devra mesurer la probabilité pour que cet événement se produise.

Ω est un ensemble non vide dont les éléments représentent tous les résultats possibles d'une expérience aléatoire donnée $\mathcal{E}$. Un tel ensemble Ω est appelé *univers* des possibles ou espace fondamental et ses éléments sont appelés *éventualités*, *épreuves* ou *événements élémentaires*.

On dit qu'une expérience aléatoire est *discrète*, si l'univers Ω est dénombrable (fini ou infini) ou qu'elle est *continue*, s'il est infini non dénombrable.

25.4 Tribus d'événements, espaces probabilisables

On se donne un univers Ω , on note $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω et pour tout $A \in \mathcal{P}(\Omega)$, $\Omega \setminus A$ désigne le complémentaire de A dans Ω .

Définition

On appelle *tribu* (ou σ -*algèbre* d'événements, ou famille d'événements observables) sur Ω toute partie $\mathcal{B}$ de $\mathcal{P}(\Omega)$ telle que :

- $\emptyset \in \mathcal{B}$;
- $\mathcal{B}$ est stable par passage au complémentaire, ce qui signifie que :
$$\forall A \in \mathcal{B}, \Omega \setminus A \in \mathcal{B}.$$
- $\mathcal{B}$ est stable par réunions dénombrables, ce qui signifie que pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{B}$, la réunion $\bigcup_{n \in \mathbb{N}} A_n$ est dans $\mathcal{B}$.

Définition

Si $\mathcal{B}$ est une tribu sur Ω , on dit alors que le couple $(\Omega, \mathcal{B})$ est un *espace probabilisable* (ou mesurable).

Les éléments de $\mathcal{B}$ sont appelés *événements*. Les opérations sur les ensembles s'interprètent, en termes d'événements, comme indiqué dans le tableau suivant :

Ensemble	Événement
un élément de $\mathcal{B}$	un événement
l'ensemble vide $\emptyset$	l'événement impossible
l'univers Ω	l'événement certain
un singleton $\{\omega\}$ appartenant à $\mathcal{B}$	un événement élémentaire
$A \subset B$ pour A et B dans $\mathcal{B}$	si A est réalisé, alors B est réalisé
le complémentaire $\Omega \setminus A$ pour $A \in \mathcal{B}$	l'événement contraire de A
$A \cap B$ pour A et B dans $\mathcal{B}$	réalisation simultanée de A et B
$A \cap B = \emptyset$ pour A et B dans $\mathcal{B}$	A et B sont incompatibles
$A \cup B$ pour A et B dans $\mathcal{B}$	réalisation de A ou B
$(A_i)_{i \in I}$ partition de Ω avec les A_i dans $\mathcal{B}$	$(A_i)_{i \in I}$ système complet d'événements

$\mathcal{B} = \{\emptyset, \Omega\}$ est la *tribu triviale*, c'est la plus petite ; $\mathcal{B} = \mathcal{P}(\Omega)$ est la *tribu grossière*, c'est la plus grande ; $\mathcal{B} = \{\emptyset, A, \Omega \setminus A, \Omega\}$ est la *tribu de Bernoulli* associée à A , où A est une partie non vide de Ω distincte de Ω .

Théorème

Si $(\Omega, \mathcal{B})$ est un espace probablisable, on a alors :

- $\Omega \in \mathcal{B}$;
- si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments de $\mathcal{B}$ alors $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{B}$ ($\mathcal{B}$ est stable par intersection dénombrable) ;
- si A, B sont dans $\mathcal{B}$, alors $A \cup B$, $A \cap B$, $A \setminus B$ et $A \Delta B$ sont dans $\mathcal{B}$.

Théorème

Une intersection de tribus sur Ω est une tribu sur Ω , ce qui signifie que si $(\mathcal{B}_i)_{i \in I}$ est une famille de tribus sur Ω , alors $\mathcal{B} = \bigcap_{i \in I} \mathcal{B}_i$ est aussi

une tribu sur Ω . Pour toute partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$, il existe une plus petite tribu sur Ω (pour l'ordre de l'inclusion) qui contient $\mathcal{A}$, à savoir la tribu $\sigma(\mathcal{A}) = \bigcap_{\substack{\mathcal{B} \text{ tribu sur } \Omega \\ \mathcal{A} \subset \mathcal{B}}} \mathcal{B}$.

Définition

Si $\mathcal{A}$ est une partie de $\mathcal{P}(\Omega)$, on dit alors que l'intersection $\sigma(\mathcal{A})$ de toutes les tribus sur Ω qui contiennent $\mathcal{A}$ est la *tribu engendrée* par $\mathcal{A}$.

Définition

La *tribu borélienne* sur $\mathbb{R}$ est la tribu sur $\mathbb{R}$ engendrée par l'ensemble des intervalles réels (finis ou non). On la note $\mathcal{B}_{\mathbb{R}}$.

La tribu borélienne sur $\mathbb{R}$ est aussi la tribu engendrée par la famille des ouverts de $\mathbb{R}$. Pour tout intervalle réel I non réduit à un point, la tribu borélienne $\mathcal{B}_I$ sur I est la tribu engendrée par l'ensemble des intervalles réels contenus dans I . On peut vérifier que $\mathcal{B}_I = \{I \cap J, J \in \mathcal{B}_{\mathbb{R}}\}$.

Théorème

La tribu borélienne sur $\mathbb{R}$ est engendrée par la famille des intervalles de la forme $]-\infty, x]$ où x décrit $\mathbb{R}$. Elle est aussi engendrée par la famille des intervalles de la forme $]x, y]$ où $x < y$ dans $\mathbb{R}$. Cela signifie que :

$$\mathcal{B}_{\mathbb{R}} = \sigma\{]-\infty, x] ; x \in \mathbb{R}\} = \sigma\{]x, y] ; x < y \text{ dans } \mathbb{R}\}.$$

25.5 Espaces probabilisés

Définition

On appelle *probabilité* (ou *mesure de probabilité*) sur un espace probabilisable $(\Omega, \mathcal{B})$ toute application $\mathbb{P} : \mathcal{B} \rightarrow [0, 1]$ telle que :

- $\mathbb{P}(\Omega) = 1$;
- pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux incompatibles dans $\mathcal{B}$, on a $\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ (σ -additivité de $\mathbb{P}$).

Avec ces conditions, on dit que le triplet $(\Omega, \mathcal{B}, \mathbb{P})$ est un *espace probabilisé*.

Dans ce qui suit, $(\Omega, \mathcal{B}, \mathbb{P})$ est un espace probabilisé.

Théorème

On a les propriétés suivantes, où A, B, A_k désignent des événements :

- $\mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$; $\mathbb{P}(\emptyset) = 0$;
- si $A \subset B$, on alors $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$ et $\mathbb{P}(A) \leq \mathbb{P}(B)$;
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- pour toute suite croissante $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{B}$ (c'est-à-dire telle que $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$), la suite $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$ est convergente et on a $\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n)$;
- pour toute suite décroissante $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{B}$ (c'est-à-dire telle que $A_{n+1} \subset A_n$ pour tout $n \in \mathbb{N}$), la suite $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$ est convergente et on a $\mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n)$;
- pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements, on a

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mathbb{P}(A_n) \quad (\text{inégalité de Boole})$$

où $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$ peut valoir $+\infty$.

Corollaire

Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{B}$, on a :

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcup_{k=0}^n A_k\right) \text{ et } \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcap_{k=0}^n A_k\right).$$

Théorème de Poincaré

Pour toute suite $(A_k)_{1 \leq k \leq n}$ d'événements dans $\mathcal{B}$, on a :

$$\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n (-1)^{k-1} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) \right).$$

Définition

On appelle *système complet d'événements* toute suite au plus dénombrable $(A_i)_{i \in I}$ d'éléments de $\mathcal{B}$ formant une partition de Ω .

Théorème

Si $(A_i)_{i \in I}$ est un système complet d'événements, on a alors $\sum_{i \in I} \mathbb{P}(A_i) = 1$ et pour tout événement $A \in \mathcal{B}$, on a $\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \cap A_i)$.

Définition

On dit qu'une propriété $\mathcal{P}$ dans $(\Omega, \mathcal{B}, \mathbb{P})$ est *presque sûre* (ou est vérifiée *presque sûrement*) si $^1 \mathbb{P}(A) = 1$ où $A = \{\omega \in \Omega, \omega \text{ vérifie } \mathcal{P}\}$.

25.6 Espaces probabilisés discrets

Pour ce paragraphe, l'ensemble $\Omega = \{\omega_i ; i \in I\}$ est dénombrable avec $I \subset \mathbb{N}$ et on utilise $\mathcal{B} = \mathcal{P}(\Omega)$ pour tribu d'événements.

1. Par exemple, avec les notions du chapitre suivant, dire que deux variables aléatoires réelles discrètes sont *presque sûrement* égales signifie $\mathbb{P}(X = Y) = 1$ ou encore que $\mathbb{P}(X \neq Y) = 0$ ou que $\mathbb{P}(|Y - X| > 0) = 0$. De même, dire qu'une variable aléatoire réelle discrète X est *presque sûrement* nulle signifie $\mathbb{P}(X = 0) = 1$.

Théorème

Une application $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0, 1]$ telle que $\sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 1$ est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ si, et seulement si, on a :

$$\forall A \in \mathcal{P}(\Omega), \mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

Pour $\Omega = \{\omega_1, \dots, \omega_n\}$ fini, il existe une unique probabilité prenant la même valeur sur tous les événements élémentaires, c'est la *probabilité uniforme* définie par :

$$\forall \omega \in \Omega, \mathbb{P}(\{\omega\}) = \frac{1}{n} = \frac{1}{\text{Card}(\Omega)}$$

ce qui signifie que pour toute partie A de Ω , on a $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$

(c'est le quotient du nombre de cas favorables par le nombre de cas possibles). Dans cette situation, on dit qu'on est dans le cadre de l'équiprobabilité et le calcul des probabilités est ramené à des calculs de dénombrement.

Pour tout $n \in \mathbb{N}^*$ et tout réel $p \in]0, 1[$ et $\Omega = \{0, \dots, n\}$ fini, on définit une unique probabilité sur $(\Omega, \mathcal{P}(\Omega))$ avec :

$$\forall k \in \{0, 1, \dots, n\}, \mathbb{P}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}$$

(loi binomiale de paramètres $p \in]0, 1[$ et $n \in \mathbb{N}^*$).

Pour tout réel $\lambda > 0$ et $\Omega = \mathbb{N}$, on définit une unique probabilité sur $(\Omega, \mathcal{P}(\Omega))$ avec :

$$\forall k \in \mathbb{N}, \mathbb{P}(\{k\}) = e^{-\lambda} \frac{\lambda^k}{k!}$$

(loi de Poisson de paramètre $\lambda > 0$).

Pour $\Omega = \mathbb{N}^*$ et tout réel $p \in]0, 1[$, on définit une unique probabilité sur $(\Omega, \mathcal{P}(\Omega))$ avec :

$$\forall k \in \mathbb{N}^*, \mathbb{P}(\{k\}) = p(1-p)^{k-1}$$

(loi géométrique de paramètre $p \in]0, 1[$).

25.7 Probabilités conditionnelles

Définition

Soient A, B deux événements dans $\mathcal{B}$ avec $\mathbb{P}(B) \neq 0$. La probabilité de A sachant B est le réel noté $\mathbb{P}(A | B)$ ou $\mathbb{P}_B(A)$ défini par

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Théorème

Pour tout événement $B \in \mathcal{B}$ de probabilité non nulle, l'application :

$$\begin{aligned} \mathbb{P}_B : \mathcal{B} &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \end{aligned}$$

est une probabilité sur $(\Omega, \mathcal{B})$.

On dit que $\mathbb{P}_B$ est la probabilité conditionnelle sur $(\Omega, \mathcal{B})$ sachant B et par définition, on a $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B)$.

Théorème (formule des probabilités composées)

Si $A_1, \dots, A_n$ sont des événements dans $\mathcal{B}$ tels que $\mathbb{P}\left(\bigcap_{k=1}^{n-1} A_k\right) \neq 0$, on a

alors $\mathbb{P}\left(\bigcap_{k=1}^p A_k\right) \neq 0$ pour tout p compris entre 1 et $n-1$ et :

$$\mathbb{P}\left(\bigcap_{k=1}^n A_k\right) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}\left(A_n \mid \bigcap_{k=1}^{n-1} A_k\right).$$

Théorème (formule des probabilités totales)

Soit $(A_i)_{i \in I}$ un système complet d'événements dans $\mathcal{B}$, où I est une partie de $\mathbb{N}$, tel que $\mathbb{P}(A_i) \in]0, 1[$ pour tout $i \in I$.

Pour tout événement $A \in \mathcal{B}$, on a $\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \mid A_i)\mathbb{P}(A_i)$.

Théorème (formules de Bayes)

Soit $(A_i)_{i \in I}$ un système complet d'événements dans $\mathcal{B}$, où I est une partie de $\mathbb{N}$, tel que $\mathbb{P}(A_i) \neq 0$ pour tout $i \in I$. Pour tout événement $A \in \mathcal{B}$ de

probabilité non nulle et tout $j \in I$, on a $\mathbb{P}(A_j \mid A) = \frac{\mathbb{P}(A \mid A_j)\mathbb{P}(A_j)}{\sum_{i \in I} \mathbb{P}(A \mid A_i)\mathbb{P}(A_i)}$.

25.8 Événements indépendants

Définition

On dit que deux événements A et B dans $\mathcal{B}$ sont *indépendants*, si on a $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Définition

Soit $(A_i)_{i \in I}$ une famille d'éléments de $\mathcal{B}$ ayant au moins deux éléments. On dit que ces événements sont *mutuellement indépendants*, si pour toute partie finie J de I , on a $\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j)$.

Théorème

Soit $(A_i)_{i \in \mathbb{N}}$ une famille d'événements dans $\mathcal{B}$. Ces événements sont mutuellement indépendants si, et seulement si, pour tout $J \subsetneq I$ fini tel que $\mathbb{P}\left(\bigcap_{j \in J} A_j\right) \neq 0$ et tout indice $i \in I \setminus J$, on a $\mathbb{P}\left(A_i \mid \bigcap_{j \in J} A_j\right) = \mathbb{P}(A_i)$.

Théorème

Soit $(A_n)_{n \in \mathbb{N}}$ une famille formée d'événements mutuellement indépendants dans $\mathcal{B}$. Pour tout $n \in \mathbb{N}$, les événements $\Omega \setminus A_0, \dots, \Omega \setminus A_n, \Omega \setminus A_{n+1}, \dots$ sont mutuellement indépendants.

Corollaire

Si $(A_k)_{1 \leq k \leq n}$ est une famille formée de $n \geq 2$ événements mutuellement indépendants dans $\mathcal{B}$, on a alors $\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = 1 - \prod_{k=1}^n (1 - \mathbb{P}(A_k))$.

26

Variables aléatoires réelles discrètes

$(\Omega, \mathcal{B}, \mathbb{P})$ est un espace probabilisé.

Définition

On dit qu'une application $X : \Omega \rightarrow \mathbb{R}$ est une *variable aléatoire réelle discrète* sur $(\Omega, \mathcal{B}, \mathbb{P})$ si $X(\Omega)$ est dénombrable et si $X^{-1}(\{x\}) \in \mathcal{B}$ pour tout réel x . Dans le cas où $X(\Omega)$ est fini, on dit que X est *étagée*.

Dans le cas où Ω est dénombrable, en prenant $\mathcal{B} = \mathcal{P}(\Omega)$ pour tribu d'événements, toute application $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire discrète.

On notera :

- $(X \in A)$ l'événement $X^{-1}(A)$;
- $(X = x)$ l'événement $X^{-1}(\{x\})$ où x est un réel ;
- $(X \leq x)$ l'événement $X^{-1}([-\infty, x])$ où x est un réel ;
- $(X < x)$ l'événement $X^{-1}([-\infty, x[)$ où x est un réel ;
- $(x < X \leq y)$ l'événement $X^{-1}([x, y])$ où x et y sont des réels.

Théorème

Si X est une variable aléatoire réelle discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$, l'application :

$$\begin{aligned} \mathbb{P}_X : \mathcal{P}(\mathbb{R}) &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}_X(A) = \mathbb{P}(X \in A) \end{aligned}$$

définit alors une probabilité sur $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$.

Définition

Soit X une variable aléatoire réelle discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$. La mesure de probabilité $\mathbb{P}_X$ définie par le théorème précédent est la *loi* de X .

La loi de X est donc uniquement déterminée par la donnée des probabilités élémentaires $p_i = \mathbb{P}(X = x_i)$ vérifiant l'égalité $\sum_{i \in I} p_i = 1$.

Théorème

Pour toute variable aléatoire discrète X , on a $\lim_{x \rightarrow +\infty} \mathbb{P}^{-1}(|X| > x) = 0$.

- Pour $n \in \mathbb{N}^*$, on dit qu'une variable aléatoire X suit une *loi uniforme discrète* si $X(\Omega) = I_n = \{1, \dots, n\}$ et $\mathbb{P}(X = k) = \frac{1}{n}$ pour tout $k \in I_n$. On note¹ $X \hookrightarrow \mathcal{U}(\{1, \dots, n\})$. Cette loi permet de décrire les expériences aléatoires où il y a équiprobabilité.
- On dit qu'une variable aléatoire X suit une *loi de Bernoulli* de paramètre $p \in]0, 1[$ si $X(\Omega) = \{0, 1\}$ et $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = 1) = p$. On note $X \hookrightarrow \mathcal{B}(p)$. Cette loi permet de décrire les expériences aléatoires à deux issues possibles (par exemple, 1 pour pile, 0 pour face).
- On dit qu'une variable aléatoire X suit une *loi binomiale* de paramètres $p \in]0, 1[$ et $n \in \mathbb{N}^*$, si $X(\Omega) = \{0, \dots, n\}$ et $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ pour tout $k \in \{0, 1, \dots, n\}$. Cette loi permet de modéliser les expériences aléatoires où l'on répète, de façons indépendantes, n épreuves de Bernoulli de même paramètre p et on s'intéresse au nombre X de succès.
- On dit qu'une variable aléatoire X suit une *loi de Poisson* de paramètre $\lambda \in \mathbb{R}_+^*$, si $X(\Omega) = \mathbb{N}$ et $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ pour tout $k \in \mathbb{N}$. On note $X \hookrightarrow \mathcal{P}(\lambda)$. Les lois de Poisson permettent de modéliser des situations où on s'intéresse au nombre d'événements d'un certain type se produisant dans un intervalle de temps donné.
- Pour $p \in]0, 1]$, on dit qu'une variable aléatoire X suit une *loi géométrique* de paramètre p , si $X(\Omega) = \mathbb{N}^*$ et $\mathbb{P}(X = k) = p(1 - p)^{k-1}$ pour tout k dans $\mathbb{N}^*$. On note $X \hookrightarrow \mathcal{G}(p)$. Cette loi permet de modéliser les expériences aléatoires où l'on répète, de façons indépendantes, une série infinie d'épreuves de Bernoulli de paramètre p et on s'intéresse au rang d'obtention du premier succès.

1. ou encore $X \sim \mathcal{U}(\{1, \dots, n\})$. Lorsqu'une variable aléatoire suit une certaine loi, la notation $\sim$ tend désormais à se répandre dans la littérature scientifique.

- Pour $M < N$ et $n < N$ dans $\mathbb{N}^*$, on dit qu'une variable aléatoire X suit une *loi hypergéométrique* de paramètres n , M et N , si $X(\Omega) = \mathbb{N}$ et

$$\mathbb{P}(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$$

pour tout $k \in \llbracket 0, \min(M, n) \rrbracket$. On note $X \hookrightarrow \mathcal{H}(n, M, N)$.

Théorème

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires discrètes telle que, pour tout entier $n \in \mathbb{N}^*$, la variable aléatoire X_n suit une loi binomiale de paramètres n et $p_n \in]0, 1[$.

S'il existe un réel $\lambda > 0$ tel que $\lim_{n \rightarrow +\infty} np_n = \lambda$, on a alors pour tout $k \in \mathbb{N}$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

ce qui s'exprime aussi en disant que la suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers la loi de Poisson de paramètre λ .

Théorème

Soient $p \in]0, 1[$, $n \in \mathbb{N}^*$ et $(X_N)_{N > n}$ une suite de variables aléatoires discrètes telle que pour tout $N > n$, $X_N \hookrightarrow \mathcal{H}(n, \lfloor pN \rfloor, N)$.

Dans ce cas, pour tout $k \in \mathbb{N}$, on a

$$\lim_{N \rightarrow +\infty} \mathbb{P}(X_N = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

(avec la convention $\binom{n}{k} = 0$ pour $k < 0$ ou $k > n$), ce qui s'exprime en disant que $(X_N)_{N > n}$ converge en loi vers la loi binomiale de paramètres n et p .

Définition

Soit X une variable aléatoire sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{N}^*$ telle que $\mathbb{P}(X \geq n) > 0$ pour tout $n \in \mathbb{N}^*$. Le *taux de panne* de X est la suite $(\lambda_n)_{n \in \mathbb{N}^*} = (\mathbb{P}((X = n) \mid (X \geq n)))_{n \in \mathbb{N}^*}$.

Théorème

Soient X une variable aléatoire sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{N}^*$ telle que $\mathbb{P}(X \geq n) > 0$ pour tout $n \in \mathbb{N}^*$ et $(\lambda_n)_{n \in \mathbb{N}^*}$ son taux de panne. On a :

- $\lambda_n \in [0, 1[$ et $\mathbb{P}(X > n) = \prod_{k=1}^n (1 - \lambda_k)$ pour tout $n \in \mathbb{N}^*$;
- $\mathbb{P}(X = 1) = \lambda_1$ et $\mathbb{P}(X = n) = \lambda_n \prod_{k=1}^{n-1} (1 - \lambda_k)$ pour tout $n \geq 2$;
- $\sum_{n=1}^{+\infty} \lambda_n = +\infty$.

Théorème

Une suite réelle $(\lambda_n)_{n \in \mathbb{N}^*}$ à valeurs dans $[0, 1[$ telle que $\sum_{n=1}^{+\infty} \lambda_n = +\infty$ est le taux de panne d'une variable aléatoire à valeurs dans $\mathbb{N}^*$.

Définition

On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{N}^*$ est *sans mémoire* si² : $\forall (n, m) \in \mathbb{N}^* \times \mathbb{N}^*, \mathbb{P}(X > n + m) = \mathbb{P}(X > n) \mathbb{P}(X > m)$.

Théorème

Soit X une variable aléatoire sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{N}^*$ telle que $\mathbb{P}(X \geq n) > 0$ pour tout $n \in \mathbb{N}^*$. Les propriétés suivantes sont équivalentes :

- X suit une loi géométrique de paramètre $p \in]0, 1[$;
- le taux de panne de X est constant ;
- X est sans mémoire.

26.1 Opérations sur les variables aléatoires réelles discrètes

Théorème

Si X est une variable aléatoire discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$ et f une application de $\mathbb{R}$ dans $\mathbb{R}$, $f \circ X$ est alors une variable aléatoire réelle discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$.

2. ce qui revient à dire que $\forall (n, m) \in \mathbb{N}^* \times \mathbb{N}^*, \mathbb{P}(X > n + m \mid X > n) = \mathbb{P}(X > m)$.

On notera $f(X)$ la variable aléatoire $f \circ X$.

La loi de $Y = f(X)$ est déterminée par celle de X . Précisément pour tout sous ensemble A de $\mathbb{R}$, on a :

$$\mathbb{P}_Y(A) = \mathbb{P}((f \circ X)^{-1}(A)) = \mathbb{P}(X^{-1}(f^{-1}(A))) = \mathbb{P}_X(f^{-1}(A)).$$

En notant $X(\Omega) = \{x_i ; i \in I\}$ et $(f \circ X)(\Omega) = \{y_j ; j \in J\}$, on a pour tout $j \in J$:

$$q_j = \mathbb{P}(f(X) = y_j) = \sum_{i \in I, f(x_i) = y_j} \mathbb{P}(X = x_i).$$

Théorème

L'ensemble $\mathcal{V}(\Omega, \mathcal{B}, \mathbb{P})$ des variables aléatoires réelles discrètes définies sur $(\Omega, \mathcal{B}, \mathbb{P})$ est un $\mathbb{R}$ -espace vectoriel. De plus, le produit de deux éléments de $\mathcal{V}(\Omega, \mathcal{B}, \mathbb{P})$ est encore dans $\mathcal{V}(\Omega, \mathcal{B}, \mathbb{P})$.

Si f est une application de $\mathbb{R}$ dans $\mathbb{R}$ et X une variable aléatoire réelle discrète sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ telle que $X(\Omega) = \{x_i ; i \in I\}$, on a alors $f(X) = \sum_{i \in I} f(x_i) \mathbf{1}_{(X=x_i)}$.

Corollaire

Pour toute suite finie $(X_k)_{1 \leq k \leq n}$ de variables aléatoires réelles discrètes sur $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$, $\min_{1 \leq k \leq n} (X_k)$ et $\max_{1 \leq k \leq n} (X_k)$ sont des variables aléatoires réelles discrètes.

26.2 Fonction de répartition d'une variable aléatoire discrète

Définition

Soit X une variable aléatoire réelle discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$. Sa *fonction de répartition* est la fonction $F_X : x \in \mathbb{R} \mapsto \mathbb{P}(X \leq x)$.

En notant $X(\Omega) = \{x_i ; i \in I\}$ avec $I \subset \mathbb{N}$ et $p_i = \mathbb{P}(X = x_i)$ pour tout $i \in I$, on a $F_X(x) = \sum_{i \in I, x_i \leq x} p_i$ pour tout réel x . Dans le cas où X est étagée avec

$X(\Omega) = \{x_1, \dots, x_n\}$ et $x_1 < \dots < x_n$, on a :

$$F_X(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \sum_{i=1}^k p_i & \text{si } x \in [x_k, x_{k+1}[\\ 1 & \text{si } x \in [x_n, +\infty[\end{cases}$$

Théorème

Soit X une variable aléatoire réelle discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$ de fonction de répartition F_X .

- La fonction F_X est croissante.
- La fonction F_X admet une limite à gauche et à droite en tout point.
- $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
- La fonction F_X est continue à droite en tout point.
- La fonction F_X est continue en un point x si, et seulement si, on a $\mathbb{P}(X = x) = 0$.

26.3 Loi conjointe et lois marginales

Définition

Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{B}, \mathbb{P})$.

On appelle *loi conjointe* la loi $\mathbb{P}_{(X,Y)}$ du couple (X, Y) où

$$\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}((X \in A) \cap (Y \in B))$$

également notée $\mathbb{P}(X \in A, Y \in B)$.

On appelle *lois marginales* du couple (X, Y) les lois $\mathbb{P}_X$ et $\mathbb{P}_Y$ de X et Y .

Théorème

Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{B}, \mathbb{P})$. Alors, pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$, on a

$$\mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y) \quad \text{et} \quad \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x, Y = y).$$

26.4 Variables aléatoires discrètes indépendantes

Définition

On dit que deux variables aléatoires discrètes X et Y sur $(\Omega, \mathcal{B}, \mathbb{P})$ sont *indépendantes* si pour toutes parties A, B de $\mathbb{R}$ les événements $(X \in A)$ et $(Y \in B)$ sont indépendants, ce qui revient à dire que :

$$\mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A) \mathbb{P}(Y \in B).$$

Si X et Y sont indépendantes, on a alors :

$$\mathbb{P}((X \in A) \mid (Y \in B)) = \mathbb{P}(X \in A) = \mathbb{P}_X(A)$$

pour tout $B \subset \mathbb{R}$ tel que $\mathbb{P}_Y(B) = \mathbb{P}(Y \in B) > 0$ et toute partie A de $\mathbb{R}$, ce qui se traduit en disant que la loi de X sachant que $(Y \in B)$ est la loi de X .

Théorème

Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{B}, \mathbb{P})$ indépendantes, alors pour toutes fonctions f, g de $\mathbb{R}$ dans $\mathbb{R}$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Théorème

Soient X et Y deux variables aléatoires réelles discrètes sur $(\Omega, \mathcal{B}, \mathbb{P})$ avec $X(\Omega) = \{x_i ; i \in I\}$ et $Y(\Omega) = \{y_j ; j \in J\}$. Ces variables aléatoires sont indépendantes si, et seulement si, pour tout $(i, j) \in I \times J$ les événements $(X = x_i)$ et $(Y = y_j)$ sont indépendants, ce qui revient à dire que :

$$\forall (i, j) \in I \times J, \mathbb{P}((X = x_i) \cap (Y = y_j)) = \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j).$$

Définition

Soit $(X_i)_{i \in I}$ une famille de variables aléatoires réelles discrètes ayant au moins deux éléments. On dit que ces variables aléatoires sont *mutuellement indépendantes* (ou *indépendantes*) si pour toute partie finie J de I et toute famille $(A_j)_{j \in J}$ de parties de $\mathbb{R}$, la famille $((X_j \in A_j))_{j \in J}$ est formée d'événements indépendants, ce qui revient à dire que pour toute partie finie K de J et toute famille $(A_k)_{k \in K}$ de parties de $\mathbb{R}$, on a

$$\mathbb{P}\left(\bigcap_{k \in K} (X_k \in A_k)\right) = \prod_{k \in K} \mathbb{P}(X_k \in A_k).$$

Théorème

Soit $(X_i)_{i \in I}$ une famille de variables aléatoires réelles discrètes ayant au moins deux éléments. Ces variables aléatoires sont mutuellement indépendantes si, et seulement si, pour toute partie finie J de I et toute suite

$$(t_j)_{j \in J} \text{ de réels, on a } \mathbb{P}\left(\bigcap_{j \in J} (X_j = t_j)\right) = \prod_{j \in J} \mathbb{P}(X_j = t_j).$$

Théorème

Soit $(X_k)_{1 \leq k \leq n}$ une suite de n variables aléatoires discrètes sur $(\Omega, \mathcal{B}, \mathbb{P})$ indépendantes.

- Pour toute suite $(f_k)_{1 \leq k \leq n}$ de fonctions de $\mathbb{R}$ dans $\mathbb{R}$, les variables aléatoires $(f_k(X_k))_{1 \leq k \leq n}$ sont indépendantes.
- Pour tout entier naturel r compris entre 1 et $n-1$ et toutes fonctions $\varphi : \mathbb{R}^r \rightarrow \mathbb{R}$ et $\psi : \mathbb{R}^{n-r} \rightarrow \mathbb{R}$, les variables aléatoires $\varphi(X_1, \dots, X_r)$ et $\psi(X_{r+1}, \dots, X_n)$ sont indépendantes.

Théorème

Soient X, Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{N}$ de lois respectives $(p_n)_{n \in \mathbb{N}}$ et $(q_n)_{n \in \mathbb{N}}$. Dans le cas où ces variables aléatoires sont indépendantes, la loi de $Z = X + Y$ est définie pour tout $n \in \mathbb{N}$

$$\text{par } \mathbb{P}(Z = n) = \sum_{k=0}^n p_k q_{n-k}.$$

La suite $(\mathbb{P}(Z = n))_{n \in \mathbb{N}}$ est le *produit de convolution* des suites $(p_n)_{n \in \mathbb{N}}$ et $(q_n)_{n \in \mathbb{N}}$.

26.5 Espérance, variance, moments d'une variable aléatoire discrète

X est une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$. On note $X(\Omega) = \{x_i ; i \in I\}$ avec $I \subset \mathbb{N}$ et $p_i = \mathbb{P}(X = x_i)$ pour tout $i \in I$.

Définition

On dit que X admet une *espérance* si la série $\sum_{i \in I} p_i x_i$ est absolument convergente et dans ce cas, l'espérance de X est le réel $\mathbb{E}(X) = \sum_{i \in I} p_i x_i$.

Pour X à valeurs réelles positives, on peut toujours considérer $\mathbb{E}(X)$ comme un élément de $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$.

Une variable aléatoire discrète presque sûrement égale à une constante λ admet une espérance $\mathbb{E}(X) = \lambda$.

Si une variable aléatoire discrète X est à valeurs positives et d'espérance nulle, elle est alors presque sûrement nulle.

Une variable aléatoire étagée (c'est-à-dire telle que $X(\Omega)$ soit fini) admet une espérance. Une variable aléatoire bornée admet une espérance.

Pour $X \hookrightarrow \mathcal{U}(\{1, \dots, n\})$, on a $\mathbb{E}(X) = \frac{n+1}{2}$.

Pour $X \hookrightarrow \mathcal{B}(p)$, on a $\mathbb{E}(X) = p$.

Pour $X \hookrightarrow \mathcal{B}(n, p)$, on a $\mathbb{E}(X) = np$.

Pour $X \hookrightarrow \mathcal{G}(p)$, on a $\mathbb{E}(X) = \frac{1}{p}$.

Pour $X \hookrightarrow \mathcal{P}(\lambda)$, on a $\mathbb{E}(X) = \lambda$.

Théorème

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs réelles positives.

- L'application $t \mapsto \mathbb{P}(X > t)$ est intégrable sur tout segment $[a, b]$ (avec $a < b$).
 - Pour X étagée, on a $\mathbb{E}(X) = \int_0^{+\infty} \mathbb{P}(X > t) dt$.
 - Pour X à valeurs dans $\mathbb{N}$, les assertions suivantes sont équivalentes :
 - X admet une espérance ;
 - la série $\sum \mathbb{P}(X > n)$ est convergente ;
 - l'intégrale $\int_0^{+\infty} \mathbb{P}(X > t) dt$ est convergente ;
- et dans ce cas, on a $\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X > n) = \int_0^{+\infty} \mathbb{P}(X > t) dt$.

Théorème

L'ensemble $L^1(\Omega, \mathcal{B}, \mathbb{P})$ des variables aléatoires discrètes admettant une espérance sur $(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel de l'espace $\mathcal{V}(\Omega, \mathcal{B}, \mathbb{P})$ des variables aléatoires réelles discrètes sur $(\Omega, \mathcal{B}, \mathbb{P})$ et l'application $\mathbb{E} :$

$$\begin{cases} L^1(\Omega, \mathcal{B}, \mathbb{P}) & \rightarrow \mathbb{R} \\ X & \mapsto \mathbb{E}(X) \end{cases} \text{ est une forme linéaire positive (c'est-à-dire que}$$

si X et Y dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$ sont telles que $X \leq Y$, on a alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$).

Théorème (formule de transfert)

Soient X une variable aléatoire réelle discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$ et f une application de $\mathbb{R}$ dans $\mathbb{R}$. La variable aléatoire $f(X)$ admet une espérance si, et seulement si, la série $\sum_{i \in I} p_i f(x_i)$ converge absolument et dans ce cas, on a

$$\mathbb{E}(f(X)) = \sum_{i \in I} p_i f(x_i).$$

Théorème (inégalité de Jensen)

Si X est une variable aléatoire réelle discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$ admettant une espérance et f une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$ telle que $f(X)$ admette une espérance, on a alors $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$.

Théorème (inégalité de Markov)

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$ admettant une espérance. Pour tout réel $\varepsilon > 0$, on a $\mathbb{P}(|X| > \varepsilon) \leq \frac{\mathbb{E}(|X|)}{\varepsilon}$.

Corollaire

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$ admettant une espérance. Pour toute fonction $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ strictement croissante, pour tout réel $\varepsilon > 0$, on a

$$\mathbb{P}(|X| \geq \varepsilon) \leq \frac{\mathbb{E}(\varphi(|X|))}{\varphi(\varepsilon)}.$$

Théorème

Si $X_1, \dots, X_n$ sont $n \geq 2$ variables aléatoires discrètes indépendantes sur $(\Omega, \mathcal{B}, \mathbb{P})$ admettant une espérance, la variable aléatoire $X = \prod_{k=1}^n X_k$ admet alors une espérance et on a $\mathbb{E}(X) = \prod_{k=1}^n \mathbb{E}(X_k)$.

Définition

On dit qu'une variable aléatoire discrète X sur $(\Omega, \mathcal{B}, \mathbb{P})$ est *de carré sommable* si la variable aléatoire X^2 admet une espérance.

Théorème

L'ensemble $L^2(\Omega, \mathcal{B}, \mathbb{P})$ des variables aléatoires de carré sommable sur $(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel de l'espace $L^1(\Omega, \mathcal{B}, \mathbb{P})$. Pour X et Y dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ la variable aléatoire XY est dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$ et l'application :

$$\begin{aligned} \varphi : (L^2(\Omega, \mathcal{B}, \mathbb{P}))^2 &\rightarrow \mathbb{R} \\ (X, Y) &\mapsto \mathbb{E}(XY) \end{aligned}$$

est une forme bilinéaire, symétrique et positive sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$.

Définition

La *variance* et l'*écart type* d'une variable aléatoire $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$ sont les réels $\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$ et $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

La variance de $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$ est définie par :

$$\begin{aligned}\mathbb{V}(X) &= \sum_{i \in I} p_i (x_i - \mathbb{E}(X))^2 = \sum_{i \in I} p_i x_i^2 - 2\mathbb{E}(X) \sum_{i \in I} p_i x_i + (\mathbb{E}(X))^2 \sum_{i \in I} p_i \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (\text{formule de König})\end{aligned}$$

et pour tous réels a, b , on a $\mathbb{V}(aX + b) = \mathbb{E}(a^2(X - \mathbb{E}(X))^2) = a^2\mathbb{V}(X)$.

Théorème

Pour $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$, on a :

$$(\mathbb{V}(X) = 0) \Leftrightarrow (\mathbb{P}(X = \mathbb{E}(X)) = 1).$$

Si $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$ est de variance non nulle, la variable aléatoire centrée réduite associée à X est alors la variable aléatoire définie par $X' = \frac{X - \mathbb{E}(X)}{\sigma(X)}$.

On a $\mathbb{E}(X') = 0$ et $\mathbb{V}(X') = \mathbb{E}((X')^2) = 1$.

Pour $X \hookrightarrow \mathcal{U}(\{1, \dots, n\})$, on a $\mathbb{V}(X) = \frac{n^2-1}{12}$.

Pour $X \hookrightarrow \mathcal{B}(p)$, on a $\mathbb{V}(X) = p(1-p)$.

Pour $X \hookrightarrow \mathcal{B}(n, p)$ avec $n \geq 2$, on a $\mathbb{V}(X) = np(1-p)$.

Pour $X \hookrightarrow \mathcal{G}(p)$, on a $\mathbb{V}(X) = \frac{1-p}{p^2}$.

Pour $X \hookrightarrow \mathcal{P}(\lambda)$, on a $\mathbb{V}(X) = \lambda$.

Théorème (inégalité de Bienaymé-Tchebychev)

Soit $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$. Pour tout réel $\varepsilon > 0$, on a :

$$\mathbb{P}(|X - \mathbb{E}(X)| > \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Définition

La *covariance* de deux variables aléatoires X et Y dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ est le réel $\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$ et, pour X et Y de variance non nulle, le *coefficient de corrélation* de X et Y est $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$.

Définition

On dit que deux variables aléatoires X et Y dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ sont *non corrélées* si $\text{Cov}(X, Y) = 0$.

Théorème

L'application $\text{Cov} : \begin{cases} (L^2(\Omega, \mathcal{B}, \mathbb{P}))^2 & \rightarrow \mathbb{R} \\ (X, Y) & \mapsto \text{Cov}(X, Y) \end{cases}$ est une forme bilinéaire, symétrique et positive sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$.

Corollaire

Soient $X_1, \dots, X_n$ dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ et $\lambda_1, \dots, \lambda_n$ des réels. On a $\mathbb{V}\left(\sum_{k=1}^n \lambda_k X_k\right) = \sum_{k=1}^n \lambda_k^2 \mathbb{V}(X_k) + 2 \sum_{1 \leq j < k \leq n} \lambda_j \lambda_k \text{Cov}(X_j, X_k)$ et dans le cas particulier où les variables aléatoires X_k sont deux à deux non corrélées, on a $\mathbb{V}\left(\sum_{k=1}^n \lambda_k X_k\right) = \sum_{k=1}^n \lambda_k^2 \mathbb{V}(X_k)$.

Définition

Étant donné un entier $r \geq 1$, on dit qu'une variable aléatoire discrète X sur $(\Omega, \mathcal{B}, \mathbb{P})$ admet un *moment d'ordre r* , si la variable aléatoire X^r admet une espérance.

Si X admet un moment d'ordre r , on note alors $m_r(X) = \mathbb{E}(X^r) = \sum_{i \in I} p_i x_i^r$ et on dit que $m_r(X)$ est le moment d'ordre r de X .

Lemme

Soient $r < s$ dans $\mathbb{N}^*$ et X une variable aléatoire discrète sur $(\Omega, \mathcal{B}, \mathbb{P})$. Si X admet un moment d'ordre s , elle en admet alors un d'ordre r .

26.6 Fonction génératrice d'une variable aléatoire discrète

Pour ce paragraphe, X est une variable aléatoire sur un espace probablisé $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{N}$ et pour tout $n \in \mathbb{N}$, on note $p_n = \mathbb{P}(X = n)$.

Lemme

La série entière $\sum p_n t^n$ converge normalement sur $[-1, 1]$ et a un rayon de convergence $R_X \geq 1$.

Définition

Avec les notations qui précèdent, on dit que $\sum p_n t^n$ est la *série génératrice* de X .

On notera $g_X(t) = \sum_{n=0}^{+\infty} p_n t^n$ pour tout $t \in [-1, 1]$ et on dit que g_X est la *fonction génératrice* de X .

Théorème

La fonction génératrice g_X est de classe $\mathcal{C}^\infty$ sur $] -1, 1[$, de dérivées successives $g_X^{(p)}(t) = \sum_{n=p}^{+\infty} \frac{n!}{(n-p)!} p_n t^{n-p}$ et elle est croissante et convexe sur $[0, 1]$.

Elle est strictement croissante sur $[0, 1]$ si, et seulement si, on a $p_0 < 1$ et elle est strictement convexe sur $[0, 1]$ si, et seulement si, on a $p_0 + p_1 < 1$.

Pour tout $n \in \mathbb{N}$, on a $p_n = \frac{g_X^{(n)}(0)}{n!}$, ce qui se traduit en disant que la fonction génératrice de X caractérise sa loi (deux variables aléatoires qui ont même fonction génératrice sont de même loi).

Pour $X \hookrightarrow \mathcal{U}(\{1, \dots, n\})$, on a $g_X(t) = \frac{1}{n} \sum_{k=1}^n s^k = \frac{t(1-t^n)}{n(1-s)}$ et $R_X = +\infty$.

Pour $X \hookrightarrow \mathcal{B}(p)$, on a $g_X(t) = (1-p) + pt$ et $R_X = +\infty$.

Pour $X \hookrightarrow \mathcal{B}(n, p)$, on a $g_X(t) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} t^k = ((1-p) + pt)^n$ et $R_X = +\infty$.

Pour $X \hookrightarrow \mathcal{G}(p)$ avec $p \in]0, 1[$, on a $g_X(t) = p \sum_{n=1}^{+\infty} (1-p)^{n-1} t^n = \frac{pt}{1-(1-p)t}$ et $R_X = \frac{1}{1-p}$.

Pour $X \hookrightarrow \mathcal{P}(\lambda)$ avec $\lambda \in \mathbb{R}_+^*$, on a $g_X(t) = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} t^n = e^{\lambda(t-1)}$ et $R_X = +\infty$.

Théorème

Si X et Y sont deux variables aléatoires indépendantes sur $(\Omega, \mathcal{B}, \mathbb{P})$, la fonction génératrice de $X + Y$ est donnée par :

$$\forall t \in [-1, 1], g_{X+Y}(t) = g_X(t)g_Y(t)$$

son rayon de convergence étant $R_{X+Y} = \min(R_X, R_Y)$.

Théorème

La variable aléatoire X admet une espérance si, et seulement si, sa fonction génératrice est dérivable à gauche en 1 et dans ce cas, on a $\mathbb{E}(X) = g'_X(1^-)$.

Théorème

La variable aléatoire X est de carré sommable si, et seulement si, sa fonction génératrice est deux fois dérivable à gauche en 1 et dans ce cas, on a :

$$\begin{cases} \mathbb{E}(X^2) = g''_X(1^-) + \mathbb{E}(X) \\ \mathbb{V}(X) = g''_X(1^-) + g'_X(1^-)(1 - g'_X(1^-)) \end{cases}.$$

Définition

Pour X admettant un moment d'ordre $n \geq 1$, le réel $\mathbb{E}(X(X-1)\cdots(X-n+1))$ est le *moment factoriel* d'ordre n .

Théorème

La variable aléatoire X admet un moment d'ordre $r \geq 1$ si, et seulement si, la fonction génératrice est r fois dérivable à gauche en 1 et dans ce cas, on a :

$$\forall t \in [-1, 1], g_X^{(r)}(t) = \sum_{n=r}^{+\infty} \frac{n!}{(n-r)!} p_n t^{n-r}$$

(en -1 , il s'agit de dérivée à droite et en 1 de dérivée à gauche) et en particulier, $\mathbb{E}(X(X-1)\cdots(X-r+1)) = g_X^{(r)}(1^-)$.

27

Variables aléatoires réelles

$(\Omega, \mathcal{B}, \mathbb{P})$ est un espace probabilisé.

27.1 Fonction de répartition d'une probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$

Définition

La *fonction de répartition* d'une mesure de probabilité $\mathbb{P}$ sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est l'application :

$$\begin{aligned} F : \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto \mathbb{P}(]-\infty, x]). \end{aligned}$$

Théorème

Soient $\mathbb{P}$ une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ et F sa fonction de répartition.

- F est croissante sur $\mathbb{R}$ telle que $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$.
- Elle est continue à droite en tout point $x \in \mathbb{R}$, c'est-à-dire que :

$$\forall x \in \mathbb{R}, F(x^+) = \lim_{t \rightarrow x^+} F(t) = F(x) = \mathbb{P}(]-\infty, x]).$$

- Elle admet une limite à gauche en tout point $x \in \mathbb{R}$ avec :

$$\forall x \in \mathbb{R}, F(x^-) = \lim_{t \rightarrow x^-} F(t) = \mathbb{P}(]-\infty, x[).$$

- Pour tous réels $x < y$, on a :

$$\mathbb{P}(\{x\}) = F(x) - F(x^-) \text{ et } \mathbb{P}(]x, y]) = F(y) - F(x).$$

Corollaire

Soient $\mathbb{P}$ une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ et F sa fonction de répartition.

- La fonction F est continue en x_0 si, et seulement si, $\mathbb{P}(\{x_0\}) = 0$.
- L'ensemble des points de discontinuité de F est au plus dénombrable.

Une probabilité $\mathbb{P}$ sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est caractérisée par sa fonction de répartition, ce qui signifie que si deux mesures de probabilités $\mathbb{P}_1$ et $\mathbb{P}_2$ sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ ont la même fonction de répartition, elles sont alors égales.

Théorème

Si F est une application de $\mathbb{R}$ dans $\mathbb{R}$ qui est croissante, continue à droite en tout point de $\mathbb{R}$ telle que $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$, il existe alors une unique mesure de probabilité $\mathbb{P}$ sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ telle que F soit la fonction de répartition de $\mathbb{P}$.

27.2 Variables aléatoires réelles

Définition

On dit qu'une application $X : \Omega \rightarrow \mathbb{R}$ est une *variable aléatoire réelle* sur l'espace probabilisable $(\Omega, \mathcal{B})$ si : $\forall A \in \mathcal{B}_{\mathbb{R}}, X^{-1}(A) \in \mathcal{B}$.

Les variables aléatoires réelles discrètes sont bien des variables aléatoires du fait que pour tout $A \in \mathcal{B}_{\mathbb{R}}, X^{-1}(A) = \bigcup_{\substack{i \in I \\ x_i \in A}} X^{-1}(\{x_i\})$ est dans $\mathcal{B}$ comme réunion

dénombrable d'éléments de $\mathcal{B}$.

Théorème

Une application $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire réelle sur $(\Omega, \mathcal{B})$ si, et seulement si, on a $X^{-1}(]-\infty, x]) \in \mathcal{B}$ pour tout $x \in \mathbb{R}$, ce qui est encore équivalent à $X^{-1}(]-\infty, x]) \in \mathcal{B}$ pour tout $x \in \mathbb{R}$, ou à $X^{-1}(]a, b]) \in \mathcal{B}$ pour tous réels $a < b$.

On note $\mathcal{V}(\Omega, \mathcal{B})$ l'ensemble des variables aléatoires réelles définies sur $(\Omega, \mathcal{B})$. Pour toute variable aléatoire $X \in \mathcal{V}(\Omega, \mathcal{B})$, on note :

- $(X \in A)$ pour $X^{-1}(A) = \{\omega \in \Omega, X(\omega) \in A\}$ où $A \in \mathcal{B}_{\mathbb{R}}$;

- $(X = x)$ pour $X^{-1}(\{x\}) = \{\omega \in \Omega, X(\omega) = x\}$ où x est un réel ;
- $(X \leq x)$ pour $X^{-1}(]-\infty, x]) = \{\omega \in \Omega, X(\omega) \leq x\}$ où x est un réel ;
- $(X < x)$ pour $X^{-1}(]-\infty, x[) = \{\omega \in \Omega, X(\omega) < x\}$ où x est un réel.

Théorème

L'ensemble $\mathcal{V}(\Omega, \mathcal{B})$ est une $\mathbb{R}$ -algèbre.

Définition

On dit qu'une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires réelles sur $(\Omega, \mathcal{B})$ *converge simplement* sur Ω , s'il existe une fonction $X : \Omega \rightarrow \mathbb{R}$ telle pour tout $\omega \in \Omega$, la suite réelle $(X_n(\omega))_{n \in \mathbb{N}}$ converge vers $X(\omega)$.

Théorème

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{B})$. Si $(X_n)_{n \in \mathbb{N}}$ converge simplement vers X sur Ω , la fonction X est alors une variable aléatoire sur $(\Omega, \mathcal{B})$.

Définition

Soit I un intervalle réel. On dit qu'une fonction $f : I \rightarrow \mathbb{R}$ est *borélienne* si :

$$\forall A \in \mathcal{B}_{\mathbb{R}}, f^{-1}(A) \in \mathcal{B}_I$$

où $\mathcal{B}_I = \{I \cap J ; J \in \mathcal{B}_{\mathbb{R}}\}$ est la tribu des boréliens de I .

Théorème

Si X est une variable aléatoire réelle sur $(\Omega, \mathcal{B})$ et f une application borélienne de $\mathbb{R}$ dans $\mathbb{R}$, alors $f \circ X$ est une variable aléatoire réelle sur $(\Omega, \mathcal{B})$.

On notera $f(X)$ pour $f \circ X$ si f est borélienne.

Une variable aléatoire X sur $(\Omega, \mathcal{B})$ permet de transporter la mesure de probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{B})$ en la mesure de probabilité $\mathbb{P}_X$ sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Théorème

Si X est une variable aléatoire réelle sur l'espace probabilisé $(\Omega, \mathcal{B}, \mathbb{P})$, l'application $\mathbb{P}_X : \begin{cases} \mathcal{B}_{\mathbb{R}} & \rightarrow [0, 1] \\ A & \mapsto \mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A)) \end{cases}$ définit alors une probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Définition

Si X est une variable aléatoire réelle sur $(\Omega, \mathcal{B})$, on dit alors que la mesure de probabilité $\mathbb{P}_X$ définie par le théorème précédent est la *loi* de X et que la fonction de répartition de $\mathbb{P}_X$ est la *fonction de répartition* de X . On la note F_X .

Si μ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, on dit qu'une variable aléatoire réelle X sur $(\Omega, \mathcal{B})$ suit la loi μ , si $\mathbb{P}_X = \mu$ et on note alors $X \hookrightarrow \mu$.

La fonction de répartition F_X dépend de la loi de X , elle est définie par :

$$\forall x \in \mathbb{R}, F_X(x) = \mathbb{P}_X([-\infty, x]) = \mathbb{P}(X \leq x)$$

et pour $a < b$, on a $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.

La fonction de répartition F_X est croissante, à valeurs dans $[0, 1]$, continue à droite en tout point de $\mathbb{R}$, telle que $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$ et l'ensemble des points de discontinuité de F_X est au plus dénombrable.

Deux variables aléatoires de même loi ont la même fonction de répartition et réciproquement si deux variables aléatoires réelles sur $(\Omega, \mathcal{B})$ ont même fonction de répartition, elles ont alors la même loi. Cela se traduit en disant que la fonction de répartition d'une variable aléatoire caractérise sa loi.

27.3 Variables aléatoires réelles indépendantes

Soient X et Y deux variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$.

Définition

On dit que X et Y sont *indépendantes*, si pour tous boréliens A, B , les événements $(X \in A)$ et $(Y \in B)$ sont indépendants dans $(\Omega, \mathcal{B}, \mathbb{P})$, ce qui revient à dire que :

$$\mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A) \mathbb{P}(Y \in B).$$

Théorème

Les variables aléatoires réelles X et Y sont indépendantes si, et seulement si, on a :

$$\forall (x, y) \in \mathbb{R}^2, \mathbb{P}((X \leq x) \cap (Y \leq y)) = \mathbb{P}(X \leq x) \mathbb{P}(Y \leq y).$$

Définition

Soit $(X_i)_{i \in I}$ une famille variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ ayant au moins deux éléments. On dit que ces variables aléatoires sont *mutuellement indépendantes* (ou *indépendantes*) si pour toute famille $(A_i)_{i \in I}$ de boréliens, la famille $((X_i \in A_i))_{i \in I}$ est formée d'événements indépendants, ce qui revient à dire que pour toute partie finie J de I , on a :

$$\mathbb{P}\left(\bigcap_{j \in J} (X_j \in A_j)\right) = \prod_{j \in J} \mathbb{P}(X_j \in A_j).$$

Théorème

Si X et Y sont deux variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ indépendantes, alors pour toutes fonctions boréliennes f et g de $\mathbb{R}$ dans $\mathbb{R}$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Théorème

Soit $(X_k)_{1 \leq k \leq n}$ une famille de n variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ indépendantes.

- Pour toute suite $(f_k)_{1 \leq k \leq n}$ de fonctions boréliennes de $\mathbb{R}$ dans $\mathbb{R}$, les variables aléatoires $(f_k(X_k))_{1 \leq k \leq n}$ sont indépendantes.
- Pour tout entier naturel r compris entre 1 et $n-1$ et toutes fonctions boréliennes $\varphi : \mathbb{R}^r \rightarrow \mathbb{R}$ et $\psi : \mathbb{R}^{n-r} \rightarrow \mathbb{R}$, les variables aléatoires $\varphi(X_1, \dots, X_r)$ et $\psi(X_{r+1}, \dots, X_n)$ sont indépendantes.

27.4 Espérance d'une variable aléatoire réelle

X est une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$.

L'application $t \mapsto \mathbb{P}(X > t) = 1 - \mathbb{P}(X \leq t) = 1 - F_X(t)$ étant décroissante à valeurs dans $[0, 1]$, elle est intégrable sur tout segment $[0, x] \subset \mathbb{R}_+$ et la fonction croissante $x \mapsto \int_0^x \mathbb{P}(X > t) dt$ admet une limite, éventuellement infinie, quand x tend vers $+\infty$.

Définition

L'espérance d'une variable aléatoire réelle positive X sur $(\Omega, \mathcal{B}, \mathbb{P})$ est l'élément de $\overline{\mathbb{R}_+} = [0, +\infty]$ défini par :

$$\mathbb{E}(X) = \int_0^{+\infty} \mathbb{P}(X > t) dt.$$

Dans le cas où $\mathbb{E}(X) < +\infty$, on dit que X est *intégrable*.

Théorème

Si X et Y sont deux variables aléatoires réelles positives sur $(\Omega, \mathcal{B}, \mathbb{P})$ telles que $X \leq Y$, on a alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$ (l'espérance est croissante).

Théorème (inégalité de Markov)

Si X est une variable aléatoire positive et intégrable sur $(\Omega, \mathcal{B}, \mathbb{P})$, on a alors $\mathbb{P}(X > \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon}$ pour tout réel $\varepsilon > 0$.

Corollaire

Soit X une variable aléatoire positive et intégrable sur $(\Omega, \mathcal{B}, \mathbb{P})$. On a $\mathbb{E}(X) = 0$ si, et seulement si, $\mathbb{P}(X = 0) = 1$ (c'est-à-dire que X est presque sûrement nulle).

Théorème

Une variable aléatoire réelle positive sur $(\Omega, \mathcal{B}, \mathbb{P})$ est limite simple d'une suite croissante de variables aléatoires réelles étagées positives.

Théorème de Beppo Levi

Si $(X_n)_{n \in \mathbb{N}}$ est une suite croissante de variables aléatoires réelles positives sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui converge vers une variable aléatoire réelle X sur $(\Omega, \mathcal{B}, \mathbb{P})$, on a alors $\mathbb{E}(X) = \lim_{n \rightarrow +\infty} \mathbb{E}(X_n)$ dans $\overline{\mathbb{R}_+}$.

Théorème

Si X et Y sont deux variables aléatoires réelles positives sur $(\Omega, \mathcal{B}, \mathbb{P})$ et λ un réel strictement positif, on a alors $\mathbb{E}(\lambda X + Y) = \lambda \mathbb{E}(X) + \mathbb{E}(Y)$ dans $\overline{\mathbb{R}}_+$.

Théorème

Soient X et Y deux variables aléatoires réelles positives et indépendantes sur $(\Omega, \mathcal{B}, \mathbb{P})$. Si ces variables aléatoires sont intégrables, il en est alors de même de XY et on a $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Si X est une variable aléatoire réelle X sur $(\Omega, \mathcal{B}, \mathbb{P})$, on lui associe alors les variables aléatoires positives X^+ et X^- sur $(\Omega, \mathcal{B}, \mathbb{P})$ définies par :

$$X^+ = \max(X, 0) = \frac{1}{2}(X + |X|) \text{ et } X^- = \max(-X, 0) = \frac{1}{2}(|X| - X).$$

On a alors $X = X^+ - X^-$ et $|X| = X^+ + X^-$.

Définition

On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{B}, \mathbb{P})$ est *intégrable*, si $\mathbb{E}(X^+)$ et $\mathbb{E}(X^-)$ sont finies.

On note $L^1(\Omega, \mathcal{B}, \mathbb{P})$ l'ensemble des variables aléatoires intégrables sur $(\Omega, \mathcal{B}, \mathbb{P})$.

Définition

Si X est une variable aléatoire sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui est intégrable, son *espérance* est alors le réel $\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-)$.

Théorème

Une variable aléatoire X sur $(\Omega, \mathcal{B}, \mathbb{P})$ est intégrable si, et seulement si, on a $\mathbb{E}(|X|) < +\infty$ et dans ce cas, on a :

$$\mathbb{E}(X) = \int_0^{+\infty} \mathbb{P}(X > t) dt - \int_0^{+\infty} \mathbb{P}(X < -t) dt.$$

Théorème

L'ensemble $L^1(\Omega, \mathcal{B}, \mathbb{P})$ des variables aléatoires réelles intégrables sur $(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel de l'espace $\mathcal{V}(\Omega, \mathcal{B})$ des variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ et l'application :

$$\begin{aligned} \mathbb{E} : L^1(\Omega, \mathcal{B}, \mathbb{P}) &\rightarrow \mathbb{R} \\ X &\mapsto \mathbb{E}(X) \end{aligned}$$

est une forme linéaire croissante (c'est-à-dire que si X et Y sont intégrables et vérifient $X \leq Y$ alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$).

Théorème

Soient X et Y deux variables aléatoires réelles indépendantes sur $(\Omega, \mathcal{B}, \mathbb{P})$. Si ces variables aléatoires sont intégrables, il en est alors de même de XY et on a $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

27.5 Variance, covariance, moments d'ordre r

X est une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$.

Définition

On dit que X est *de carré intégrable*, si la variable aléatoire positive X^2 est intégrable.

On note $L^2(\Omega, \mathcal{B}, \mathbb{P})$ l'ensemble des variables aléatoires réelles de carré intégrable sur $(\Omega, \mathcal{B}, \mathbb{P})$.

Théorème

L'ensemble $L^2(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel de l'espace $L^1(\Omega, \mathcal{B}, \mathbb{P})$ des variables aléatoires réelles intégrables sur $(\Omega, \mathcal{B}, \mathbb{P})$. Pour X et Y dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ la variable aléatoire XY est dans $L^1(\Omega, \mathcal{B}, \mathbb{P})$ et l'application :

$$\begin{aligned} \varphi : (L^2(\Omega, \mathcal{B}, \mathbb{P}))^2 &\rightarrow \mathbb{R} \\ (X, Y) &\mapsto \mathbb{E}(XY) \end{aligned}$$

est une forme bilinéaire, symétrique et positive sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$.

Définition

La *variance* et l'*écart type* d'une variable aléatoire $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$ sont les réels respectivement définis par :

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \text{ et } \sigma(X) = \sqrt{\mathbb{V}(X)}.$$

Théorème de König

Pour tout $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$, on a $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ et, pour tous réels a et b , $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$.

Théorème

Pour tout $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$, on a $\mathbb{V}(X) = 0$ si, et seulement si, $\mathbb{P}(X = \mathbb{E}(X)) = 1$ (X est presque sûrement constante).

Définition

La *covariance* de deux variables aléatoires X et Y dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ est le réel défini par $\text{Cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right)$ et, dans le cas où $\mathbb{V}(X)$ et $\mathbb{V}(Y)$ sont non nuls, le *coefficient de corrélation* de X et Y est la réel $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$.

Définition

On dit que deux variables aléatoires X et Y dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ sont *non corrélées*, si on a $\text{Cov}(X, Y) = 0$.

Théorème

Pour X et Y dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$, on a $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ et l'application :

$$\begin{aligned} \psi : (L^2(\Omega, \mathcal{B}, \mathbb{P}))^2 &\rightarrow \mathbb{R} \\ (X, Y) &\mapsto \text{Cov}(X, Y) \end{aligned}$$

est une forme bilinéaire, symétrique et positive sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$.

Théorème

Soient $X_1, \dots, X_n$ dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ et $\lambda_1, \dots, \lambda_n$ des réels. On a :

$$\mathbb{V}\left(\sum_{k=1}^n \lambda_k X_k\right) = \sum_{k=1}^n \lambda_k^2 \mathbb{V}(X_k) + 2 \sum_{1 \leq j < k \leq n} \lambda_j \lambda_k \text{Cov}(X_j, X_k)$$

et dans le cas particulier où les variables aléatoires X_k sont deux à deux non corrélées, on a $\mathbb{V}\left(\sum_{k=1}^n \lambda_k X_k\right) = \sum_{k=1}^n \lambda_k^2 \mathbb{V}(X_k)$.

Théorème (inégalité de Bienaymé-Tchebychev)

Soit $X \in L^2(\Omega, \mathcal{B}, \mathbb{P})$. Pour tout réel $\varepsilon > 0$, on a :

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq \varepsilon\right) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Plus généralement si $X_1, \dots, X_n$ dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ sont deux à deux non corrélées, on a alors pour tout réel $\varepsilon > 0$:

$$\mathbb{P}\left(\left|\sum_{k=1}^n (X_k - \mathbb{E}(X_k))\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{k=1}^n \mathbb{V}(X_k).$$

Définition

On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{B}, \mathbb{P})$ admet un *moment d'ordre* $r \in \mathbb{N}^*$, si la variable aléatoire X^r est intégrable.

On note $L^r(\Omega, \mathcal{B}, \mathbb{P})$ l'ensemble des variables aléatoires sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui admettent un moment d'ordre r .

Théorème

Pour tout $r \in \mathbb{N}^*$, l'ensemble $L^r(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel de l'espace $\mathcal{V}(\Omega, \mathcal{B}, \mathbb{P})$ des variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$. Pour $r < s$ dans $\mathbb{N}^*$, $L^s(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel de $L^r(\Omega, \mathcal{B}, \mathbb{P})$.

Définition

Soient $r \in \mathbb{N}^*$, $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles dans $L^r(\Omega, \mathcal{B}, \mathbb{P})$ et X une variable aléatoire sur $(\Omega, \mathcal{B}, \mathbb{P})$. On dit que $(X_n)_{n \in \mathbb{N}}$ converge en moyenne d'ordre r vers X si $\lim_{n \rightarrow +\infty} \mathbb{E}(|X_n - X|^r) = 0$.

Pour $r = 1$, on dira simplement « convergence en moyenne ».

27.6 Loi faible des grands nombres

Définition

Soient $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$.

On dit que la suite $(X_n)_{n \in \mathbb{N}}$ *converge en probabilité* vers X si :

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0.$$

Théorème

Si $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui converge en probabilité vers les variables aléatoires réelles X et Y , on a alors $X = Y$ presque sûrement.

Théorème

Pour tout $r \in \mathbb{N}^*$, la convergence en moyenne d'ordre r entraîne la convergence en probabilité.

Théorème

Soient $r < s$ dans $\mathbb{N}^*$ et $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles dans $L^s(\Omega, \mathcal{B}, \mathbb{P})$.

Si $(X_n)_{n \in \mathbb{N}}$ converge en moyenne d'ordre s vers X , elle converge alors en moyenne d'ordre r vers X .

Théorème

Si $(X_n)_{n \in \mathbb{N}}$ et $(Y_n)_{n \in \mathbb{N}}$ sont deux suites de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui convergent respectivement en probabilité vers les variables aléatoires réelles X et Y , alors les suites $(X_n + Y_n)_{n \in \mathbb{N}}$ et $(X_n Y_n)_{n \in \mathbb{N}}$ convergent en probabilité respectivement vers $X + Y$ et XY .

Théorème (loi faible des grands nombres)

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ indépendantes, toutes de carré intégrable et telles que :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}(X_k) = \mu \text{ et } \lim_{n \rightarrow +\infty} \frac{1}{n^2} \sum_{k=1}^n \mathbb{V}(X_k) = 0$$

La suite $\left(\frac{1}{n} \sum_{k=1}^n X_k \right)_{n \in \mathbb{N}^*}$ converge en probabilité vers μ , ce qui signifie que,

pour tout réel $\varepsilon > 0$, on a $\lim_{n \rightarrow +\infty} \mathbb{P} \left(\left| \frac{1}{n} \sum_{k=1}^n X_k - \mu \right| > \varepsilon \right) = 0$.

28

Variables aléatoires à densité

$(\Omega, \mathcal{B}, \mathbb{P})$ est un espace probabilisé.

28.1 Densité de probabilité

Définition

Une *densité de probabilité* est une fonction $f : \mathbb{R} \setminus S \rightarrow \mathbb{R}_+$, où S est une partie finie de $\mathbb{R}$, intégrable sur $\mathbb{R}$ et telle que $\int_{\mathbb{R}} f(t) dt = 1$.

Définition

On dit qu'une variable aléatoire réelle X sur $(\Omega, \mathcal{B}, \mathbb{P})$ est à *densité* (ou qu'elle est *continue*), s'il existe une densité de probabilité f sur $\mathbb{R}$, telle que pour tous réels $a < b$, on ait $\mathbb{P}(a < X \leq b) = \int_a^b f(t) dt$.

Théorème

Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$ admettant une densité f .

— La fonction de répartition de X est définie par :

$$\forall x \in \mathbb{R}, F_X(x) = \int_{-\infty}^x f(t) dt.$$

— Cette fonction F_X est croissante et continue sur $\mathbb{R}$ telle que $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$, dérivable en tout $x_0 \in \mathbb{R} \setminus S$ où f est continue avec $F'_X(x_0) = f(x_0)$.

Théorème

Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$ admettant une densité f .

— Pour tout $x \in \mathbb{R}$, on a :

$$\mathbb{P}(X = x) = 0, \quad \mathbb{P}(X < x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t) dt$$

$$\mathbb{P}(X > x) = \mathbb{P}(X \geq x) = \int_x^{+\infty} f(t) dt$$

et pour tous réels $a < b$, on a :

$$\begin{aligned} \mathbb{P}(a < X < b) &= \mathbb{P}(a \leq X < b) = \mathbb{P}(a \leq X \leq b) \\ &= \mathbb{P}(a < X \leq b) = \int_a^b f(x) dx. \end{aligned}$$

Théorème

Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$. Si sa fonction de répartition F_X est continue et de classe $\mathcal{C}^1$ par morceaux sur $\mathbb{R}$, cette variable aléatoire est alors de densité $f = F'_X$ (cette fonction est définie sur $\mathbb{R}$ privé de l'ensemble des points où F_X n'est pas dérivable).

Théorème

Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$ possédant une densité f . Pour toute fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$ telle que $\varphi'(x) \neq 0$ pour tout $x \in \mathbb{R}$, la variable aléatoire $Y = \varphi(X)$ a pour densité la fonction $g = f \circ \varphi^{-1} \left| (\varphi^{-1})' \right| \mathbf{1}_{\varphi(\mathbb{R})}$.

Théorème

Si X et Y sont deux variables aléatoires réelles à densité sur $(\Omega, \mathcal{B}, \mathbb{P})$ indépendantes de densités respectives f et g , la variable aléatoire $X + Y$ est alors de densité h définie par :

$$h(x) = \int_{\mathbb{R}} f(t)g(x-t) dt = \int_{\mathbb{R}} g(t)f(x-t) dt.$$

La fonction h qui intervient dans le théorème précédent est le *produit de convolution* des fonctions f et g .

28.2 Variables aléatoires à densité classiques

Définition

On dit qu'une variable aléatoire réelle X suit la *loi uniforme* sur le segment $[a, b]$, avec $a < b$, si elle possède une densité définie par :

$$\forall t \in \mathbb{R}, f(t) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(t)$$

On note $X \hookrightarrow \mathcal{U}([a, b])$.

Sa fonction de répartition est définie sur $\mathbb{R}$ par :

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \int_a^x \frac{dt}{b-a} = \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}$$

La loi uniforme sur $[a, b]$ est aussi la loi uniforme sur $]a, b[$, $]a, b]$ ou $[a, b[$. Cette loi décrit une expérience aléatoire où on choisit au hasard un point dans $[a, b]$ de sorte que la probabilité que x soit dans un intervalle $J = [\alpha, \beta] \subset [a, b]$ est proportionnelle à la longueur de J , ce qui donne $\mathbb{P}(\alpha < X \leq \beta) = \frac{\beta-\alpha}{b-a}$.

Définition

On dit qu'une variable aléatoire réelle X suit une *loi exponentielle* de paramètre $\lambda > 0$ si elle possède une densité définie par :

$$\forall t \in \mathbb{R}, f(t) = \lambda e^{-\lambda t} \mathbf{1}_{\mathbb{R}_+}(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases}.$$

On note $X \hookrightarrow \mathcal{E}(\lambda)$.

Sa fonction de répartition est définie sur $\mathbb{R}$ par :

$$F_X(x) = \begin{cases} \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Définition

On dit qu'une variable aléatoire X sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{R}_+^*$ est *sans mémoire*, si on a :

$$\forall (x, y) \in \mathbb{R}_+ \times \mathbb{R}_+, \mathbb{P}(X > x + y) = \mathbb{P}(X > x) \mathbb{P}(X > y).$$

La condition d'être sans mémoire est équivalente à :

$$\mathbb{P}((X > x + y) \mid (X > y)) = \frac{\mathbb{P}(X > x + y)}{\mathbb{P}(X > y)} = \mathbb{P}(X > x).$$

Théorème

Une variable aléatoire X sur $(\Omega, \mathcal{B}, \mathbb{P})$ à valeurs dans $\mathbb{R}_+^*$ est sans mémoire si, et seulement si, elle suit une loi exponentielle.

Définition

Soient σ un réel strictement positif et μ réel. On dit qu'une variable aléatoire réelle X suit une *loi normale* (ou *loi de Gauss* ou *loi de Laplace-Gauss*) de paramètres μ et σ si elle possède une densité définie par :

$$\forall t \in \mathbb{R}, f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

On note $X \hookrightarrow \mathcal{N}(\mu, \sigma)$.

Sa fonction de répartition est définie sur $\mathbb{R}$ par :

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Théorème

Si la variable aléatoire X suit une loi normale de paramètres μ et σ , alors pour tous réels $a \in \mathbb{R}^*$ et $b \in \mathbb{R}$, la variable aléatoire $Y = aX + b$ suit une loi de Gauss de paramètres $b + a\mu$ et $|a|\sigma$. En particulier, $\frac{X-\mu}{\sigma}$ suit une *loi normale centrée réduite*, c'est-à-dire de paramètres 0 et 1.

Si $X \hookrightarrow \mathcal{N}(\mu, \sigma)$, on a alors $\mathbb{P}(a < X \leq b) = \mathbb{P}(\alpha < Y \leq \beta) = F_Y(\beta) - F_Y(\alpha)$, où $Y \hookrightarrow \mathcal{N}(0, 1)$, $\alpha = \frac{a-\mu}{\sigma}$, $\beta = \frac{b-\mu}{\sigma}$ et pour tout réel x , on a :

$$\begin{aligned} \mathbb{P}(Y \leq x) &= F_Y(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+\infty} e^{-\frac{t^2}{2}} dt \\ &= 1 - F_Y(-x) = \mathbb{P}(Y \geq -x) \end{aligned}$$

Pour tout réel $x > 0$, on a :

$$\mathbb{P}(|Y| \leq x) = F_Y(x) - F_Y(-x) = 2F_Y(x) - 1$$

$$\mathbb{P}(|Y| \geq x) = 1 - F_Y(x) + F_Y(-x) = 2(1 - F_Y(x)).$$

Théorème

Si $(X_k)_{1 \leq k \leq n}$ est une suite de variables aléatoires indépendantes telle que, pour tout k compris entre 1 et n , X_k suit une loi normale de paramètres $(\mu_k, \sigma_k) \in \mathbb{R} \times \mathbb{R}_+^*$, alors la variable aléatoire $\sum_{k=1}^n X_k$ suit une loi normale de paramètres $\left(\sum_{k=1}^n \mu_k, \sqrt{\sum_{k=1}^n \sigma_k^2} \right)$.

28.3 Espérance des variables aléatoires à densité

Théorème

Soit X une variable aléatoire réelle positive sur $(\Omega, \mathcal{B}, \mathbb{P})$. La suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires étagées définie par $X_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbf{1}_{A_{n,k}}$ pour tout $n \in \mathbb{N}^*$, où $A_{n,k} = \left(\frac{k}{2^n} \leq X < \frac{k+1}{2^n} \right)$ pour $0 \leq k \leq n2^n - 1$, converge en croissant vers X et on a l'égalité dans $\overline{\mathbb{R}}_+$:

$$\mathbb{E}(X) = \lim_{n \rightarrow +\infty} \mathbb{E}(X_n) = \lim_{n \rightarrow +\infty} \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{P} \left(\frac{k}{2^n} \leq X < \frac{k+1}{2^n} \right).$$

À toute variable aléatoire réelle X sur $(\Omega, \mathcal{B}, \mathbb{P})$, on associe les variables aléatoires positives $X^+ = \max(X, 0)$ et $X^- = \max(-X, 0)$.

Corollaire

Si X est une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$ possédant une densité f , on a alors dans $\overline{\mathbb{R}}_+$ les égalités :

$$\mathbb{E}(X^+) = \int_0^{+\infty} t f(t) dt; \quad \mathbb{E}(X^-) = \int_{-\infty}^0 -t f(t) dt;$$

$$\mathbb{E}(|X|) = \int_{-\infty}^{+\infty} |t| f(t) dt.$$

Théorème

Une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$ possédant une densité f est intégrable si, et seulement si, l'intégrale $\int_{-\infty}^{+\infty} t f(t) dt$ est absolument convergente et dans ce cas, on a $\mathbb{E}(X) = \int_{-\infty}^{+\infty} t f(t) dt$.

Pour $X \hookrightarrow \mathcal{U}([a, b])$ avec $a < b$, on a $\mathbb{E}(X) = \frac{a+b}{2}$.

Pour $X \hookrightarrow \mathcal{G}(\lambda)$ avec $\lambda \in \mathbb{R}_+^*$, on a $\mathbb{E}(X) = \frac{1}{\lambda}$.

Pour $X \hookrightarrow \mathcal{N}(\mu, \sigma)$ avec $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$, on a $\mathbb{E}(X) = \mu$.

28.4 Formule de transfert, moments d'ordre r **Théorème (formule de transfert)**

Soient X une variable aléatoire réelle sur $(\Omega, \mathcal{B}, \mathbb{P})$ possédant une densité f et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction borélienne. La variable aléatoire $\varphi(X)$ admet une espérance si, et seulement si, l'intégrale $\int_{\mathbb{R}} \varphi(t) f(t) dt$ est absolument convergente et dans ce cas, on a $\mathbb{E}(\varphi(X)) = \int_{\mathbb{R}} \varphi(t) f(t) dt$.

Prenant pour fonction φ la fonction $t \mapsto t^r$ où $r \in \mathbb{N}^*$, on en déduit que la variable aléatoire à densité X admet un moment d'ordre r si, et seulement si, l'intégrale

$\int_{-\infty}^{+\infty} t^r f(t) dt$ est absolument convergente et dans ce cas, ce moment est donné

par $\mathbb{E}(X^r) = \int_{-\infty}^{+\infty} t^r f(t) dt$.

Si X est de carré intégrable, elle admet alors une espérance et sa variance est donnée par :

$$\mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \int_{-\infty}^{+\infty} (t - \mathbb{E}(X))^2 f(t) dt$$

Le théorème de König nous dit aussi que $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.

Pour $X \hookrightarrow \mathcal{U}([a, b])$ avec $a < b$, on a $\mathbb{V}(X) = \frac{(b-a)^2}{12}$.

Pour $X \hookrightarrow \mathcal{G}(\lambda)$ avec $\lambda \in \mathbb{R}_+^*$, on a $\mathbb{V}(X) = \frac{1}{\lambda^2}$.

Pour $X \hookrightarrow \mathcal{N}(\mu, \sigma)$ avec $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$, on a $\mathbb{V}(X) = \sigma^2$.

28.5 Convergence en loi, théorème de la limite centrale

On se donne une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ (ou de manière plus générale, chaque X_n est une variable aléatoire réelle sur un espace probabilisé $(\Omega_n, \mathcal{B}_n, \mathbb{P}_n)$) et pour tout $n \in \mathbb{N}$, on note F_n la fonction de répartition de X_n .

Définition

On dit que la suite $(X_n)_{n \in \mathbb{N}}$ *converge en loi* vers une variable aléatoire X si, en tout point de continuité x de la fonction de répartition F_X , on a $\lim_{n \rightarrow +\infty} F_n(x) = F_X(x)$, soit $\lim_{n \rightarrow +\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x)$.

Théorème

Si $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui converge en loi vers une variable aléatoire X dont la fonction de répartition F_X est continue, la suite $(F_n)_{n \in \mathbb{N}}$ converge alors uniformément vers F_X sur tout segment.

Théorème

Si X et toutes les X_n pour $n \in \mathbb{N}$ sont à valeurs dans $\mathbb{N}$, alors la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers X si, et seulement si, pour tout $k \in \mathbb{N}$ on a $\lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)$.

Théorème

Si $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui converge en probabilité vers une variable aléatoire X , elle converge alors en loi vers X .

Corollaire

Si $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{B}, \mathbb{P})$ qui converge en moyenne d'ordre $r \in \mathbb{N}^*$ vers une variable aléatoire X , elle converge alors en loi vers X .

Théorème de la limite centrale

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant une même loi X d'espérance μ et d'écart type σ . En notant, pour tout $n \in \mathbb{N}^*$:

$$Y_n = \frac{1}{\sigma\sqrt{n}} \left(\sum_{k=1}^n X_k - n\mu \right)$$

la suite de variables aléatoires $(Y_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire Y qui suit la loi normale centrée réduite $\mathcal{N}(0, 1)$, c'est-à-dire que pour tout réel x , on a :

$$\lim_{n \rightarrow +\infty} \mathbb{P}(Y_n \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Pour n grand, une loi binomiale, qui est une somme de n lois de Bernoulli de même paramètre p indépendantes, peut être approchée par une loi normale centrée réduite de même espérance et de même écart type (théorème de Moivre-Laplace).

ANNEXE

Fonctions trigonométriques

x en degrés	0	30	45	60	90
x en radians	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin(x)$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\cos(x)$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\tan(x)$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	non défini

Fonctions trigonométriques réciproques

x	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arcsin(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\arccos(x)$	$\frac{\pi}{2}$	$\frac{\pi}{3}$	$\frac{\pi}{4}$	$\frac{\pi}{6}$	0

x	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$
$\arctan(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$

Formules de trigonométrie

Passage aux complexes : pour tout $x \in \mathbb{R}$, $\cos(x) = \operatorname{Re}(e^{ix})$ et $\sin(x) = \operatorname{Im}(e^{ix})$.

Formules d'Euler : pour tout $x \in \mathbb{R}$,
$$\begin{cases} \cos(x) = \frac{e^{ix} + e^{-ix}}{2} \\ \sin(x) = \frac{e^{ix} - e^{-ix}}{2i} \end{cases}.$$

*Addition des angles*¹ :

$$\begin{cases} \text{pour tout } (a, b) \in \mathbb{R}^2, \begin{cases} \cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \\ \sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a) \end{cases} ; \\ \text{pour tout } (a, b) \in \mathbb{R}^2 \text{ tel que } a \not\equiv \frac{\pi}{2} [\pi], b \not\equiv \frac{\pi}{2} [\pi] \text{ et } a+b \not\equiv \frac{\pi}{2} [\pi], \\ \qquad \qquad \qquad \tan(a+b) = \frac{\tan(a) + \tan(b)}{1 - \tan(a)\tan(b)} \end{cases}.$$

Transformation de produits en sommes :

$$\text{pour tout } (a, b) \in \mathbb{R}^2, \begin{cases} \cos(a)\cos(b) = \frac{1}{2}(\cos(a+b) + \cos(a-b)) \\ \sin(a)\sin(b) = \frac{1}{2}(\cos(a-b) - \cos(a+b)) \\ \cos(a)\sin(b) = \frac{1}{2}(\sin(a+b) - \sin(a-b)) \end{cases}.$$

Transformation de sommes en produits :

$$\text{pour tout } (p, q) \in \mathbb{R}^2, \begin{cases} \cos(p) + \cos(q) = 2\cos\left(\frac{p+q}{2}\right)\cos\left(\frac{p-q}{2}\right) \\ \cos(p) - \cos(q) = -2\sin\left(\frac{p+q}{2}\right)\sin\left(\frac{p-q}{2}\right) \\ \sin(p) + \sin(q) = 2\cos\left(\frac{p-q}{2}\right)\sin\left(\frac{p+q}{2}\right) \\ \sin(p) - \sin(q) = 2\cos\left(\frac{p+q}{2}\right)\sin\left(\frac{p-q}{2}\right) \end{cases}.$$

Fonctions trigonométriques en fonction de $t = \tan\left(\frac{x}{2}\right)$:

$$\text{pour tout } x \in \mathbb{R} \text{ tel que } x \not\equiv \frac{\pi}{4} \left[\frac{\pi}{2}\right] \text{ et } x \not\equiv \frac{\pi}{2} [\pi], \begin{cases} \cos(x) = \frac{1-t^2}{1+t^2} \\ \sin(x) = \frac{2t}{1+t^2} \\ \tan(x) = \frac{2t}{1-t^2} \end{cases}.$$

1. Les relations $\cos(a-b)$, $\sin(a-b)$ et $\tan(a-b)$ s'obtiennent en remplaçant b par $-b$ et en utilisant la parité de ces fonctions trigonométriques (paire ou impaire).

Dérivées usuelles

α désigne un réel.

domaine de définition	fonction f	domaine de dérivabilité	fonction f'
$\mathbb{R}_+^*$	$x \mapsto x^\alpha$	$\mathbb{R}_+^*$	$x \mapsto \alpha x^{\alpha-1}$
$\mathbb{R}$	$x \mapsto e^x$	$\mathbb{R}$	$x \mapsto e^x$
$\mathbb{R}_+^*$	$x \mapsto \ln(x)$	$\mathbb{R}_+^*$	$x \mapsto \frac{1}{x}$
$\mathbb{R}$	$x \mapsto \sin(x)$	$\mathbb{R}$	$x \mapsto \cos(x)$
$\mathbb{R}$	$x \mapsto \cos(x)$	$\mathbb{R}$	$x \mapsto -\sin(x)$
$\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi; k \in \mathbb{Z} \right\}$	$x \mapsto \tan(x)$	$\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi; k \in \mathbb{Z} \right\}$	$x \mapsto \frac{1}{\cos^2(x)}$
$[-1, 1]$	$x \mapsto \arcsin(x)$	$] -1, 1[$	$x \mapsto \frac{1}{\sqrt{1-x^2}}$
$[-1, 1]$	$x \mapsto \arccos(x)$	$] -1, 1[$	$x \mapsto -\frac{1}{\sqrt{1-x^2}}$
$\mathbb{R}$	$x \mapsto \arctan(x)$	$\mathbb{R}$	$x \mapsto \frac{1}{1+x^2}$
$\mathbb{R}$	$x \mapsto \operatorname{sh}(x)$	$\mathbb{R}$	$x \mapsto \operatorname{ch}(x)$
$\mathbb{R}$	$x \mapsto \operatorname{ch}(x)$	$\mathbb{R}$	$x \mapsto \operatorname{sh}(x)$
$\mathbb{R}$	$x \mapsto \operatorname{th}(x)$	$\mathbb{R}$	$x \mapsto \frac{1}{\operatorname{ch}^2(x)}$

Primitives usuelles

α désigne un réel différent de -1 .

k désigne un entier relatif.

Pour chaque fonction f , une primitive de f sur l'intervalle I est donnée à une constante additive près.

Intervalle I	Fonction f	Une primitive de f sur I
$\mathbb{R}_+^*$	$x \mapsto x^\alpha$	$x \mapsto \frac{x^{\alpha+1}}{\alpha+1}$
$\mathbb{R}_+^*$ ou $\mathbb{R}_-^*$	$x \mapsto \frac{1}{x}$	$x \mapsto \ln x $
$\mathbb{R}$	$x \mapsto e^x$	$x \mapsto e^x$
$\mathbb{R}_+^*$	$x \mapsto \ln(x)$	$x \mapsto x \ln(x) - x$
$\mathbb{R}$	$x \mapsto \sin(x)$	$x \mapsto -\cos(x)$
$\mathbb{R}$	$x \mapsto \cos(x)$	$x \mapsto \sin(x)$
$] -\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[$	$x \mapsto \tan(x)$	$x \mapsto -\ln \cos(x) $
$] -\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[$	$x \mapsto \frac{1}{\cos^2(x)} = 1 + \tan^2(x)$	$x \mapsto \tan(x)$
$] -1, 1[$	$x \mapsto \frac{1}{\sqrt{1-x^2}}$	$x \mapsto \arcsin(x)$
$\mathbb{R}$	$x \mapsto \frac{1}{1+x^2}$	$x \mapsto \arctan(x)$
$\mathbb{R}$	$x \mapsto \operatorname{sh}(x)$	$x \mapsto \operatorname{ch}(x)$
$\mathbb{R}$	$x \mapsto \operatorname{ch}(x)$	$x \mapsto \operatorname{sh}(x)$

Développements limités usuels

Les développements limités ci-dessous sont tous au voisinage de 0.

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

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

$$e^x = \sum_{k=0}^n \frac{x^k}{k!} + o(x^n)$$

$$\ln(1+x) = \sum_{k=1}^n (-1)^{k-1} \frac{x^k}{k} + o(x^n)$$

$$\cos(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2n})$$

$$\sin(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+1})$$

$$\tan(x) = x + \frac{x^3}{3} + o(x^3)$$

$$\arctan(x) = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} + o(x^{2n+1})$$

$$\operatorname{sh}(x) = \sum_{k=0}^n \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+1})$$

$$\operatorname{ch}(x) = \sum_{k=0}^n \frac{x^{2k}}{(2k)!} + o(x^{2n})$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots + \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n + o(x^n)$$

Développements en série entière usuels

$$e^z = \sum_{n=0}^{+\infty} \frac{z^n}{n!} \text{ sur } \mathbb{C}$$

$$\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n \text{ sur } \mathcal{D}(0,1)$$

$$\sin(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \text{ sur } \mathbb{R}$$

$$\cos(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n} \text{ sur } \mathbb{R}$$

$$\operatorname{sh}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!} \text{ sur } \mathbb{R}$$

$$\operatorname{ch}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!} \text{ sur } \mathbb{R}$$

$$\ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n \text{ sur }]-1,1] \quad \arctan(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \text{ sur } [-1,1]$$

$$(1+x)^\alpha = 1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} x^n \text{ sur }]-1,1[\text{ où } \alpha \in \mathbb{R} \setminus \mathbb{N}$$

Index

A

- Abel
 - lemme d'—, 162
 - règle d'— pour les séries numériques, 119
 - théorème d'—, 162
 - théorème radial d'—, 169
- abélien (groupe —), 229
- absolue
 - convergence — d'une intégrale généralisée, 132
 - convergence — d'une série de fonctions, 159
 - convergence — d'une série numérique, 117
- absorbant, 242
- accroissement(s)
 - égalité des — finis, 55
 - inégalité des — finis pour les fonctions numériques, 55
 - inégalité des — finis pour les fonctions vectorielles, 191
 - taux d'—, 50
- adhérence, 91
 - valeur d'—, 86
- adhérent(s)
 - caractérisation séquentielle des points —, 91
 - point —, 36, 91
- adjacentes (suites —), 21
- adjoint(e)
 - d'un endomorphisme, 332
 - matrice —, 347
- affine
 - application —, 351
 - droite —, 349
 - espace —, 349
 - espace — euclidien, 355
 - isométrie —, 358
 - plan —, 349
 - projection —, 353
 - projection — orthogonale, 358
 - réflexion —, 358
 - repère —, 350
 - sous-espace —, 350
 - symétrie —, 353
 - symétrie — orthogonale, 358
- aléatoire(s)
 - espérance d'une variable — discrète, 383
 - espérance d'une variable — réelle, 396
 - espérance d'une variable — réelle positive, 395
 - expérience —, 368
 - loi d'une variable —, 377, 393
 - variable — à densité, 402
 - variable — continue, 402
 - variable — discrète, 376
 - variable — étagée, 376
 - variable — réelle, 391
 - variable — réelle intégrable, 396
 - variable — réelle positive intégrable, 395
 - variables — indépendantes, 381, 393

- variables — mutuellement indépen-
dantes, 382, 394
- variables — non corrélées, 387, 398
- Alembert
 - règle de d' — pour les séries entières,
164
 - règle de d' — pour les séries numé-
riques, 115
- Alembert-Gauss (théorème de d' —), 301
- algèbre, 265
 - morphisme d' —, 265
 - sous- —, 265
- algorithme d'Euclide, 297
- alterné(e)(s)
 - critère spécial des séries —, 118
 - forme bilinéaire —, 317
 - groupe —, 239
 - série —, 118
- angle géométrique, 329
- anneau(x), 239
 - automorphisme d'un —, 241
 - caractéristique d'un —, 242
 - commutatif, 239
 - endomorphisme d'un —, 241
 - idéal d'un —, 242
 - image d'un morphisme d' —, 241
 - intègre, 240
 - isomorphisme d' —, 241
 - morphisme d' —, 241
 - noyau d'un morphisme d' —, 241
 - sous- —, 241
- annulateur
 - idéal —, 311
 - polynôme —, 311
- anté-duale (base —), 274
- antécédent, 224
- anti-déplacement, 358
- antisymétrique
 - forme bilinéaire —, 317
 - matrice —, 282
 - relation —, 226
- application(s), 223
- affine, 351
- bijective, 225
- composée de deux —, 224
- continue dans des espaces vectoriels
normés, 96
- contractante, 98
- différentiable, 198, 199
- différentielle d'une —, 199
- différentielle d'une — en un point,
199
- gradient d'une — numérique diffé-
rentiable, 206
- image d'une —, 224
- image d'une — linéaire, 265
- injective, 225
- linéaire, 264
- linéaire canoniquement associée à
une matrice, 281
- linéaire tangente en un point, 199
- lipschitzienne dans des espaces vec-
toriels normés, 98
- matrice d'une — linéaire, 277
- multilinéaire, 101
- noyau d'une — linéaire, 265
- polynomiale définie sur un espace
vectoriel normé, 107
- rang d'une — linéaire, 276
- réciproque, 225
- surjective, 225
- transposée d'une — linéaire, 267
- uniformément continue dans des es-
paces vectoriels normés, 98
- arc(s), 103
 - composantes connexes par —, 104
 - connexe par —, 104
 - de classe $\mathbb{C}^k$, 192
 - dérivée le long d'un —, 201
 - extrémité d'un —, 103
 - origine d'un —, 103
 - paramétré du plan, 192–196
- arccosinus, 34
- arcsinus, 34

arctangente, 34
 arithmético-géométrique (suite —), 26
 arithmétique (suite —), 25
 assertion(s), 217
 conjonction de deux —, 219
 disjonction de deux —, 219
 équivalence de deux —, 219
 négation d'une —, 219
 associative (loi —), 227
 asymptote, 196
 horizontale, 196
 oblique, 196
 verticale, 196
 autoadjoint(e)
 endomorphisme —, 337
 matrice —, 282
 automorphisme
 d'un anneau, 241
 d'un corps, 244
 d'un espace vectoriel, 266
 d'un groupe, 233
 orthogonal négatif, 340
 orthogonal positif, 340
 axiome, 217

B

barycentre, 360
 base
 anté-duale, 274
 d'un espace vectoriel, 270
 duale, 274
 formule de changement de —, 283
 orthogonale pour une forme quadratique, 321
 théorème de la — incomplète, 272
 Bayes (formules de —), 374
 Beppo Levi (théorème de —), 395
 Bernoulli (loi de —), 377
 Bertrand (séries de —), 115
 Bessel (inégalité de —), 180, 333
 Bézout

 identité de —, 246, 297
 théorème de —, 247, 298
 Bienaymé-Tchebychev (inégalité de —), 386, 399
 bijection réciproque, 225
 bijective (application —), 225
 bilinéaire
 discriminant d'une forme —, 319
 forme —, 317
 forme — antisymétrique, 317
 forme — non dégénérée, 322
 forme — symétrique, 317
 forme — symétrique définie, 327
 forme — symétrique définie positive, 324
 forme — symétrique positive, 324, 327
 matrice d'une forme —, 318
 noyau d'une forme —, 322
 rang d'une forme — symétrique, 323
 binaire (relation —), 226
 binôme (formule du — de Newton), 240, 279
 binomiale (loi —), 373, 377
 Bolzano-Weierstrass (théorème de —), 25, 106
 Boole (inégalité de —), 371
 bordant, 291
 borélienne
 fonction —, 392
 tribu —, 370
 borne(s)
 inférieure, 15
 supérieure, 15
 théorème des — atteintes, 46
 bornée
 fonction —, 31
 partie —, 88
 suite —, 17
 boule
 fermée, 87
 ouverte, 87

branche

infinie, 195

parabolique, 196

C

Cantor (théorème de —), 368

caractérisation

d'un fermé relatif, 93

d'un ouvert relatif, 93

d'un voisinage relatif, 93

de la continuité pour une application linéaire, 100

de la continuité pour une application multilinéaire, 101

séquentielle de la continuité, 44, 97

séquentielle de la limite, 39, 95

séquentielle des fermés, 92

séquentielle des fermés relatifs, 94

séquentielle des points adhérents, 91

caractéristique

d'un anneau, 242

équation — d'une équation différentielle linéaire d'ordre 2 à coefficients constants, 141

équation — d'une suite récurrente linéaire d'ordre 2, 27

polynôme — d'un endomorphisme, 307

polynôme — d'une matrice, 308

sous-espace —, 314

cardinal, 365

cartésien (produit —), 223

Cauchy, 204

critère intégral de —, 130

problème de —, 140

produit de — de séries entières, 167

produit de — de séries numériques, 120

règle de — pour les séries entières, 165

règle de — pour les séries numériques, 116

suite de —, 107

Cauchy-Schwarz (théorème de —), 328, 346

Cayley-Hamilton (théorème de —), 312

changement de variable

pour les intégrales, 79

pour les intégrales généralisées, 131

Chasles (relation de —), 349

chemin, 103

classe

de — $\mathcal{C}^1$ par morceaux, 182

fonction de — $\mathcal{C}^k$, 52

fonction de plusieurs variables de — $\mathcal{C}^1$, 207

fonction de plusieurs variables de — $\mathcal{C}^k$, 212

fonction de plusieurs variables de — $\mathcal{C}^\infty$, 212

classe d'équivalence, 226

coefficient de corrélation, 386, 398

coefficients

de Fourier, 177

de Fourier exponentiels, 178

cofacteur, 286

colinéaire, 262

comatrice, 288

combinaison linéaire, 262

commutatif

anneau —, 239

groupe —, 229

commutative

convergence — d'une série numérique, 124

loi —, 227

théorème de convergence —, 124

compact, 102

comparaison

critères de — pour les séries entières, 165

critères de — pour les séries numériques, 114

critère de — logarithmique, 115

- sommation des relations de —, 119
- théorème de — série-intégrale, 120, 130
- complémentaire d'une partie d'un ensemble, 222
- composantes, 270
- composée
 - de deux applications, 224
 - de deux polynômes, 295
- concave (fonction —), 58
- conditionnelle (probabilité —), 373
- cône isotrope, 322
- congruence (relation de —), 242, 245
- conjointe (loi —), 381
- conjonction de deux assertions, 219
- conjugaison (théorème de — pour les intégrales généralisées), 128
- conjugué (d'un nombre complexe), 256
- connexe par arcs, 104
- constant (polynôme —), 294
- continue
 - application — dans des espaces vectoriels normés, 96
 - application uniformément — dans des espaces vectoriels normés, 98
 - expérience aléatoire —, 368
 - fonction —, 42
 - fonction — à gauche ou à droite, 43
 - fonction numérique — par morceaux, 47
 - fonction uniformément —, 48
 - fonction vectorielle — par morceaux, 189
- continuité
 - à gauche ou à droite d'une fonction, 43
 - caractérisation de la — pour une application linéaire, 100
 - caractérisation de la — pour une application multilinéaire, 101
 - caractérisation séquentielle de la —, 44, 97
 - d'une fonction en un point, 42
 - d'une fonction vectorielle, 185
 - en un point dans des espaces vectoriels normés, 96
 - par morceaux d'une fonction numérique, 47
 - par morceaux d'une fonction vectorielle, 189
 - prolongement par —, 43
 - uniforme dans des espaces vectoriels normés, 98
 - uniforme pour une fonction numérique, 48
- contractante (application —), 49, 98
- contraposée, 220
- convergence
 - absolue d'une intégrale généralisée, 132
 - absolue d'une série de fonctions, 159
 - absolue d'une série numérique, 117
 - commutative d'une série numérique, 124
 - d'une série numérique, 110
 - d'une série vectorielle, 143
 - d'une suite d'éléments d'un espace vectoriel normé, 85
 - disque ouvert de —, 164
 - domaine de — d'une série entière, 164
 - en loi, 408
 - en moyenne d'ordre r , 399
 - en moyenne quadratique, 181
 - en probabilité, 400
 - intervalle de — d'une série entière, 164
 - normale d'une série de fonctions, 159
 - norme de la — en moyenne, 83
 - norme de la — en moyenne quadratique, 83
 - norme de la — uniforme, 151

- rayon de —, 163
 - semi- — d'une intégrale généralisée, 133
 - semi- — d'une série numérique, 117
 - simple d'une suite de fonctions, 150
 - simple d'une série de fonctions, 156
 - théorème de — dominée de Lebesgue, 135
 - uniforme d'une suite de fonctions, 150
 - uniforme d'une série de fonctions, 157
 - convexe
 - fonction —, 58
 - partie —, 361
 - coordonnées, 270
 - coordonnées barycentriques, 361
 - corde, 58
 - corps, 243
 - automorphisme d'un —, 244
 - endomorphisme d'un —, 244
 - extension d'un —, 243
 - isomorphisme de —, 244
 - morphisme de —, 244
 - sous- —, 243
 - cosinus hyperbolique, 33
 - courbe
 - de niveau, 209
 - paramétrée, 192
 - représentative d'une fonction réelle, 30
 - covariance, 386, 398
 - Cramer
 - formules de —, 290
 - système de —, 290
 - critère(s)
 - d'Abel pour les séries numériques, 119
 - de comparaison
 - pour les séries entières, 165
 - pour les séries numériques, 114
 - de comparaison logarithmique, 115
 - de d'Alembert pour les séries entières, 164
 - de d'Alembert pour les séries numériques, 115
 - de Leibniz, 118
 - de Riemann, 114
 - de Riemann généralisé, 115
 - intégral de Cauchy, 130
 - spécial des séries alternées, 118
 - critique
 - point — pour les fonctions d'une variable, 54
 - point — pour les fonctions de plusieurs variables, 210
 - croissante
 - fonction réelle —, 31
 - suite —, 18
 - cycle(s), 236
 - disjoints, 237
 - support d'un —, 236
 - cyclique (groupe —), 231
- ## D
- décomposition
 - en éléments simples, 305
 - en facteurs irréductibles, 303
 - en facteurs premiers, 251
 - lemme de — des noyaux, 312
 - primaire, 251
 - décroissante
 - fonction réelle —, 31
 - suite —, 18
 - défini(e)
 - endomorphisme symétrique — positif, 338
 - forme bilinéaire — positive, 324
 - forme bilinéaire symétrique —, 327
 - forme quadratique — positive, 324
 - matrice hermitienne — positive, 347
 - matrice symétrique — positive, 338
 - degré (d'un polynôme), 294

- démonstration
 - par contraposition, 220
 - par contre-exemple, 220
 - par l'absurde, 220
- dénombrable, 365
- dense
 - partie — dans $\mathbb{R}$, 21
 - partie — dans un espace vectoriel normé, 92
- densité
 - dans $\mathbb{R}$, 21
 - dans un espace vectoriel normé, 92
- densité de probabilité, 402
- déplacement, 358
- dérivable
 - fonction numérique —, 50
 - fonction numérique — à gauche ou à droite, 51
 - fonction vectorielle —, 185
 - fonction vectorielle — à gauche ou à droite, 186
- dérivée(s)
 - à gauche ou à droite d'une fonction numérique, 51
 - à gauche ou à droite d'une fonction vectorielle, 186
 - d'une fonction numérique, 50
 - d'une fonction vectorielle, 185
 - directionnelle, 201
 - le long d'un arc, 201
 - partielles d'ordre k , 211
 - partielles d'une application, 202
 - partielles secondes, 211
 - selon un vecteur, 201
 - usuelles, 415
- déterminant
 - d'un endomorphisme, 289
 - d'une famille de vecteurs, 288
 - d'une matrice carrée, 286
 - d'une matrice triangulaire, 287
 - de Vandermonde, 288
 - extrait, 291
 - principal, 292
- développable (fonction — en série entière), 171
- développement(s)
 - en séries entières usuels, 174, 417
 - limité, 68
 - limités usuels, 71, 72, 417
- diagonale (matrice —), 282
- diagonalisable
 - endomorphisme —, 309
 - matrice —, 309
- diamètre, 356
- différence
 - de deux parties d'un ensemble, 222
 - symétrique de deux parties d'un ensemble, 222
- différentielle
 - d'une application, 199
 - d'une application en un point, 199
- dimension
 - d'un espace vectoriel, 271
 - finie, 271
 - infinie, 271
- directe
 - base orthonormée —, 341
 - image — d'un ensemble par une application, 224
 - somme —, 268
- Dirichlet
 - condition de —, 179
 - espace de —, 179
 - fonction de —, 42, 68
 - intégrale de —, 133, 135
 - théorème de —, 183
- discriminant
 - d'un polynôme de degré 2 à coefficients complexes, 257
 - d'une forme bilinéaire, 319
- disjonction de deux assertions, 219
- disque (ouvert de convergence), 164
- distance, 84
 - associée à une norme, 84

d'un vecteur à un sous-espace, 85
 divergence (grossière), 112
 dividende, 245
 diviseur, 245
 d'un entier relatif, 245
 de 0, 240
 division euclidienne
 de deux entiers relatifs, 245
 de deux polynômes, 296
 domaine de définition, 224
 domination
 relation de — pour les fonctions, 60
 relation de — pour les suites, 65
 droite(s)
 affine, 349
 numérique achevée, 36
 perpendiculaires, 355
 vectorielle, 271
 dual (espace —), 266
 Duhamel
 règle de —, 116
 règle de Raabe —, 116
 Dunford (théorème de —), 315

E

écart type, 386, 398
 échelonnés (polynômes —), 295
 égalité
 de Parseval, 182
 des accroissements finis, 55
 du parallélogramme, 328
 élément
 inversible, 228
 neutre, 227
 ordre d'un — dans un groupe, 234
 régulier, 227
 simplifiable, 227
 encadrement (limite finie par —), 20, 40
 endomorphisme
 adjoint d'un —, 332
 autoadjoint, 337

d'un anneau, 241
 d'un corps, 244
 d'un espace vectoriel, 264
 d'un groupe, 233
 diagonalisable, 309
 nilpotent, 313
 spectre d'un —, 307
 symétrique, 337
 symétrique défini positif, 338
 symétrique positif, 338
 trace d'un —, 280
 trigonalisable, 309
 valeur propre d'un —, 307
 vecteur propre d'un —, 307
 engendré(e)
 idéal —, 296
 sous-espace vectoriel —, 262
 sous-groupe — par une partie, 230
 tribu —, 370
 ensemble
 au plus dénombrable, 365
 complémentaire d'une partie d'un —, 222
 d'arrivée, 224
 de définition, 224
 de départ, 224
 dénombrable, 365
 différence de deux parties d'un —, 222
 différence symétrique de deux parties d'un —, 222
 image d'un — par une application, 224
 image réciproque d'un — par une application, 224
 intersection de deux parties d'un —, 222
 ordonné, 226
 partition d'un —, 223
 permutation d'un —, 225
 quotient, 226
 recouvrement disjoint d'un —, 223

- réunion de deux parties d'un —, 222
- totalelement ordonné, 226
- vide, 220
- épreuve, 368
- équation
 - caractéristique associée a une équation homogène, 141
 - caractéristique d'une suite récurrente linéaire d'ordre 2, 27
 - différentielle linéaire scalaire d'ordre 2 à coefficients constants, 141
 - différentielle linéaire scalaire d'ordre 1, 139
 - différentielle linéaire scalaire d'ordre 2, 146
 - différentielle linéaire scalaire d'ordre n , 148
 - différentielle linéaire vectorielle d'ordre 1, 147
 - homogène, 139, 141, 146–148
- équipotents, 365
- équivalence
 - classe d'—, 226
 - de deux assertions, 219
 - relation d'—, 226
- équivalentes
 - fonctions —, 63
 - suites —, 67
- escalier (fonction en —), 80
- espace(s)
 - affine, 349
 - affine euclidien, 355
 - automorphisme d'un — vectoriel, 266
 - base d'un — vectoriel, 270
 - dimension d'un — vectoriel, 271
 - dual, 266
 - endomorphisme d'un — vectoriel, 264
 - euclidien, 327
 - hermitien, 345
 - isomorphisme d'— vectoriels, 266
 - métrique, 84
 - préhilbertien, 327
 - probabilisable, 369
 - probabilisé, 371
 - sous- — vectoriel, 262
 - sous- — vectoriel engendré, 262
 - vectoriel, 261
 - vectoriel normé, 83
 - vectoriel normé produit, 85
- espérance, 407
 - d'une variable aléatoire discrète, 383
 - d'une variable aléatoire réelle, 396
 - d'une variable aléatoire réelle positive, 395
- étagé(e)(s)
 - polynômes —, 295
 - variable aléatoire —, 376
- étoilée
 - par rapport à un point, 104
 - partie —, 104
- Euclide
 - algorithme d'—, 249
 - théorème d'—, 251, 303
- euclidien
 - espace —, 327
 - espace affine —, 355
- Euler
 - formules d'—, 414
 - indicatrice d'—, 253
 - intégrales d'—, 135
 - théorème d'—, 253
- événement(s), 368
 - certain, 369
 - contraire, 369
 - élémentaire, 368
 - impossible, 369
 - indépendants, 374
 - mutuellement indépendants, 375
 - système complet d'—, 372
- éventualité, 368
- expérience aléatoire, 368
 - continue, 368
 - discrète, 368

exponentielle
 d'un endomorphisme, 144
 d'une matrice carrée, 144
 fonction — complexe, 118, 172
 loi —, 404
 extension (d'un corps), 243
 extractrice, 24
 extrait(e)
 déterminant —, 291
 matrice —, 291
 suite —, 24
 suite — dans un espace vectoriel
 normé, 86
 extremum
 condition nécessaire d'un — local,
 54
 condition suffisante d'un — local,
 54, 73
 local pour les fonctions de plusieurs
 variables, 210
 local pour une fonction numérique,
 54

F

famille
 génératrice, 270
 libre, 270
 liée, 270
 orthogonale, 330
 orthonormale, 330
 rang d'une — de vecteurs, 275
 sommable de réels ou complexes, 122
 sommable de réels positifs, 121
 Fermat (petit théorème de —), 252
 fermé(e)(s)
 boule —, 87
 caractérisation séquentielle des —,
 92
 caractérisation séquentielle des —
 relatifs, 94
 ensemble —, 89
 intervalle —, 36
 relatif, 93
 finie (dimension —), 271
 fixe
 point —, 28, 109
 théorème du point —, 109
 fonction(s), 223
 arccosinus, 34
 arcsinus, 34
 arctangente, 34
 borélienne, 392
 bornée, 31
 complexe, 29
 concave, 58
 contractante, 49
 convergence normale d'une série de
 —, 159
 convergence simple d'une suite de
 —, 150
 convergence simple d'une série de
 —, 156
 convergence uniforme d'une suite de
 —, 150
 convergence uniforme d'une série de
 —, 157
 convexe, 58
 cosinus hyperbolique, 33
 de classe $\mathcal{C}^1$ par morceaux, 182
 de classe $\mathcal{C}^k$, 52
 de Dirichlet, 42, 68
 de Lejeune-Dirichlet, 42, 68
 de plusieurs variables de classe $\mathcal{C}^1$,
 207
 de plusieurs variables de classe $\mathcal{C}^k$,
 212
 de plusieurs variables de classe $\mathcal{C}^\infty$,
 212
 de répartition, 380, 390, 393, 402
 développable en série entière, 171
 en escalier, 80
 équivalentes, 63
 exponentielle complexe, 118, 172
 génératrice, 388

- graphe d'une — réelle, 30
 - indicatrice, 367
 - indicatrice d'Euler, 253
 - intégrable au sens de Riemann, 81
 - intégrable pour les intégrales généralisées, 132
 - lipschitzienne, 49
 - majorée, 31
 - maximum d'une — réelle, 31
 - minimum d'une — réelle, 31
 - minorée, 31
 - numérique impaire, 30
 - numérique paire, 30
 - numérique périodique, 30
 - partie entière, 32
 - polynomiale, 299
 - polynomiale définie sur un espace vectoriel normé, 107
 - prolongement d'une —, 30
 - rationnelle, 304
 - réciproques des fonctions trigonométriques, 34, 413
 - réelle, 29
 - réelle croissante, 31
 - réelle décroissante, 31
 - réelle monotone, 31
 - restriction d'une —, 30
 - sinus hyperbolique, 33
 - somme d'une série de fonctions, 157
 - tangente hyperbolique, 33
 - trigonométriques, 413
- forme
 - base orthogonale pour une — quadratique, 321
 - bilinéaire, 317
 - bilinéaire alternée, 317
 - bilinéaire antisymétrique, 317
 - bilinéaire non dégénérée, 322
 - bilinéaire symétrique, 317
 - bilinéaire symétrique définie, 327
 - bilinéaire symétrique définie positive, 324
 - bilinéaire symétrique positive, 324, 327
 - linéaire, 266
 - matrice d'une — bilinéaire, 318
 - noyau d'une — bilinéaire, 322
 - polaire, 319
 - rang d'une — bilinéaire symétrique, 323
- forme quadratique, 319
 - base orthogonale pour une —, 321
 - définie positive, 324
 - non dégénérée, 322
 - positive, 324
 - rang d'une —, 323
 - signature d'une —, 324
- formule(s)
 - d'Euler, 414
 - de Bayes, 374
 - de changement de base, 283
 - de Cramer, 290
 - de Grassmann, 273
 - de König, 386
 - de Leibniz, 53
 - de Moivre, 258
 - de Stirling, 68
 - de Taylor, 302
 - de Taylor avec reste intégral pour les fonctions numériques, 80
 - de Taylor avec reste intégral pour les fonctions vectorielles, 191
 - de Taylor-Lagrange, 80
 - de transfert, 384, 407
 - de trigonométrie, 414
 - de Viète, 301
 - des probabilités composées, 374
 - des probabilités totales, 374
 - du binôme de Newton, 240, 279
- Fourier
 - coefficients de —, 177
 - coefficients de — exponentiels, 178
 - série de —, 177
- fraction rationnelle, 304

décomposition en éléments simples
 d'une —, 305
 pôle d'une —, 305
 Fresnel (intégrales de —), 132, 135
 frontière, 91
 Fubini (théorème de —), 125

G

Gauss

intégrale de —, 130, 135
 loi de —, 405
 méthode du pivot de —, 290
 théorème de —, 248, 298
 théorème de réduction de —, 320

génératrice (famille —), 270

géométrie

angle —, 329
 loi —, 373, 377
 série —, 111
 suite —, 25

gradient (d'une application numérique différentiable), 206

Gram-Schmidt (théorème d'orthonormalisation de —), 331, 346

grand O, 60, 65

graphe

d'une application, 224
 d'une fonction réelle, 30

Grassmann (formule de —), 273

grossière (divergence —), 112

groupe(s), 229

abélien, 229
 alterné, 239
 automorphisme d'un —, 233
 commutatif, 229
 cyclique, 231
 des permutations, 235
 endomorphisme d'un —, 233
 image d'un morphisme de —, 233
 isomorphisme de —, 233
 linéaire, 266

monogène, 231
 morphisme de —, 233
 noyau d'un morphisme de —, 233
 ordre d'un élément dans un —, 234
 produit, 231
 sous- —, 229
 sous- — engendré par une partie, 230
 spécial orthogonal, 340
 symétrique, 235

H

harmonique (série —), 112

Heine (théorème de —), 49

hermitien(ne)

espace —, 345
 matrice —, 347
 matrice — définie positive, 347
 matrice — positive, 347
 produit scalaire —, 345

hessienne (matrice —), 212

homogène (équation —), 139, 141, 146–148

homothétie, 352

hypergéométrique (loi —), 377

hyperplan vectoriel, 263, 266

I

idéal

annulateur, 311
 d'un anneau, 242
 engendré, 296
 principal, 296

identité

de Bézout, 246, 297
 matrice —, 279

image

d'un ensemble par une application, 224
 d'un morphisme d'anneaux, 241
 d'un morphisme de groupes, 233

- d'une application, 224
- d'une application linéaire, 265
- directe d'un ensemble par une application, 224
- réciproque d'un ensemble par une application, 224
- impaire (fonction —), 30
- implication, 219
- inclusion, 221
- incomplète (théorème de la base —), 272
- indépendant(e)(s)
 - événements —, 374
 - événements mutuellement —, 375
 - variables aléatoires —, 381, 382, 393
 - variables aléatoires mutuellement —, 382, 394
- indicatrice
 - d'Euler, 253
 - fonction —, 367
- indirecte (base orthonormée —), 341
- inégalité
 - de Bessel, 180, 333
 - de Bienaymé-Tchebychev, 386, 399
 - de Boole, 371
 - de Jensen, 59, 385
 - de Markov, 385, 395
 - de Taylor-Lagrange pour les fonctions numériques, 80
 - de Taylor-Lagrange pour les fonctions vectorielles, 192
 - des accroissements finis pour les fonctions numériques, 55
 - des accroissements finis pour les fonctions vectorielles, 191
 - des pentes, 59
 - triangulaire pour les fonctions vectorielles, 190
 - triangulaire pour une distance, 84
 - triangulaire pour une norme, 83
- infimum (d'une partie de $\mathbb{R}$), 15
- infinie (dimension —), 271
- inflexion (point d'—), 60
- injective (application —), 225
- intégrabilité
 - par comparaison, 133
 - par domination, 133
- intégrable
 - fonction — au sens de Riemann, 81
 - fonction — pour les intégrales généralisées, 132
 - variable aléatoire réelle—, 396
 - variable aléatoire réelle positive —, 395
- intégrale(s)
 - d'Euler, 135
 - d'une fonction en escalier sur un segment, 81
 - d'une fonction vectorielle, 189
 - de Dirichlet, 133, 135
 - de Fresnel, 132, 135
 - de Gauss, 130, 135
 - de Riemann, 129
 - généralisées dépendant d'un paramètre, 137–138
 - généralisée semi-convergente, 133
- intégration
 - par changement de variable, 79
 - par changement de variable pour les intégrales généralisées, 131
 - par parties, 79
 - par parties pour les intégrales généralisées, 131
 - théorème fondamental de —, 76
 - théorème d'— terme à terme d'une série de fonctions, 136
 - théorème d'— terme à terme d'une série de fonctions sur un segment, 161
 - théorème d'— terme à terme d'une série entière, 170
- intègre (anneau —), 240
- intérieur, 91
 - point —, 36, 91
- interne (loi de composition —), 227

- intersection de deux parties d'un ensemble, 222
 - intervalle, 35
 - fermé, 36
 - image d'un — par une fonction continue, 105
 - ouvert, 36
 - semi-ouvert, 36
 - inverse
 - d'un élément, 228
 - d'une matrice inversible, 280
 - inversible
 - élément —, 228
 - matrice —, 280
 - irréductible (polynôme —), 303
 - isobarycentre, 360
 - isométrie
 - affine, 358
 - directe, 340
 - vectorielle, 339
 - isomorphisme
 - d'anneaux, 241
 - d'espaces vectoriels, 266
 - de corps, 244
 - de groupes, 233
 - isotrope
 - cône —, 322
 - vecteur —, 322
- ## J
- jacobien, 204
 - jacobienne (matrice —), 204
 - Jensen (inégalité de —), 59, 385
- ## K
- König (formule de —), 386, 398
 - Kronecker (symbole de —), 293
- ## L
- Lagrange
 - formule de Taylor- —, 80
 - inégalité de Taylor- — pour les fonctions numériques, 80
 - inégalité de Taylor- — pour les fonctions vectorielles, 192
 - polynôme d'interpolation de —, 306
 - théorème de —, 232
 - Laplace-Gauss (loi de —), 405
 - Lebesgue
 - lemme de Riemann- —, 181
 - théorème de convergence dominée de —, 135
 - Leibniz
 - formule de —, 53
 - règle de —, 118
 - théorème de —, 302
 - Lejeune-Dirichlet
 - condition de —, 179
 - espace de —, 179
 - fonction de —, 42, 68
 - intégrale de —, 133, 135
 - théorème de —, 183
 - lemme
 - d'Abel, 162
 - de décomposition des noyaux, 312
 - de Riemann-Lebesgue, 181
 - des trois pentes, 59
 - libre (famille —), 270
 - liée (famille —), 270
 - limite
 - caractérisation séquentielle de la —, 39, 95
 - d'une fonction vectorielle, 185
 - finie d'une fonction en un point, 37
 - finie d'une suite complexe, 23
 - finie d'une suite réelle, 18
 - finie par encadrement, 20, 40
 - finie à droite d'une fonction en un point, 38
 - finie à gauche d'une fonction en un point, 38

- infinie d'une fonction en un point, 37
 - infinie d'une suite réelle, 18
 - infinie par minoration ou majoration, 20, 40
 - simple d'une suite de fonctions, 150
 - théorème de la — monotone, 41
 - théorème de la double —, 153, 160
 - unicité de la —, 19, 38
 - uniforme d'une suite de fonctions, 150
 - linéaire
 - application —, 264
 - combinaison —, 262
 - forme —, 266
 - groupe —, 266
 - image d'une application —, 265
 - matrice d'une application —, 277
 - noyau d'une application —, 265
 - rang d'une application —, 276
 - transposée d'une application —, 267
 - lipschitzienne
 - application — dans des espaces vectoriels normés, 98
 - fonction —, 49
 - loi
 - associative, 227
 - binomiale, 373, 377
 - commutative, 227
 - conjointe, 381
 - convergence en —, 408
 - d'une variable aléatoire, 377, 393
 - de Bernoulli, 377
 - de composition interne, 227
 - de Gauss, 405
 - de Laplace-Gauss, 405
 - de Poisson, 373, 377
 - exponentielle, 404
 - faible des grands nombres, 401
 - géométrique, 373, 377
 - hypergéométrique, 377
 - marginale, 381
 - normale, 405
 - normale centrée réduite, 405
 - uniforme, 404
 - uniforme discrète, 377
- M
- majorant
 - d'une fonction, 31
 - d'une partie de $\mathbb{R}$, 15
 - d'une suite, 17
 - majoration (limite infinie par —), 20, 40
 - majorée
 - fonction —, 31
 - suite —, 17
 - marginale (loi —), 381
 - Markov (inégalité de —), 385, 395
 - matrice(s)
 - adjointe, 347
 - antisymétrique, 282
 - linéaire canoniquement associée à une —, 281
 - autoadjointe, 282
 - d'une application linéaire, 277
 - d'une forme bilinéaire, 318
 - de dilatation, 285
 - de passage, 283
 - de transvection, 285
 - diagonale, 282
 - diagonalisable, 309
 - élémentaire, 285
 - extraite, 291
 - hermitienne, 347
 - hermitienne définie positive, 347
 - hermitienne positive, 347
 - hessienne, 212
 - invertible, 280
 - jacobienne, 204
 - nilpotente, 313
 - normale, 347
 - orthogonale, 331, 337
 - orthogonale négative, 340

- orthogonale positive, 340
- polynôme caractéristique d'une —, 308
- produit de —, 279
- rang d'une —, 281
- semblables, 308
- sous- —, 291
- spectre d'une —, 308
- symétrique, 282
- symétrique définie positive, 338
- symétrique positive, 338
- trace d'une — carrée, 280
- transposée d'une —, 281
- triangulaire, 282
- trigonalisable, 310
- unitaire, 347
- valeur propre d'une —, 308
- vecteur propre d'une —, 308
- maximum
 - d'une fonction réelle, 31
 - local pour les fonctions de plusieurs variables, 210
 - local pour une fonction numérique, 54
- mesure de probabilité, 371
- méthode
 - de la variation de la constante, 140
 - de la variation des constantes, 147
 - du pivot de Gauss, 290
- métrique (espace —), 84
- mineur, 286
- minimal (polynôme —), 311
- minimum
 - d'une fonction réelle, 31
 - local pour les fonctions de plusieurs variables, 210
 - local pour une fonction numérique, 54
- Minkowski (théorème de —), 328, 346
- minorant
 - d'une fonction, 31
 - d'une partie de $\mathbb{R}$, 15
 - d'une suite, 17
- minoration (limite infinie par —), 20, 40
- minorée
 - fonction —, 31
 - suite —, 17
- mixte (produit —), 341
- module (d'un nombre complexe), 256
- Moivre (formule de —), 258
- Moivre-Laplace (théorème de —), 409
- moment (d'ordre r), 387, 399, 407
- Monge (notations de —), 214
- monogène (groupe —), 231
- monôme, 294
- monotone
 - fonction réelle —, 31
 - suite —, 18
- morceaux
 - continue par —, 47
 - de classe $\mathcal{C}^1$ par —, 182
- morphisme
 - d'algèbres, 265
 - d'anneaux, 241
 - de corps, 244
 - de groupes, 233
 - image d'un — d'anneaux, 241
 - image d'un — de groupes, 233
 - noyau d'un — d'anneaux, 241
 - noyau d'un — de groupes, 233
- moyenne
 - convergence en — quadratique, 181
 - norme de la convergence en —, 83
 - norme de la convergence en — quadratique, 83
 - valeur — d'une fonction sur un segment, 78
- multilinéaire (application —), 101
- multiple (d'un entier relatif), 245
- multiplicativité
 - sous- — de la norme d'opérateur, 101
 - sous- — de normes matricielles, 143
- multiplicité (d'une racine), 299

N

négation d'une assertion, 219
 négligeabilité
 relation de — pour les fonctions, 62
 relation de — pour les suites, 66
 neutre (élément —), 227
 Newton (formule du binôme de —), 240, 279
 nilpotent(e)
 endomorphisme —, 313
 matrice —, 313
 nombre complexe
 argument d'un —, 258
 conjugué d'un —, 256
 module d'un —, 256
 partie imaginaire d'un —, 256
 partie réelle d'un —, 256
 racine n -ième d'un —, 259
 nombre premier, 250
 non dégénérée (forme bilinéaire —), 322
 normale
 convergence — d'une série de fonctions, 159
 loi —, 405
 loi — centrée réduite, 405
 matrice —, 347
 normé
 espace vectoriel —, 83
 espace vectoriel — produit, 85
 norme(s), 83
 d'opérateur, 100
 de la convergence en moyenne, 83
 de la convergence en moyenne quadratique, 83
 de la convergence uniforme, 83, 151
 équivalentes, 88
 euclidienne, 84
 linéaire, 100
 matricielle, 143
 matricielle sous-multiplicative, 143

sous-multiplicativité de la — d'opérateur, 101
 subordonnée, 100

noyau(x)
 d'un morphisme d'anneaux, 241
 d'un morphisme de groupes, 233
 d'une application linéaire, 265
 d'une forme bilinéaire, 322
 lemme de décomposition des —, 312

O

optimalité
 condition nécessaire d'—, 213
 condition suffisante d'—, 214
 optimisation (théorème d'— sous contrainte), 211
 orbites (d'une permutation), 238
 ordonné
 ensemble —, 226
 ensemble totalement —, 226
 ordre
 d'un élément dans un groupe, 234
 racine d'— m , 299
 relation d'—, 226
 relation d'— total, 226
 orthogonal(e), 321
 automorphisme — négatif, 340
 automorphisme — positif, 340
 base — pour une forme quadratique, 321
 d'une partie d'un espace préhilbertien, 330, 346
 d'une partie d'un espace vectoriel, 267
 d'une partie du dual d'un espace vectoriel, 267
 famille —, 330
 groupe spécial —, 340
 matrice —, 331, 337
 matrice — négative, 340
 matrice — positive, 340
 projection —, 333

projection affine —, 358
 symétrie —, 336
 symétrie affine —, 358
 orthogonalité
 d'une forme linéaire et d'un vecteur, 267
 de deux vecteurs, 321, 329, 346
 orthonormale (famille —), 330
 orthonormalisation de Gram-Schmidt, 331, 346
 orthonormé(e)
 famille —, 330
 repère —, 355
 ouvert(e)
 boule —, 87
 disque — de convergence, 164
 ensemble —, 89
 intervalle —, 36
 intervalle — de convergence, 164
 relatif, 93

P

p-combinaison, 367
p-liste, 367
p-uplet, 367
 paire (fonction —), 30
 paquets (sommation par —), 124
 parallélisme (relation de —), 350
 parallélogramme (égalité du —), 328
 Parseval (égalité de —), 182
 partie entière, 32
 partition d'un ensemble, 223
 passage (matrice de —), 283
 pentes (inégalité des —), 59
 périodique (fonction numérique —), 30
 permutation(s)
 d'un ensemble, 225
 groupe des —, 235
 impaire, 239
 orbites d'une —, 238
 paire, 239
 signature d'une —, 238
 support d'une —, 236
 perpendiculaires (droites —), 355
 petit o, 62, 66
 PGCD, 246, 297
 pivot de Gauss (méthode du —), 290
 plan
 affine, 349
 tangent, 210
 vectoriel, 271
 plus grand commun diviseur
 de r polynômes, 297
 de deux entiers, 246
 plus petit commun multiple (de deux entiers), 248
 Poincaré (théorème de —), 366, 372
 point(s)
 adhérent, 36, 91
 caractérisation séquentielle des —
 adhérents, 91
 critique pour les fonctions de plusieurs variables, 210
 critique pour les fonctions numériques, 54
 d'inflexion, 60, 194
 de rebroussement de deuxième espèce, 194
 de rebroussement de première espèce, 194
 fixe, 28, 109
 intérieur, 36, 91
 ordinaire, 194
 régulier, 193
 singulier, 193
 stationnaire, 193
 théorème du — fixe, 109
 Poisson (loi de —), 373, 377
 polaire (forme —), 319
 pôle (d'une fraction rationnelle), 305
 polynôme(s), 293
 annulateur, 311
 associés, 296

- caractéristique d'un endomorphisme, produit
 - 307
 - caractéristique d'une matrice, 308
 - composée de deux —, 295
 - constant, 294
 - d'interpolation de Lagrange, 306
 - degré d'un —, 294
 - dérivé, 301
 - discriminant d'un — de degré 2 à
 - coefficients complexes, 257
 - échelonnés en degrés, 295
 - étagés, 295
 - irréductible, 303
 - minimal, 311
 - premiers entre eux, 298
 - racine d'un —, 299
 - scindé, 300
 - simplement scindé, 300
 - unitaire, 295
 - valuation d'un —, 294
- postulat, 217
 - PPCM, 248
 - préhilbertien (espace —), 327
 - premier(s)
 - entiers — entre eux, 247
 - nombre —, 250
 - polynômes — entre eux, 298
 - presque sûrement, 372
 - primitive(s)
 - d'une fonction, 74
 - d'une fonction vectorielle, 190
 - usuelles, 75, 416
 - principe de superposition, 140, 142
 - probabilisé (espace —), 371
 - probabilité(s), 371
 - conditionnelle, 373
 - convergence en —, 400
 - densité de —, 402
 - formule des — composées, 374
 - formule des — totales, 374
 - uniforme, 373
 - problème de Cauchy, 140
 - cartésien, 223
 - d'espaces vectoriels normés, 85
 - de Cauchy de séries entières, 167
 - de Cauchy de séries numériques, 120
 - de convolution, 383, 403
 - de matrices, 279
 - groupe —, 231
 - mixte, 341
 - scalaire, 327
 - scalaire hermitien, 345
 - vectoriel, 342
- projecteur, 269
 - projecteurs spectraux, 315
 - projection
 - affine, 353
 - affine orthogonale, 358
 - orthogonale, 333
 - théorème de — orthogonale, 333
 - vectorielle, 269
 - prolongement
 - d'une fonction, 30
 - par continuité, 43
 - propre
 - sous-espace —, 307, 308
 - valeur — d'un endomorphisme, 307
 - valeur — d'une matrice, 308
 - vecteur — d'un endomorphisme, 307
 - vecteur — d'une matrice, 308
 - Pythagore (théorème de —), 330
- ## Q
- quadratique
 - base orthogonale pour une forme —, 321
 - convergence en moyenne —, 181
 - forme —, 319
 - forme — définie positive, 324
 - forme — non dégénérée, 322
 - forme — positive, 324
 - norme de la convergence en moyenne —, 83

rang d'une forme —, 323
 signature d'une forme —, 324
 quadrique, 326
 quantificateur
 existentiel, 221
 universel, 221
 quotient, 245
 ensemble —, 226

R

Raabe-Duhamel (règle de —), 116
 racine
 n -ième d'un nombre complexe, 259
 n -ième de l'unité, 260
 carrée d'un nombre complexe, 257
 d'ordre m d'un polynôme, 299
 d'un polynôme, 299
 radial (théorème — d'Abel), 169
 rang
 d'une application linéaire, 276
 d'une famille de vecteurs, 275
 d'une forme bilinéaire symétrique, 323
 d'une matrice, 281
 théorème du —, 276
 rationnelle
 fonction —, 304
 fraction —, 304
 pôle d'une fraction —, 305
 rayon de convergence, 163
 réciproque
 bijection —, 225
 image — d'un ensemble par une application, 224
 recouvrement disjoint, 223
 récurrence
 théorème de — double, 218
 théorème de — forte, 218
 théorème de — simple, 218
 réduction de Gauss, 320
 réflexion, 336
 réflexion affine, 358
 réflexive (relation —), 226
 règle(s)
 d'Abel pour les séries numériques, 119
 de Cauchy pour les séries entières, 165
 de Cauchy pour les séries numériques, 116
 de d'Alembert pour les séries entières, 164
 de d'Alembert pour les séries numériques, 115
 de Duhamel, 116
 de la chaîne, 200
 de Leibniz, 118
 de Raabe-Duhamel, 116
 de Riemann, 114
 de Riemann généralisée, 115
 régularisée (d'une fonction), 179
 régulier (élément —), 227
 régulière (partie —), 68
 relatif(s)
 caractérisation séquentielle des fermés —, 94
 fermé —, 93
 ouvert —, 93
 voisinage —, 93
 relation
 antisymétrique, 226
 binaire, 226
 d'équivalence, 226
 d'ordre, 226
 d'ordre total, 226
 de Chasles, 349
 de congruence, 242, 245
 de divisibilité, 245
 de parallélisme, 350
 réflexive, 226
 symétrique, 226
 transitive, 226

- répartition (fonction de —), 380, 390, 393
 - repère
 - affine, 350
 - orthonormé, 355
 - reste, 245
 - d'une série numérique convergente, 111
 - d'une série de fonctions, 157
 - formule de Taylor avec — intégral pour les fonctions numériques, 80
 - formule de Taylor avec — intégral pour les fonctions vectorielles, 191
 - restriction (d'une fonction), 30
 - retournement, 336
 - réunion de deux parties d'un ensemble, 222
 - Riemann
 - fonction — -intégrable, 81
 - fonction intégrable au sens de —, 81
 - intégrales de —, 129
 - lemme de — -Lebesgue, 181
 - règle de —, 114
 - règle de — généralisée, 115
 - séries de —, 114
 - sommes de — pour une fonction numérique, 79
 - sommes de — pour une fonction vectorielle, 190
 - Rolle (théorème de —), 55
 - rotation, 344
 - rotation vectorielle, 340, 344
 - réflexion, 344
- S
- scalaire
 - produit —, 327
 - produit — hermitien, 345
 - Schwarz (théorème de —), 212
 - scindé
 - polynôme —, 300
 - polynôme simplement —, 300
 - sécante, 58
 - segment, 35
 - semblables (matrices —), 308
 - semi-ouvert (intervalle —), 36
 - séquentielle
 - caractérisation — de la continuité, 44, 97
 - caractérisation — de la limite, 39, 95
 - caractérisation — des fermés, 92
 - caractérisation — des fermés relatifs, 94
 - caractérisation — des points adhérents, 91
 - série(s)
 - à termes positifs, 113
 - absolument convergente, 117
 - alternée, 118
 - commutativement convergente, 124
 - comparaison — -intégrale, 120, 130
 - critères de comparaison, 114, 165
 - de Bertrand, 115
 - de Riemann, 114
 - de Taylor, 173
 - dérivée d'une — entière, 167
 - développements en — entières usuels, 174, 417
 - fonction développable en — entière, 171
 - géométrique, 111
 - harmonique, 112
 - harmonique alternée, 117
 - numérique, 110
 - primitive d'une — entière, 167
 - reste d'une — de fonctions, 157
 - reste d'une — numérique convergente, 111
 - somme d'une — numérique convergente, 111

- sommes de — convergentes classiques, 120
- télescopique, 113
- trigonométrique, 176
- vectorielle, 143
- signature
 - d'une forme quadratique, 324
 - d'une permutation, 238
- simple
 - convergence — d'une suite de fonctions, 150
 - convergence — d'une série de fonctions, 156
 - limite — d'une suite de fonctions, 150
- simplifiable (élément —), 227
- singleton, 220
- sinus hyperbolique, 33
- solutions
 - d'une équation différentielle linéaire d'ordre 1, 139
 - d'une équation différentielle linéaire d'ordre 2 à coefficients constants, 142
 - de l'équation homogène associée à une équation différentielle linéaire d'ordre 2, 141
- sommable
 - famille — de réels ou complexes, 122
 - famille — de réels positifs, 121
 - variable aléatoire discrète de carré —, 385
- sommation
 - des relations de comparaison, 119
 - par paquets, 124
- somme directe
 - de p sous-espaces vectoriels, 269
 - de deux sous-espaces vectoriels, 268
- somme(s)
 - d'une famille sommable de réels, 123
 - d'une famille sommable de réels positifs, 121
 - d'une famille sommable de complexes, 123
 - d'une série numérique convergente, 111
 - d'une série de fonctions, 157
 - de la série harmonique alternée, 120
 - de Riemann pour une fonction numérique, 79
 - de Riemann pour une fonction vectorielle, 190
 - de séries convergentes classiques, 120
 - partielles d'une série numérique, 110
 - partielles d'une série vectorielle, 143
- sous-algèbre, 265
- sous-anneau, 241
- sous-corps, 243
- sous-espace(s)
 - affine, 350
 - caractéristiques, 314
 - propre, 307, 308
 - somme directe de p — vectoriels, 269
 - somme directe de deux — vectoriels, 268
 - supplémentaire d'un — vectoriel, 268
 - vectoriel, 262
 - vectoriel stable, 308
- sous-groupe, 229
- sous-matrice, 291
- spectral (théorème —), 337
- spectre
 - d'un endomorphisme, 307
 - d'une matrice, 308
- sphère, 87, 355
- stable
 - intervalle — par une fonction, 28
 - sous-espace vectoriel —, 308
- stationnaire (suite —), 17
- Stirling (formule de —), 68
- subdivision, 47

- adaptée à une fonction en escalier, 80
 - pas d'une —, 47
 - régulière, 47
 - suite(s)
 - adjacentes, 21
 - arithmético-géométrique, 26
 - arithmétique, 25
 - bornée, 17
 - bornée dans un espace vectoriel normé, 86
 - constante, 17
 - de Cauchy, 107
 - équivalentes, 67
 - extraite, 24
 - extraite dans un espace vectoriel normé, 86
 - géométrique, 25
 - majorée, 17
 - minorée, 17
 - monotone, 18
 - récurrente linéaire d'ordre 2, 27
 - sous- —, 24
 - stationnaire, 17
 - support d'une —, 293
 - supplémentaire (d'un sous-espace vectoriel), 268
 - support
 - d'un arc paramétré, 192
 - d'un cycle, 236
 - d'une permutation, 236
 - d'une suite, 293
 - supremum (d'une partie de $\mathbb{R}$), 15
 - surjective (application —), 225
 - symétrie
 - affine, 353
 - affine orthogonale, 358
 - orthogonale, 336
 - vectorielle, 269
 - symétrique
 - endomorphisme — défini positif, 338
 - endomorphisme — positif, 338
 - endomorphisme —, 337
 - forme bilinéaire —, 317
 - groupe —, 235
 - matrice —, 282
 - matrice — définie positive, 338
 - matrice — positive, 338
 - relation —, 226
 - système
 - complet d'événements, 372
 - de Cramer, 290
- ## T
- tangent
 - plan —, 210
 - vecteur — à une partie d'un espace normé, 208
 - tangente hyperbolique, 33
 - tautologie, 217
 - taux de panne, 378
 - Taylor
 - formule de —, 302
 - formule de — avec reste intégral pour les fonctions numériques, 80
 - formule de — avec reste intégral pour les fonctions vectorielles, 191
 - séries de —, 173
 - théorème de —, 302
 - Taylor-Lagrange
 - formule de —, 80
 - inégalité de — pour les fonctions numériques, 80
 - inégalité de — pour les fonctions vectorielles, 192
 - Taylor-Young
 - formule de — pour les fonctions de plusieurs variables, 213
 - formule de — pour les fonctions numériques, 71
 - formule de — pour les fonctions vectorielles, 192
 - Thalès (théorème de —), 353
 - théorème(s)

- d'Abel, 162
- d'approximation uniforme, 155
- d'Euclide, 251, 303
- d'Euler, 253
- d'intégrabilité par comparaison, 133
- d'intégrabilité par domination, 133
- d'intégration terme à terme d'une série de fonctions, 136
- d'intégration terme à terme d'une série de fonctions sur un segment, 161
- d'intégration terme à terme d'une série entière, 170
- d'optimisation sous contrainte, 211
- d'orthonormalisation de Gram-Schmidt, 331, 346
- de Bayes, 374
- de Beppo Levi, 395
- de Bézout, 247, 298
- de Bienaymé-Tchebychev, 386, 399
- de Bolzano-Weierstrass, 25, 106
- de Cantor, 368
- de Cauchy-Schwarz, 328, 346
- de Cayley-Hamilton, 312
- de comparaison série-intégrale, 120, 130
- de conjugaison pour les intégrales généralisées, 128
- de convergence commutative, 124
- de convergence dominée de Lebesgue, 135
- de croissance pour les intégrales généralisées, 128
- de d'Alembert-Gauss, 301
- de décomposition des noyaux, 312
- de Dirichlet, 183
- de Dunford, 315
- de dérivation terme à terme d'une série de fonctions, 161
- de dérivation terme à terme d'une série entière, 170
- de Fubini, 125
- de Gauss, 248, 298
- de Heine, 49
- de Jensen, 385
- de König, 398
- de la base incomplète, 272
- de la double limite, 153, 160
- de la limite centrale, 409
- de la limite de la dérivée, 57
- de la limite monotone, 41
- de Lagrange, 232
- de Leibniz, 118, 302
- de Lejeune-Dirichlet, 183
- de linéarité pour les intégrales généralisées, 128
- de Markov, 385, 395
- de Minkowski, 328, 346
- de Moivre-Laplace, 409
- de Parseval, 182
- de Poincaré, 366, 372
- de projection orthogonale, 333
- de Pythagore, 330
- de Raabe-Duhamel, 116
- de récurrence double, 218
- de récurrence forte, 218
- de récurrence simple, 218
- de réduction de Gauss, 320
- de Rolle, 55
- de Schwarz, 212
- de sommation par paquets, 124
- de Taylor, 302
- de Taylor avec reste intégral pour les fonctions numériques, 80
- de Taylor avec reste intégral pour les fonctions vectorielles, 191
- de Taylor-Young pour les fonctions de plusieurs variables, 213
- de Taylor-Young pour les fonctions numériques, 71
- de Taylor-Young pour les fonctions vectorielles, 192
- de Thalès, 353
- de transfert, 384, 407

de Weierstrass, 155
 de Wilson, 252
 des accroissements finis, 55
 des bornes atteintes, 46, 103
 des croissances comparées, 63
 des restes chinois, 254
 des valeurs intermédiaires, 44, 105
 du point fixe, 109
 du rang, 276
 fondamental de l'algèbre, 301
 fondamental de l'analyse, 76, 191
 fondamental de l'intégration, 76
 petit — de Fermat, 252
 radial d'Abel, 169
 spectral, 337
 total (relation d'ordre —), 226
 trace
 d'un endomorphisme, 280
 d'une matrice carrée, 280
 trajectoire (d'un arc paramétré), 192
 transfert (formule de —), 384, 407
 transitive (relation —), 226
 translation, 350
 transposée
 d'une application linéaire, 267
 d'une matrice, 281
 transposition, 236
 triangulaire (matrice —), 282
 tribu, 369
 borélienne, 370
 de Bernoulli, 369
 engendrée, 370
 grossière, 369
 triviale, 369
 trigonalisable
 endomorphisme —, 309
 matrice —, 310
 trigonométrie (formules de —), 414
 troncature (d'un développement limité),
 69

U

uniforme
 approximation —, 155
 continuité —, 48
 continuité — dans des espaces vectoriels normés, 98
 convergence — d'une suite de fonctions, 150
 convergence — d'une série de fonctions, 157
 limite — d'une suite de fonctions, 150
 loi —, 404
 loi — discrète, 377
 norme de la convergence —, 83
 unitaire
 matrice —, 347
 polynôme —, 295
 univers, 368

V

valeur
 d'adhérence, 86
 moyenne d'une fonction sur un segment, 78
 propre d'un endomorphisme, 307
 propre d'une matrice, 308
 valuation
 p -adique, 251
 d'un polynôme, 294
 Vandermonde (déterminant de —), 288
 variable(s) aléatoire(s)
 à densité, 402
 continue, 402
 de carré intégrable, 397
 discrète, 376
 discrète de carré sommable, 385
 espérance d'une — discrète, 383
 espérance d'une — réelle, 396

- espérance d'une — réelle positive, 395
 - étagée, 376
 - indépendantes, 381, 382, 393
 - loi d'une —, 377, 393
 - mutuellement indépendantes, 382, 394
 - non corrélées, 387, 398
 - réelle, 391
 - réelle intégrable, 396
 - réelle positive intégrable, 395
 - sans mémoire, 379, 404
 - variance, 386, 398
 - variation
 - méthode de la — de la constante, 140
 - méthode de la — des constantes, 147
 - vecteur(s)
 - isotrope, 322
 - orthogonaux, 321, 329, 346
 - propre d'un endomorphisme, 307
 - propre d'une matrice, 308
 - tangent à une partie d'un espace normé, 208
 - vectériel(s)
 - automorphisme d'un espace —, 266
 - base d'un espace —, 270
 - dimension d'un espace —, 271
 - endomorphisme d'un espace —, 264
 - espace —, 261
 - espace — normé, 83
 - espace — normé produit, 85
 - hyperplan —, 263, 266
 - isomorphisme d'espaces —, 266
 - produit —, 342
 - somme directe de p sous-espaces —, 269
 - somme directe de deux sous-espaces —, 268
 - sous-espace —, 262
 - sous-espace — engendré, 262
 - sous-espace — stable, 308
 - supplémentaire d'un sous-espace —, 268
 - Viète (formules de —), 301
 - voisinage
 - d'un point d'un espace vectoriel normé, 89
 - d'un réel, 36
 - de $+\infty$, 36
 - de $-\infty$, 36
 - relatif, 93
- ## W
- Weierstrass
 - théorème de —, 155
 - théorème de Bolzano- —, 25, 106
 - Wilson (théorème de —), 252
 - wronskien, 146
- ## Y
- Young
 - formule de Taylor- — pour les fonctions de plusieurs variables, 213
 - formule de Taylor- — pour les fonctions numériques, 71
 - formule de Taylor- — pour les fonctions vectorielles, 192

Enrichi par des **résumés de cours**, ce formulaire sera d'une aide précieuse pour l'étudiant souhaitant s'assurer de l'assimilation des **nouveaux programmes de mathématiques des classes préparatoires aux grandes écoles** comme ceux des **deux premières années de Licences scientifiques**.

Il s'avérera également très utile aux **candidats au Capes de mathématiques**.

Le livre est divisé en trois parties : analyse, algèbre et probabilités. Un **index détaillé** permet à l'étudiant de trouver immédiatement l'information recherchée. En fin d'ouvrage sont rassemblés un formulaire de trigonométrie, les dérivées et primitives usuelles ainsi que les développements limités et les développements en séries entières classiques. Remarques, exemples et contre-exemples, viennent compléter l'ensemble.

Les auteurs

Olivier Rodot est directeur du département de mathématiques des classes préparatoires de l'EPITA, école d'ingénieurs en informatique.

Jean-Étienne Rombaldi a enseigné à l'université Grenoble-Alpes. Il a été préparateur et membre du jury du Capes externe et de l'agrégation interne de mathématiques.

ISBN : 978-2-8073-3988-0



9 782807 339880

deboeck **B**
SUPÉRIEUR

www.deboecksuperieur.com



