

Maths – MPI

Romain Bricout

19 octobre 2025

Introduction

Ce document réunit l'ensemble de mes cours de Mathématiques de MPI, ainsi que les exercices les accompagnant. Le professeur était M. Walbron. J'ai adapté certaines formulations me paraissant floues ou ne me plaisant pas mais le contenu pur des cours est strictement équivalent.

Les éléments des tables des matières initiale et présentes au début de chaque chapitre sont cliquables (amenant directement à la partie cliquée). C'est également le cas des références à des éléments antérieurs de la forme, par exemple, « Démonstration 5.22 ».

Cette version contient, en plus des cours imprimés distribués durant l'année, toutes les démonstrations qui vont avec. Voir l'autre version pour n'avoir que les cours bruts.

Table des matières

I	Cours	17
1	Espaces vectoriels normés	18
1.1	Bornes supérieures, bornes inférieures	19
1.1.1	Borne supérieure d'une partie de \mathbb{R}	19
1.1.2	Borne supérieure d'une application à valeurs dans \mathbb{R}	21
1.1.3	Règles pratiques	21
1.2	Normes.	22
1.2.1	Définition	22
1.2.2	Exemples fondamentaux	23
1.2.3	Normes équivalentes	25
1.2.4	Boules	28
1.2.5	Parties bornées	31
1.3	Convergence des suites	36
1.3.1	Définition	36
1.3.2	Propriétés usuelles	37
1.3.3	Cas particulier en dimension finie	39
1.3.4	Point adhérent à une partie	41
1.4	Limites de fonctions	44
1.4.1	Définition	44
1.4.2	Caractérisation séquentielle de la limite	44
1.4.3	Propriétés usuelles	45
1.4.4	Cas particulier de la dimension finie	46
1.4.5	Composition des limites	47

1.4.6	Extensions des définitions	47
1.5	Fonctions continues	48
1.5.1	Continuité en un point	48
1.5.2	Continuité sur une partie	49
1.5.3	Cas particulier de la dimension finie	49
1.5.4	Fonctions lipschitziennes	50
1.5.5	Continuité des applications linéaires et n -linéaires	52
1.5.6	Norme subordonnée	59
1.6	Topologie d'un espace vectoriel normé.	63
1.6.1	Intérieur d'une partie, voisinage d'un point	63
1.6.2	Parties ouvertes	65
1.6.3	Parties fermées	66
1.6.4	Ouverts ou fermés relatifs à une partie	71
1.6.5	Image réciproque d'un ouvert ou d'un fermé par une fonction continue	71
1.6.6	Frontière d'une partie	74
1.7	Compacité	74
1.7.1	Valeurs d'adhérence d'une suite	74
1.7.2	Théorème de Bolzano-Weierstrass	76
1.7.3	Parties compactes	78
1.7.4	Théorème des bornes atteintes	82
1.8	Connexité par arcs	87
1.8.1	Chemin	87
1.8.2	Parties connexes par arcs	88
1.8.3	Théorème des valeurs intermédiaires	89

2 Séries numériques et vectorielles : révisions et compléments 91

2.1	Rappels	91
2.1.1	Définitions et notations	91
2.1.2	Convergence d'une série	92
2.1.3	Lien entre convergence de suites et convergence de séries	93

2.2	Séries réelles à termes positifs	94
2.2.1	Théorème de Cesàro	98
2.2.2	Théorème de comparaison par domination de séries à termes positifs	99
2.2.3	Théorème de comparaison par équivalence de séries à termes positifs	101
2.2.4	Théorème de comparaison série - intégrale	103
2.3	Séries absolument convergentes	107
2.3.1	Lien entre absolue convergence et convergence	107
2.3.2	Un exemple fondamental : l'exponentielle de matrice	108
2.3.3	Extension des résultats par comparaison	108
2.3.4	Produit de Cauchy de deux séries absolument convergentes	110
2.4	Séries alternées	112
3	Familles sommables	114
3.1	Sommes finies	114
3.1.1	Définition	114
3.1.2	Propriétés	117
3.2	Conventions de calcul dans $\mathbb{R}_+ \cup \{+\infty\}$	118
3.3	Somme d'une famille de réels positifs	119
3.3.1	Propriétés	120
3.3.2	Théorème de sommation par paquets	121
3.3.3	Théorème de Fubini	121
3.4	Familles sommables dans un espace vectoriel normé de dimension finie	122
3.4.1	Définitions	122
3.4.2	Propriétés	124
3.4.3	Théorème de sommation par paquets	125
3.4.4	Théorème de Fubini	127
3.4.5	Produit de Cauchy de deux séries	129
4	Rappels et compléments d'algèbre linéaire	131

4.1	Sommes de sous-espaces vectoriels	132
4.1.1	Généralités	132
4.1.2	Sommes directes	132
4.1.3	Sous-espaces supplémentaires	134
4.1.4	Cas particulier de deux sous-espaces	138
4.1.5	Applications linéaires et sommes directes	138
4.2	Somme de sous-espaces vectoriels en dimension finie.	139
4.2.1	Base adaptée à un sous-espace	139
4.2.2	Sommes directes et bases	140
4.2.3	Dimension d'une somme de sous-espaces vectoriels	140
4.2.4	Sous-espaces supplémentaires	142
4.2.5	Dimension d'une somme de deux sous-espaces vectoriels	144
4.3	Polynômes d'endomorphismes et de matrices.	144
4.3.1	\mathbb{K} -algèbres	144
4.3.2	Cas particulier des algèbres $\mathcal{L}(E)$ ou $\mathcal{M}_n(\mathbb{K})$	146
4.3.3	Polynôme annulateur d'une matrice ou d'un endomorphisme	146
4.3.4	Utilisation pratique d'un polynôme annulateur	149
4.4	Matrices semblables, trace.	156
4.4.1	Trace d'une matrice	156
4.4.2	Matrices semblables	156
4.4.3	Trace d'un endomorphisme	157
4.5	Opérations par blocs	158
4.5.1	Cas général	158
4.5.2	Cas particuliers des matrices carrées	159
4.5.3	Interprétation des blocs	159

5 Réduction des endomorphismes 162

5.1	Éléments propres d'un endomorphisme	163
5.1.1	Valeurs propres et vecteurs propres	163
5.1.2	Lien avec les polynômes annulateurs	164

5.1.3	Sous-espaces propres	166
5.2	Polynôme caractéristique d'un endomorphisme	168
5.2.1	Caractérisation des valeurs propres en dimension finie	168
5.2.2	Définition et lien avec les valeurs propres	168
5.2.3	Ordre de multiplicité et dimension du sous-espace propre	171
5.2.4	Endomorphisme scindé	172
5.3	Éléments propres d'une matrice carrée	173
5.3.1	Valeurs propres et vecteurs propres	173
5.3.2	Lien avec les polynômes annulateurs	174
5.3.3	Sous-espaces propres	174
5.4	Polynôme caractéristique d'une matrice carrée	175
5.4.1	Définition et lien avec les valeurs propres	175
5.4.2	Ordre de multiplicité et dimension du sous-espace propre	177
5.4.3	Matrice scindée	177
5.5	Endomorphismes diagonalisables, matrices diagonalisables	178
5.5.1	Définition	178
5.5.2	Caractérisations équivalentes	179
5.5.3	Lien avec le polynôme caractéristique	181
5.6	Lien entre diagonalisabilité et polynômes annulateurs	182
5.6.1	Racines du polynôme minimal	182
5.6.2	Lemme des noyaux	183
5.6.3	Application à la diagonalisabilité	185
5.6.4	Diagonalisabilité d'un endomorphisme induit	188
5.7	Quelques applications de la diagonalisation	189
5.7.1	Puissances d'une matrice, suites récurrentes linéairement	189
5.7.2	Systèmes d'équations différentielles	190
5.8	Endomorphismes trigonalisables, matrices trigonalisables	190
5.8.1	Définition et propriétés	190
5.8.2	Caractérisation équivalente	191

5.8.3	Théorème de Cayley-Hamilton	194
5.8.4	Sous-espaces caractéristiques	195
5.9	Endomorphismes nilpotents, matrices nilpotentes	197
5.9.1	Généralités	197
5.9.2	Éléments propres d'un nilpotent	199
5.9.3	Application aux sous-espaces caractéristiques d'un endomorphisme	199
6	Intégrales généralisées	202
	Fonctions continues par morceaux sur un intervalle	203
6.1	Intégrales généralisées sur $[a ; +\infty[$	203
6.1.1	Définition et exemples fondamentaux	203
6.1.2	Propriétés	205
6.1.3	Cas des fonctions réelles positives	207
6.1.4	Théorème de comparaison entre fonctions positives	208
6.1.5	Lien avec les séries	211
6.2	Intégrales généralisées sur d'autres types d'intervalles	215
6.2.1	Intégrales généralisées sur $[a ; b[$	215
6.2.2	Intégrales généralisées sur $]a ; b]$	215
6.2.3	Intégrales généralisées sur $]a ; b[$	217
6.2.4	Propriétés communes à toutes ces intégrales	219
6.3	Résumé pour étudier la convergence d'une intégrale	225
6.4	Fonctions intégrables sur un intervalle.	227
6.4.1	Intégrales absolument convergentes	227
6.4.2	Fonctions intégrables	229
6.4.3	Théorème de comparaison des fonctions intégrables	230
6.5	Intégration des relations de comparaison	231
6.5.1	Théorème de comparaison par domination	232
6.5.2	Théorème de comparaison par équivalence	233
7	Intégrales à paramètre	235

7.1	Introduction	235
7.2	Convergence simple	236
7.2.1	Convergence simple d'une suite de fonctions	236
7.2.2	Convergence simple d'une série de fonctions	238
7.3	Suites et séries de fonctions intégrables	239
7.3.1	Théorème de convergence dominée	239
7.3.2	Théorème d'intégration terme à terme	242
7.4	Fonctions définies par une intégrale à paramètre	244
7.4.1	Continuité	244
7.4.2	Dérivabilité	245
7.5	Domination sur des sous-intervalles	249
7.6	Complément : la fonction Γ d'Euler.	252
8	Espaces préhilbertiens réels	254
8.1	Généralités	254
8.1.1	Produit scalaire	254
8.1.2	Exemples fondamentaux	255
8.1.3	Norme euclidienne	256
8.1.4	Vecteurs orthogonaux	259
8.2	Bases orthonormées	260
8.2.1	Familles orthonormées	260
8.2.2	Existence de bases orthonormées	261
8.2.3	Calculs en base orthonormée	263
8.3	Sous-espaces orthogonaux	263
8.3.1	Orthogonalité de deux sous-espaces vectoriels	263
8.3.2	Orthogonal d'un sous-espace vectoriel	264
8.4	Projection orthogonale sur un sous-espace vectoriel de dimension finie	266
8.4.1	Projection orthogonale	266
8.4.2	Distance à un sous-espace vectoriel	267

9 Endomorphismes dans un espace euclidien 269

9.1	Adjoint d'un endomorphisme	269
9.1.1	Représentation des formes linéaires	269
9.1.2	Adjoint	270
9.1.3	Matrice de l'adjoint	274
9.1.4	Stabilité de sous-espaces vectoriels	276
9.2	Orientation d'un \mathbb{R} -espace vectoriel de dimension finie	276
9.3	Isométries vectorielles	277
9.4	Matrices orthogonales	280
9.4.1	Déterminant d'une isométrie vectorielle	283
9.4.2	Changements de bases orthonormées	284
9.4.3	Produit mixte	285
9.4.4	Produit vectoriel en dimension 3	286
9.5	Étude en dimension 2.	287
9.6	Réduction des isométries vectorielles ou des matrices orthogonales.	289
9.6.1	Réduction des isométries vectorielles	289
9.6.2	Réduction des matrices orthogonales	294
9.6.3	Étude en dimension 3	294
9.7	Endomorphismes auto-adjoints.	300
9.7.1	Définition et propriétés	300
9.7.2	Théorème spectral	301
9.8	Endomorphismes auto-adjoints positifs, définis-positifs.	304
9.8.1	Endomorphismes auto-adjoints positifs	304
9.8.2	Matrices symétriques positives	306

10 Fonctions vectorielles 307

10.1	Dérivée en un point	307
10.1.1	Dérivabilité en un point	307
10.1.2	Interprétation géométrique, développement limité d'ordre 1, continuité	308
10.1.3	Dérivées à gauche, dérivées à droite	309

10.1.4	Lien avec les coordonnées	309
10.1.5	Théorèmes opératoires	310
10.2	Fonction dérivée	312
10.3	Dérivées successives	315
10.3.1	Définitions et exemples	315
10.3.2	Théorèmes opératoires pour les dérivées successives	316
10.4	Intégrales	317
10.4.1	Définition	317
10.4.2	Propriétés	318
10.4.3	Primitives d'une fonction continue	319
10.4.4	Formules de Taylor	320
11	Suites et séries de fonctions	321
11.1	Convergence d'une suite de fonctions	321
11.1.1	Convergence simple	322
11.1.2	Convergence uniforme	322
11.2	Convergence d'une série de fonctions	329
11.2.1	Convergence simple	329
11.2.2	Convergence uniforme	331
11.2.3	Convergence normale	333
11.3	Propriétés de la fonction limite.	337
11.3.1	Monotonie	337
11.3.2	Continuité	338
11.3.3	Interversion de limite et d'intégrale	346
11.3.4	Interversion de limites	351
11.3.5	Dérivabilité	353
11.3.6	Dérivation à un ordre plus élevé	356
11.4	Généralisation	357
11.4.1	Convergence simple	357
11.4.2	Convergence uniforme	357

11.4.3	Convergence normale des séries	358
11.4.4	Résultats préservés	358
11.5	Approximation uniforme	358
11.5.1	Densité des fonctions en escaliers dans les fonctions continues par morceaux .	358
11.5.2	Densité des polynômes sur un segment dans les fonctions continues	360
12	Séries entières	362
12.1	Convergence simple d'une série entière	362
12.1.1	Rayon de convergence	362
12.1.2	Détermination du rayon de convergence	364
12.1.3	Comparaison de séries entières	368
12.1.4	Opérations sur les séries entières	371
12.2	Propriétés de la fonction somme d'une série entière	373
12.2.1	Convergence uniforme et continuité	373
12.2.2	Primitivation et dérivation	374
12.2.3	Convergence radiale	378
12.3	Fonction développable en série entière.	379
12.3.1	Généralités	379
12.3.2	Unicité du développement en série entière	382
12.3.3	Série de Taylor d'une fonction	383
12.3.4	Développements en série entière usuels	386
13	Probabilités	388
13.1	Dénombrabilité	388
13.1.1	Vocabulaire	388
13.1.2	Exemples	389
13.1.3	Quelques propriétés	392
13.2	Espace probabilisé	393
13.2.1	Univers d'une expérience aléatoire	393
13.2.2	Tribu d'événements	393

13.2.3	Probabilité	395
13.2.4	Propriétés	396
13.2.5	Probabilité discrète	402
13.3	Probabilités conditionnelles	404
13.3.1	Généralités	404
13.3.2	Systèmes complets d'événements	407
13.3.3	Formule des probabilités totales	408
13.3.4	Formule de Bayes	409
13.4	Indépendance	412
13.4.1	Indépendance de deux événements	412
13.4.2	Indépendance mutuelle	414
14	Variables aléatoires discrètes	416
14.1	Variables aléatoires discrètes.	417
14.1.1	Définition	417
14.1.2	Probabilité-image d'une variable aléatoire discrète	417
14.1.3	Loi d'une variable aléatoire discrète	418
14.1.4	Cas des variables aléatoires discrètes réelles ou complexes	421
14.2	Espérance	423
14.2.1	Définitions	423
14.2.2	Propriétés	425
14.2.3	Théorème de transfert	428
14.2.4	Inégalité de Markov	430
14.3	Variance d'une variable réelle	431
14.3.1	Moments d'ordre 2	431
14.3.2	Variance et écart-type	433
14.3.3	Inégalité de Bienaymé-Tchebychev	434
14.3.4	Généralisation	436
14.4	Lois classiques.	437
14.4.1	Loi uniforme	437

14.4.2	Loi de Bernoulli	437
14.4.3	Loi binomiale	438
14.4.4	Loi géométrique	438
14.4.5	Loi de Poisson	439
14.5	Couples de variables aléatoires	440
14.5.1	Généralités	440
14.5.2	Lois marginales	440
14.5.3	Lois conditionnelles	443
14.5.4	Covariance	443
14.6	Indépendance de variables aléatoires	445
14.6.1	Généralités	445
14.6.2	Espérance et indépendance	446
14.6.3	Généralisation	446
14.6.4	Théorème de réalisation	448
14.6.5	Somme de variables indépendantes identiquement distribuées	449
14.7	Loi faible des grands nombres	450
14.8	Fonction génératrice d'une variable aléatoire à valeurs entières	451
14.8.1	Généralités	451
14.8.2	Lien entre espérance et fonction génératrice	454
14.8.3	Fonction génératrice d'une somme de variables indépendantes	455

15 Équations différentielles linéaires 457

15.1	Équations et systèmes différentiels linéaires du premier ordre	457
15.1.1	Généralités	457
15.1.2	Problème de Cauchy	459
15.1.3	Équation différentielle linéaire homogène	460
15.1.4	Cas général	463
15.2	Exponentielle d'un endomorphisme, d'une matrice	464
15.2.1	Définition	464
15.2.2	Propriétés algébriques	464

15.2.3	Propriétés fonctionnelles	465
15.3	Équations linéaires homogènes à coefficients constants	467
15.3.1	Forme générale de la solution	467
15.3.2	Cas praticable	467
15.4	Équations différentielles linéaires scalaires d'ordre n	471
15.4.1	Généralités	471
15.4.2	Représentation matricielle	472
15.4.3	Équation différentielle linéaire homogène	473
15.4.4	Cas général	473
15.4.5	Diverses idées pour résoudre une équation différentielle linéaire d'ordre n . . .	474
16	Calcul différentiel	481
16.1	Dérivées partielles	482
16.1.1	Dérivée selon un vecteur	482
16.1.2	Dérivées partielles dans une base	482
16.1.3	Absence de lien entre la continuité et l'existence de dérivées selon tout vecteur	483
16.2	Différentielle	486
16.2.1	Application différentiable	486
16.2.2	Différentielle	492
16.2.3	Différentiabilité sur un ouvert	493
16.2.4	Lien avec les dérivées partielles	493
16.2.5	Caractérisation des fonctions à dérivée partielle nulle	495
16.2.6	Matrice jacobienne	497
16.2.7	Cas particulier où $F = \mathbb{R}$	498
16.3	Opérations sur les fonctions différentiables	499
16.3.1	Combinaison linéaire	499
16.3.2	Composition par une application linéaire	499
16.3.3	Composition par une application k -linéaire	500
16.3.4	Composition d'applications différentiables	500
16.3.5	Dérivation le long d'un chemin	504

16.4	Fonctions de classe \mathcal{C}^1	505
16.4.1	Définition	505
16.4.2	Caractérisation	505
16.4.3	Opérations sur les fonctions de classe \mathcal{C}^1	507
16.4.4	Caractérisation des fonctions constantes parmi les \mathcal{C}^1	510
16.5	Vecteurs tangents à une partie	510
16.6	Optimisation au premier ordre	513
16.6.1	Vocabulaire	513
16.6.2	Points critiques, extrema locaux d'une fonction sur un ouvert	514
16.6.3	Extrema locaux d'une fonction sur une partie	517
16.7	Fonctions de classe \mathcal{C}^k	522
16.7.1	Dérivées partielles d'ordre supérieur	522
16.7.2	Fonctions de classe \mathcal{C}^k	522
16.7.3	Théorème de Schwarz	522
16.8	Optimisation au second ordre	523
16.8.1	Hessienne	523
16.8.2	Développement limité à l'ordre 2	523
16.8.3	Application à l'étude des points critiques	524
17	Structures algébriques	530
17.1	Groupes	531
17.1.1	Rappels	531
17.1.2	Ordre d'un élément dans un groupe	534
17.1.3	Sous-groupe engendré par une partie	537
17.1.4	Théorème de Lagrange	541
17.2	Anneaux : rappels	542
17.2.1	Généralités	542
17.2.2	Sous-anneaux	543
17.2.3	Pièges calculatoires classiques	543

17.3	Anneau $\mathbb{Z}/n\mathbb{Z}$	545
17.3.1	Définition	545
17.3.2	Propriétés	545
17.3.3	Théorème chinois des restes	547
17.3.4	Indicatrice d'Euler	550
17.4	Idéaux d'un anneau commutatif	551
17.4.1	Généralités	551
17.4.2	Arithmétique dans un anneau principal	556
17.5	\mathbb{K} -algèbre	559
17.5.1	Généralités	559
17.5.2	Sous-algèbres	560
17.5.3	Morphismes d'algèbres	560
17.5.4	Polynôme minimal d'un élément	560
II	Exercices	562
1	Espaces vectoriels normés	563
2	Séries numériques et vectorielles : révisions et compléments	564
3	Familles sommables	565
4	Rappels et compléments d'algèbre linéaire	566
5	Réduction des endomorphismes	567
6	Intégrales généralisées	568
7	Intégrales à paramètre	569
8	Espaces préhilbertiens réels	570
9	Endomorphismes dans un espace euclidien	571
10	Fonctions vectorielles	572

11 Suites et séries de fonctions	573
12 Séries entières	574
13 Probabilités	575
14 Variables aléatoires discrètes	576
15 Équations différentielles linéaires	577
16 Calcul différentiel	578
17 Structures algébriques	579

Première partie

Cours

Chapitre 1

Espaces vectoriels normés

Sommaire

1.1	Bornes supérieures, bornes inférieures	19
1.1.1	Borne supérieure d'une partie de \mathbb{R}	19
1.1.2	Borne supérieure d'une application à valeurs dans \mathbb{R}	21
1.1.3	Règles pratiques	21
1.2	Normes	22
1.2.1	Définition	22
1.2.2	Exemples fondamentaux	23
1.2.3	Normes équivalentes	25
1.2.4	Boules	28
1.2.5	Parties bornées	31
1.3	Convergence des suites	36
1.3.1	Définition	36
1.3.2	Propriétés usuelles	37
1.3.3	Cas particulier en dimension finie	39
1.3.4	Point adhérent à une partie	41
1.4	Limites de fonctions	44
1.4.1	Définition	44
1.4.2	Caractérisation séquentielle de la limite	44
1.4.3	Propriétés usuelles	45
1.4.4	Cas particulier de la dimension finie	46
1.4.5	Composition des limites	47
1.4.6	Extensions des définitions	47
1.5	Fonctions continues	48
1.5.1	Continuité en un point	48
1.5.2	Continuité sur une partie	49
1.5.3	Cas particulier de la dimension finie	49
1.5.4	Fonctions lipschitziennes	50
1.5.5	Continuité des applications linéaires et n -linéaires	52
1.5.6	Norme subordonnée	59
1.6	Topologie d'un espace vectoriel normé	63
1.6.1	Intérieur d'une partie, voisinage d'un point	63
1.6.2	Parties ouvertes	65
1.6.3	Parties fermées	66
1.6.4	Ouverts ou fermés relatifs à une partie	71

1.6.5	Image réciproque d'un ouvert ou d'un fermé par une fonction continue . . .	71
1.6.6	Frontière d'une partie	74
1.7	Compacité.	74
1.7.1	Valeurs d'adhérence d'une suite	74
1.7.2	Théorème de Bolzano-Weierstrass	76
1.7.3	Parties compactes	78
1.7.4	Théorème des bornes atteintes	82
1.8	Connexité par arcs	87
1.8.1	Chemin	87
1.8.2	Parties connexes par arcs	88
1.8.3	Théorème des valeurs intermédiaires	89

Dans ce chapitre, la lettre \mathbb{K} désigne \mathbb{R} ou \mathbb{C} .

1.1 Bornes supérieures, bornes inférieures

1.1.1 Borne supérieure d'une partie de \mathbb{R}

On rappelle le théorème fondamental, dit « théorème (ou axiome) de la borne supérieure ».

Théorème 1.1

Toute partie A de \mathbb{R} , non-vide et majorée, possède une borne supérieure, notée $\sup A$.

Toute partie A de \mathbb{R} , non-vide et minorée, possède une borne inférieure, notée $\inf A$.

On dispose de caractérisations équivalentes de la borne supérieure.

Proposition 1.2

Soient A une partie de \mathbb{R} , non-vide et majorée, et s un réel.

Alors il y a équivalence entre les propositions suivantes :

$$(\alpha) \quad s = \sup A$$

$$(\beta) \quad \begin{cases} \forall a \in A, \quad a \leq s \\ \forall \varepsilon > 0, \quad \exists x \in A, \quad s - \varepsilon < x \leq s \end{cases}$$

$$(\gamma) \quad \begin{cases} \forall a \in A, \quad a \leq s \\ \exists (x_n) \in A^{\mathbb{N}}, \quad x_n \xrightarrow[n \rightarrow +\infty]{} s \end{cases}$$

Démonstration 1.3 $((\alpha) \implies (\beta))$

► $s = \sup A$ est le plus petit majorant de A donc c'est un majorant de A :

$$\forall a \in A, \quad a \leq s.$$

► s est le plus petit majorant de A donc

$$\forall \varepsilon > 0, \quad s - \varepsilon < s$$

donc $s - \varepsilon$ n'est pas un majorant de A .

Donc il existe $x \in A$ tel que $s - \varepsilon < x \leq s$. ■

Démonstration 1.4 $((\beta) \implies (\alpha))$

s est un majorant de A et tout réel strictement inférieur à s n'est pas un majorant *i.e.* tout majorant est supérieur ou égal à s .

Donc $s = \sup A$. ■

Démonstration 1.5 $((\beta) \implies (\gamma))$

On spécialise $\varepsilon \leftarrow \frac{1}{n+1}$ pour $n \in \mathbb{N}$. On a

$$\forall n \in \mathbb{N}, \quad \exists x_n \in A, \quad s - \frac{1}{n+1} < x_n \leq s.$$

De cette façon, on construit une suite $(x_n) \in A^{\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}, \quad s - \frac{1}{n+1} < x_n \leq s.$$

D'après le théorème des gendarmes, on a $x_n \xrightarrow[n \rightarrow +\infty]{} s$. ■

Démonstration 1.6 $((\gamma) \implies (\beta))$

Soit $\varepsilon > 0$.

Il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \quad |x_n - s| \leq \frac{\varepsilon}{2}$$

$$\text{donc } s - \frac{\varepsilon}{2} \leq x_n.$$

Or $x_n \in A$ donc $s - \varepsilon < s - \frac{\varepsilon}{2} \leq x_n \leq s$.

D'où

$$\forall \varepsilon > 0, \quad \exists x \in A, \quad s - \varepsilon < x \leq s. \quad \blacksquare$$

On a évidemment les caractérisations associées à la borne inférieure.

1.1.2 Borne supérieure d'une application à valeurs dans \mathbb{R}

Définition 1.7

Soient X un ensemble non-vidé et $f : X \longrightarrow \mathbb{R}$.

Si f est majorée sur X , alors on appelle borne supérieure de f sur X le réel $\sup f(X) = \sup_X f = \sup_{x \in X} f(x)$.

Si f est minorée sur X , alors on appelle borne inférieure de f sur X le réel $\inf f(X) = \inf_X f = \inf_{x \in X} f(x)$.

On déduit de la Proposition 1.2 les caractérisations suivantes.

Proposition 1.8

Soient X un ensemble non-vidé, $f : X \longrightarrow \mathbb{R}$ majorée sur X et s un réel.

Alors il y a équivalence entre les propositions suivantes :

- ▷ $s = \sup_X f$
- ▷ $\begin{cases} \forall x \in X, f(x) \leq s \\ \forall \varepsilon > 0, \exists x \in X, s - \varepsilon < f(x) \leq s \end{cases}$
- ▷ $\begin{cases} \forall x \in X, f(x) \leq s \\ \exists (x_n) \in X^{\mathbb{N}}, f(x_n) \xrightarrow[n \rightarrow +\infty]{} s \end{cases}$

1.1.3 Règles pratiques

D'abord, des évidences auxquelles on ne pense pas toujours.

Proposition 1.9

Soit A une partie de \mathbb{R} , non-vidé et majorée. Alors $\forall a \in A, a \leq \sup A$.

Soient X un ensemble non-vidé et $f : X \longrightarrow \mathbb{R}$ majorée sur X . Alors $\forall x \in X, f(x) \leq \sup_X f$.

En pratique, on n'a pas souvent besoin de connaître la valeur exacte d'une borne supérieure, on a plus souvent besoin de la majorer.

Proposition 1.10

- ▷ Soient A une partie de \mathbb{R} , non-vidé et majorée, et M un réel.
Pour montrer $\sup A \leq M$, il suffit de montrer $\forall a \in A, a \leq M$.
- ▷ Soient X un ensemble non-vidé, $f : X \longrightarrow \mathbb{R}$ majorée sur X et M un réel.
Pour montrer $\sup_X f \leq M$, il suffit de montrer $\forall x \in X, f(x) \leq M$.

Multiplication par un réel positif.

Proposition 1.11

Soient X un ensemble non-vide et $f : X \longrightarrow \mathbb{R}$ majorée sur X .

Alors pour tout $\lambda \geq 0$, $\sup_X (\lambda f) = \lambda \sup_X f$.

Démonstration 1.12

Soit $\lambda \geq 0$. On pose $s = \sup_X f$.

On veut montrer $\sup_X (\lambda f) = \lambda s$.

On a $\forall x \in X$, $f(x) \leq s$ et $\lambda \geq 0$ donc

$$\forall x \in X, \lambda f(x) \leq \lambda s.$$

Donc λs est un majorant de λf .

Comme $s = \sup_X f$, il existe $(x_n) \in X^{\mathbb{N}}$ telle que $f(x_n) \xrightarrow{n \rightarrow +\infty} s$.

D'après les théorèmes d'opération sur les limites, on a

$$\lambda f(x_n) \xrightarrow{n \rightarrow +\infty} \lambda s.$$

D'où $\lambda s = \sup_X (\lambda f)$ d'après la Proposition 1.2. ■

Attention ! C'est bien sûr faux si $\lambda < 0$.

1.2 Normes

1.2.1 Définition

Définition 1.13

Soit E un \mathbb{K} -espace vectoriel.

On appelle norme sur E toute application $N : E \longrightarrow \mathbb{R}_+$ telle que :

- ▷ pour tout $x \in E$, $N(x) = 0 \iff x = 0$ (séparation)
- ▷ pour tout $x \in E$, pour tout $\lambda \in \mathbb{K}$, $N(\lambda x) = |\lambda| N(x)$ (homogénéité)
- ▷ pour tout $(x, y) \in E^2$, $N(x + y) \leq N(x) + N(y)$ (inégalité triangulaire).

Un espace vectoriel est dit espace vectoriel normé quand on lui associe une norme.

On déduit de l'inégalité triangulaire une inégalité classique (souvent appelée aussi inégalité triangulaire) :

$$\text{pour tout } (x, y) \in E^2, \quad |N(x) - N(y)| \leq N(x - y).$$

Démonstration 1.14

Soit $(x, y) \in E^2$.

D'après l'inégalité triangulaire, on a

$$N(x - y + y) \leq N(x - y) + N(y)$$

$$N(x) \leq N(x - y) + N(y)$$

$$N(x) - N(y) \leq N(x - y).$$

De même, en échangeant x et y :

$$N(y) - N(x) \leq N(y - x).$$

Or

$$N(x - y) = N(-(y - x)) = |-1| N(y - x) = N(y - x).$$

D'où

$$\begin{cases} N(x) - N(y) \leq N(x - y) \\ N(y) - N(x) \leq N(x - y) \end{cases}$$

donc

$$|N(x) - N(y)| \leq N(x - y). \quad \blacksquare$$

Si N est une norme sur E , alors on peut définir une distance entre deux vecteurs de E : $d(u, v) = N(u - v)$.

On définit ainsi une application $d : E^2 \longrightarrow \mathbb{R}_+$ telle que :

- ▷ pour tout $(x, y) \in E^2$, $d(y, x) = d(x, y)$ (symétrie)
- ▷ pour tout $(x, y) \in E^2$, $d(x, y) = 0 \iff x = y$ (séparation)
- ▷ pour tout $(x, y, z) \in E^3$, $d(x, z) \leq d(x, y) + d(y, z)$ (inégalité triangulaire).

1.2.2 Exemples fondamentaux

- ▷ La valeur absolue dans \mathbb{R} et le module dans \mathbb{C} sont des normes.
- ▷ La norme euclidienne habituelle en géométrie plane ou spatiale est une norme.

- Plus généralement, si $\langle \cdot | \cdot \rangle$ est un produit scalaire sur E , la norme euclidienne associée $x \mapsto \sqrt{\langle x | x \rangle}$ est une norme au sens précédent.
- Soit E un \mathbb{K} -espace vectoriel de dimension finie. On choisit une base de E $\mathcal{B} = (e_1, \dots, e_n)$. Si v est un vecteur de E , on note (v_1, \dots, v_n) les coordonnées de v dans la base \mathcal{B} . On définit classiquement trois normes sur E :

$$\|v\|_\infty = \max_{i \in \llbracket 1; n \rrbracket} |v_i| \quad \|v\|_1 = \sum_{i=1}^n |v_i| \quad \|v\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2}$$

appelées respectivement norme infinie ou norme sup, norme 1 et norme 2.

Cas particulier : $E = \mathbb{R}^n$ muni de la base canonique.

Cas particulier : $E = \mathcal{M}_{n,p}(\mathbb{K})$ muni de la base canonique. Si $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ est une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, alors

$$\|A\|_\infty = \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} |a_{ij}| \quad \|A\|_1 = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} |a_{ij}| \quad \|A\|_2 = \sqrt{\sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} |a_{ij}|^2}$$

- Soient X un ensemble et E l'ensemble des applications bornées de X dans \mathbb{K} . La norme sup sur E est définie par $\|f\|_\infty = \sup_{x \in X} |f(x)|$.
- Cas particulier : si $X = \mathbb{N}$, E est l'ensemble des suites bornées et $\|u\|_\infty = \sup_{n \in \mathbb{N}} |u_n|$.

Démonstration 1.15 ($\|\cdot\|_\infty$ est une norme sur E)

- Si $\|v\|_\infty = 0$ alors par définition d'un maximum

$$\forall i \in \llbracket 1 ; n \rrbracket, \quad 0 \leq |v_i| \leq \|v\|_\infty = 0 \text{ donc } v_i = 0$$

donc $v = 0$.

- Soit $\lambda \in \mathbb{K}$. On a

$$\begin{aligned} \|\lambda v\|_\infty &= \sup_{i \in \llbracket 1; n \rrbracket} |\lambda v_i| \\ &= \sup_{i \in \llbracket 1; n \rrbracket} |\lambda| \times |v_i| \\ &= |\lambda| \sup_{i \in \llbracket 1; n \rrbracket} |v_i| \\ &= |\lambda| \|v\|_\infty. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} |\lambda| \geq 0 \text{ (Proposition 1.11)}$$

- Soit $(v, w) \in E^2$. On veut montrer $\|v + w\|_\infty \leq \|v\|_\infty + \|w\|_\infty$.

D'après la Proposition 1.10, il suffit de montrer

$$\forall i \in \llbracket 1 ; n \rrbracket, \quad |v_i + w_i| \leq \|v\|_\infty + \|w\|_\infty.$$

Pour $i \in \llbracket 1 ; n \rrbracket$, on a

$$\begin{aligned} |v_i + w_i| &\leq |v_i| + |w_i| \\ &\leq \|v\|_\infty + \|w\|_\infty. \end{aligned} \quad \left. \begin{array}{l} \\ \end{array} \right\} \begin{cases} |v_i| \leq \|v\|_\infty \\ |w_i| \leq \|w\|_\infty \end{cases}$$

Donc $\|v + w\|_\infty \leq \|v\|_\infty + \|w\|_\infty$. ■

Démonstration 1.16 ($\|\cdot\|_1$ est une norme sur E)

► Si $\|v\|_1 = 0$ alors $\sum_{i=1}^n |v_i| = 0$. Or une somme de réels positifs est nulle ssi tous les réels sont nuls.

Donc

$$\forall i \in \llbracket 1 ; n \rrbracket, |v_i| = 0$$

donc $v = 0$.

► Soit $\lambda \in \mathbb{R}$. On a

$$\begin{aligned}\|\lambda v\|_1 &= \sum_{i=1}^n |\lambda v_i| \\ &= \sum_{i=1}^n |\lambda| \times |v_i| \\ &= |\lambda| \sum_{i=1}^n |v_i| \\ &= |\lambda| \|v\|_1.\end{aligned}$$

► Soit $(v, w) \in E^2$. On a

$$\begin{aligned}\|v + w\|_1 &= \sum_{i=1}^n |v_i + w_i| \\ &\leq \sum_{i=1}^n (|v_i| + |w_i|) \\ &= \|v\|_1 + \|w\|_1.\end{aligned}$$

■

Proposition 1.17

Soient E, F deux espaces vectoriels normés.

L'application de $E \times F$ dans \mathbb{R}_+ qui à (x, y) associe $\max(\|x\|_E, \|y\|_F)$ est une norme.

Autrement dit, le produit de deux espaces vectoriels normés est encore un espace vectoriel normé, résultat qui se généralise par récurrence à un nombre quelconque (fini) d'espaces vectoriels normés.

1.2.3 Normes équivalentes

Définition 1.18

Soient E un \mathbb{K} -espace vectoriel et N_1, N_2 deux normes sur E .

On dit que N_1 et N_2 sont équivalentes quand il existe deux constantes strictement positives a, b telles que pour tout $v \in E$, $aN_1(v) \leq N_2(v) \leq bN_1(v)$.

Proposition 1.19

On note $\mathcal{N}(E)$ l'ensemble des normes sur E et

$$\forall (N_1, N_2) \in \mathcal{N}(E)^2, N_1 \sim N_2 \iff N_1 \text{ et } N_2 \text{ sont équivalentes.}$$

La relation \sim est alors une relation d'équivalence sur $\mathcal{N}(E)$.

Démonstration 1.20

▷ Soit $N \in \mathcal{N}(E)$. On a $N \leq N \leq N$ donc $N \sim N$.

Donc \sim est réflexive.

▷ Soit $(N_1, N_2) \in \mathcal{N}(E)^2$ tel que $N_1 \sim N_2$.

Il existe $a, b > 0$ tels que $aN_1 \leq N_2 \leq bN_1$.

Donc $\frac{1}{b}N_2 \leq N_1 \leq \frac{1}{a}N_2$ i.e. $N_2 \sim N_1$.

Donc \sim est symétrique.

▷ Soit $(N_1, N_2, N_3) \in \mathcal{N}(E)^3$ tel que $N_1 \sim N_2$ et $N_2 \sim N_3$.

Il existe $a, b, c, d > 0$ tels que
$$\begin{cases} aN_1 \leq N_2 \leq bN_1 \\ cN_2 \leq N_3 \leq dN_2 \end{cases}$$

Donc $acN_1 \leq N_3 \leq bdN_1$ i.e. $N_1 \sim N_3$.

Donc \sim est transitive.

Finalement, \sim est une relation d'équivalence. ■

Exercice 1.21

Montrez que si E est de dimension finie, les trois normes $\|\cdot\|_\infty$, $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes.

Correction 1.22

Soient E un espace vectoriel normé de dimension finie et \mathcal{B} une base de E .

Soit $v = (v_1, \dots, v_n)_{\mathcal{B}}$.

On a

$$\|v\|_\infty = \sup_{i \in \llbracket 1; n \rrbracket} |v_i| \quad \|v\|_1 = \sum_{i=1}^n |v_i| \quad \|v\|_2 = \sqrt{\sum_{i=1}^n |v_i|^2}.$$

► On a $\|v\|_\infty \leq \|v\|_1 \leq n \|v\|_\infty$.

En effet, il existe $j \in \llbracket 1 ; n \rrbracket$ tel que $\|v\|_\infty = |v_j|$ donc

$$\|v\|_\infty = |v_j| \leq |v_1| + \cdots + |v_n| = \|v\|_1.$$

De plus, pour tout $i \in \llbracket 1 ; n \rrbracket$, $|v_i| \leq \|v\|_\infty$ donc en additionnant les inégalités :

$$\sum_{i=1}^n |v_i| = \|v\|_1 \leq \sum_{i=1}^n \|v\|_\infty = n \|v\|_\infty.$$

Donc $\|\cdot\|_\infty \sim \|\cdot\|_1$.

► En mettant des carrés partout on arrive à

$$\|v\|_\infty \leq \|v\|_2 \leq \sqrt{n} \|v\|_\infty.$$

Donc $\|\cdot\|_\infty \sim \|\cdot\|_2$.

► Par transitivité, on a $\|\cdot\|_1 \sim \|\cdot\|_2$.

Exercice 1.23

Soient $E = \mathbb{R}[X]$ et $P = \sum_{i=0}^n a_i X^i \in E$. On pose $N_1(P) = \sum_{i=0}^n |a_i|$ et $N_\infty(P) = \max_{0 \leq i \leq n} |a_i|$.

Montrez que N_1 et N_∞ sont des normes sur E .

Montrez qu'elles ne sont pas équivalentes en considérant la suite des polynômes $P_n = \sum_{i=0}^n X^i$.

Correction 1.24

► N_1 et N_∞ sont clairement des normes (cf. Démonstration 1.15 et Démonstration 1.16).

► Pour $n \in \mathbb{N}$, on pose $P_n = \sum_{i=0}^n X^i = 1 + X + \cdots + X^n$.

Par l'absurde, on suppose $N_1 \sim N_\infty$.

Il existe $a, b > 0$ tels que $a N_1 \leq N_\infty \leq b N_1$.

Donc

$$\forall n \in \mathbb{N}, \quad \underbrace{a N_1(P_n)}_{=n+1} \leq \underbrace{N_\infty(P_n)}_{=1} \leq \underbrace{b N_1(P_n)}_{=n+1}$$

donc

$$\forall n \in \mathbb{N}, \quad a(n+1) \leq 1$$

or $\lim_{n \rightarrow +\infty} a(n+1) = +\infty$: contradiction.

Donc N_1 et N_∞ ne sont pas équivalentes.

Remarque 1.25

Soit $(N_1, N_2) \in \mathcal{N}(E)^2$. On a

$$N_1 \sim N_2 \iff v \mapsto \frac{N_1(v)}{N_2(v)} \text{ est bornée sur } E \setminus \{0\}.$$

Ainsi, pour montrer que N_1 et N_2 ne sont pas équivalentes, on cherche une suite $(v_n) \in (E \setminus \{0\})^{\mathbb{N}}$ telle que

$$\frac{N_1(v_n)}{N_2(v_n)} \xrightarrow{n \rightarrow +\infty} +\infty \quad \text{ou} \quad \frac{N_1(v_n)}{N_2(v_n)} \xrightarrow{n \rightarrow +\infty} 0.$$

Le résultat suivant est fondamental.

Théorème 1.26

Si E est un \mathbb{K} -espace vectoriel de dimension finie, alors toutes les normes sur E sont équivalentes.

Quand on est en dimension finie, cela signifie que tous les résultats qu'on peut démontrer pour une norme sont à facteurs près valables pour n'importe quelle norme, autrement dit cela nous permettra de choisir la norme que l'on préfère si on ne nous l'impose pas.

Dans toute la suite, E est un espace vectoriel normé par la norme $\|\cdot\|$.

1.2.4 Boules

Définition 1.27

Soient $a \in E$ et $r \in \mathbb{R}_+^*$.

On appelle boule ouverte de centre a et de rayon r l'ensemble noté $B(a, r)$ défini de la façon suivante :

$$B(a, r) = \{v \in E \mid \|v - a\| < r\}.$$

On appelle boule fermée de centre a et de rayon r l'ensemble noté (généralement) $\overline{B}(a, r)$:

$$\overline{B}(a, r) = \{v \in E \mid \|v - a\| \leq r\}.$$

On appelle sphère de centre a et de rayon r l'ensemble (généralement) noté $S(a, r)$:

$$S(a, r) = \{v \in E \mid \|v - a\| = r\}.$$

On appelle boule-unité la boule de centre 0 et de rayon 1, sphère-unité la sphère de centre 0 et de rayon 1.

Exercice 1.28

Que sont les boules dans \mathbb{R} ? Que sont les sphères dans \mathbb{R} ?

Correction 1.29

$|\cdot|$ est une norme sur \mathbb{R} .

Soient $a \in \mathbb{R}$ et $r > 0$.

On a

$$B(a, r) =]a - r ; a + r[\quad \overline{B}(a, r) = [a - r ; a + r] \quad S(a, r) = \{a - r, a + r\}.$$

Exercice 1.30

On prend $E = \mathbb{R}^2$ et on définit les normes infinie, 1 et 2 relativement à la base canonique.

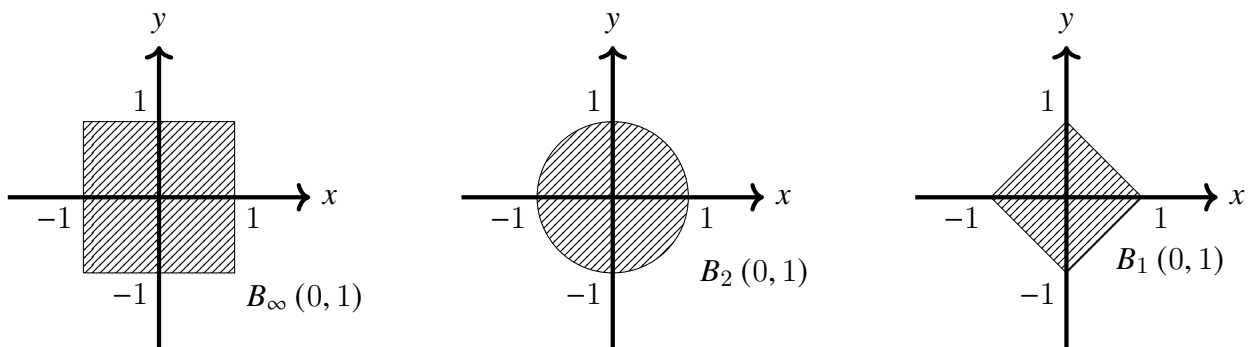
Représentez graphiquement les boules-unités pour chacune de ces trois normes.

Correction 1.31

On a

$$\|(x, y)\|_{\infty} = \max(|x|, |y|) \quad \|(x, y)\|_1 = |x| + |y| \quad \|(x, y)\|_2 = \sqrt{x^2 + y^2}.$$

On en déduit les boules suivantes :



Exercice 1.32

Montrez que toute boule ouverte est contenue dans une boule fermée et contient une boule fermée de mêmes centres.

Montrez la même chose en inversant les mots « ouverte » et « fermée ».

Correction 1.33

Soient $a \in E$ et $r > 0$.

On a

$$\overline{B}\left(a, \frac{r}{2}\right) \subseteq B(a, r) \subseteq \overline{B}(a, r)$$

et

$$B(a, r) \subseteq \overline{B}(a, r) \subseteq B(a, 2r).$$

Définition 1.34

Soit $(x, y) \in E^2$. On note $[xy] = \{tx + (1-t)y \mid t \in [0; 1]\}$, appelé segment (géométrique) d'extrémités x et y .

Une partie A de E est dite convexe quand pour tout $(x, y) \in A^2$, $[xy] \subseteq A$.

On a :

$$A \text{ est convexe} \iff \forall (x, y) \in A^2, \forall t \in [0; 1], tx + (1-t)y \in A.$$

Proposition 1.35

Les boules (ouvertes ou fermées) sont des parties convexes.

Les sphères ne sont jamais convexes.

Dans \mathbb{R} , les convexes sont les intervalles.

Démonstration 1.36 (Les boules sont convexes)

Soient $a \in E$ et $r > 0$.

Soient $(x, y) \in B(a, r)^2$ et $t \in [0; 1]$.

On veut montrer que $tx + (1-t)y \in B(a, r)$ i.e. $\|tx + (1-t)y - a\| < r$.

On a

$$\begin{aligned} \|tx + (1-t)y - a\| &= \|t(x-a) + at + (1-t)y - a\| \\ &= \|t(x-a) + (t-1)a + (1-t)y\| \\ &= \|t(x-a) + (1-t)(y-a)\| \\ &\leq \|t(x-a)\| + \|(1-t)(y-a)\| && \left. \begin{array}{l} \text{inégalité triangulaire} \\ \text{homogénéité et } t \geq 0 \text{ et} \\ 1-t \geq 0 \end{array} \right\} \\ &= t\|x-a\| + (1-t)\|y-a\| && \left. \begin{array}{l} \\ \\ x, y \in B(a, r) \end{array} \right\} \\ &< tr + (1-t)r \\ &= r. \end{aligned}$$

Ceci prouve

$$\forall (x, y) \in B(a, r), [xy] \subseteq B(a, r)$$

i.e. $B(a, r)$ est convexe.

De même, $\overline{B}(a, r)$ est convexe. ■

Démonstration 1.37 (Les sphères ne sont pas convexes)

Soient $a \in E$ et $r > 0$.

On veut montrer que $S(a, r)$ n'est pas convexe *i.e.*

$$\exists (x, y) \in S(a, r)^2, \exists t \in [0 ; 1], tx + (1 - t)y \notin S(a, r).$$

On choisit $x \in S(a, r)$ puis $y = 2a - x$ (diamétralement opposé).

On a

$$a \in [xy]$$

$$\text{car } a = \frac{1}{2}x + \frac{1}{2}y \text{ et}$$

$$a \notin S(a, r).$$

Donc $S(a, r)$ n'est pas convexe. ■

1.2.5 Parties bornées

Définition 1.38

On dit qu'une partie A de E est bornée quand il existe une boule qui la contient.

Exercice 1.39

Montrez que A est bornée ssi A est contenue dans une boule de centre 0.

Plus généralement, on choisit arbitrairement un point de E , noté x . Montrez l'équivalence A est bornée ssi A est contenue dans une boule de centre x .

Correction 1.40 (Première équivalence)

$\boxed{\Leftarrow}$ Immédiat.

$\boxed{\Rightarrow}$

Soit A une partie bornée de E .

Il existe $a \in E$ et $r > 0$ tels que $A \subseteq B(a, r)$.

On pose $r' = \|a\| + r > 0$.

Soit $x \in A$.

On a $x \in B(a, r)$ donc $\|x - a\| < r$.

Donc $r > \|\|x\| - \|a\|\|$.

Donc $r > \|x\| - \|a\|$.

Donc $r + \|a\| > \|x\|$.

Donc $A \subseteq B(0, r')$.

Correction 1.41 (Seconde équivalence)

Plus généralement, soit $x \in E$.

Avec $r'' = \|x - a\| + r$, on a de même $A \subseteq B(x, r'')$.

Exercice 1.42

Montrez qu'en dimension finie, cette définition ne dépend pas de la norme.

Correction 1.43

On veut montrer que si E est de dimension finie et N_1, N_2 sont des normes sur E , alors

$$A \text{ est bornée pour } N_1 \iff A \text{ est bornée pour } N_2.$$

Soit A une partie bornée pour N_1 .

Alors il existe $r_1 > 0$ tel que $A \subseteq B_1(0, r_1)$.

Comme E est de dimension finie, les normes N_1 et N_2 sont équivalentes.

Il existe donc $a, b > 0$ tels que $aN_2 \leq N_1 \leq bN_2$.

On pose $r_2 = \frac{r_1}{a} > 0$.

Soit $x \in A$.

Alors $x \in B_1(0, r_1)$ donc

$$\begin{aligned} N_1(x) &< r_1 \\ N_2(x) &\leq \frac{1}{a} N_1(x) < r_2. \end{aligned}$$

Donc $x \in B_2(0, r_2)$.

Donc $A \subseteq B_2(0, r_2)$.

Donc A est bornée pour N_2 .

Et réciproquement.

Proposition 1.44

Une partie A de E n'est pas bornée ssi il existe une suite (v_n) à termes dans A telle que $\|v_n\| \xrightarrow{n \rightarrow +\infty} +\infty$.

Exercice 1.45

Dans $E = \mathbb{R}^2$, on pose $A = \{(x, y) \mid x^4 + y^4 = 20\}$: A est-elle bornée ? Si oui, pour chacune des normes infinie, 1 et 2, donnez un rayon d'une boule centrée en 0 qui contient A .

Correction 1.46

E est de dimension finie donc les normes sur E sont toutes équivalentes.

On choisit la norme sup.

Soit $(x, y) \in A$.

On a $x^4 + y^4 = 20$ donc $x^4 \leq x^4 + y^4 = 20$ car $x^4 \geq 0$.

Donc $|x| \leq \sqrt[4]{20}$ et, de même, $|y| \leq \sqrt[4]{20}$.

Donc $(x, y) \in \overline{B_\infty}(0, \sqrt[4]{20})$.

Donc $A \subseteq \overline{B_\infty}(0, \sqrt[4]{20})$.

D'après l'Exercice 1.21, on a

$$\|\cdot\|_\infty \leq \|\cdot\|_1 \leq 2 \|\cdot\|_\infty$$

et

$$\|\cdot\|_\infty \leq \|\cdot\|_2 \leq \sqrt{2} \|\cdot\|_\infty.$$

Donc $A \subseteq \overline{B_1}(0, 2\sqrt[4]{20})$ et $A \subseteq \overline{B_2}(0, \sqrt{2}\sqrt[4]{20})$.

Exercice 1.47

Même question avec $E = \mathbb{C}^2$.

Correction 1.48

E est de dimension finie donc les normes sur E sont équivalentes.

x étant quelconque dans \mathbb{C} , existe-t-il $y \in \mathbb{C}$ tel que $x^4 + y^4 = 20$?

Dans \mathbb{C} , tout nombre possède une racine quatrième donc en posant y une racine quatrième de $20 - x^4$, on obtient $(x, y) \in A$.

On pose $v_n = (n, y_n)$ où y_n est une racine quatrième de $20 - n^4$.

On a $\|v_n\|_\infty = \max(n, |y_n|) \geq n$ donc

$$\|v_n\|_\infty \xrightarrow{n \rightarrow +\infty} +\infty.$$

Or $(v_n) \in A^{\mathbb{N}}$ donc A n'est pas bornée.

Exercice 1.49

Dans $E = \mathbb{R}^3$, on pose $B = \{(x, y, z) \mid x^2 + 3y^2 + 4z^2 + 2xy + 2xz - 2yz \leq 42\}$: B est-elle bornée ? Si oui, pour chacune des normes infinie, 1 et 2, donnez un rayon d'une boule centrée en 0 qui contient B .

Correction 1.50

E est de dimension finie donc les normes sur E sont équivalentes.

On a

$$\begin{aligned} \forall (x, y, z) \in E, \quad x^2 + 3y^2 + 4z^2 + 2xy + 2xz - 2yz &= (x + y + z)^2 - y^2 - z^2 - 2yz + 3y^2 + 4z^2 - 2yz \\ &= (x + y + z)^2 + 2y^2 + 3z^2 - 4yz \\ &= (x + y + z)^2 + 2(y - z)^2 + z^2. \end{aligned}$$

Si $(x, y, z) \in B$ alors $(x + y + z)^2 + 2(y - z)^2 + z^2 \leq 42$.

$$\text{Donc } \begin{cases} z^2 \leq 42 \\ 2(y - z)^2 \leq 42 \\ (x + y + z)^2 \leq 42 \end{cases}$$

$$\text{Donc } \begin{cases} |z| \leq 7 \\ |y - z| \leq 5 \\ |x + y + z| \leq 7 \end{cases}$$

$$\text{Donc } \begin{cases} |z| \leq 7 \\ |y| = |y - z + z| \leq |y - z| + |z| \leq 12 \\ |x| = |x + y + z - y - z| \leq |x + y + z| + |y| + |z| \leq 26 \end{cases}$$

Donc $\|(x, y, z)\|_\infty \leq 26$ et on a

$$B \subseteq B_\infty(0, 26) \quad B \subseteq B_1(0, 78) \quad B \subseteq B_2(0, 26\sqrt{3}).$$

Exercice 1.51

Dans $\mathcal{M}_2(\mathbb{R})$, on note \mathcal{P} l'ensemble des matrices de projecteurs : \mathcal{P} est-il borné ?

Correction 1.52

$\mathcal{M}_2(\mathbb{R})$ est de dimension finie donc les normes sont équivalentes.

On choisit la norme sup :

$$\left\| \begin{pmatrix} x & z \\ y & t \end{pmatrix} \right\|_\infty = \max(|x|, |y|, |z|, |t|).$$

On a $\mathcal{P} = \{M \in \mathcal{M}_2(\mathbb{R}) \mid M^2 = M\}$.

On pose $M = \begin{pmatrix} x & z \\ y & t \end{pmatrix}$ et on a

$$M = M^2 \iff \begin{cases} x^2 + yz = x \\ z(x+t) = z \\ y(x+t) = y \\ t^2 + yz = t \end{cases}$$

On impose $x+t=1$.

On a

$$t - t^2 = (1-x) - (1-x)^2 = x - x^2.$$

Donc avec $y \neq 0$, on a $M = \begin{pmatrix} x & \frac{x-x^2}{y} \\ y & 1-x \end{pmatrix}$.

On en déduit que \mathcal{P} contient les matrices $M_n = \begin{pmatrix} n & n-n^2 \\ 1 & 1-n \end{pmatrix}$ et on a $\|M_n\|_\infty \geq n$.

Donc \mathcal{P} n'est pas bornée.

Définition 1.53

On dit qu'une suite v à termes dans E est bornée quand l'ensemble de ses valeurs est borné, autrement dit quand il existe $M > 0$ tel que pour tout $n \in \mathbb{N}$, $\|v_n\| \leq M$.

On dit qu'une fonction f d'un ensemble X dans E est bornée quand l'ensemble de ses valeurs prises sur X est borné, autrement dit quand il existe $M > 0$ tel que pour tout $x \in X$, $\|f(x)\| \leq M$.

Exercice 1.54

Soit u une suite complexe arithmético-géométrique de raison a . À quelle condition est-elle bornée ?

Correction 1.55

Soit $(u_n) \in \mathbb{C}^{\mathbb{N}}$ telle que $\forall n \in \mathbb{N}$, $u_{n+1} = au_n + b$.

Si $a \neq 1$, (u_n) est de la forme $\left(\lambda a^n + \frac{b}{1-a}\right)_{n \in \mathbb{N}}$ où $\lambda \in \mathbb{C}$.

Si $a = 1$, (u_n) est de la forme $(\lambda + nb)_{n \in \mathbb{N}}$ où $\lambda \in \mathbb{C}$.

On choisit $\lambda \neq 0$.

On a alors

$$(u_n) \text{ est bornée } \iff \begin{cases} |a| \leq 1 \text{ et } a \neq 1 \\ \text{ou} \\ a = 1 \text{ et } b = 0 \end{cases}$$

Exercice 1.56

Soient B, B' deux boules de E . Si $(x, x') \in E^2$, on pose $f(x, x') = d(x, x')$. Montrez que f est bornée sur $B \times B'$.

Correction 1.57

Soient $(a, b) \in E^2$ et $r, s > 0$ tels que $B = B(a, r)$ et $B' = B(b, s)$.

On a

$$\forall (x, x') \in B \times B', \quad 0 \leq f(x, x') \leq \|b - a\| + r + s$$

car

$$\begin{aligned} f(x, x') &= \|x - x'\| \\ &= \|x - a + a - b + b - x'\| \\ &\leq \|x - a\| + \|a - b\| + \|b - x'\| \\ &\leq r + s + \|b - a\|. \end{aligned}$$

1.3 Convergence des suites

Dans cette section, E désigne un espace vectoriel normé par la norme $\|\cdot\|$.

1.3.1 Définition

Définition 1.58

Soient $u = (u_n)$ une suite à termes dans E et $\ell \in E$.

On dit que la suite u converge vers ℓ quand toute boule ouverte de centre ℓ contient tous les termes de la suite à partir d'un certain rang :

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N}, \quad \forall n \geq n_0, \quad u_n \in B(\ell, \varepsilon).$$

Proposition 1.59

Dans la définition, on peut remplacer les boules ouvertes par des boules fermées.

On peut réécrire la définition sous deux formes équivalentes :

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N}, \quad \forall n \geq n_0, \quad \|u_n - \ell\| < \varepsilon$$

ou

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N}, \quad \forall n \geq n_0, \quad \|u_n - \ell\| \leq \varepsilon.$$

On peut donc se ramener aux suites réelles positives : la suite vectorielle u converge vers ℓ ssi la suite réelle $(\|u_n - \ell\|)$ converge vers 0.

Une suite qui ne converge vers aucun élément de E est dite divergente.

1.3.2 Propriétés usuelles

Proposition 1.60 (Unicité de la limite)

Si une suite $u \in E^{\mathbb{N}}$ converge vers $\ell \in E$, elle ne peut converger vers un autre point de E .

On peut donc noter classiquement $\ell = \lim u = \lim_{n \rightarrow +\infty} u_n$ ou $u_n \xrightarrow[n \rightarrow +\infty]{} \ell$.

Démonstration 1.61

Par l'absurde, soit $(\ell, \ell') \in E^2$ tel que u converge vers ℓ et ℓ' , et $\ell \neq \ell'$.

$$\text{Soit } \varepsilon = \frac{\|\ell - \ell'\|}{2}.$$

Il existe $N_1 \in \mathbb{N}$ et $N_2 \in \mathbb{N}$ tels que

$$\forall n \geq N_1, \quad \|u_n - \ell\| < \varepsilon$$

et

$$\forall n \geq N_2, \quad \|u_n - \ell'\| < \varepsilon.$$

On pose $N = \max(N_1, N_2)$.

On a alors

$$\begin{aligned} \forall n \geq N, \quad \varepsilon &= \frac{\|\ell - \ell'\|}{2} \\ &= \frac{\|\ell - u_n + u_n - \ell'\|}{2} \\ &\leq \frac{\|\ell - u_n\| + \|u_n - \ell'\|}{2} \\ &< \varepsilon. \end{aligned}$$

Contradiction donc $\ell = \ell'$. ■

Proposition 1.62

Si une suite $u \in E^{\mathbb{N}}$ converge, alors elle est bornée.

Démonstration 1.63

On pose $\varepsilon = 42$.

Il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \quad \|u_n - \ell\| \leq 42$$

i.e. $\forall n \geq N, \quad u_n \in \overline{B}(\ell, 42)$.

Parmi les N premiers termes de la suite u , on détermine le plus lointain de ℓ : on pose

$$r = \max_{0 \leq k \leq N-1} \|u_k - \ell\|.$$

Puis on pose $R = \max(42, r) > 0$.

On a alors $\forall n \in \mathbb{N}, \|u_n - \ell\| \leq R$.

Donc $\forall n \in \mathbb{N}, u_n \in \overline{B}(\ell, R)$.

Donc la suite est bornée. ■

Théorème 1.64 (Opérations sur les suites convergentes)

Soient $u, v \in E^{\mathbb{N}}$ convergeant respectivement vers ℓ et m deux éléments de E .

Alors pour tout $(a, b) \in \mathbb{K}^2$, la suite $au + bv$ converge vers $a\ell + bm$.

Soit $\alpha \in \mathbb{K}^{\mathbb{N}}$ convergeant vers $\lambda \in \mathbb{K}$.

Alors la suite αu converge vers $\lambda\ell$.

Démonstration 1.65 (αu converge vers $\lambda\ell$)

On a

$$\begin{aligned}\forall n \in \mathbb{N}, \|\alpha_n u_n - \lambda\ell\| &= \|\alpha_n u_n - \lambda u_n + \lambda u_n - \lambda\ell\| \\ &= \|(\alpha_n - \lambda) u_n + \lambda(u_n - \ell)\| \\ &\leq \|(\alpha_n - \lambda) u_n\| + \|\lambda(u_n - \ell)\| \\ &= |\alpha_n - \lambda| \|u_n\| + |\lambda| \|u_n - \ell\| \\ &\leq |\alpha_n - \lambda| \|u_n\| + (|\lambda| + 1) \|u_n - \ell\|.\end{aligned}$$

Soit $\varepsilon > 0$.

Il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \|u_n - \ell\| \leq \frac{\varepsilon}{|\lambda| + 1}.$$

u converge donc est bornée : il existe $K \in \mathbb{R}_+^*$ tel que

$$\forall n \in \mathbb{N}, \|u_n\| \leq K.$$

Donc il existe $n_1 \in \mathbb{N}$ tel que

$$\forall n \geq n_1, |\alpha_n - \lambda| \leq \frac{\varepsilon}{K}.$$

Donc

$$\begin{aligned}\forall n \geq \max(n_0, n_1), \|\alpha_n u_n - \lambda\ell\| &\leq K \frac{\varepsilon}{K} + (|\lambda| + 1) \frac{\varepsilon}{|\lambda| + 1} \\ &= 2\varepsilon.\end{aligned}$$

D'où $\alpha_n u_n \xrightarrow[n \rightarrow +\infty]{} \lambda\ell$. ■

Proposition 1.66

Toute suite extraite d'une suite convergente converge vers la même limite.

Quasi-réciproque : si u est une suite telle que les deux suites extraites (u_{2n}) et (u_{2n+1}) convergent vers la même limite ℓ , alors u converge vers ℓ .

Proposition 1.67

Dans un produit de deux espaces vectoriels normés $E \times F$, une suite $(u_n) = ((a_n, b_n))$ converge ssi les suites (a_n) et (b_n) convergent dans E , respectivement F .

Dans ce cas, $\lim (a_n, b_n) = (\lim a_n, \lim b_n)$.

Ce résultat se généralise sans difficulté par récurrence à un nombre quelconque (fini) d'espaces vectoriels normés.

1.3.3 Cas particulier en dimension finie

Dans cette partie, on suppose que E est de dimension finie.

Définition 1.68

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E .

Pour $i \in \llbracket 1 ; n \rrbracket$, on appelle i -ème forme coordonnée (relative à la base \mathcal{B}), notée souvent d_i , la forme linéaire qui à un vecteur associe sa i -ème coordonnée dans la base \mathcal{B} :

$$\text{pour tout } v \in E, \quad v = \sum_{i=1}^n d_i(v) e_i.$$

Théorème 1.69

Soit \mathcal{B} une base de E .

Une suite $u \in E^{\mathbb{N}}$ converge vers $\ell \in E$ ssi pour toute forme coordonnée d relative à \mathcal{B} , la suite $(d(u_n))$ converge vers $d(\ell)$.

Autrement dit, une suite converge ssi ses suites-coordonnées dans n'importe quelle base convergent.

Dans ce cas, la limite de la suite u est le vecteur ℓ tel que pour toute forme coordonnée d , $d(\ell) = \lim_{n \rightarrow +\infty} d(u_n)$.

Démonstration 1.70

E est de dimension finie donc toutes les normes sur E sont équivalentes. On note $p = \dim E$.

\mathcal{B} étant une base de E , on choisit la norme $\|\cdot\|_\infty$ relative à \mathcal{B} .

Il existe $a, b > 0$ tels que $a \|\cdot\|_\infty \leq \|\cdot\| \leq b \|\cdot\|_\infty$.

\Rightarrow

Si (u_n) converge vers ℓ , alors $\|u_n - \ell\| \xrightarrow{n \rightarrow +\infty} 0$ donc d'après l'inégalité ci-dessus, on a

$$0 \leq \|u_n - \ell\|_\infty \leq \frac{1}{a} \|u_n - \ell\|_\infty.$$

Donc, d'après le théorème des gendarmes, on a

$$\|u_n - \ell\|_\infty \xrightarrow{n \rightarrow +\infty} 0.$$

En notant $u_n = (u_{n,1}, \dots, u_{n,p})_{\mathcal{B}}$ et $\ell = (\ell_1, \dots, \ell_p)_{\mathcal{B}}$, on a

$$\|u_n - \ell\|_\infty = \max_{k \in \llbracket 1; p \rrbracket} |u_{n,k} - \ell_k|.$$

Donc $\forall k \in \llbracket 1; p \rrbracket$, $|u_{n,k} - \ell_k| \leq \|u_n - \ell\|_\infty$ et, par théorème des gendarmes, on a

$$u_{n,k} \xrightarrow{n \rightarrow +\infty} \ell_k.$$

\Leftarrow

Si pour tout $k \in \llbracket 1; p \rrbracket$, $(u_{n,k})_{n \in \mathbb{N}}$ converge vers ℓ_k , on a $\|u_n - \ell\|_\infty \xrightarrow{n \rightarrow +\infty} 0$ et d'après l'inégalité précédente, on a

$$\|u_n - \ell\| \leq b \|u_n - \ell\|_\infty$$

donc $\|u_n - \ell\| \xrightarrow{n \rightarrow +\infty} 0$ i.e. (u_n) converge vers ℓ . ■

Exemple 1.71

Si $M_n = \begin{pmatrix} 1 & e^{-n} \\ 1/n & n \sin(1/n) \end{pmatrix}$, alors la suite de matrices (M_n) converge vers la matrice I_2 .

Corollaire 1.72

Si E est de dimension finie, la convergence d'une suite ne dépend pas du choix de la norme. On peut donc choisir la norme qu'on veut.

1.3.4 Point adhérent à une partie

Définition 1.73

Soient A une partie de E et $x \in E$.

On dit que x est un point adhérent à A quand il existe une suite $u \in A^{\mathbb{N}}$ qui converge vers x .

L'adhérence de A est l'ensemble de ses points adhérents, noté \overline{A} .

Intuitivement, l'adhérence d'une partie est elle-même à laquelle on ajoute tous les points qui se trouvent sur son bord.

Remarque 1.74

On remarque qu'on a $A \subseteq \overline{A}$ car pour tout $a \in A$, la suite constante égale à a converge vers a et est à termes dans A .

Exercice 1.75

Quelle est l'adhérence d'une boule ouverte ?

Correction 1.76

Soient $a \in E$ et $r > 0$. Montrons que $\overline{B(a, r)} = \overline{B}(a, r)$.

\subseteq

Soit $b \in \overline{B(a, r)}$.

Il existe $(u_n) \in B(a, r)^{\mathbb{N}}$ qui converge vers b .

D'après la deuxième inégalité triangulaire, on a

$$\begin{aligned} ||u_n - a|| - ||b - a|| &\leq ||(u_n - a) - (b - a)|| \\ &= ||u_n - b||. \end{aligned}$$

Or $||u_n - b|| \xrightarrow{n \rightarrow +\infty} 0$ donc d'après le théorème d'encadrement, on a

$$||u_n - a|| \xrightarrow{n \rightarrow +\infty} ||b - a||.$$

Or $\forall n \in \mathbb{N}, u_n \in B(a, r)$ donc $||u_n - a|| < r$.

Par passage à la limite, $||b - a|| \leq r$.

Donc $b \in \overline{B}(a, r)$.

\supseteq

Soit $b \in \overline{B}(a, r)$.

Alors $\|b - a\| \leq r$.

Si $\|b - a\| < r$, alors $b \in B(a, r) \subseteq \overline{B(a, r)}$.

Si $\|b - a\| = r$, on pose, pour $n \in \mathbb{N}^*$, $u_n = \frac{1}{n}a + \left(1 - \frac{1}{n}\right)b$.

La suite (u_n) converge vers b par opérations sur les limites et on a

$$\begin{aligned}\forall n \in \mathbb{N}^*, \quad \|u_n - a\| &= \left\| \left(\frac{1}{n} - 1\right)a + \left(1 - \frac{1}{n}\right)b \right\| \\ &= \left| 1 - \frac{1}{n} \right| \|b - a\| \\ &= \frac{n-1}{n} r \\ &< r\end{aligned}$$

donc $(u_n) \in B(a, r)^{\mathbb{N}}$.

On a ainsi trouvé une suite à termes dans $B(a, r)$ qui converge vers b donc $b \in \overline{B(a, r)}$.

Remarque 1.77

On a montré au passage que si $u_n \xrightarrow[n \rightarrow +\infty]{} \ell$ alors pour tout $a \in E$, $\|u_n - a\| \xrightarrow[n \rightarrow +\infty]{} \|\ell - a\|$.

Exercice 1.78

Quelle est l'adhérence de \mathbb{Z} dans \mathbb{R} ?

Correction 1.79

\mathbb{R} est muni de la norme $|\cdot|$.

Montrons que $\mathbb{Z} = \overline{\mathbb{Z}}$.

$\boxed{\subseteq}$ Trivial.

$\boxed{\supseteq}$

Soit $b \in \overline{\mathbb{Z}}$.

Il existe une suite $(u_n) \in \mathbb{Z}^{\mathbb{N}}$ qui converge vers b .

On pose $\varepsilon = \frac{1}{2}$.

Il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \quad |u_n - b| < \frac{1}{2}$$

$$i.e. \quad u_n \in \left] b - \frac{1}{2}; b + \frac{1}{2} \right[.$$

Or l'intervalle $\left] b - \frac{1}{2} ; b + \frac{1}{2} \right[$ est de longueur 1 et est ouvert donc il contient au plus un entier.

Or il en contient un donc il en contient un et un seul.

On en déduit que (u_n) est stationnaire en cet entier à partir de n_0 et donc $\lim u = b$ est un entier.

Donc $b \in \mathbb{Z}$.

Proposition 1.80

Soient A une partie de E et $x \in E$.

Alors x est adhérent à A ssi toute boule centrée en x rencontre A .

De manière formalisée : $x \in \overline{A} \iff \forall r > 0, \exists y \in A, y \in B(x, r)$.

Démonstration 1.81

$\boxed{\Leftarrow}$

On a pour tout $n \in \mathbb{N}^*$, $B\left(x, \frac{1}{n}\right) \cap A \neq \emptyset$.

On peut choisir $y_n \in B\left(x, \frac{1}{n}\right) \cap A$.

On construit ainsi une suite $(y_n) \in A^{\mathbb{N}^*}$ telle que

$$\forall n \in \mathbb{N}^*, \|x - y_n\| < \frac{1}{n}.$$

D'après le théorème d'encadrement, on a donc $y_n \xrightarrow[n \rightarrow +\infty]{} x$.

Donc $x \in \overline{A}$.

$\boxed{\Rightarrow}$

Soient $x \in \overline{A}$ et $(u_n) \in A^{\mathbb{N}}$ telle que $u_n \xrightarrow[n \rightarrow +\infty]{} x$.

Soit $r > 0$.

Il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \|u_n - x\| < r$$

donc $u_{n_0} \in A \cap B(x, r)$ donc $A \cap B(x, r) \neq \emptyset$. ■

On peut donner la définition de la densité d'une partie.

Définition 1.82

On dit qu'une partie A est dense dans E quand $\overline{A} = E$, c'est-à-dire qu'on peut trouver des éléments de A aussi proches de n'importe quel point.

Exemple 1.83

- ▷ Dans \mathbb{R} , \mathbb{Q} et $\mathbb{R} \setminus \mathbb{Q}$ sont denses (*cf.* cours de première année).
- ▷ $\text{GL}_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})$ (démonstration ultérieure).

1.4 Limites de fonctions

Dans cette section, E et F sont deux espaces vectoriels normés par les normes $\|\cdot\|_E$ et $\|\cdot\|_F$.

1.4.1 Définition

Définition 1.84

Soient f une fonction de E dans F , D son ensemble de définition, $a \in \overline{D}$ et $\ell \in F$.

On dit que f a pour limite ℓ en a quand

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in D, \|x - a\|_E < \eta \implies \|f(x) - \ell\|_F < \varepsilon.$$

Remarque 1.85

On peut remplacer les inégalités strictes sur les normes par des inégalités larges.

On peut réécrire la définition à l'aide de boules ouvertes (ou fermées) :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in D \cap B(a, \eta), f(x) \in B(\ell, \varepsilon).$$

Si E et F sont de dimension finie, cette définition ne dépend pas du choix des normes.

1.4.2 Caractérisation séquentielle de la limite

Théorème 1.86

Soient f une fonction de E dans F , D son ensemble de définition, $a \in \overline{D}$ et $\ell \in F$.

f a pour limite ℓ en a ssi pour toute suite u à termes dans D convergeant vers a , la suite $f \circ u = (f(u_n))$ converge vers ℓ .

Démonstration 1.87

\Rightarrow

Supposons que f a pour limite ℓ en a .

Soit $u \in D^{\mathbb{N}}$ qui converge vers a .

Pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$\forall x \in D, \|x - a\|_E < \eta \implies \|f(x) - \ell\|_F < \varepsilon.$$

On a $\eta > 0$ et $u_n \xrightarrow[n \rightarrow +\infty]{} a$ donc il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \|u_n - a\|_E < \eta.$$

Donc pour tout $n \geq N$, comme $u_n \in D$ et $\|u_n - a\|_E < \eta$, on a

$$\|f(u_n) - \ell\|_F < \varepsilon.$$

Donc $(f(u_n))$ converge vers ℓ .

\Leftarrow

Par contraposée, montrons que si f n'a pas pour limite ℓ en a alors il existe $u \in D^{\mathbb{N}}$ telle que $u_n \xrightarrow[n \rightarrow +\infty]{} a$ et $f(u_n) \not\xrightarrow[n \rightarrow +\infty]{} \ell$.

Si f n'a pas pour limite ℓ en a , alors il existe $\varepsilon > 0$ tel que pour tout $\eta > 0$, il existe $x \in D$ tel que $\|x - a\|_E < \eta$ et $\|f(x) - \ell\|_F \geq \varepsilon$.

Donc pour tout $n \in \mathbb{N}$, il existe $u_n \in D$ tel que
$$\begin{cases} \|u_n - a\|_E < \frac{1}{n+1} \\ \|f(u_n) - \ell\|_F \geq \varepsilon \end{cases}.$$

On construit ainsi une suite $(u_n) \in D^{\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}, \begin{cases} \|u_n - a\|_E < \frac{1}{n+1} \\ \|f(u_n) - \ell\|_F \geq \varepsilon \end{cases}$$

Par encadrement, (u_n) converge vers a mais $(f(u_n))$ ne converge pas vers ℓ . ■

En pratique, on utilise beaucoup plus souvent le sens direct de l'équivalence précédente.

1.4.3 Propriétés usuelles

Proposition 1.88 (Unicité de la limite)

Soient f une fonction de E dans F , D son ensemble de définition, $a \in \overline{D}$ et $\ell \in F$.

Si f a pour limite ℓ en a , alors elle ne peut avoir d'autre limite que ℓ en a .

On peut donc noter classiquement $\ell = \lim_a f = \lim_{x \rightarrow a} f(x)$ ou $f(x) \xrightarrow{x \rightarrow a} \ell$.

Proposition 1.89

Si f a pour limite ℓ en a , alors elle est bornée au voisinage de a .

Théorème 1.90 (Opérations sur les limites)

Soient f et g deux fonctions de E dans F , définies sur la même partie D et ayant respectivement pour limites ℓ et m deux éléments de F en $a \in \overline{D}$.

Alors pour tout $(\lambda, \mu) \in \mathbb{K}^2$, la fonction $\lambda f + \mu g$ a pour limite $\lambda \ell + \mu m$ en a .

Soient α une fonction de E dans \mathbb{K} et f une fonction définie de E dans F , définies sur la même partie D et ayant respectivement pour limites $\beta \in \mathbb{K}$ et $\ell \in F$ en $a \in \overline{D}$.

Alors αf a pour limite $\beta \ell$ en a .

Proposition 1.91

Une fonction $f = (g, h)$ à valeurs dans un produit d'espaces vectoriels normés a une limite ssi g et h ont chacune une limite.

Dans ce cas, $\lim_a f = \left(\lim_a g, \lim_a h \right)$.

Ce résultat se généralise sans difficulté par récurrence à un nombre quelconque (fini) d'espaces vectoriels normés.

1.4.4 Cas particulier de la dimension finie

Théorème 1.92

On suppose que F est de dimension finie. Soit \mathcal{B} une base de F .

Soit f une fonction de E dans F , D son ensemble de définition, $a \in \overline{D}$ et $\ell \in F$.

La fonction f a pour limite ℓ en a ssi pour toute forme coordonnée d relative à \mathcal{B} , la fonction $d \circ f$ a pour limite $d(\ell)$ en a .

Autrement dit, une fonction a une limite en a ssi ses fonctions-coordonnées dans n'importe quelle base ont chacune une limite en a .

Dans ce cas, la limite de la fonction f en a est le vecteur ℓ tel que pour toute forme coordonnée d , $d(\ell) = \lim_{x \rightarrow a} d(f(x))$.

1.4.5 Composition des limites

G désigne un troisième espace vectoriel normé.

Théorème 1.93

Soient f une fonction de E dans F et D_f son ensemble de définition. Soient g une fonction de F dans G et D_g son ensemble de définition. On suppose que $f(D_f) \subseteq D_g$ (condition qui permet de définir la composée $g \circ f$ sur D_f).

Soient $a \in \overline{D_f}$, $b \in \overline{D_g}$ et $\ell \in G$.

Si f a pour limite b en a et g a pour limite ℓ en b , alors $g \circ f$ a pour limite ℓ en a .

Autrement dit, si $\begin{cases} f(x) \xrightarrow{x \rightarrow a} b \\ g(y) \xrightarrow{y \rightarrow b} \ell \end{cases}$ alors $g \circ f(x) \xrightarrow{x \rightarrow a} \ell$.

1.4.6 Extensions des définitions

D'abord les limites infinies en un point dans le cas où l'espace d'arrivée est \mathbb{R} .

Définition 1.94

Soient f une fonction de E dans \mathbb{R} , D son ensemble de définition et $a \in \overline{D}$.

On dit que f a pour limite $+\infty$ en a quand

$$\forall M > 0, \exists \eta > 0, \forall x \in D, \|x - a\|_E \leq \eta \implies f(x) \geq M.$$

On dit que f a pour limite $-\infty$ en a quand

$$\forall M < 0, \exists \eta > 0, \forall x \in D, \|x - a\|_E \leq \eta \implies f(x) \leq M.$$

Puis les limites en « l'infini ».

Définition 1.95

Soient f une application de E dans F et $\ell \in F$.

On dit que f a pour limite ℓ quand $\|x\|$ tend vers l'infini quand

$$\forall \varepsilon > 0, \exists B > 0, \forall x \in E, \|x\|_E \geq B \implies \|f(x) - \ell\|_F \leq \varepsilon.$$

Dans le cas où $F = \mathbb{R}$, on dit que $f(x)$ a pour limite $+\infty$ quand $\|x\|$ tend vers l'infini quand

$$\forall M > 0, \exists B > 0, \forall x \in E, \|x\|_E \geq B \implies f(x) \geq M.$$

(Définition semblable pour la limite $-\infty$).

Enfin, dans le cas où l'espace de départ est \mathbb{R} , on peut parler de limite en l'infini au sens habituel.

Définition 1.96

Soient f une fonction de \mathbb{R} dans F , définie sur un ouvert $]B; +\infty[$ et $\ell \in F$.

On dit que $f(x)$ a pour limite ℓ quand x tend vers $+\infty$ quand

$$\forall \varepsilon > 0, \exists B > 0, \forall x \geq B, \|f(x) - \ell\| \leq \varepsilon.$$

(Définition semblable pour la limite x tend vers $-\infty$).

1.5 Fonctions continues

Dans cette section, E et F sont des espaces vectoriels normés par les normes $\|\cdot\|_E$ et $\|\cdot\|_F$.

1.5.1 Continuité en un point

Proposition 1.97

Soient f une fonction de E dans F , D son ensemble de définition, $a \in \overline{D}$ et $\ell \in F$.

Si f a pour limite ℓ en a et si $a \in D$, alors $\ell = f(a)$.

Dans ce cas, on dit que la fonction f est continue en a .

Définition 1.98

Soient f une fonction de E dans F , D son ensemble de définition et $a \in D$.

On dit que f est continue en a quand f a pour limite $f(a)$ en a .

On déduit de cette définition et des théorèmes précédents

- la caractérisation séquentielle de la continuité en un point ;
- le fait qu'une fonction continue en un point est bornée au voisinage de ce point ;
- les théorèmes d'opérations et de compositions des fonctions continues en un point ;
- l'équivalence entre la continuité d'une fonction et celle de ses fonctions-coordonnées dans une certaine base de F dans le cas où F est de dimension finie.

1.5.2 Continuité sur une partie

Définition 1.99

Soient f une fonction de E dans F , D son ensemble de définition et $A \subseteq D$.

On dit que f est continue sur A quand f est continue en tout point de A .

On déduit de cette définition et des théorèmes précédents

- les théorèmes d'opérations et de compositions des fonctions continues sur une partie ;
- l'équivalence entre la continuité d'une fonction et celle de ses fonctions-coordonnées dans une certaine base de F dans le cas où F est de dimension finie.

Proposition 1.100

Soient f et g deux fonctions de E dans F définies sur D et $A \subseteq D$.

Si A est dense dans D , f et g sont continues sur D et $f = g$ sur A , alors $f = g$ sur D .

Démonstration 1.101

On suppose que A est dense dans D et que f et g sont continues.

Soit $x \in D$.

Il existe $(u_n) \in A^{\mathbb{N}}$ telle que $u_n \xrightarrow{n \rightarrow +\infty} x$.

f est continue en x donc d'après la caractérisation séquentielle de la continuité, on a

$$f(u_n) \xrightarrow{n \rightarrow +\infty} f(x).$$

De même, on a

$$g(u_n) \xrightarrow{n \rightarrow +\infty} g(x).$$

Or $f = g$ sur A donc $(f(u_n)) = (g(u_n))$ donc $f(x) = g(x)$ par unicité de la limite. ■

1.5.3 Cas particulier de la dimension finie

On suppose que E et F sont de dimensions finies.

Dans une base donnée, les formes coordonnées relatives à cette base sont en particulier des applications continues.

Donc toute fonction f de E dans F dont les fonctions-coordonnées (f_1, \dots, f_n) dans une base de F sont définies polynomialement à partir des formes coordonnées dans une base de E est continue.

Exemple 1.102

- La fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ telle que $f(x, y) = (x^2 + y^2, xy - (1+x)^3)$ est continue sur \mathbb{R}^2 .
- Les applications trace et déterminant définies sur $\mathcal{M}_n(\mathbb{K})$ sont continues.

Exercice 1.103

Montrez que l'application $A \mapsto A^2$ est continue de $\mathcal{M}_n(\mathbb{K})$ dans lui-même.

Correction 1.104

On a $A^2 = \left(\sum_{k=1}^n a_{i,k} a_{k,j} \right)_{(i,j) \in \llbracket 1;n \rrbracket^2}$ donc $A \mapsto A^2$ est continue.

Exercice 1.105

En admettant (momentanément) que $\mathrm{GL}_n(\mathbb{K})$ est un ouvert, montrez que l'application $A \mapsto A^{-1}$ est continue de $\mathrm{GL}_n(\mathbb{K})$ dans lui-même.

Correction 1.106

Si $A \in \mathrm{GL}_n(\mathbb{K})$, alors $A^{-1} = \frac{1}{\det A} (\mathrm{Com} A)^\top$.

Les coefficients de $(\mathrm{Com} A)^\top$ sont des déterminants calculés à partir des coefficients de A donc dépendent polynomialement de ces coefficients, donc $A \mapsto (\mathrm{Com} A)^\top$ est continue.

Donc $A \mapsto A^{-1}$ est le produit de deux fonctions continues et est donc continue.

1.5.4 Fonctions lipschitziennes

Définition 1.107

Soient f une application de E dans F , A une partie de E et $K \in \mathbb{R}_+$.

On dit que f est K -lipschitzienne sur A (ou lipschitzienne de rapport K) quand

$$\forall (x, y) \in A^2, \quad \|f(y) - f(x)\|_F \leq K \|y - x\|_E.$$

On dit que f est lipschitzienne sur A quand il existe $K \in \mathbb{R}_+$ tel que f soit K -lipschitzienne sur A .

Remarque 1.108

Si f est K -lipschitzienne sur A , alors le rapport K n'est pas unique, puisque pour tout $L \geq K$, on a encore f L -lipschitzienne sur A .

Proposition 1.109

Toute fonction lipschitzienne est continue.

Mais la réciproque est fautive (contre-exemple : la fonction $\sqrt{\cdot}$ sur $[0 ; +\infty[$).

Un exemple fondamental : la fonction $x \mapsto d(x, A)$.

Définition 1.110

Soit A une partie de E .

Pour $x \in E$, on appelle distance de x à A le réel $\inf_{a \in A} d(x, a)$.

Proposition 1.111

Pour toute partie A de E , la fonction $x \mapsto d(x, A)$ est 1-lipschitzienne.

L'adhérence de A est l'ensemble des points à distance nulle de A , i.e. tels que $d(x, A) = 0$.

Démonstration 1.112 (1-lipschitzianité de la fonction)

On veut montrer

$$\forall (x, y) \in E^2, \quad |d(x, A) - d(y, A)| \leq \|x - y\|.$$

On montre d'abord

$$\forall (x, y) \in E^2, \quad d(x, A) - d(y, A) \leq \|x - y\|$$

ce qui est équivalent à

$$\forall (x, y) \in E^2, \quad d(x, A) \leq d(y, A) + \|x - y\|.$$

On a $d(x, A) = \inf_{a \in A} \|x - a\|$ donc pour tout $a \in A$, on a

$$d(x, A) \leq \|x - a\|.$$

Or $\|x - a\| = \|x - y + y - a\| \leq \|x - y\| + \|y - a\|$ donc

$$d(x, A) \leq \|x - y\| + \|y - a\|.$$

De plus, on a $d(y, A) = \inf_{a \in A} \|y - a\|$ donc il existe $(u_n) \in A^{\mathbb{N}}$ telle que

$$\|y - u_n\| \xrightarrow{n \rightarrow +\infty} d(y, A).$$

Alors pour tout $n \in \mathbb{N}$, $d(x, A) \leq \|x - y\| + \|y - u_n\|$, donc par passage à la limite quand $n \rightarrow +\infty$:

$$d(x, A) \leq \|x - y\| + d(y, A)$$

donc $d(x, A) - d(y, A) \leq \|x - y\|$.

En échangeant x et y , on obtient

$$\forall (x, y) \in E^2, \quad d(y, A) - d(x, A) \leq \|y - x\| = \|x - y\|.$$

D'où

$$\forall (x, y) \in E^2, \quad |d(x, A) - d(y, A)| \leq \|x - y\|.$$

■

Démonstration 1.113 ($x \in \overline{A} \iff d(x, A) = 0$)

\implies

Si $x \in \overline{A}$ alors il existe une suite $(u_n) \in A^{\mathbb{N}}$ qui converge vers x .

Or $x \mapsto d(x, A)$ est lipschitzienne donc continue, donc

$$d(u_n, A) \xrightarrow[n \rightarrow +\infty]{} d(x, A).$$

Or pour tout $n \in \mathbb{N}$, $d(u_n, A) = 0$ car $(u_n) \in A^{\mathbb{N}}$, donc

$$d(x, A) = 0.$$

\impliedby

Si $d(x, A) = 0$, alors $\inf_{a \in A} \|x - a\| = 0$.

Donc il existe $(a_n) \in A^{\mathbb{N}}$ telle que $\|x - a_n\| \xrightarrow[n \rightarrow +\infty]{} 0$ i.e. (a_n) converge vers x i.e. $x \in \overline{A}$.

■

1.5.5 Continuité des applications linéaires et n -linéaires

Proposition 1.114

Soit $f \in \mathcal{L}(E, F)$.

Il y a équivalence entre les propositions suivantes :

- (1) f est continue en 0 ;
- (2) f est continue en un point x ;
- (3) f est continue sur E ;

(4) f est lipschitzienne sur E ;

(5) il existe $K \geq 0$ tel que pour tout $x \in E$, $\|f(x)\|_F \leq K \|x\|_E$;

(6) f est bornée sur la boule-unité ;

(7) f est bornée sur une boule.

Démonstration 1.115

On a clairement (4) \implies (3), (3) \implies (2) et (3) \implies (1). ■

Démonstration 1.116 ((1) \implies (2))

Soit $x \in E$.

f est continue en 0 donc $\lim_{h \rightarrow 0} f(h) = f(0) = 0$.

Donc $f(x+h) = f(x) + f(h) \xrightarrow{h \rightarrow 0} f(x)$.

Donc f est continue en x . ■

Démonstration 1.117 ((2) \implies (3))

Soit $y \in E$.

f est continue en x donc $\lim_{t \rightarrow x} f(t) = f(x)$.

Donc

$$\begin{aligned} f(t) &= f(t - y + x - x + y) \\ &= f(t - y + x) - f(x) + f(y) \\ &\xrightarrow{t \rightarrow y} f(y). \end{aligned}$$

Donc f est continue en y .

Donc f est continue sur E . ■

Démonstration 1.118 ((3) \implies (7))

Soit $x \in E$.

f est continue en x donc f est bornée au voisinage de x i.e. il existe $r > 0$ tel que f soit bornée sur $B(x, r)$. ■

Démonstration 1.119 ((7) \implies (6))

f est bornée sur $B(x, r)$: il existe $M > 0$ tel que

$$\forall t \in B(x, r), \quad \|f(t)\| \leq M.$$

Pour tout $y \in B(0, r)$, on a $x + y \in B(x, r)$.

Donc $\|f(x + y)\| \leq M$.

Donc $\|f(x) + f(y)\| \leq M$ donc $\|f(y)\| - \|f(x)\| \leq \|f(y) + f(x)\| \leq M$.

Donc $\|f(y)\| \leq M + \|f(x)\| = M'$.

Pour tout $z \in B(0, 1)$, on a $rz \in B(0, r)$.

Donc d'après ce qui précède :

$$\|f(rz)\| \leq M'$$

$$\|rf(z)\| \leq M'$$

$$r\|f(z)\| \leq M'$$

$$\|f(z)\| \leq \frac{M'}{r}.$$

Donc f est bornée sur $B(0, 1)$. ■

Démonstration 1.120 ((6) \implies (5))

Pour tout $x \in E \setminus \{0\}$, $\left\| \frac{x}{\|x\|} \right\| = 1$.

Donc $\frac{x}{2\|x\|} \in B(0, 1)$.

f est bornée sur $B(0, 1)$: il existe $M > 0$ tel que

$$\forall t \in B(0, 1), \quad \|f(t)\| \leq M.$$

Donc pour tout $x \neq 0$, $\left\| f\left(\frac{x}{2\|x\|}\right) \right\| \leq M$.

Donc $\left\| \frac{1}{2\|x\|} f(x) \right\| \leq M$.

Donc $\frac{1}{2\|x\|} \|f(x)\| \leq M$.

Donc $\|f(x)\| \leq 2M\|x\|$.

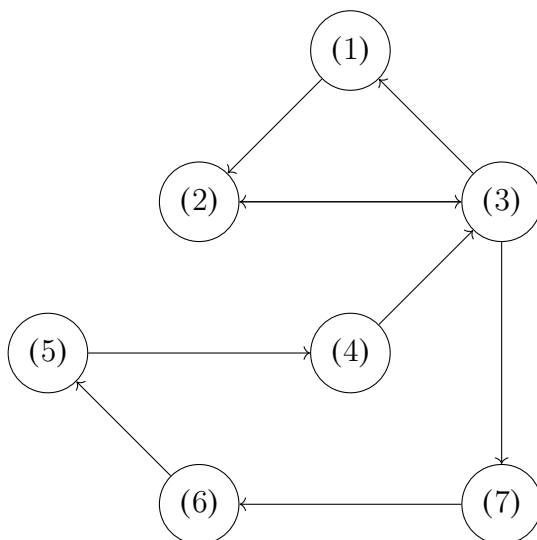
Ceci est vrai aussi pour $x = 0$: $\|f(0)\| = \|0\| = 0 \leq 2M\|0\|$. ■

Démonstration 1.121 ((5) \implies (4))

On a

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

On a donc montré toutes les équivalences :



■

Exercice 1.122

On pose $E = \mathcal{C}^0([0; 1], \mathbb{R})$ muni de la norme infinie.

L'application $f \mapsto \int_0^1 f(t) dt$ est-elle continue sur E ?

Correction 1.123

Pour $f \in E$, on a $\|f\|_\infty = \sup_{[0;1]} |f| \in \mathbb{R}$ car $|f|$ est continue sur le segment $[0; 1]$.

On note $I : f \mapsto \int_0^1 f(t) dt$. I est linéaire.

Pour $f \in E$, on a $|I(f)| \leq \int_0^1 |f|$.

Or $\forall t \in [0; 1], \quad |f(t)| \leq \|f\|_\infty$.

Donc

$$|I(f)| \leq \int_0^1 |f| \leq \int_0^1 \|f\|_\infty dt = \|f\|_\infty.$$

Donc d'après la Proposition 1.114, I est continue sur E .

Exercice 1.124

E désigne le même espace et on pose $\|f\|_1 = \int_0^1 |f(t)| dt$.

Montrez que $\|\cdot\|_1$ est une norme sur E .

L'application $f \mapsto f(1)$ est-elle continue sur E ?

Correction 1.125 ($\|\cdot\|_1$ est une norme sur E)

► Soit $f \in E$. Si $\|f\|_1 = 0$ alors $\int_0^1 |f| = 0$. Or $|f|$ est continue et positive donc d'après le théorème de stricte positivité de l'intégrale, $|f| = 0$ donc $f = 0$.

► Soient $f \in E$ et $\lambda \in \mathbb{R}$. On a

$$\|\lambda f\|_1 = \int_0^1 |\lambda f| = \int_0^1 |\lambda| |f| = |\lambda| \int_0^1 |f| = |\lambda| \|f\|_1.$$

► Soit $(f, g) \in E^2$. On a

$$\|f + g\|_1 = \int_0^1 |f + g| \leq \int_0^1 (|f| + |g|) = \int_0^1 |f| + \int_0^1 |g| = \|f\|_1 + \|g\|_1.$$

► Donc $\|\cdot\|_1$ est une norme sur E .

Correction 1.126 (Continuité de l'application ?)

On pose $V : E \rightarrow \mathbb{R}$
 $f \mapsto f(1)$

Pour $n \in \mathbb{N}^*$, on pose $f_n : x \mapsto \begin{cases} 0 & \text{si } x \in \left[0 ; 1 - \frac{1}{n}\right] \\ n^2 \left(x - \left(1 - \frac{1}{n}\right)\right) & \text{sinon} \end{cases}$

Pour $n \in \mathbb{N}^*$, on a $\int_0^1 |f| = \int_0^1 f_n = \frac{1}{2}$ donc $f_n \in B(0, 1)$.

Or $|V(f_n)| = |f_n(1)| = n \xrightarrow{n \rightarrow +\infty} +\infty$.

On a ainsi trouvé une suite (f_n) à termes dans $B(0, 1)$ telle que $V(f_n) \xrightarrow{n \rightarrow +\infty} +\infty$.

Donc V n'est pas bornée sur $B(0, 1)$.

Donc comme V est linéaire, V n'est pas continue sur $(E, \|\cdot\|_1)$.

Remarque : on a $\forall f \in E, |V(f)| = |f(1)| \leq \|f\|_\infty$ donc V est continue sur $(E, \|\cdot\|_\infty)$.

Définition 1.127

On note $\mathcal{L}_c(E, F)$ l'ensemble des applications linéaires continues de E dans F .

Proposition 1.128

$\mathcal{L}_c(E, F)$ est un sous-espace vectoriel de $\mathcal{L}(E, F)$, en général distinct de $\mathcal{L}(E, F)$.

Cas particulier en dimension finie.

Théorème 1.129

On suppose que E est de dimension finie.

Toute application linéaire de E dans F est lipschitzienne sur E , donc continue.

Autrement dit, si E est de dimension finie, alors $\mathcal{L}_c(E, F) = \mathcal{L}(E, F)$.

Démonstration 1.130

On note $p = \dim E$ et $\mathcal{B} = (e_1, \dots, e_p)$ une base de E .

Pour $x = (x_1, \dots, x_p)_{\mathcal{B}}$, on a $\|x\|_{\infty} = \sup_{1 \leq i \leq p} |x_i| = \max_{1 \leq i \leq p} |x_i|$.

Soit $f \in \mathcal{L}(E, F)$ et N une norme sur F .

Pour $x = (x_1, \dots, x_p)_{\mathcal{B}}$, on a $x = \sum_{k=1}^p x_k e_k$.

Donc, f étant linéaire, on a $f(x) = \sum_{k=1}^p x_k f(e_k)$.

Donc

$$\begin{aligned} N(f(x)) &= N\left(\sum_{k=1}^p x_k f(e_k)\right) \\ &\leq \sum_{k=1}^p N(x_k f(e_k)) \\ &= \sum_{k=1}^p |x_k| N(f(e_k)). \end{aligned}$$

De plus, on a

$$\begin{aligned} \forall i \in \llbracket 1 ; p \rrbracket, \quad |x_i| &\leq \|x\|_{\infty} \\ |x_i| N(f(e_i)) &\leq \|x\|_{\infty} N(f(e_i)). \end{aligned}$$

Donc $N(f(x)) \leq \|x\|_{\infty} \underbrace{\sum_{k=1}^p N(f(e_k))}_K$.

Ceci prouve d'après la Proposition 1.114 que f est continue de $(E, \|\cdot\|_\infty)$ dans (F, N) .

Soit maintenant $\|\cdot\|$ une norme quelconque sur E .

Comme E est de dimension finie, toutes les normes sur E sont équivalentes, donc il existe $a, b > 0$ tels que $a \|\cdot\| \leq \|\cdot\|_\infty \leq b \|\cdot\|$.

Donc pour tout $x \in E$, $N(f(x)) \leq bK \|x\|$.

Donc f est continue de $(E, \|\cdot\|)$ dans (F, N) . ■

Remarque 1.131

L'hypothèse de dimension finie de E est indispensable. Dans le cas contraire, c'est faux en général.

Le résultat précédent s'étend aux applications multilinéaires.

Théorème 1.132

Soient E_1, \dots, E_n des espaces vectoriels normés de dimensions finies et $f : E_1 \times \dots \times E_n \longrightarrow F$ une application n -linéaire.

Il existe alors une constante $K \geq 0$ telle que

$$\text{pour tout } (x_1, \dots, x_n) \in E_1 \times \dots \times E_n, \quad \|f(x_1, \dots, x_n)\| \leq K \|x_1\|_{E_1} \dots \|x_n\|_{E_n}.$$

Démonstration 1.133

Pour tout $i \in \llbracket 1 ; n \rrbracket$, on note $p_i = \dim E_i$ et $\mathcal{B}_i = (e_{i,1}, \dots, e_{i,p_i})$ une base de E_i .

Soit $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$. Pour tout $i \in \llbracket 1 ; n \rrbracket$, on note $x_i = (x_{i,1}, \dots, x_{i,p_i})_{\mathcal{B}_i}$.

On a

$$\begin{aligned} f(x_1, \dots, x_n) &= f\left(\sum_{j_1=1}^{p_1} x_{1,j_1} e_{1,j_1}, \dots, \sum_{j_n=1}^{p_n} x_{n,j_n} e_{n,j_n}\right) \\ &= \sum_{j_1=1}^{p_1} \dots \sum_{j_n=1}^{p_n} x_{1,j_1} \dots x_{n,j_n} f(e_{1,j_1}, \dots, e_{n,j_n}). \end{aligned}$$

Donc

$$\begin{aligned} N(f(x_1, \dots, x_n)) &\leq \sum_{\substack{1 \leq j_1 \leq p_1 \\ \vdots \\ 1 \leq j_n \leq p_n}} |x_{1,j_1}| \dots |x_{n,j_n}| N(f(e_{1,j_1}, \dots, e_{n,j_n})) \\ &\leq \|x_1\|_{1,\infty} \dots \|x_n\|_{n,\infty} \underbrace{\sum_{\substack{1 \leq j_1 \leq p_1 \\ \vdots \\ 1 \leq j_n \leq p_n}} N(f(e_{1,j_1}, \dots, e_{n,j_n}))}_{K}. \end{aligned}$$

On conclut de la même façon que dans la démonstration précédente. ■

Corollaire 1.134

Soient E_1, \dots, E_n des espaces vectoriels normés de dimensions finies.

Toute application $f : E_1 \times \dots \times E_n \longrightarrow F$ qui est n -linéaire est continue sur $E_1 \times \dots \times E_n$.

Exemple 1.135

- Le produit matriciel de $\mathcal{M}_{np}(\mathbb{K}) \times \mathcal{M}_{pq}(\mathbb{K})$ dans $\mathcal{M}_{nq}(\mathbb{K})$ est bilinéaire, donc continu.
- Un produit scalaire dans un espace euclidien est bilinéaire, donc continu.
- Le déterminant dans $\mathcal{M}_n(\mathbb{K})$ est n -linéaire par rapport aux colonnes, donc il est continu.

1.5.6 Norme subordonnée

On définit sur l'espace vectoriel $\mathcal{L}_c(E, F)$ des applications linéaires continues de E dans F la notion de norme subordonnée (relative aux deux normes sur E et F) ou norme triple.

Définition 1.136

Soit $f \in \mathcal{L}_c(E, F)$.

On pose $\|f\| = \sup_{x \in B(0,1)} \|f(x)\|$, appelée la norme subordonnée de f .

Remarque 1.137

Cette définition a un sens car f étant linéaire de E dans F et continue, elle est bornée sur $B(0,1)$ d'après la Proposition 1.114.

Remarque 1.138

On a

$$\|f\| = \sup_{x \in B(0,1)} \|f(x)\| = \sup_{x \in S(0,1)} \|f(x)\| = \sup_{x \in \overline{B}(0,1)} \|f(x)\|.$$

Proposition 1.139

Soit $f \in \mathcal{L}_c(E, F)$.

Alors $\|f\|$ est

- égal à $\sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|}$, mais aussi à $\sup_{x \in S(0,1)} \|f(x)\|$;
- le plus petit réel positif M tel que pour tout $x \in E$, $\|f(x)\| \leq M \|x\|$.

Démonstration 1.140

On note $N_1(f) = \sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|}$ et $N_2(f) = \sup_{x \in S(0,1)} \|f(x)\|$. ■

Démonstration 1.141 ($N_1(f) = N_2(f)$)

Pour tout $x \neq 0$, $\frac{x}{\|x\|} \in S(0,1)$.

Donc $\left\| f\left(\frac{x}{\|x\|}\right) \right\| \leq N_2(f)$.

Donc $\frac{1}{\|x\|} \|f(x)\| \leq N_2(f)$ i.e. $N_1(f) \leq N_2(f)$.

De plus, pour tout $x \in S(0,1)$, $\|x\| = 1$ donc $\frac{\|f(x)\|}{\|x\|} = \|f(x)\| \leq N_1(f)$.

Donc $N_2(f) \leq N_1(f)$.

Finalement, on a $N_1(f) = N_2(f)$. ■

Démonstration 1.142 ($N_2(f) = \|f\|$)

Pour tout $x \in B(0,1) \setminus \{0\}$, on a $\frac{x}{\|x\|} \in S(0,1)$ donc

$$\left\| f\left(\frac{x}{\|x\|}\right) \right\| \leq N_2(f)$$
$$\frac{1}{\|x\|} \|f(x)\| \leq N_2(f).$$

Or $\|x\| \leq 1$ donc $\|f(x)\| \leq \frac{1}{\|x\|} \|f(x)\| \leq N_2(f)$.

Ceci est encore vrai pour $x = 0$ donc $\|f\| \leq N_2(f)$.

De plus, soient $x \in S(0,1)$ et $\lambda \in [0; 1[$.

On a $\|\lambda x\| = \lambda < 1$ donc $\lambda x \in B(0,1)$.

Donc $\|f(\lambda x)\| = \lambda \|f(x)\| \leq \|f\|$.

Donc, par passage à la limite quand $\lambda \rightarrow 1$:

$$\|f(x)\| \leq \|f\|$$

i.e. $N_2(f) \leq \|f\|$.

Donc $N_2(f) = \|f\|$. ■

Démonstration 1.143 (Second point)

On a

$$\begin{aligned}\|f\| &= \sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|} \\ &= \min \left\{ K \in \mathbb{R} \mid \forall x \neq 0, \frac{\|f(x)\|}{\|x\|} \leq K \right\} \\ &= \min \{ K \in \mathbb{R} \mid \forall x \in E, \|f(x)\| \leq K \|x\| \}.\end{aligned}$$

■

Exemple 1.144

Dans l'Exercice 1.122, on avait montré $\forall f \in E, |I(f)| \leq \|f\|_\infty$.

On a $|I(1)| = 1 = 1 \times \|1\|_\infty$.

Donc $\|I\| = 1$.

Méthode 1.145

Si

$$\forall x \in E, \|f(x)\| \leq K \|x\|$$

et s'il existe $x_0 \in E$ tel que

$$\|f(x_0)\| = K \|x_0\|$$

alors $\|f\| = K$.

Proposition 1.146

Les normes subordonnées sont des normes sur les espaces $\mathcal{L}_c(E, F)$.

Elles sont dites sous-multiplicatives : pour toutes applications linéaires continues et composables f et g ,

$$\|f \circ g\| \leq \|f\| \times \|g\|.$$

Démonstration 1.147 ($\|\cdot\|$ est une norme sur $\mathcal{L}_c(E, F)$)

► Soit $f \in \mathcal{L}_c(E, F)$.

Si $\|f\| = 0$ alors $\forall x \in E, \|f(x)\| \leq 0 \times \|x\|$ donc $f = 0$.

► Soient $f \in \mathcal{L}_c(E, F)$ et $\lambda \in \mathbb{K}$.

On a

$$\|\lambda f\| = \sup_{x \neq 0} \frac{\|\lambda f(x)\|}{\|x\|} = \sup_{x \neq 0} |\lambda| \frac{\|f(x)\|}{\|x\|} = |\lambda| \sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|} = |\lambda| \|f\|.$$

► Soit $(f, g) \in \mathcal{L}_c(E, F)^2$.

On a

$$\begin{aligned} \forall x \in B(0, 1), \quad \|(f + g)(x)\| &= \|f(x) + g(x)\| \\ &\leq \|f(x)\| + \|g(x)\| \\ &\leq \|f\| + \|g\| \end{aligned}$$

donc $\|f + g\| \leq \|f\| + \|g\|$. ■

Démonstration 1.148 (Sous-multiplicativité)

On a

$$\begin{aligned} \forall x \in E, \quad \|f \circ g(x)\| &= \|f(g(x))\| \\ &\leq \|f\| \|g(x)\| \\ &\leq \|f\| \|g\| \|x\|. \end{aligned}$$

Donc $\|f \circ g\| \leq \|f\| \|g\|$. ■

Comme en dimension finie, on peut représenter par choix de bases les applications linéaires par des matrices, on définit de manière semblable la notion de norme sous-multiplicative de matrices (relativement aux normes) ou norme triple.

Définition 1.149

Soit $(n, p) \in (\mathbb{N}^*)^2$. On choisit deux normes sur \mathbb{K}^p et \mathbb{K}^n (espaces identifiés à ceux des matrices-colonnes).

Pour toute matrice $A \in \mathcal{M}_{np}(\mathbb{K})$, on pose $\|A\| = \sup_{\|X\|=1} \|AX\|$.

Proposition 1.150

Des normes étant choisies sur les espaces \mathbb{K}^p et \mathbb{K}^n , les normes subordonnées sont des normes sur tous les espaces $\mathcal{M}_{np}(\mathbb{K})$.

Elles sont dites sous-multiplicatives : pour toutes matrices multipliables A et B ,

$$\|AB\| \leq \|A\| \times \|B\|.$$

Remarque 1.151

Dans le cas où un espace vectoriel normé E est aussi un \mathbb{K} -algèbre, on dit qu'il est une algèbre normée quand la norme vérifie en plus la propriété de sous-multiplicativité : $\forall (x, y) \in E^2, \quad \|xy\| \leq \|x\| \cdot \|y\|$.

Remarque 1.152

En dimension finie, toute \mathbb{K} -algèbre A possède des normes sous-multiplicatives.

Démonstration 1.153

Soit A une \mathbb{K} -algèbre de dimension finie.

L'application $A^2 \longrightarrow A$ est bilinéaire donc continue.
 $(a, b) \longmapsto ab$

Il existe donc $K > 0$ tel que

$$\forall (a, b) \in A^2, \quad \|ab\| \leq K \|a\| \|b\|.$$

On pose $N = K \|\cdot\|$.

On a alors

$$\forall (a, b) \in A^2, \quad N(ab) \leq N(a) N(b)$$

et N est une norme sur A . ■

1.6 Topologie d'un espace vectoriel normé

Dans cette section, E est un espace vectoriel normé.

1.6.1 Intérieur d'une partie, voisinage d'un point

Définition 1.154

Soient A une partie de E et $a \in A$.

On dit que a est un point intérieur à A quand on peut trouver un rayon $r > 0$ tel que $B(a, r)$ soit incluse dans A . On dit aussi dans ce cas que A est un voisinage de a .

L'intérieur de A est l'ensemble de ses points intérieurs, noté $\overset{\circ}{A}$.

On a :

$$a \in \overset{\circ}{A} \iff \exists r > 0, \quad B(a, r) \subseteq A.$$

Exercice 1.155

Dans \mathbb{R} , quels sont les intérieurs des parties suivantes : $[0 ; 1]$, $[0 ; +\infty[$, \mathbb{Q} ?

Correction 1.156

► Si $A = [0 ; 1]$, alors $\overset{\circ}{A} =]0 ; 1[$.

En effet, pour $x \in]0 ; 1[$, on peut poser $r = \min\left(\frac{x}{2}, \frac{1-x}{2}\right) > 0$ pour avoir $B(x, r) \subseteq [0 ; 1]$.

► Si $A = [0 ; +\infty[$, alors $\overset{\circ}{A} =]0 ; +\infty[$ (même idée).

► Si $A = \mathbb{Q}$, alors $\overset{\circ}{A} = \emptyset$.

En effet, pour tout $x \in \mathbb{Q}$, pour tout $r > 0$, il existe $y \in \mathbb{R} \setminus \mathbb{Q}$ tel que $|x - y| < r$ i.e. $B(x, r) \not\subseteq \mathbb{Q}$.

Exercice 1.157

Quel est l'intérieur d'une boule de centre a et de rayon $r > 0$?

Correction 1.158

Soient $a \in E$ et $r > 0$.

Si $A = B(a, r)$, alors $\overset{\circ}{A} = A$.

En effet, pour tout $x \in B(a, r)$, on pose $p = \frac{r - \|x - a\|}{2} > 0$ et on a

$$B(x, p) \subseteq A.$$

Remarque 1.159

Cette notion dépend a priori de la norme utilisée. En dimension finie, ce n'est pas le cas : l'intérieur d'une partie d'un espace vectoriel normé de dimension finie ne dépend pas du choix de la norme (pourquoi ?).

Démonstration 1.160

Si N_1, N_2 sont deux normes équivalentes sur E , A est une partie de E et $a \in E$, alors a est intérieur à A pour N_1 ssi a est intérieur à A pour N_2 .

Il existe $\alpha, \beta > 0$ tels que $\alpha N_2 \leq N_1 \leq \beta N_2$.

Si a est intérieur à A pour N_1 , alors il existe $r > 0$ tel que $B_1(a, r) \subseteq A$.

On pose $p = \frac{r}{\beta} > 0$ et on montre $B_2(a, p) \subseteq A$.

Soit $x \in B_2(a, p)$.

On a $N_2(a - x) < p = \frac{r}{\beta}$.

Donc $N_1(a-x) \leq \beta N_2(a-x) < r$.

Donc $x \in B_1(a, r) \subseteq A$.

Donc x est intérieur à A pour N_2 .

On montre la réciproque de même, en montrant $B_1(a, ar) \subseteq B_2(a, r)$. ■

Proposition 1.161

Soient $u \in E^{\mathbb{N}}$ et $\ell \in E$.

La suite u converge vers ℓ ssi tout voisinage de ℓ contient tous les termes de la suite à partir d'un certain rang.

1.6.2 Parties ouvertes

Définition 1.162

On dit qu'une partie A de E est ouverte (ou est un ouvert) quand à tout point de $a \in A$, on peut associer un rayon $r > 0$ tel que la boule de centre a et de rayon r soit incluse dans A :

$$\forall a \in A, \exists r > 0, B(a, r) \subseteq A.$$

Autrement dit, A est ouverte quand tout point de A est intérieur à A : $A = \mathring{A}$, ou, autrement dit, quand A est un voisinage de chacun de ses points.

Proposition 1.163

L'ensemble vide et E sont des parties ouvertes. Toute boule ouverte est une partie ouverte. Tout produit (fini) de parties ouvertes est ouvert.

Démonstration 1.164

Soient E, F deux espaces vectoriels normés par $\|\cdot\|_E$ et $\|\cdot\|_F$.

On pose $N(x, y) = \max(\|x\|_E, \|y\|_F)$ pour obtenir une norme N sur $E \times F$.

Montrons que si A est un ouvert de E et B un ouvert de F , alors $A \times B$ est un ouvert de $E \times F$.

Soit $(a, b) \in A \times B$.

$a \in A$ et A est un ouvert donc il existe $r > 0$ tel que $B_E(a, r) \subseteq A$.

$b \in B$ et B est un ouvert donc il existe $s > 0$ tel que $B_F(b, s) \subseteq B$.

On pose $p = \min(r, s) > 0$.

Montrons que $B_{E \times F}((a, b), p) \subseteq A \times B$.

Soit $(x, y) \in B_{E \times F}((a, b), p)$.

On a $N((x, y) - (a, b)) < p$ i.e. $N(x - a, y - b) < p$.

Donc $\|x - a\|_E < p$ et $\|y - b\|_F < p$.

Donc $x \in B_E(a, p)$ et $y \in B_F(b, p)$.

Or $p \leq r$ donc $B_E(a, p) \subseteq B_E(a, r) \subseteq A$ et $p \leq s$ donc $B_F(b, p) \subseteq B_F(b, s) \subseteq B$.

Donc $(x, y) \in A \times B$.

Donc $B_{E \times F}((a, b), p) \subseteq A \times B$.

On généralise à un produit de plusieurs ouverts par récurrence. ■

La topologie de E est l'ensemble de tous les ouverts de E .

Remarque 1.165

La topologie dépend a priori de la norme utilisée. En dimension finie, ce n'est pas le cas : dans un espace vectoriel normé de dimension finie, le fait d'être un ouvert ne dépend pas du choix de la norme.

1.6.3 Parties fermées

On rappelle la notion de point adhérent à une partie.

Définition 1.166

Soient A une partie de E et $x \in E$.

On dit que x est un point adhérent à A quand il existe une suite $u \in A^{\mathbb{N}}$ qui converge vers x , ou, ce qui revient au même, quand toute boule centrée en x rencontre A , ou encore quand $d(x, A) = 0$.

L'adhérence de A est l'ensemble de ses points adhérents, noté \overline{A} .

On a montré

Définition 1.167

On dit qu'une partie A de E est fermée (ou est un fermé) quand tout point adhérent à A est dans A , autrement dit quand la propriété suivante est vraie :

si une suite quelconque à termes dans A converge vers un point x de E , alors $x \in A$.

Ou encore : A est fermée quand $A = \overline{A}$.

Proposition 1.168

L'ensemble vide et E sont des parties fermées. Toute boule fermée est une partie fermée. Tout produit (fini) de parties fermées est fermé.

On note le lien avec les parties ouvertes.

Proposition 1.169

Soit A une partie de E .

Alors A est une partie ouverte ssi son complémentaire est une partie fermée.

Démonstration 1.170

\Rightarrow

On suppose A ouverte. On veut montrer que $E \setminus A$ est fermée.

Soit $(x_n) \in (E \setminus A)^{\mathbb{N}}$ qui converge vers ℓ .

Par l'absurde, supposons $\ell \in A$.

A est ouverte donc il existe $\varepsilon > 0$ tel que $B(\ell, \varepsilon) \subseteq A$.

Or $x_n \xrightarrow[n \rightarrow +\infty]{} \ell$ donc il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, x_n \in B(\ell, \varepsilon).$$

Donc pour tout $n \geq N$, $x_n \in A$: contradiction.

Donc $\ell \in E \setminus A$.

Donc $E \setminus A$ est un fermé.

\Leftarrow

Supposons que $E \setminus A$ est fermée. On veut montrer que A est ouverte.

Soit $a \in A$.

Par l'absurde, on suppose $\forall r > 0, \exists x \in B(a, r), x \notin A$.

Alors pour tout $n \in \mathbb{N}$, il existe $x_n \in B\left(a, \frac{1}{n+1}\right)$ tel que $x_n \notin A$.

On a construit une suite $(x_n) \in (E \setminus A)^{\mathbb{N}}$ telle que $\forall n \in \mathbb{N}, \|a - x_n\| < \frac{1}{n+1}$.

Par théorème d'encadrement, on a $x_n \xrightarrow[n \rightarrow +\infty]{} a$.

Or $a \notin E \setminus A$: contradiction car $E \setminus A$ est un fermé.

Donc il existe $r > 0$ tel que $B(a, r) \subseteq A$.

Donc A est un ouvert. ■

Encore une fois, le fait d'être un fermé en dimension finie ne dépend pas de la norme.

Proposition 1.171

- *Toute réunion de parties ouvertes est ouverte. Toute intersection finie de parties ouvertes est ouverte.*
 - *Toute intersection de parties fermées est fermée. Toute réunion finie de parties fermées est fermée.*
-

Démonstration 1.172 (Réunion d'ouverts)

Soit $(A_i)_{i \in I}$ une famille de parties ouvertes.

Montrons que $\bigcup_{i \in I} A_i$ est ouverte.

Soit $x \in \bigcup_{i \in I} A_i$.

Il existe $i \in I$ tel que $x \in A_i$.

Or A_i est ouverte donc il existe $r > 0$ tel que $B(x, r) \subseteq A_i$.

Donc $B(x, r) \subseteq \bigcup_{i \in I} A_i$. ■

Démonstration 1.173 (Intersection finie d'ouverts)

Soient A_1, \dots, A_n des parties ouvertes.

Montrons que $\bigcap_{i=1}^n A_i$ est ouverte.

Soit $x \in \bigcap_{i=1}^n A_i$.

Pour tout $i \in \llbracket 1 ; n \rrbracket$, il existe $r_i > 0$ tel que $B(x, r_i) \subseteq A_i$.

On pose $r = \min_{1 \leq i \leq n} r_i > 0$.

Pour tout $i \in \llbracket 1 ; n \rrbracket$, $B(x, r) \subseteq B(x, r_i) \subseteq A_i$.

Donc $B(x, r) \subseteq \bigcap_{i=1}^n A_i$. ■

Remarque 1.174

Si la famille d'ouverts n'est pas finie, on ne peut rien dire sur l'intersection.

Par exemple, pour $n \in \mathbb{N}$, on pose les ouverts $A_n = \left] \frac{-1}{n+1} ; \frac{1}{n+1} \right[$.

Alors $\bigcap_{n \in \mathbb{N}} A_n = \{0\}$ n'est pas ouverte.

Exercice 1.175

Montrez que pour tout $a \in E$, $E \setminus \{a\}$ est un ouvert. D duisez-en que si A est une partie finie de E , alors $E \setminus A$ est un ouvert.

Correction 1.176

Pour tout $x \in E \setminus \{a\}$, on pose $r = \frac{\|x - a\|}{2}$.

Alors $B(x, r) \subseteq E \setminus \{a\}$.

Donc $E \setminus \{a\}$ est un ouvert.

Si $A = \{a_1, \dots, a_n\}$, alors $E \setminus A$ est le compl mentaire de $\bigcup_{i=1}^n \{a_i\}$, qui est un ferm  par union finie de ferm s, et est donc un ouvert.

Exercice 1.177

Quels sont les sous-espaces vectoriels de E qui sont ouverts ?

Correction 1.178

Soit F un sous-espace vectoriel de E ouvert dans E .

$0 \in F$ et F est un ouvert donc il existe $r > 0$ tel que $B(0, r) \subseteq F$.

Soit $x \in E \setminus \{0\}$.

On a $\frac{r}{2} \frac{x}{\|x\|} \in B(0, r)$ donc

$$x = \frac{2\|x\|}{r} \left(\frac{r}{2} \frac{x}{\|x\|} \right) \in F.$$

Donc $E = F$: E est le seul sous-espace vectoriel de E ouvert dans E .

Exercice 1.179

Montrez que $F = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0 \text{ et } xy = 1\}$ est un fermé de \mathbb{R}^2 .

Correction 1.180

Soit $((x_n, y_n)) \in F^{\mathbb{N}}$ qui converge vers (a, b) .

Montrons que $(a, b) \in F$.

On a $x_n \xrightarrow[n \rightarrow +\infty]{} a$, $y_n \xrightarrow[n \rightarrow +\infty]{} b$ et $\forall n \in \mathbb{N}$, $x_n y_n = 1$.

Donc par passage à la limite quand $n \rightarrow +\infty$, on a $a \geq 0$ et $ab = 1$.

Donc $(a, b) \in F$.

Donc F est un fermé.

Exercice 1.181

On note S l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ telles que tous les coefficients soient positifs et sur chaque ligne la somme des coefficients vaut 1.

Montrez que S est un fermé.

NB : S est l'ensemble des matrices dites stochastiques.

Remarque 1.182

A priori, une partie de E n'est ni ouverte ni fermée : par exemple, dans \mathbb{R} , l'ensemble $]0 ; 1]$ n'est ni ouvert ni fermé.

Donc ne pas confondre « complémentaire » et « contraire » : on peut dire qu'une partie est un fermé quand son complémentaire est un ouvert, mais pas que le contraire d'être un ouvert c'est être un fermé.

Remarque 1.183

Il est souvent assez facile de montrer qu'une partie est un fermé grâce à la caractérisation séquentielle. Donc pour montrer qu'une partie est un ouvert, on montre souvent de cette façon que son complémentaire est un fermé.

Les fermés sont souvent définis par des égalités ou des inégalités larges. Par complémentaire, les ouverts sont souvent définis par des inégalités strictes ou des différences.

1.6.4 Ouverts ou fermés relatifs à une partie

Les définitions précédentes parlent d'ouverts et de fermés de E . On peut définir ces notions relativement à une partie.

Définition 1.184

Soient A une partie de E et U un sous-ensemble de A .

On dit que U est un ouvert de A quand il existe un ouvert V de E tel que $U = A \cap V$.

On dit que U est un fermé de A quand il existe un fermé V de E tel que $U = A \cap V$.

On remarque que les fermés de A sont les complémentaires dans A des ouverts de A . On peut caractériser de même une partie U fermée de A par l'égalité entre U et l'ensemble de ses points adhérents dans A .

1.6.5 Image réciproque d'un ouvert ou d'un fermé par une fonction continue

Rappel 1.185

Si f est une fonction de E dans F définie sur D_f et $B \subseteq F$, l'image réciproque de B par f est

$$f^{-1}(B) = \{x \in D_f \mid f(x) \in B\}.$$

Théorème 1.186

Soit f une fonction de E dans F définie sur D .

Alors on a équivalence entre les propositions suivantes :

- (1) f est continue sur D ;
- (2) pour tout fermé B de F , son image réciproque $f^{-1}(B)$ est un fermé de D ;
- (3) pour tout ouvert B de F , son image réciproque $f^{-1}(B)$ est un ouvert de D .

Ceci est valable en particulier quand f est une application continue de E dans F , auquel cas on peut se passer des notions d'ouvert ou fermé relatif.

Démonstration 1.187 ((1) \implies (2))

Soient B un fermé de F et $(u_n) \in f^{-1}(B)^{\mathbb{N}}$ telle que $u_n \xrightarrow{n \rightarrow +\infty} \ell \in D$.

f étant continue sur D et donc en ℓ , on a $f(u_n) \xrightarrow{n \rightarrow +\infty} f(\ell)$.

De plus, on a $(f(u_n)) \in B^{\mathbb{N}}$ or B est un fermé donc $f(\ell) \in B$.

Donc $\ell \in f^{-1}(B)$.

Donc $f^{-1}(B)$ est un fermé de D . ■

Démonstration 1.188 ((2) \implies (3))

Soit A un ouvert de F .

Alors $F \setminus A$ est un fermé de F .

Donc $f^{-1}(F \setminus A)$ est un fermé de D .

Or $f^{-1}(F \setminus A) = D \setminus f^{-1}(A)$.

Donc $f^{-1}(A)$ est un ouvert de D . ■

Démonstration 1.189 ((3) \implies (1))

On suppose que pour tout A ouvert de F , $f^{-1}(A)$ est un ouvert de D .

Soit $d \in D$. Montrons que f est continue en d .

Soit $\varepsilon > 0$.

La boule $B(f(d), \varepsilon)$ est un ouvert de F .

Donc $f^{-1}(B(f(d), \varepsilon))$ est un ouvert de D .

Or $f(d) \in B(f(d), \varepsilon)$ donc $d \in f^{-1}(B(f(d), \varepsilon))$.

Donc il existe $\alpha > 0$ tel que $D \cap B(d, \alpha) \subseteq f^{-1}(B(f(d), \varepsilon))$.

Donc pour tout $x \in D$ tel que $x \in B(d, \alpha)$, on a $f(x) \in B(f(d), \varepsilon)$ i.e.

$$\forall x \in D, \|x - d\| < \alpha \implies \|f(x) - f(d)\| < \varepsilon.$$

Donc f est continue en d . ■

Exemple 1.190 (Cas particuliers fondamentaux)

Si f est continue sur E et à valeurs réelles, alors pour tout $a \in \mathbb{R}$, les ensembles suivants sont des fermés de E :

$$\{x \in E \mid f(x) \geq a\} \quad \{x \in E \mid f(x) \leq a\} \quad \{x \in E \mid f(x) = a\}.$$

Exemple 1.191

► Les courbes de fonctions continues de \mathbb{R} dans \mathbb{R} sont des fermés de \mathbb{R}^2 .

► L'ensemble des matrices de trace nulle est un fermé de $\mathcal{M}_n(\mathbb{K})$.

Démonstration 1.192 (Courbes des fonctions continues)

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ est continue, on pose $\Gamma_f = \{(x, y) \in \mathbb{R}^2 \mid y = f(x)\}$.

On a alors

$$\Gamma_f = \varphi^{-1}(\{0\})$$

où $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ est continue sur \mathbb{R}^2 car f est continue sur \mathbb{R} .
 $(x, y) \mapsto y - f(x)$

Or $\{0\}$ est un fermé de \mathbb{R} donc Γ_f est un fermé de \mathbb{R}^2 . ■

Démonstration 1.193 (Ensemble des matrices de trace nulle)

L'ensemble des matrices de trace nulle dans $\mathcal{M}_n(\mathbb{K})$ est

$$T = \{M \in \mathcal{M}_n(\mathbb{K}) \mid \text{tr}(M) = 0\}.$$

Or $\mathcal{M}_n(\mathbb{K})$ est de dimension finie et tr est linéaire donc tr est continue.

Donc T est l'image réciproque du fermé $\{0\}$ par l'application continue tr .

Donc T est un fermé de $\mathcal{M}_n(\mathbb{K})$. ■

Par passage au complémentaire, si f est continue sur E et à valeurs réelles, alors pour tout $a \in \mathbb{R}$, les ensembles suivants sont des ouverts de E :

$$\{x \in E \mid f(x) < a\} \quad \{x \in E \mid f(x) > a\} \quad \{x \in E \mid f(x) \neq a\}.$$

Exemple 1.194

- L'ensemble des couples $(x, y) \in \mathbb{R}^2$ tels que $x > 0$ et $y > x$ est un ouvert de \mathbb{R}^2 .
- $\text{GL}_n(\mathbb{K})$ est un ouvert de $\mathcal{M}_n(\mathbb{K})$: si une matrice A est inversible, alors toutes les matrices proches de A le sont aussi.

Démonstration 1.195 (Ensemble des couples susmentionnés)

On pose

$$\begin{aligned} A &= \{(x, y) \in \mathbb{R}^2 \mid x > 0 \text{ et } y > x\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid x \in]0 ; +\infty[\} \cap \{(x, y) \in \mathbb{R}^2 \mid y - x \in]0 ; +\infty[\}. \end{aligned}$$

Or $(x, y) \mapsto x$ et $(x, y) \mapsto y - x$ sont continues.

Donc A est un ouvert de \mathbb{R}^2 . ■

Démonstration 1.196 ($\text{GL}_n(\mathbb{K})$)

On a $\text{GL}_n(\mathbb{K}) = \det^{-1}(\mathbb{K} \setminus \{0\})$.

Or \det est continue et $\mathbb{K} \setminus \{0\}$ est un ouvert de \mathbb{K} donc $\text{GL}_n(\mathbb{K})$ est un ouvert de $\mathcal{M}_n(\mathbb{K})$. ■

1.6.6 Frontière d'une partie

Définition 1.197

Soit A une partie de E . On appelle frontière de A l'ensemble $\overline{A} \setminus \overset{\circ}{A}$.

Exemple 1.198

- Si B est une boule, alors son intérieur est la boule ouverte de même centre et de même rayon, son adhérence est la boule fermée et sa frontière est la sphère.
- L'ensemble des rationnels est d'intérieur vide, d'adhérence égale à \mathbb{R} et donc de frontière \mathbb{R} .

1.7 Compacité

Dans cette section, E est un espace vectoriel normé.

1.7.1 Valeurs d'adhérence d'une suite

Définition 1.199

Soient $u = (u_n) \in E^{\mathbb{N}}$ et $a \in E$.

On dit que a est une valeur d'adhérence de la suite u quand il existe une extractrice φ telle que la suite extraite $(u_{\varphi(n)})$ converge vers a .

Une suite peut avoir une ou plusieurs valeurs d'adhérence ou ne pas avoir de valeur d'adhérence :

- la suite $(n)_{n \in \mathbb{N}}$ n'a pas de valeur d'adhérence ;
- toute suite convergente possède une seule valeur d'adhérence : sa limite ;
- la suite u définie par $u_{2n} = \frac{1}{n+1}$ et $u_{2n+1} = 1 - \frac{1}{n+1}$ possède deux valeurs d'adhérence : 0 et 1 ;
- il est possible de numérotter les rationnels, autrement dit de créer une suite u qui prend exactement toutes les valeurs rationnelles dans \mathbb{R} : cette suite a pour valeurs d'adhérence tous les réels.

On peut donner une caractérisation équivalente sans passer par la notion de suite extraite.

Proposition 1.200

Soient $u = (u_n) \in E^{\mathbb{N}}$ et $a \in E$.

Alors a est une valeur d'adhérence de u ssi pour tout $\varepsilon > 0$, $\{n \in \mathbb{N} \mid u_n \in B(a, \varepsilon)\}$ est infini.

Démonstration 1.201



Supposons que pour tout $\varepsilon > 0$, $\{n \in \mathbb{N} \mid u_n \in B(a, \varepsilon)\}$ est infini.

On spécialise $\varepsilon \leftarrow \frac{1}{k+1}$ pour $k \in \mathbb{N}$.

L'ensemble $\{n \in \mathbb{N} \mid u_n \in B(a, 1)\}$ est infini donc non-vidé. On choisit $\varphi(0)$ un élément de cet ensemble.

L'ensemble $\left\{n \in \mathbb{N} \mid u_n \in B\left(a, \frac{1}{2}\right)\right\}$ est infini donc il contient des entiers strictement supérieurs à $\varphi(0)$; on en choisit un, qu'on note $\varphi(1)$.

Si on suppose avoir construit $\varphi(0) < \varphi(1) < \dots < \varphi(k)$ tels que $u_{\varphi(0)} \in B(a, 1)$, $u_{\varphi(1)} \in B\left(a, \frac{1}{2}\right)$, ..., $u_{\varphi(k)} \in B\left(a, \frac{1}{k+1}\right)$, comme l'ensemble

$$\left\{n \in \mathbb{N} \mid u_n \in B\left(a, \frac{1}{k+2}\right)\right\}$$

est infini, on peut choisir $\varphi(k+1)$ dans cet ensemble tel que $\varphi(k+1) > \varphi(k)$.

Par récurrence, on construit une suite $(\varphi(k))_{k \in \mathbb{N}}$ strictement croissante d'entiers naturels tels que

$$\forall k \in \mathbb{N}, u_{\varphi(k)} \in B\left(a, \frac{1}{k+1}\right)$$

$$i.e. \|u_{\varphi(k)} - a\| < \frac{1}{k+1}.$$

Par théorème d'encadrement, on a $u_{\varphi(k)} \xrightarrow[k \rightarrow +\infty]{} a$: a est une valeur d'adhérence de la suite (u_n) .



Supposons que a est une valeur d'adhérence de la suite (u_n) .

Il existe alors une extractrice φ telle que $u_{\varphi(k)} \xrightarrow[k \rightarrow +\infty]{} a$.

Donc pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geq N, \|u_{\varphi(n)} - a\| < \varepsilon.$$

Donc $\{n \in \mathbb{N} \mid u_n \in B(a, \varepsilon)\}$ contient $\varphi(N), \varphi(N+1), \dots$ i.e. c'est un ensemble infini. ■

Ceci peut encore être réécrit de la façon suivante.

Proposition 1.202

Soient $u = (u_n) \in E^{\mathbb{N}}$ et $a \in E$.

Alors a est une valeur d'adhérence de u ssi $\forall \varepsilon > 0, \forall N \in \mathbb{N}, \exists n \geq N, \|u_n - a\| < \varepsilon$.

Démonstration 1.203

Soit I une partie de \mathbb{N} .

On a

$$\begin{aligned} I \text{ est infini} &\iff I \text{ n'est pas majorée} \\ &\iff \neg (\exists N \in \mathbb{N}, \forall n \in I, n \leq N) \\ &\iff \forall N \in \mathbb{N}, \exists n \in I, n > N. \end{aligned}$$

Pour tout $\varepsilon > 0$, on pose $I_\varepsilon = \{n \in \mathbb{N} \mid u_n \in B(a, \varepsilon)\}$.

On a alors, d'après la Proposition 1.200 :

$$\begin{aligned} a \text{ est une valeur d'adhérence de } u &\iff \forall \varepsilon > 0, I_\varepsilon \text{ est infini} \\ &\iff \forall \varepsilon > 0, \forall N \in \mathbb{N}, \exists n \in I_\varepsilon, n > N \\ &\iff \forall \varepsilon > 0, \forall N \in \mathbb{N}, \exists n > N, \|u_n - a\| < \varepsilon. \quad \blacksquare \end{aligned}$$

Exercice 1.204

Soit $u = (u_n) \in E^{\mathbb{N}}$. Montrez que l'ensemble V des valeurs d'adhérence de la suite u est un fermé de E en utilisant les ensembles $U_p = \{u_n \mid n \geq p\}$.

Correction 1.205

Soit $a \in E$.

On a

$$\begin{aligned} a \in V &\iff \forall \varepsilon > 0, \forall N \in \mathbb{N}, \exists x \in U_N, \|x - a\| < \varepsilon \\ &\iff \forall \varepsilon > 0, \forall N \in \mathbb{N}, U_N \cap B(a, \varepsilon) \neq \emptyset \\ &\iff \forall N \in \mathbb{N}, \forall \varepsilon > 0, U_N \cap B(a, \varepsilon) \neq \emptyset \\ &\iff \forall N \in \mathbb{N}, a \in \overline{U_N} \\ &\iff a \in \bigcap_{N \in \mathbb{N}} \overline{U_N}. \end{aligned}$$

$$\text{Donc } V = \bigcap_{N \in \mathbb{N}} \overline{U_N}.$$

V est donc un fermé par intersection de fermés.

1.7.2 Théorème de Bolzano-Weierstrass

Théorème 1.206

Si E est de dimension finie, alors toute suite bornée de E possède une valeur d'adhérence.

Démonstration 1.207

On note $\mathcal{P}(k)$ le prédicat « si E est de dimension k , alors toute suite bornée de E possède une valeur d'adhérence ».

► Pour $k = 1$:

On pose $E = \text{Vect}(e_1)$.

Si (u_n) est une suite bornée de E , en notant $(u_n) = (\lambda_n e_1)$ où (λ_n) est une suite bornée de \mathbb{K} , d'après le théorème de Bolzano-Weierstrass dans \mathbb{R} ou \mathbb{C} , (λ_n) possède une valeur d'adhérence et donc (u_n) aussi.

D'où $\mathcal{P}(1)$.

► Soit $k \in \mathbb{N}^*$ tel que $\mathcal{P}(k)$ soit vraie.

Soit E de dimension $k + 1$.

On choisit une base $\mathcal{B} = (e_1, \dots, e_{k+1})$ de E .

Soit $(u_n) \in E^{\mathbb{N}}$ une suite bornée.

Alors les suites-coordonnées associées sont bornées.

Pour tout $n \in \mathbb{N}$, on note $u_n = (u_{1,n}, \dots, u_{k+1,n})_{\mathcal{B}}$.

$(u_{k+1,n})_{n \in \mathbb{N}}$ est une suite bornée de \mathbb{K} donc (même théorème) il existe une extractrice φ telle que $(u_{k+1,\varphi(n)})_{n \in \mathbb{N}}$ converge.

Pour tout $n \in \mathbb{N}$, on pose $v_n = (u_{1,n}, \dots, u_{k,n}, 0)_{\mathcal{B}}$.

$(v_{\varphi(n)})_{n \in \mathbb{N}}$ est une suite de vecteurs de $\text{Vect}(e_1, \dots, e_k)$ et bornée donc par hypothèse de récurrence, il existe une extractrice ψ telle que $(v_{\varphi \circ \psi(n)})_{n \in \mathbb{N}}$ converge.

De plus, $(u_{k+1,\varphi \circ \psi(n)})_{n \in \mathbb{N}}$ converge car c'est une suite extraite d'une suite convergente.

Donc $(u_{\varphi \circ \psi(n)})_{n \in \mathbb{N}}$ converge.

Donc $\mathcal{P}(k + 1)$ est vrai.

► Donc pour tout $k \in \mathbb{N}^*$, $\mathcal{P}(k)$ est vrai. ■

Remarque 1.208

Ce théorème est faux en dimension infinie donc il faut bien mettre en valeur la dimension finie.

On peut ajouter une précision au théorème précédent.

Proposition 1.209

Si E est de dimension finie, alors toute suite bornée de E qui ne possède qu'une seule valeur d'adhérence est convergente vers cette valeur d'adhérence.

Démonstration 1.210

Supposons E de dimension finie.

Soit $(u_n) \in E^{\mathbb{N}}$ une suite bornée qui admet une unique valeur d'adhérence ℓ .

Par l'absurde, on suppose que (u_n) ne converge pas vers ℓ .

On a

$$\exists \varepsilon > 0, \quad \underbrace{\forall N \in \mathbb{N}, \quad \exists n \geq N, \quad \|u_n - \ell\| \geq \varepsilon}_{\{n \in \mathbb{N} \mid u_n \notin B(\ell, \varepsilon)\} \text{ infini}}.$$

En ordonnant les éléments de cet ensemble et en les notant $\varphi(0) < \varphi(1) < \dots$, on construit une extractrice φ telle que

$$\forall n \in \mathbb{N}, \quad u_{\varphi(n)} \notin B(\ell, \varepsilon).$$

Or $(u_{\varphi(n)})$ est bornée et E est de dimension finie donc d'après le théorème de Bolzano-Weierstrass, il existe une extractrice ψ et $\ell' \in E$ tels que

$$u_{\varphi \circ \psi(n)} \xrightarrow{n \rightarrow +\infty} \ell'.$$

Or pour tout $n \in \mathbb{N}$, $u_{\varphi(n)}$ appartient au fermé $E \setminus B(\ell, \varepsilon)$ donc $\ell' \in E \setminus B(\ell, \varepsilon)$.

Donc $\ell' \neq \ell$.

Donc $\ell' = \lim_{n \rightarrow +\infty} u_{\varphi \circ \psi(n)}$ est une autre valeur d'adhérence de (u_n) : contradiction.

Donc $u_n \xrightarrow{n \rightarrow +\infty} \ell$. ■

1.7.3 Parties compactes

Définition 1.211

Soit A une partie de E .

On dit que A est une partie compacte de E (ou un compact de E) quand toute suite à termes dans A possède une valeur d'adhérence dans A (propriété dite de Bolzano-Weierstrass).

Exemple 1.212

- ▶ Tout segment $[a ; b]$ de \mathbb{R} est un compact et ce sont les seuls intervalles compacts. $[0 ; 1] \cup [2 ; 3]$ est compact.
- ▶ Dans \mathbb{K}^n , tout pavé $\prod_{i=1}^n [a_i ; b_i]$ est un compact. Plus généralement, un produit (fini) de compacts est compact.

Les parties compactes sont donc celles dont on peut extraire des sous-suites convergentes. Un résultat précédent se généralise alors.

Proposition 1.213

Si A est une partie compacte, alors toute suite de A qui ne possède qu'une seule valeur d'adhérence est convergente vers cette valeur d'adhérence.

Un compact étant connu, il est facile d'en construire d'autres.

Proposition 1.214

Si A est une partie compacte de E , alors toute partie B fermée dans A est aussi compacte.

Démonstration 1.215

Soient A une partie compacte de E , B un fermé de A et $(u_n) \in B^{\mathbb{N}}$.

Comme $B \subseteq A$, on a $(u_n) \in A^{\mathbb{N}}$.

A est compacte donc il existe une extractrice φ et $\ell \in A$ tels que $u_{\varphi(n)} \xrightarrow[n \rightarrow +\infty]{} \ell$.

La suite $(u_{\varphi(n)})$ est à termes dans B et converge vers ℓ donc comme B est un fermé, on a $\ell \in B$.

Ainsi, toute suite de $B^{\mathbb{N}}$ possède une valeur d'adhérence dans B i.e. B est un compact. ■

Reconnaître si une partie est compacte n'est pas toujours facile. On dispose d'une condition nécessaire, qui est suffisante en dimension finie.

Proposition 1.216

Soit A une partie de E .

Si A est compacte, alors A est une partie fermée et bornée.

Démonstration 1.217

► Si A n'est pas bornée, alors pour tout $n \in \mathbb{N}$, il existe $a_n \in A$ tel que $\|a_n\| \geq n$.

Si (a_n) possède une valeur d'adhérence dans A , alors il existe une extractrice φ et $\ell \in A$ tels que $a_{\varphi(n)} \xrightarrow{n \rightarrow +\infty} \ell$.

Alors $\|a_{\varphi(n)}\| \xrightarrow{n \rightarrow +\infty} \|\ell\|$: contradiction.

Donc A n'est pas compacte.

► Supposons que A est compacte.

Soit $(u_n) \in A^{\mathbb{N}}$ telle que $u_n \xrightarrow{n \rightarrow +\infty} \ell \in E$.

A étant compacte, il existe φ une extractrice et $\ell' \in A$ tels que $u_{\varphi(n)} \xrightarrow{n \rightarrow +\infty} \ell'$.

La suite $(u_{\varphi(n)})$ est extraite de la suite convergente (u_n) donc $u_{\varphi(n)} \xrightarrow{n \rightarrow +\infty} \ell$.

Donc par unicité de la limite, on a $\ell' = \ell \in A$.

Donc A est fermée. ■

La réciproque est hélas fausse en général. Néanmoins, en dimension finie, elle est vraie.

Proposition 1.218

Si E est de dimension finie, alors une partie de E est compacte ssi elle est fermée et bornée.

Remarque 1.219

En fait, il n'y a qu'en dimension finie que ce résultat est vrai. Un théorème de Riesz affirme que la boule-unité fermée d'un espace vectoriel normé est compacte ssi l'espace est de dimension finie, ce qui revient à dire que l'équivalence précédente n'est valable que dans un espace de dimension finie.

En dimension infinie, il se passe des choses vraiment étranges : les compacts sont des parties très petites et plates, par exemple, un compact est forcément d'intérieur vide. Heureusement, il est plus courant de travailler à notre niveau en dimension finie.

Exemple 1.220

► L'ensemble des matrices stochastiques de $\mathcal{M}_n(\mathbb{R})$ est un compact.

► La boule-unité fermée de $E = \mathcal{C}^0([0; 1], \mathbb{R})$ pour la norme infinie n'est pas compacte, car la suite des fonctions $(x \mapsto x^n)$ a pour seule valeur d'adhérence possible la fonction $x \mapsto 0$ si $x \neq 1$ et $1 \mapsto 1$, qui n'est même pas dans l'espace E .

Démonstration 1.221 (Matrices stochastiques)

On note

$$S_n = \left\{ M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{R}) \mid \forall (i,j) \in \llbracket 1 ; n \rrbracket^2, m_{i,j} \geq 0 \text{ et } \forall i \in \llbracket 1 ; n \rrbracket, \sum_{j=1}^n m_{i,j} = 1 \right\}.$$

Soit $M = (m_{i,j}) \in S_n$.

Pour tout $i \in \llbracket 1 ; n \rrbracket$, $\sum_{j=1}^n m_{i,j}$ est une somme de réels positifs qui vaut 1 donc pour tout $j \in \llbracket 1 ; n \rrbracket$, $0 \leq m_{i,j} \leq 1$.

Donc $\|M\|_\infty \leq 1$.

Donc S_n est bornée.

Soit $(M_k) = \left((m_{i,j}^k)_{i,j} \right)_k$ une suite de matrices de S_n qui converge vers $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$:

$$\forall (i,j) \in \llbracket 1 ; n \rrbracket^2, m_{i,j}^k \xrightarrow[k \rightarrow +\infty]{} a_{i,j}.$$

Par passage à la limite quand $k \rightarrow +\infty$ dans les deux conditions qui définissent S_n , on obtient

$$\forall (i,j) \in \llbracket 1 ; n \rrbracket^2, a_{i,j} \geq 0 \quad \text{et} \quad \forall i \in \llbracket 1 ; n \rrbracket, \sum_{j=1}^n a_{i,j} = 1.$$

Donc $A \in S_n$.

Donc S_n est fermée.

On aurait aussi pu considérer les fonctions continues sur $\mathcal{M}_n(\mathbb{R})$

$$c_{i,j} : (m_{i,j}) \mapsto m_{i,j} \quad \text{et} \quad s_i : (m_{i,j}) \mapsto \sum_{j=1}^n m_{i,j}$$

et remarquer que

$$S_n = \bigcap_{1 \leq i,j \leq n} c_{i,j}^{-1}([0 ; +\infty[) \cap \bigcap_{i=1}^n s_i^{-1}(\{1\})$$

ce qui montre que S_n est un fermé par intersection de fermés.

Alors, comme $\mathcal{M}_n(\mathbb{R})$ est de dimension finie, on en déduit que S_n est un compact de $\mathcal{M}_n(\mathbb{R})$. ■

Démonstration 1.222 (Deuxième point)

On pose $f_n : x \mapsto x^n$.

Pour tout $n \in \mathbb{N}$, $\|f\|_\infty = 1$.

Si (f_n) a une valeur d'adhérence $g \in \overline{B}(0, 1)$, alors il existe une extractrice φ telle que $f_{\varphi(n)} \xrightarrow{n \rightarrow +\infty} g$ i.e. $\|f_{\varphi(n)} - g\|_\infty \xrightarrow{n \rightarrow +\infty} 0$.

Or pour tout $x \in [0 ; 1]$, $|f_{\varphi(n)}(x) - g(x)| \leq \|f_{\varphi(n)} - g\|_\infty$.

Donc par encadrement, on a $f_{\varphi(n)}(x) \xrightarrow{n \rightarrow +\infty} g(x)$.

Si $x \in [0 ; 1[$, alors $f_{\varphi(n)}(x) = x^{\varphi(n)} \xrightarrow{n \rightarrow +\infty} 0$.

Si $x = 1$, alors $f_{\varphi(n)}(x) = 1 \xrightarrow{n \rightarrow +\infty} 1$.

Donc $g : x \mapsto \begin{cases} 0 & \text{si } x \in [0 ; 1[\\ 1 & \text{sinon} \end{cases}$

Or $g \notin E$: contradiction de la compacité. ■

Une application importante de la notion de compacité est le théorème suivant.

Théorème 1.223

Tout sous-espace vectoriel de dimension finie de E est fermé.

Démonstration 1.224

Soit F un sous-espace vectoriel de E de dimension finie.

Soit $(u_n) \in F^{\mathbb{N}}$ une suite convergente vers $\ell \in E$.

Alors (u_n) est bornée : il existe $R > 0$ tel que $\forall n \in \mathbb{N}, u_n \in \overline{B}(0, R)$.

Donc $\forall n \in \mathbb{N}, u_n \in \overline{B}(0, R) \cap F = \{x \in F \mid \|x\| \leq R\} = \overline{B_F}(0, R)$.

Donc $\overline{B_F}(0, R)$ est un fermé borné de F et donc un compact de F .

Il existe donc une extractrice φ et $a \in \overline{B_F}(0, R)$ tels que $u_{\varphi(n)} \xrightarrow{n \rightarrow +\infty} a$.

Donc $\ell = a \in F$.

Donc F est fermé. ■

En dimension infinie, là encore il peut se passer des choses étranges : un sous-espace de E de dimension infinie peut être dense (et donc non-fermé s'il est différent de E).

1.7.4 Théorème des bornes atteintes

Le principal intérêt des compacts est de pouvoir généraliser un théorème de première année.

Théorème 1.225

Soient E, F deux espaces vectoriels normés, A une partie de E et $f : A \longrightarrow F$.

Si f est continue sur A et A est compacte, alors $f(A)$ est compacte.

Démonstration 1.226

On suppose que f est continue et que A est compacte.

Soit $(u_n) \in f(A)^{\mathbb{N}}$.

Pour tout $n \in \mathbb{N}$, $u_n \in f(A)$ donc il existe $v_n \in A$ tel que $u_n = f(v_n)$.

$(v_n) \in A^{\mathbb{N}}$ et A est compacte donc il existe une extractrice φ et $a \in A$ tels que $v_{\varphi(n)} \xrightarrow[n \rightarrow +\infty]{} a$.

f est continue en a donc $u_{\varphi(n)} = f(v_{\varphi(n)}) \xrightarrow[n \rightarrow +\infty]{} f(a) \in f(A)$.

Donc $f(A)$ est compacte. ■

On résume en disant que l'image continue d'un compact est un compact.

En particulier, toute fonction continue sur un compact est donc bornée. Dans le cas des fonctions numériques (*i.e.* à valeurs dans \mathbb{R}), on peut même être plus précis.

Théorème 1.227

Toute fonction continue sur un compact et à valeurs réelles est bornée et atteint ses bornes.

Autrement dit, si $f : A \longrightarrow \mathbb{R}$ est continue sur A et A est une partie compacte de E , alors il existe $(a, b) \in A^2$ tel que pour tout $x \in A$, $f(a) \leq f(x) \leq f(b)$, ce qui revient à dire que f possède un minimum et un maximum sur A .

Démonstration 1.228

$f(A)$ est un fermé borné de \mathbb{R} donc possède un minimum et un maximum. ■

Remarque 1.229

Pour toute partie X bornée de \mathbb{R} non-vide, $\sup X$ et $\inf X$ sont dans l'adhérence de X .

Remarque 1.230

Ce théorème est à rapprocher du théorème vu en première année : toute fonction de \mathbb{R} dans \mathbb{R} continue sur un segment est bornée et atteint ses bornes.

Néanmoins, le théorème de l'an dernier donnait un résultat un peu plus précis que celui de cette année car il donnait aussi l'image du segment, en précisant qu'il s'agissait aussi d'un segment, car il faisait aussi intervenir le théorème des valeurs intermédiaires.

Ici, dans la version proposée cette année, on ne peut rien dire de plus.

Exercice 1.231

Un exercice classique, à savoir refaire ! C'est la base de nombreux exercices.

Soient E de dimension finie et $f : E \rightarrow \mathbb{R}$ continue et telle que $f(x)$ tende vers $+\infty$ quand $\|x\|$ tend vers $+\infty$. Montrez que f possède un minimum.

Exemple : dans le plan euclidien géométrique, on choisit trois points A, B, C ; montrez alors qu'il existe un point M du plan tel que la somme $AM + BM + CM$ soit minimale.

Correction 1.232 (Cas général)

On a $f(0) \in \mathbb{R}$ donc il existe $A > 0$ tel que $\forall x \in E, \|x\| > A \implies f(x) \geq f(0)$.

Sur $\overline{B}(0, A)$, fermé borné d'un espace de dimension finie donc un compact, f est continue et y admet donc un minimum en x_0 d'après le théorème des bornes atteintes.

Pour tout $x \in E$,

▷ si $x \notin \overline{B}(0, A)$, alors $f(x) \geq f(0) \geq f(x_0)$ car $0 \in \overline{B}(0, A)$

▷ si $x \in \overline{B}(0, A)$, alors $f(x) \geq f(x_0)$.

Donc $f(x_0) = \min_E f$.

Correction 1.233 (Exemple)

On note \mathcal{P} le plan considéré.

On pose $f : \mathcal{P} \rightarrow \mathbb{R}$ qui est une fonction continue.

$$M \begin{pmatrix} x \\ y \end{pmatrix} \mapsto AM + BM + CM$$

Par inégalité triangulaire, on a $f(M) \geq 3OM + \text{cte}$ donc

$$f(M) \xrightarrow{\|\vec{OM}\| \rightarrow +\infty} +\infty.$$

D'où l'existence d'un minimum d'après la propriété démontrée précédemment.

Exercice 1.234

Soit $f : (x, y) \mapsto xy\sqrt{1-x^2-2y^2}$.

Justifiez que l'ensemble de définition D de f est un compact de \mathbb{R}^2 .

Déterminez les points critiques de f dans l'ouvert $\overset{\circ}{D}$, puis les maxima et minima de f .

Correction 1.235

On a $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + 2y^2 \leq 1\}$.

Pour tout $(x, y) \in D$, on a

$$x^2 \leq 1 \text{ donc } |x| \leq 1$$

et

$$2y^2 \leq 1 \text{ donc } |y| \leq \frac{1}{\sqrt{2}} \leq 1$$

donc $\|(x, y)\|_\infty \leq 1$ donc D est borné.

De plus, $D = \varphi^{-1}([-\infty; 1])$ où $\varphi : (x, y) \mapsto x^2 + 2y^2$ continue sur \mathbb{R}^2 donc D est un fermé.

\mathbb{R}^2 est de dimension finie donc D est un compact.

f est continue sur D donc d'après le théorème des bornes atteintes, $\max_D f$ et $\min_D f$ existent.

Sur \mathring{D} , f est de classe \mathcal{C}^1 .

$$(x, y) \in \mathring{D} \text{ est un point critique de } f \text{ ssi } \nabla f(x, y) = 0 \text{ i.e. } \begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases}$$

Or

$$\begin{aligned} \begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases} &\iff \begin{cases} y\sqrt{1-x^2-2y^2} - xy\frac{x}{\sqrt{1-x^2-2y^2}} = 0 \\ x\sqrt{1-x^2-2y^2} - xy\frac{2y}{\sqrt{1-x^2-2y^2}} = 0 \end{cases} \\ &\iff \begin{cases} y(1-x^2-2y^2) - x^2y = 0 \\ x(1-x^2-2y^2) - 2xy^2 = 0 \end{cases} \\ &\iff (S) \begin{cases} y(1-x^2-2y^2) = 0 \\ x(1-x^2-4y^2) = 0 \end{cases} \end{aligned}$$

Si $x = 0$ alors $y = 0$ donc une solution : $(0, 0)$.

Si $x \neq 0$, alors

$$\begin{aligned} (S) &\iff \begin{cases} x^2 = 1 - 4y^2 \\ y(1 - 2(1 - 4y^2) - 2y^2) = 0 \end{cases} \\ &\iff \begin{cases} x^2 = 1 - 4y^2 \\ y = 0 \text{ ou } y^2 = \frac{1}{6} \end{cases} \\ &\iff \begin{cases} x^2 = \frac{1}{3} \\ y^2 = \frac{1}{6} \end{cases} \end{aligned}$$

On en déduit quatre autres solutions : $\left(\frac{t}{\sqrt{3}}, \frac{t}{\sqrt{6}}\right)$ où $t \in \{-1, 1\}$.

★★ À finir ★★

On retrouve aussi le théorème de Heine en conséquence de la compacité.

Définition 1.236

Soient E, F deux espaces vectoriels normés, A une partie de E et $f : A \longrightarrow F$.

On dit que f est uniformément continue sur A quand

$$\forall \varepsilon > 0, \exists \eta > 0, \forall (x, y) \in A^2, \|x - y\| \leq \eta \implies \|f(x) - f(y)\| \leq \varepsilon.$$

Théorème 1.237

Soient E, F deux espaces vectoriels normés, A une partie de E et $f : A \longrightarrow F$.

Si f est continue sur A et A est compacte, alors f est uniformément continue sur A .

Démonstration 1.238

Par l'absurde, on suppose

$$\exists \varepsilon > 0, \forall \eta > 0, \exists (x, y) \in A^2, \begin{cases} \|x - y\| \leq \eta \\ \|f(x) - f(y)\| > \varepsilon. \end{cases}$$

On spécialise $\eta \leftarrow \frac{1}{n+1}$ pour $n \in \mathbb{N}$.

Pour $n \in \mathbb{N}$, il existe $(x_n, y_n) \in A^2$ tel que $\|x - y\| \leq \frac{1}{n+1}$ et $\|f(x) - f(y)\| > \varepsilon$.

On a ainsi construit deux suites $(x_n), (y_n) \in A^{\mathbb{N}}$ telles que

$$\forall n \in \mathbb{N}, \|x_n - y_n\| \leq \frac{1}{n+1} \quad \text{et} \quad \|f(x_n) - f(y_n)\| > \varepsilon.$$

A étant compacte, (x_n) possède une valeur d'adhérence dans A donc il existe une extractrice φ et $\ell \in A$ tels que $x_{\varphi(n)} \xrightarrow[n \rightarrow +\infty]{} \ell$.

On a

$$\begin{aligned} \forall n \in \mathbb{N}, \|y_{\varphi(n)} - \ell\| &\leq \|y_{\varphi(n)} - x_{\varphi(n)}\| + \|x_{\varphi(n)} - \ell\| \\ &\leq \frac{1}{\varphi(n) + 1} + \|x_{\varphi(n)} - \ell\|. \end{aligned}$$

Or $\varphi(n) \xrightarrow[n \rightarrow +\infty]{} +\infty$ et $\|x_{\varphi(n)} - \ell\| \xrightarrow[n \rightarrow +\infty]{} 0$ donc par encadrement, on a $y_{\varphi(n)} \xrightarrow[n \rightarrow +\infty]{} \ell$.

Par continuité de f en ℓ , on a
$$\begin{cases} f(x_{\varphi(n)}) \xrightarrow{n \rightarrow +\infty} f(\ell) \\ f(y_{\varphi(n)}) \xrightarrow{n \rightarrow +\infty} f(\ell) \end{cases}$$

Donc $f(x_{\varphi(n)}) - f(y_{\varphi(n)}) \xrightarrow{n \rightarrow +\infty} 0$, ce qui contredit l'inégalité

$$\forall n \in \mathbb{N}, \|f(x_{\varphi(n)}) - f(y_{\varphi(n)})\| > \varepsilon.$$

■

1.8 Connexité par arcs

Dans cette section, E est un espace vectoriel normé.

1.8.1 Chemin

Définition 1.239

Soient A une partie de E et $a, b \in A$.

On appelle chemin (ou arc) dans A de a à b toute application continue $\varphi : [0 ; 1] \longrightarrow A$ telle que $\varphi(0) = a$ et $\varphi(1) = b$. Le support du chemin est l'image de φ .

On peut définir une relation d'équivalence sur une partie de E en mettant en relation les points joignables par un chemin.

Définition 1.240

Soient A une partie de E et $a, b \in A$.

On pose $a\mathcal{R}b$ quand il existe un chemin dans A de a à b .

Proposition 1.241

Avec les notations précédentes, la relation \mathcal{R} est une relation d'équivalence sur A .

Démonstration 1.242

► Soit $a \in A$.

La fonction
$$\begin{array}{ccc} \varphi : [0 ; 1] & \longrightarrow & A \\ t & \longmapsto & a \end{array}$$
 est continue sur $[0 ; 1]$ et on a $\varphi(0) = a$ et $\varphi(1) = a$.

Donc $a\mathcal{R}a$: \mathcal{R} est réflexive.

► Soit $(a, b) \in A^2$ tel que $a\mathcal{R}b$.

Il existe une fonction continue $\varphi : [0 ; 1] \longrightarrow A$ telle que $\varphi(0) = a$ et $\varphi(1) = b$.

On pose $\psi : [0 ; 1] \longrightarrow A$
 $t \longmapsto \varphi(1 - t)$

ψ est une fonction continue sur $[0 ; 1]$ telle que $\psi(0) = b$ et $\psi(1) = a$.

Donc $b\mathcal{R}a : \mathcal{R}$ est symétrique.

► Soit $(a, b, c) \in A^3$ tel que $a\mathcal{R}b$ et $b\mathcal{R}c$.

Il existe $(\varphi, \psi) \in \mathcal{C}^0([0 ; 1], A)$ tel que $\varphi(0) = a$, $\varphi(1) = b$, $\psi(0) = b$ et $\psi(1) = c$.

On pose $\theta : [0 ; 1] \longrightarrow A$

$$x \longmapsto \begin{cases} \varphi(2x) & \text{si } 0 \leq x \leq \frac{1}{2} \\ \psi(2x - 1) & \text{sinon} \end{cases}$$

θ est une fonction continue sur $[0 ; 1]$ telle que $\theta(0) = a$ et $\theta(1) = c$.

Donc $a\mathcal{R}c : \mathcal{R}$ est transitive.

► Finalement, \mathcal{R} est une relation d'équivalence. ■

1.8.2 Parties connexes par arcs

Définition 1.243

Soit A une partie de E .

On dit que A est connexe par arcs quand tout couple de points $(a, b) \in A^2$ est joignable par un chemin.

Exemple 1.244

- Les parties convexes de E sont connexes par arcs.
- Les parties étoilées de E sont connexes par arcs.
- \mathbb{C}^* et $\mathbb{C} \setminus D$ où D est la demi-droite des réels négatifs sont connexes par arcs.

Démonstration 1.245

- Une partie convexe est une partie dont tous les points sont reliables en ligne droite donc, en particulier, est une partie connexe par arcs.

- Une partie A est dite étoilée quand il existe $c \in A$ tel que pour tout $b \in A$, $[cb] \subseteq A$. Alors A est clairement connexe par arcs. ■

Les classes d'équivalences de la relation notée \mathcal{R} précédemment s'appellent les composantes connexes par arcs de A : ce sont par définition des parties connexes par arcs.

Proposition 1.246

Les seules parties connexes par arcs de \mathbb{R} sont les intervalles.

Remarque 1.247

Il existe une notion plus générale, celle de partie connexe : une partie A de E est dite connexe quand les seules parties de A à la fois ouvertes et fermées sont \emptyset et A . Elle est plus délicate à aborder et est hors-programme, c'est pourquoi on s'en tient à la notion de connexité par arcs (toute partie connexe par arcs est connexe).

1.8.3 Théorème des valeurs intermédiaires

Là encore, la notion de connexité par arcs permet de généraliser des résultats de première année.

Théorème 1.248

Soient E, F deux espaces vectoriels normés, A une partie de E et $f : A \longrightarrow F$.

Si f est continue par A et A est connexe par arcs, alors $f(A)$ est connexe par arcs.

Démonstration 1.249

Supposons que A est connexe par arcs et que f est continue.

Soit $(x, y) \in f(A)^2$.

Il existe $(a, b) \in A^2$ tel que $f(a) = x$ et $f(b) = y$.

Or A est connexe par arcs donc il existe $\varphi : [0 ; 1] \longrightarrow A$ continue telle que $\varphi(0) = a$ et $\varphi(1) = b$.

$f \circ \varphi$ est donc un chemin qui relie x et y (par composition de fonctions continues).

Donc $f(A)$ est connexe par arcs. ■

On résume en disant que l'image continue d'un connexe par arcs est un connexe par arcs.

Dans le cas des fonctions numériques (*i.e.* à valeurs dans \mathbb{R}), on peut même être plus précis.

Théorème 1.250

Toute fonction continue sur un connexe par arcs et à valeurs réelles vérifie la propriété des valeurs intermédiaires.

Autrement dit, si $f : A \longrightarrow F$ est continue sur A une partie connexe par arcs de E , alors $f(A)$ est un intervalle.

Ou encore :

$$\forall (y, z) \in f(A)^2, \forall w \in [yz], \exists t \in A, f(t) = w.$$

Démonstration 1.251

Évident à partir du Théorème 1.248. ■

Chapitre 2

Séries numériques et vectorielles : révisions et compléments

Sommaire

2.1	Rappels	91
2.1.1	Définitions et notations	91
2.1.2	Convergence d'une série	92
2.1.3	Lien entre convergence de suites et convergence de séries	93
2.2	Séries réelles à termes positifs.	94
2.2.1	Théorème de Cesàro	98
2.2.2	Théorème de comparaison par domination de séries à termes positifs	99
2.2.3	Théorème de comparaison par équivalence de séries à termes positifs	101
2.2.4	Théorème de comparaison série - intégrale	103
2.3	Séries absolument convergentes	107
2.3.1	Lien entre absolue convergence et convergence	107
2.3.2	Un exemple fondamental : l'exponentielle de matrice	108
2.3.3	Extension des résultats par comparaison	108
2.3.4	Produit de Cauchy de deux séries absolument convergentes	110
2.4	Séries alternées	112

Dans ce chapitre, E désigne un espace vectoriel normé (qui peut être \mathbb{R} ou \mathbb{C}) et $\|\cdot\|$ la norme associée (qui est dans ces cas la valeur absolue ou le module).

2.1 Rappels

2.1.1 Définitions et notations

Définition 2.1 (Série vectorielle)

Soit u une suite de E .

On associe à cette suite la suite s définie de la façon suivante : pour tout $n \in \mathbb{N}$, $s_n = \sum_{k=0}^n u_k$.

La suite s est appelée série de terme général u_n et notée $\sum_{n \geq 0} u_n$ ou $\sum u$.

Chaque nombre s_n est appelé somme partielle d'indice n de la série.

L'adjectif « numérique » associé au mot « série » signifie que les termes généraux de la série sont en fait des nombres réels ou complexes.

2.1.2 Convergence d'une série

Définition 2.2

Soit u une suite de E .

On dit que la série $\sum u$ converge ssi la suite des sommes partielles $(s_n) = \left(\sum_{k=0}^n u_k \right)$ converge.

Dans ce cas, si $\ell = \lim_{n \rightarrow +\infty} s_n$, alors ℓ est appelée somme de la série $\sum u$ et on note $\ell = \sum_{n=0}^{+\infty} u_n$.

On appelle aussi reste partiel d'indice n de la série le nombre $r_n = \sum_{k=n+1}^{+\infty} u_k$, de sorte que $r_n + s_n = \ell$.

La suite des restes partiels converge donc vers 0.

Dans le cas contraire, on dit que la série $\sum u$ diverge.

Exemple 2.3

► Soit $x \in \mathbb{C}$. La série $\sum_{n \geq 0} x^n$ converge ssi $|x| < 1$ et, dans ce cas, $\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$. Cette série est appelée série géométrique de raison x .

► Les séries de Riemann : $\sum_{n \geq 1} \frac{1}{n^\alpha}$ converge ssi $\alpha > 1$.

► Pour tout $z \in \mathbb{C}$, la série $\sum \frac{z^n}{n!}$ converge et $\sum_{n=0}^{+\infty} \frac{z^n}{n!} = e^z$.

On peut bien sûr généraliser aux séries quelques théorèmes d'opérations.

Proposition 2.4

Soient u, v deux suites de E et λ un scalaire.

Si les séries $\sum u$ et $\sum v$ convergent, alors la série $\sum (u + \lambda v)$ converge et $\sum_{n=0}^{+\infty} (u_n + \lambda v_n) = \sum_{n=0}^{+\infty} u_n + \lambda \sum_{n=0}^{+\infty} v_n$.

Ceci prouve aussi que l'ensemble des séries convergentes est un espace vectoriel.

Remarque 2.5

La somme d'une série divergente et d'une série convergente est une série divergente.

En revanche, il n'y a rien à dire a priori à propos de la somme de deux séries divergentes.

2.1.3 Lien entre convergence de suites et convergence de séries

Proposition 2.6

Soit u une suite de E .

Si la série $\sum u$ converge, alors la suite u converge vers 0.

Remarque 2.7

- La réciproque est fausse.
- Par contraposition, si une suite u ne tend pas vers 0, alors la série associée diverge : on dit que la série $\sum u$ diverge grossièrement.

Exemple 2.8

On appelle série harmonique la série $\sum_{n \geq 1} \frac{1}{n}$.

Cette série diverge, pourtant son terme général tend vers 0.

Définition 2.9

Soit u une suite de E . On pose $v_n = u_{n+1} - u_n$.

La série $\sum v$ est appelée la série télescopique (ou série domino, ou série différence) associée à u .

Proposition 2.10

Une suite converge ssi sa série télescopique associée converge.

Exercice 2.11

On pose $u_n = \sum_{k=1}^n \frac{1}{k} - \ln n$. Montrez que la suite u converge.

Correction 2.12

u converge ssi $\sum (u_{n+1} - u_n)$ converge.

On a

$$\begin{aligned}u_{n+1} - u_n &= \frac{1}{n+1} - \ln(n+1) + \ln n \\&= \frac{1}{n+1} + \ln \frac{n}{n+1} \\&= \frac{1}{n+1} + \ln \left(1 - \frac{1}{n+1}\right).\end{aligned}$$

Or $\ln(1+u) \underset{u \rightarrow 0}{=} u - \frac{u^2}{2} + o(u^2)$ et $\frac{1}{n+1} \xrightarrow{n \rightarrow +\infty} 0$ donc

$$\begin{aligned}u_{n+1} - u_n &\underset{n \rightarrow +\infty}{=} \frac{1}{n+1} + \left(\frac{-1}{n+1} + \frac{1}{2(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right) \right) \\&\underset{n \rightarrow +\infty}{=} \frac{1}{2(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right) \\&\underset{n \rightarrow +\infty}{\sim} \frac{1}{2n^2}.\end{aligned}$$

Les séries $\sum (u_{n+1} - u_n)$ et $\sum \frac{1}{2n^2}$ sont à termes positifs à partir d'un certain rang et $\sum \frac{1}{n^2}$ converge donc d'après le théorème de comparaison des séries à termes positifs, $\sum (u_{n+1} - u_n)$ converge *i.e.* u converge.

Note culturelle : on appelle constante d'Euler-Mascheroni la limite de cette suite, notée γ et on a le développement asymptotique

$$\sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{=} \ln n + \gamma + o(1).$$

On remarque que l'on a

$$\sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln n.$$

2.2 Séries réelles à termes positifs

Dans cette section, on s'intéresse uniquement aux séries dont le terme général est un réel positif.

On appelle un premier théorème issu du cours de première année.

Théorème 2.13

Soient u et v deux suites réelles positives.

► Si $0 \leq u \leq v$ et si la série $\sum v$ converge, alors la série $\sum u$ converge.

- Si $0 \leq u \leq v$ et si la série $\sum u$ diverge, alors la série $\sum v$ diverge.
- Si $u \sim v$, alors les séries $\sum u$ et $\sum v$ sont de même nature.

Une application classique : la règle de d'Alembert.

Proposition 2.14

Soit u une suite réelle strictement positive telle que $\frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} \ell$. Alors

- si $\ell < 1$, la série $\sum u$ converge ;
- si $\ell > 1$, la série $\sum u$ diverge ;
- si $\ell = 1$, on ne peut rien conclure.

Démonstration 2.15

- Si $\ell < 1$, on pose $K = \frac{1+\ell}{2}$.

Comme $\ell < K < 1$ et $\frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} \ell$, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $0 < \frac{u_{n+1}}{u_n} \leq K$.

Par récurrence, on montre $\forall n \geq N$, $u_n \leq u_N K^{n-N} = \frac{u_N}{K^N} K^n$.

Comme $0 < K < 1$, la série $\sum K^n$ converge donc d'après le théorème de comparaison des séries à termes positifs, $\sum u_n$ converge.

- Si $\ell > 1$, à partir d'un certain rang, on a $u_n \geq \left(\frac{1+\ell}{2}\right)^n \times \text{cte} \xrightarrow{n \rightarrow +\infty} \pm\infty$.

Donc $\sum u_n$ diverge grossièrement.

- Si $\ell = 1$, on ne peut rien dire. Deux exemples :

— En posant $u_n = \frac{1}{n}$, on a $\frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} 1$ mais $\sum u_n$ diverge.

— En posant $u_n = \frac{1}{n^2}$, on a $\frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} 1$ mais $\sum u_n$ converge. ■

Exercice 2.16

Soient $x, y > 0$. Représentez graphiquement l'ensemble des couples (x, y) tels que la série $\sum \frac{x^n}{y^n + n^x}$ converge.

Correction 2.17

Pour tout $n \in \mathbb{N}$, on pose $u_n = \frac{x^n}{y^n + n^x} \geq 0$.

► Si $x = 1$, alors $u_n = \frac{1}{y^n + n} \xrightarrow{n \rightarrow +\infty} 0$.

— Si $y > 1$:

On a $n \underset{n \rightarrow +\infty}{=} o(y^n)$ donc $y^n + n \underset{n \rightarrow +\infty}{\sim} y^n$.

Donc $u_n \underset{n \rightarrow +\infty}{\sim} \frac{1}{y^n} = \left(\frac{1}{y}\right)^n$.

Or $0 \leq \frac{1}{y} < 1$ donc la série géométrique $\sum \left(\frac{1}{y}\right)^n$ converge.

Ainsi, d'après le théorème de comparaison des séries à termes positifs, $\sum u_n$ converge.

— Sinon, on a $y^n \underset{n \rightarrow +\infty}{=} o(n)$ donc $u_n \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$.

Or $\sum \frac{1}{n}$ diverge donc d'après le théorème de comparaison des séries à termes positifs, $\sum u_n$ diverge.

► Si $y < 1$, on a $n \underset{n \rightarrow +\infty}{=} o(y^n)$ donc $y^n + n \underset{n \rightarrow +\infty}{\sim} y^n$.

Donc $u_n \underset{n \rightarrow +\infty}{\sim} \frac{x^n}{y^n} = \left(\frac{x}{y}\right)^n$.

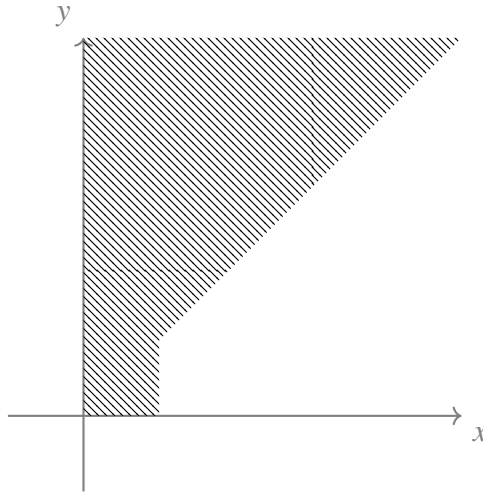
Ainsi, d'après le théorème de comparaison des séries à termes positifs, $\sum u_n$ converge ssi $\frac{x}{y} < 1$ i.e. $x < y$.

► Si $y \leq 1$, on a $y^n \underset{n \rightarrow +\infty}{=} o(n^x)$ donc $u_n \underset{n \rightarrow +\infty}{\sim} \frac{x^n}{n^x}$.

— Si $x > 1$, on a $n^x \underset{n \rightarrow +\infty}{=} o(x^n)$ donc $\lim_{n \rightarrow +\infty} u_n = +\infty$: $\sum u_n$ diverge grossièrement.

— Si $x < 1$, on a $0 \leq \frac{x^n}{n^x} \leq x^n$ donc, comme $0 < x < 1$, $\sum x^n$ converge et par théorème de comparaison des séries à termes positifs, $\sum \frac{x^n}{n^x}$ converge, donc $\sum u_n$ converge.

Finalement, on a la représentation graphique suivante :



Exercice 2.18

Montrez que la suite (u_n) définie par $u_0 \in [0 ; 1]$ et $u_{n+1} = \frac{1}{2} (u_n + u_n^2)$ converge vers 0 et donnez la nature de la série $\sum u_n$.

Correction 2.19

Localisation : on montre par récurrence triviale que $\forall n \in \mathbb{N}, u_n \in [0 ; 1]$.

Monotonie : pour tout $n \in \mathbb{N}$, on a

$$u_{n+1} - u_n = \frac{u_n + u_n^2}{2} - u_n = \frac{u_n^2 - u_n}{2} \leq \frac{u_n(u_n - 1)}{2} \leq 0$$

donc (u_n) est décroissante.

Or (u_n) est minorée par 0 donc (u_n) converge d'après le théorème de la limite monotone.

On pose $\ell = \lim_n u_n \in [0 ; 1]$.

Par opérations sur les limites, on a

$$u_{n+1} = \frac{u_n + u_n^2}{2} \xrightarrow{n \rightarrow +\infty} \frac{\ell + \ell^2}{2}.$$

Or (u_{n+1}) est extraite de (u_n) donc $u_{n+1} \xrightarrow{n \rightarrow +\infty} \ell$.

Par unicité de la limite, on a donc $\ell = \frac{\ell + \ell^2}{2}$.

Donc $\ell(\ell - 1) = 0$, or $\ell \in [0 ; 1]$ donc $\ell = 0$.

Enfin, on a $u_n \xrightarrow{n \rightarrow +\infty} 0$ donc il existe $N \in \mathbb{N}$ tel que $\forall n \geq N, u_n^2 \leq \frac{u_n}{2}$.

Donc pour $n \geq N$, on a $u_{n+1} \leq \frac{3}{4}u_n$.

Par récurrence, on montre que pour tout $n \geq N, u_n \leq \left(\frac{3}{4}\right)^{-N} u_N \times \left(\frac{3}{4}\right)^n$.

Or la série géométrique $\sum \left(\frac{3}{4}\right)^n$ converge donc d'après le théorème de comparaison des séries à termes positifs, $\sum u_n$ converge.

On donne quelques versions plus élaborées du théorème de comparaison.

2.2.1 Théorème de Cesàro

Théorème 2.20

Soit u une suite numérique qui converge vers ℓ . Alors $\frac{u_0 + \cdots + u_n}{n} \xrightarrow{n \rightarrow +\infty} \ell$.

Démonstration 2.21

Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ tel que $\forall n \geq N, |u_n - \ell| \leq \varepsilon$.

On a

$$\begin{aligned} \frac{u_0 + \cdots + u_n}{n} - \ell &= \frac{u_0 + \cdots + u_N}{n} + \frac{u_{N+1} + \cdots + u_n}{n} - \ell \\ &= \frac{u_0 + \cdots + u_N}{n} + \frac{u_{N+1} - \ell + \cdots + u_n - \ell}{n} + \frac{(n - N)\ell}{n} - \ell \\ &= \frac{u_0 + \cdots + u_N}{n} - \frac{N\ell}{n} + \frac{u_{N+1} - \ell + \cdots + u_n - \ell}{n}. \end{aligned}$$

Donc

$$\begin{aligned} \left| \frac{u_0 + \cdots + u_n}{n} - \ell \right| &\leq \overbrace{\left| \frac{u_0 + \cdots + u_N - N\ell}{n} \right|}^K + \frac{|u_{N+1} - \ell| + \cdots + |u_n - \ell|}{n} \\ &\leq \frac{K}{n} + \frac{(n - N)\varepsilon}{n} \\ &\leq \frac{K}{n} + \varepsilon. \end{aligned} \quad \left. \vphantom{\frac{K}{n}} \right) \frac{n - N}{n} \leq 1$$

Or $\frac{K}{n} \xrightarrow{n \rightarrow +\infty} 0$ donc il existe $N' \in \mathbb{N}$ tel que $\forall n \geq N', \frac{K}{n} \leq \varepsilon$.

On pose $N'' = \max(N, N')$.

Donc pour tout $n \geq N''$, $\left| \frac{u_0 + \cdots + u_n}{n} - \ell \right| \leq 2\varepsilon$.

On a montré

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \left| \frac{u_0 + \cdots + u_n}{n} - \ell \right| \leq \varepsilon$$

donc $\frac{u_0 + \cdots + u_n}{n} \xrightarrow{n \rightarrow +\infty} \ell$. ■

Dans le cas où $\ell \neq 0$, la série $\sum u$ diverge grossièrement et $\sum_{k=0}^n u_k \underset{n \rightarrow +\infty}{\sim} n\ell$.

Dans le cas où $\ell = 0$, on peut juste dire $\sum_{k=0}^n u_k = o(n)$.

Exercice 2.22

Soit u la suite définie par récurrence par $u_0 > 0$ et pour tout $n \in \mathbb{N}$, $u_{n+1} = u_n + \frac{u_n + 1}{u_n + 2}$.

Étudiez la convergence ou divergence de la suite u , puis donnez un équivalent simple de u_n quand n tend vers $+\infty$.

Correction 2.23

Par récurrence immédiate, (u_n) est à termes strictement positifs.

On en déduit que pour tout $n \in \mathbb{N}$, $u_{n+1} - u_n = \frac{u_n + 1}{u_n + 2} > 0$ donc (u_n) est croissante.

Par l'absurde, on suppose que (u_n) converge vers $\ell \in \mathbb{R}_+^*$.

De la même façon que dans l'Exercice 2.18, par passage à la limite, on obtient $\ell - \ell = \frac{\ell + 1}{\ell + 2}$ donc $\ell = -1$: contradiction.

Donc (u_n) est croissante et diverge donc elle diverge vers $+\infty$.

De plus, on a $\sum_{k=0}^{n-1} (u_{k+1} - u_k) = u_n - u_0 = \sum_{k=0}^{n-1} \frac{u_k + 1}{u_k + 2}$.

Or $\frac{u_n + 1}{u_n + 2} \xrightarrow{n \rightarrow +\infty} 1$ donc d'après le théorème de Cesàro, on a

$$\frac{\sum_{k=0}^{n-1} \frac{u_k + 1}{u_k + 2}}{n} \xrightarrow{n \rightarrow +\infty} 1$$

i.e. $u_n - u_0 \underset{n \rightarrow +\infty}{\sim} n$.

Donc $u_n \underset{n \rightarrow +\infty}{\sim} n$.

2.2.2 Théorème de comparaison par domination de séries à termes positifs

Dans le cas convergent d'abord, les restes partiels suivent la même relation de comparaison.

Théorème 2.24

Soient u, v deux suites réelles positives.

Si $u = \mathcal{O}(v)$ et la série $\sum v$ converge, alors la série $\sum u$ converge. De plus, $\sum_{k=n+1}^{+\infty} u_k = \mathcal{O}\left(\sum_{k=n+1}^{+\infty} v_k\right)$.

Si $u = o(v)$ et la série $\sum v$ converge, alors la série $\sum u$ converge. De plus, $\sum_{k=n+1}^{+\infty} u_k = o\left(\sum_{k=n+1}^{+\infty} v_k\right)$.

Démonstration 2.25

► Si $u_n \underset{n \rightarrow +\infty}{=} \mathcal{O}(v_n)$, il existe $K > 0$ et $N \in \mathbb{N}$ tels que $\forall n \geq N, 0 \leq u_n \leq K v_n$.

La série $\sum v_n$ converge donc d'après le théorème de comparaison des séries à termes positifs, $\sum u_n$ converge.

Pour $n \geq N$, on a $\sum_{k=n+1}^{+\infty} u_k \leq K \sum_{k=n+1}^{+\infty} v_k$.

D'où $\sum_{k=n+1}^{+\infty} u_k \underset{n \rightarrow +\infty}{=} \mathcal{O}\left(\sum_{k=n+1}^{+\infty} v_k\right)$.

► Si $u_n \underset{n \rightarrow +\infty}{=} o(v_n)$, pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que $\forall n \geq N, 0 \leq u_n \leq \varepsilon v_n$.

La série $\sum v_n$ converge donc d'après le théorème de comparaison des séries à termes positifs, $\sum u_n$ converge.

Pour $n \geq N$, on a $\sum_{k=n+1}^{+\infty} u_k \leq \varepsilon \sum_{k=n+1}^{+\infty} v_k$.

D'où $\sum_{k=n+1}^{+\infty} u_k \underset{n \rightarrow +\infty}{=} o\left(\sum_{k=n+1}^{+\infty} v_k\right)$. ■

Dans le cas divergent ensuite, les sommes partielles suivent aussi la même relation de comparaison.

Théorème 2.26

Soient u, v deux suites réelles positives.

Si $u = \mathcal{O}(v)$ et la série $\sum u$ diverge, alors la série $\sum v$ diverge. De plus, $\sum_{k=0}^n u_k = \mathcal{O}\left(\sum_{k=0}^n v_k\right)$.

Si $u = o(v)$ et la série $\sum u$ diverge, alors la série $\sum v$ diverge. De plus, $\sum_{k=0}^n u_k = o\left(\sum_{k=0}^n v_k\right)$.

Démonstration 2.27

► Si $u_n \underset{n \rightarrow +\infty}{=} \mathcal{O}(v_n)$, il existe $K > 0$ et $N \in \mathbb{N}$ tels que $\forall n \geq N, 0 \leq u_n \leq K v_n$.

Par comparaison de séries à termes positifs, $\sum v_n$ diverge.

u et v étant deux suites positives telles que $\sum u_n$ et $\sum v_n$ divergent, on en déduit que

$$\sum_{k=0}^n u_k \xrightarrow{n \rightarrow +\infty} +\infty \text{ et } \sum_{k=0}^n v_k \xrightarrow{n \rightarrow +\infty} +\infty.$$

Pour $n \geq N$, on a

$$\begin{aligned} \sum_{k=0}^n u_k &= \sum_{k=0}^N u_k + \sum_{k=N+1}^n u_k \\ &\leq \sum_{k=0}^N u_k + K \sum_{k=N+1}^n v_k \\ &\leq \underbrace{\sum_{k=0}^N u_k - K \sum_{k=0}^N v_k}_L + K \sum_{k=0}^n v_k. \end{aligned}$$

Or $\sum_{k=0}^n v_k \xrightarrow{n \rightarrow +\infty} +\infty$ donc il existe $N' \in \mathbb{N}$ tel que $\forall n \geq N', L \leq \sum_{k=0}^n v_k$.

Alors, pour $n \geq \max(N, N')$, on a $\sum_{k=0}^n u_k \leq (K+1) \sum_{k=0}^n v_k$.

On a montré : $\sum_{k=0}^n u_k \underset{n \rightarrow +\infty}{=} \mathcal{O}\left(\sum_{k=0}^n v_k\right)$.

► Idem. ■

2.2.3 Théorème de comparaison par équivalence de séries à termes positifs

Théorème 2.28

Soient u, v deux suites réelles positives.

Si $u \sim v$, alors les séries $\sum u$ et $\sum v$ sont de même nature ; l'une converge ssi l'autre converge.

De plus,

► si les séries convergent, alors les restes partiels sont équivalents : $\sum_{k=n+1}^{+\infty} u_k \underset{n \rightarrow +\infty}{\sim} \sum_{k=n+1}^{+\infty} v_k$;

► si les séries divergent, alors les sommes partielles divergent vers $+\infty$ et sont équivalentes :

$$\sum_{k=0}^n u_k \underset{n \rightarrow +\infty}{\sim} \sum_{k=0}^n v_k.$$

Démonstration 2.29

► Si les séries convergent, comme $u_n \underset{n \rightarrow +\infty}{\sim} v_n$, on a

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, (1 - \varepsilon) v_n \leq u_n \leq (1 + \varepsilon) v_n$$

donc

$$\forall n \geq N, (1 - \varepsilon) \sum_{k=n+1}^{+\infty} v_k \leq \sum_{k=n+1}^{+\infty} u_k \leq (1 + \varepsilon) \sum_{k=n+1}^{+\infty} v_k.$$

$$\text{Donc } \sum_{k=n+1}^{+\infty} u_k \underset{n \rightarrow +\infty}{\sim} \sum_{k=n+1}^{+\infty} v_k.$$

► Si les séries divergent, la démonstration est similaire à la Démonstration 2.27. ■

Exercice 2.30

Soit $a > 0$. On pose $u_n = \sin \frac{a^n}{n}$ pour $n \in \mathbb{N}^*$. Selon la valeur de a , déterminez la nature de la série $\sum_{n \geq 1} u_n$.

Montrez que si $a = 1$, alors $\sum_{k=1}^n u_k \sim \ln n$ et si $a < 1$, $\sum_{k=n}^{+\infty} u_k = o(a^n)$.

Correction 2.31

► Si $a = 1$, on a $u_n = \sin \frac{1}{n} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$.

Or $\sum \frac{1}{n}$ diverge donc par théorème de comparaison des séries à termes positifs, $\sum u_n$ diverge.

► Si $a < 1$, on a $u_n = \sin \frac{a^n}{n} \underset{n \rightarrow +\infty}{\sim} \frac{a^n}{n} \underset{n \rightarrow +\infty}{=} o(a^n)$.

Or $\sum a^n$ converge car $a < 1$ donc par théorème de comparaison des séries à termes positifs, $\sum u_n$ converge.

De plus, d'après le Théorème 2.28, si $a = 1$, on a

$$\sum_{k=1}^n u_k \underset{n \rightarrow +\infty}{\sim} \sum_{k=1}^n \frac{1}{n} \underset{n \rightarrow +\infty}{\sim} \ln n$$

et d'après le Théorème 2.24, si $a < 1$, on a

$$\begin{aligned} \sum_{k=n+1}^{+\infty} u_k &\underset{n \rightarrow +\infty}{=} o\left(\sum_{k=n+1}^{+\infty} a^k\right) \\ &\underset{n \rightarrow +\infty}{=} o\left(\frac{a^{n+1}}{1-a}\right) \\ &\underset{n \rightarrow +\infty}{=} o(a^n). \end{aligned}$$

2.2.4 Théorème de comparaison série - intégrale

Proposition 2.32

Soit f une fonction continue, positive et décroissante sur \mathbb{R}_+ .

Alors la série de terme général $f(n)$ et la suite de terme général $\int_0^n f$ sont de même nature.

Démonstration 2.33

f étant décroissante et continue sur \mathbb{R}_+ , pour $k \in \mathbb{N}^*$, pour tout $t \in [k-1; k]$, $f(k) \leq f(t)$.

$$\text{Donc } \int_{k-1}^k f(k) dt \leq \int_{k-1}^k f(t) dt \text{ i.e. } f(k) \leq \int_{k-1}^k f(t) dt.$$

$$\text{De même, } \int_k^{k+1} f(t) dt \leq f(k).$$

$$\text{Donc pour tout } n \in \mathbb{N}^*, \sum_{k=1}^n \int_k^{k+1} f(t) dt \leq \sum_{k=1}^n f(k) \leq \sum_{k=1}^n \int_{k-1}^k f(t) dt.$$

$$\text{Donc } \int_1^{n+1} f(t) dt \leq \sum_{k=1}^n f(k) \leq \int_0^n f(t) dt.$$

f étant positive, les suites $\left(\sum_{k=1}^n f(k)\right)$ et $\left(\int_0^n f(t) dt\right)$ sont croissantes.

Donc si $\left(\int_0^n f\right)$ converge, elle est majorée, donc $\left(\sum_{k=1}^n f(k)\right)$ est majorée et donc convergente, i.e. $\sum f(n)$ converge.

Réciproquement, si $\sum f(n)$ converge, $\left(\sum_{k=1}^n f(k)\right)$ est majorée donc $\left(\int_1^{n+1} f\right)$ est majorée et croissante donc convergente. Donc $\left(\int_0^n f\right)$ converge. ■

Méthode 2.34 (À retenir)

La technique d'encadrement des sommes partielles d'une série $\sum f(n)$ (ou des restes partiels) par des intégrales quand f est continue, positive et monotone.

Exemple 2.35

$$\triangleright \sum_{k=1}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln n \text{ (à connaître).}$$

► Si $\alpha > 1$, un équivalent simple de $\sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha}$ quand $n \longrightarrow +\infty$ est $\frac{1}{(\alpha-1)n^{\alpha-1}}$.

Rappel 2.36

Si $u_n \underset{n \rightarrow +\infty}{\sim} v_n$ et $u_n \xrightarrow[n \rightarrow +\infty]{} +\infty$ ou 0, alors

$$\ln u_n \underset{n \rightarrow +\infty}{\sim} \ln v_n.$$

Démonstration 2.37 (Premier point)

La fonction $f : t \mapsto \frac{1}{t}$ est positive, décroissante et continue sur $[1 ; +\infty[$.

On a donc l'inégalité :

$$\underbrace{\int_2^{n+1} \frac{1}{t} dt}_{=\ln(n+1)-\ln 2} \leq \sum_{k=2}^n \frac{1}{k} \leq \underbrace{\int_1^n \frac{1}{t} dt}_{=\ln n}.$$

Or on a

$$\ln(n+1) - \ln 2 \underset{n \rightarrow +\infty}{\sim} \ln(n+1) \underset{n \rightarrow +\infty}{\sim} \ln n$$

donc d'après le théorème d'encadrement, on a

$$\sum_{k=2}^n \frac{1}{k} \underset{n \rightarrow +\infty}{\sim} \ln n. \quad \blacksquare$$

Démonstration 2.38 (Deuxième point)

Soit $\alpha > 1$.

La fonction $f : t \mapsto \frac{1}{t^\alpha}$ est continue, décroissante et positive sur $[1 ; +\infty[$.

On a

$$\int_{n+1}^{+\infty} f(t) dt \leq \sum_{k=n+1}^{+\infty} f(k) \leq \underbrace{\int_n^{+\infty} f(t) dt}_{=\lim_{X \rightarrow +\infty} \int_n^X f(t) dt}.$$

Soient $X \geq n \geq 1$.

On a

$$\int_{n+1}^{X+1} \frac{1}{t^\alpha} dt \leq \sum_{k=n+1}^X \frac{1}{k^\alpha} \leq \int_n^X \frac{1}{t^\alpha} dt.$$

Or

$$\int_n^X \frac{1}{t^\alpha} dt = \left[\frac{1}{-\alpha+1} t^{-\alpha+1} \right]_n^X = \frac{1}{1-\alpha} \left(\frac{1}{X^{\alpha-1}} - \frac{1}{n^{\alpha-1}} \right).$$

Donc

$$\frac{1}{\alpha-1} \left(\frac{1}{(n+1)^{\alpha-1}} - \frac{1}{(X+1)^{\alpha-1}} \right) \leq \sum_{k=n+1}^X \frac{1}{k^\alpha} \leq \frac{1}{\alpha-1} \left(\frac{1}{n^{\alpha-1}} - \frac{1}{X^{\alpha-1}} \right).$$

Or $\frac{1}{X^{\alpha-1}} \xrightarrow{X \rightarrow +\infty} 0$ et $\frac{1}{(X+1)^{\alpha-1}} \xrightarrow{X \rightarrow +\infty} 0$.

Donc, par passage à la limite quand $X \rightarrow +\infty$, on a

$$\frac{1}{(\alpha-1)(n+1)^{\alpha-1}} \leq \sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha} \leq \frac{1}{(\alpha-1)n^{\alpha-1}}.$$

Or $\frac{1}{(n+1)^{\alpha-1}} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n^{\alpha-1}}$ donc par encadrement, on a

$$\sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha} \underset{n \rightarrow +\infty}{\sim} \frac{1}{(\alpha-1)n^{\alpha-1}}.$$

■

Exercice 2.39

Pour $n \in \mathbb{N}^*$, on pose $u_n = \sum_{k=n}^{+\infty} \frac{\ln k}{k^2}$.

Justifiez l'existence de u_n , puis montrez la divergence de la série $\sum u_n$.

Montrez que $\sum_{k=1}^n u_k \underset{n \rightarrow +\infty}{\sim} \frac{\ln^2 n}{2}$.

Correction 2.40

Pour $k \geq 1$, on a

$$0 \leq \frac{\ln k}{k^2} = \frac{\ln k}{k^{1/2} \times k^{3/2}} = \frac{\ln k}{k^{1/2}} \times \frac{1}{k^{3/2}}.$$

Or $\frac{\ln k}{k^{1/2}} \xrightarrow{k \rightarrow +\infty} 0$ donc $\frac{\ln k}{k^{1/2}} \underset{k \rightarrow +\infty}{=} o\left(\frac{1}{k^{3/2}}\right)$.

Or $\sum \frac{1}{k^{3/2}}$ converge donc d'après le théorème de comparaison des séries à termes positifs, $\sum \frac{\ln k}{k^2}$ converge.

Donc pour tout $n \geq 1$, $\sum_{k=n}^{+\infty} \frac{\ln k}{k^2}$ existe.

La fonction $f : t \mapsto \frac{\ln t}{t^2}$ est de classe \mathcal{C}^∞ et positive sur $[1; +\infty[$.

On a $f' : t \mapsto \frac{t - 2t \ln t}{t^4}$.

Or, pour $t \geq 2$, on a $t \leq 2t \ln t$ et $t^4 \geq 0$ donc

$$\forall t \in [2; +\infty[, \quad f'(t) \leq 0$$

donc f est décroissante sur $[2; +\infty[$.

Soient $X \geq n \geq 3$.

On a

$$\int_n^{X+1} \frac{\ln t}{t^2} dt \leq \sum_{k=n}^X \frac{\ln k}{k^2} \leq \int_{n-1}^X \frac{\ln t}{t^2} dt.$$

Or, par intégration par parties, on a

$$\int_{n-1}^X \frac{\ln t}{t^2} dt = \frac{\ln(n-1) + 1}{n-1} - \frac{\ln X + 1}{X}.$$

Donc

$$\frac{\ln n + 1}{n} - \frac{\ln(X+1) + 1}{X+1} \leq \sum_{k=n}^X \frac{\ln k}{k^2} \leq \frac{\ln(n-1) + 1}{n-1} - \frac{\ln X + 1}{X}.$$

Donc, par passage à la limite quand $X \rightarrow +\infty$, on a

$$\frac{\ln n + 1}{n} \leq u_n \leq \frac{\ln(n-1) + 1}{n-1}.$$

Or $\ln(n-1) \underset{n \rightarrow +\infty}{\sim} \ln n$ donc

$$\frac{\ln(n-1) + 1}{n-1} \underset{n \rightarrow +\infty}{\sim} \frac{\ln n + 1}{n} \underset{n \rightarrow +\infty}{\sim} \frac{\ln n}{n}.$$

D'où $u_n \underset{n \rightarrow +\infty}{\sim} \frac{\ln n}{n}$.

Pour $n \geq 3$, on a $\frac{\ln n}{n} \geq \frac{1}{n} \geq 0$.

Or $\sum \frac{1}{n}$ diverge donc d'après le théorème de comparaison des séries à termes positifs, $\sum \frac{\ln n}{n}$ diverge et donc $\sum u_n$ diverge.

De plus, d'après le Théorème 2.28, on a

$$\sum_{k=1}^n u_k \underset{n \rightarrow +\infty}{\sim} \sum_{k=1}^n \frac{\ln k}{k}.$$

La fonction $g : t \mapsto \frac{\ln t}{t}$ est de classe \mathcal{C}^∞ , positive et décroissante sur $[3; +\infty[$.

On pose $n \geq 4$. On a

$$\int_4^{n+1} \frac{\ln t}{t} dt \leq \sum_{k=4}^n g(k) \leq \int_3^n \frac{\ln t}{t} dt.$$

Or

$$\int_3^n \frac{\ln t}{t} dt = \int_3^n \frac{1}{t} \ln t dt = \left[\frac{\ln^2 t}{2} \right]_3^n = \frac{\ln^2 n}{2} - \frac{\ln^2 3}{2}.$$

Donc

$$\underbrace{\frac{\ln^2(n+1)}{2} + K_2}_{\underset{n \rightarrow +\infty}{\sim} \frac{\ln^2 n}{2}} \leq \sum_{k=4}^n \frac{\ln k}{k} \leq \underbrace{\frac{\ln^2 n}{2} + K_1}_{\underset{n \rightarrow +\infty}{\sim} \frac{\ln^2 n}{2}}.$$

Par encadrement, on a $\sum_{k=4}^n u_k \underset{n \rightarrow +\infty}{\sim} \frac{\ln^2 n}{2}$ puis $\sum_{k=1}^n u_k \underset{n \rightarrow +\infty}{\sim} \frac{\ln^2 n}{2}$.

2.3 Séries absolument convergentes

Définition 2.41

Soit u une suite de E .

On dit que la série $\sum u$ est absolument convergente ssi la série à termes positifs $\sum \|u\|$ est convergente.

2.3.1 Lien entre absolue convergence et convergence

Théorème 2.42

Si E est de dimension finie, alors toute série absolument convergente est convergente.

Démonstration 2.43

Soient E un espace vectoriel normé de dimension finie $p \in \mathbb{N}$ et \mathcal{B} une base de E .

On choisit la norme sup $\|\cdot\|_\infty$ associée à cette base.

Toutes les normes sont équivalentes sur E donc si $\|\cdot\|$ est une norme telle que $\sum \|u_n\|$ converge, alors $\sum \|u_n\|_\infty$ converge d'après le théorème de comparaison des séries à termes positifs.

On note $(u_{n,1}, \dots, u_{n,p})$ les coordonnées de u_n dans \mathcal{B} .

On a $\forall k \in \llbracket 1 ; p \rrbracket$, $|u_{n,k}| \leq \|u_n\|_\infty$ donc d'après le théorème de comparaison des séries à termes positifs, $\sum_n |u_{n,k}|$ converge, donc $\sum_n u_{n,k}$ est absolument convergente et donc convergente.

Donc les suites des sommes partielles $\left(\sum_{j=0}^n u_j \right) = \left(\left(\sum_{j=0}^n u_{j,1}, \dots, \sum_{j=0}^n u_{j,p} \right) \right)$ convergent i.e. $\sum u_n$ converge. ■

Remarque 2.44

- La réciproque est fautive : la série $\sum_{n \geq 1} \frac{(-1)^n}{n}$ converge (on l'appelle la série harmonique alternée) mais ne converge pas absolument.
- L'hypothèse de la dimension finie est indispensable. En dimension infinie, ce résultat est faux en général.

Exercice 2.45

Soit $x > 0$. Montrez que les séries suivantes convergent :

$$\sum_{n \geq 2} \frac{\ln(n^2 + (-1)^n n)}{n^2 + (-1)^n x^n} \quad \sum_{n \geq 0} \sqrt{n} \cos(x) \sin^n(x) \quad \sum_{n \geq 0} \frac{(-1)^n \sqrt{n+x}}{x^n + n^{2/x}}.$$

2.3.2 Un exemple fondamental : l'exponentielle de matrice

Soit $p \in \mathbb{N}^*$.

On choisit comme norme sur $E = \mathcal{M}_p(\mathbb{C})$ une norme sous-multiplicative.

Alors pour tout $n \in \mathbb{N}^*$, $\|A^n\| \leq \|A\|^n$, donc $\left\| \frac{A^n}{n!} \right\| \leq \frac{\|A\|^n}{n!}$.

Or la série $\sum \frac{\|A\|^n}{n!}$ converge (et sa somme vaut $\exp \|A\|$), donc par comparaison de séries à termes positifs, la série $\sum \frac{A^n}{n!}$ est absolument convergente.

On pose alors $\exp A = \sum_{n=0}^{+\infty} \frac{A^n}{n!}$.

2.3.3 Extension des résultats par comparaison

Définition 2.46

Soit u une suite de E et v une suite réelle positive.

On dit que $u = \mathcal{O}(v)$ quand $\exists M > 0$, $\exists n_0 \in \mathbb{N}$, $\forall n \geq n_0$, $\|u_n\| \leq M v_n$.

On dit que $u = o(v)$ quand $\forall \varepsilon > 0$, $\exists n_0 \in \mathbb{N}$, $\forall n \geq n_0$, $\|u_n\| \leq \varepsilon v_n$.

Proposition 2.47

Soient u une suite de E et v une suite réelle positive.

Si E est de dimension finie, $u_n = \mathcal{O}(v_n)$ quand n tend vers $+\infty$ et la série $\sum v$ converge, alors la série $\sum u$ est absolument convergente.

De plus,
$$\sum_{k=n+1}^{+\infty} u_k = \mathcal{O}\left(\sum_{k=n+1}^{+\infty} v_k\right).$$

Démonstration 2.48

On a $u_n \underset{n \rightarrow +\infty}{=} \mathcal{O}(v_n) \iff \|u_n\| \underset{n \rightarrow +\infty}{=} \mathcal{O}(v_n).$

Si $\sum v_n$ converge alors $\sum \|u_n\|$ converge d'après le théorème de comparaison des séries à termes positifs donc, comme E est de dimension finie, $\sum u_n$ converge.

De plus, on a $\left\| \sum_{k=n+1}^{+\infty} u_k \right\| \leq \sum_{k=n+1}^{+\infty} \|u_k\|$ donc d'après le Théorème 2.24, on a

$$\sum_{k=n+1}^{+\infty} \|u_k\| \underset{n \rightarrow +\infty}{=} \mathcal{O}\left(\sum_{k=n+1}^{+\infty} v_k\right).$$

Donc

$$\left\| \sum_{k=n+1}^{+\infty} u_k \right\| \underset{n \rightarrow +\infty}{=} \mathcal{O}\left(\sum_{k=n+1}^{+\infty} v_k\right).$$

■

Ceci est encore valable si $u_n = o(v_n)$.

Proposition 2.49

Soient u une suite de E et v une suite réelle positive.

Si E est de dimension finie, $u_n = o(v_n)$ quand n tend vers $+\infty$ et la série $\sum v$ converge, alors la série $\sum u$ est absolument convergente.

De plus,
$$\sum_{k=n+1}^{+\infty} u_k = o\left(\sum_{k=n+1}^{+\infty} v_k\right).$$

Démonstration 2.50

Idem.

■

2.3.4 Produit de Cauchy de deux séries absolument convergentes

Définition 2.51

Une \mathbb{K} -algèbre est un \mathbb{K} -espace vectoriel muni d'un produit interne bilinéaire.

Définition 2.52

Soient E une algèbre normée de dimension finie, $\sum_{n \geq 0} a_n$ et $\sum_{n \geq 0} b_n$ deux séries à termes dans E .

On appelle produit de Cauchy des deux séries la série $\sum_{n \geq 0} c_n$ où pour tout $n \in \mathbb{N}$, $c_n = \sum_{k=0}^n a_k b_{n-k}$.

Remarque 2.53

Quand les séries ne commencent pas à partir du rang 0, il faut se méfier ! Une idée simple est de se ramener au cas précédent en décalant les indices.

Exemple très courant : les séries commencent au rang 1. Dans ce cas, le produit de Cauchy des séries

$\sum_{n \geq 1} a_n$ et $\sum_{n \geq 1} b_n$ est la série $\sum_{n \geq 1} c_n$ où pour tout $n \in \mathbb{N}^*$, $c_n = \sum_{k=1}^n a_k b_{n+1-k}$.

Théorème 2.54

Avec les mêmes hypothèses sur E .

Si les séries $\sum_{n \geq 0} a_n$ et $\sum_{n \geq 0} b_n$ convergent absolument, alors leur produit de Cauchy est aussi absolument convergent et

$$\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} a_n \times \sum_{k=0}^{+\infty} b_k.$$

Démonstration 2.55 (Premier cas : $E = \mathbb{R}$ et $(a_n), (b_n)$ positives)

On pose $A_n = \sum_{k=0}^n a_k$, $B_n = \sum_{k=0}^n b_k$ et $C_n = \sum_{k=0}^n c_k$.

On a $A_n B_n = \sum_{k=0}^n a_k \times \sum_{k=0}^n b_k = \sum_{\substack{0 \leq k \leq n \\ 0 \leq \ell \leq n}} a_k b_\ell$.

On a $(\star) \ C_n \leq A_n B_n \leq C_{2n}$.

Les séries $\sum a_n$ et $\sum b_n$ convergent donc les suites (A_n) et (B_n) convergent et sont donc majorées.

Comme ce sont des suites positives, la suite $(A_n B_n)$ est majorée (produit d'inégalités entre positifs).

Donc (C_n) , qui est croissante, est majorée donc converge *i.e.* $\sum c_n$ converge.

$$\text{Donc } A_n \xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{+\infty} a_k, B_n \xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{+\infty} b_k, C_n \xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{+\infty} c_k \text{ et } C_{2n} \xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{+\infty} c_k.$$

Donc, par passage à la limite dans (\star) , on a

$$\sum_{k=0}^{+\infty} c_k = \sum_{k=0}^{+\infty} a_k \times \sum_{k=0}^{+\infty} b_k. \quad \blacksquare$$

Démonstration 2.56 (Cas général)

On a $\dim E \in \mathbb{N}$ donc toutes les normes sont équivalentes sur E ; on en choisit donc une qui est sous-multiplicative.

Avec les mêmes notations, on a $A_n B_n - C_n = \sum_{k=0}^n \sum_{j=n-k+1}^n c_j$, d'où

$$\begin{aligned} \|A_n B_n - C_n\| &\leq \sum_{k=0}^n \sum_{j=n-k+1}^n \|c_j\| \\ &= \sum_{k=0}^n \sum_{j=n-k+1}^n \left\| \sum_{\ell=0}^j a_{j-\ell} b_\ell \right\| \\ &\leq \sum_{k=0}^n \sum_{j=n-k+1}^n \sum_{\ell=0}^j \|a_{j-\ell} b_\ell\| \\ &\leq \sum_{k=0}^n \sum_{j=n-k+1}^n \sum_{\ell=0}^j \|a_{j-\ell}\| \times \|b_\ell\|. \end{aligned}$$

On pose $A'_n = \sum_{k=0}^n \|a_k\|$, $B'_n = \sum_{k=0}^n \|b_k\|$ et $C'_n = \sum_{k=0}^n c'_k = \sum_{k=0}^n \sum_{j=0}^k \|a_{k-j}\| \times \|b_j\|$.

On a de même $\|A_n B_n - C_n\| \leq A'_n B'_n - C'_n$.

D'après le premier cas, (C'_n) converge vers $\sum_{k=0}^{+\infty} \|a_k\| \times \sum_{k=0}^{+\infty} \|b_k\| = \lim_n A'_n B'_n$.

Donc $A'_n B'_n - C'_n \xrightarrow{n \rightarrow +\infty} 0$ donc par encadrement $A_n B_n - C_n \xrightarrow{n \rightarrow +\infty} 0$.

Or $A_n \xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{+\infty} a_k$ et $B_n \xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{+\infty} b_k$ donc

$$C_n \xrightarrow{n \rightarrow +\infty} \sum_{k=0}^{+\infty} a_k \times \sum_{k=0}^{+\infty} b_k. \quad \blacksquare$$

2.4 Séries alternées

Définition 2.57

Une série alternée est une série réelle $\sum u_n$ telle que pour tout $n \in \mathbb{N}$, u_{n+1} est de signe opposé à u_n .

En général, les séries alternées sont reconnaissables à la présence d'un facteur $(-1)^n$ dans l'expression du terme général.

On dispose d'une condition suffisante de convergence d'une série alternée qu'on appelle le critère spécial des séries alternées.

Théorème 2.58

Soit $\sum (-1)^n u_n$ une série alternée.

Si la suite u

- ▷ est positive,
- ▷ est décroissante,
- ▷ et converge vers 0,

alors la série $\sum (-1)^n u_n$ converge.

Dans ce cas, la somme de la série est positive, et si on note $R_n = \sum_{k=n+1}^{+\infty} (-1)^k u_k$ le reste partiel d'indice n , alors pour tout $n \in \mathbb{N}$, R_n est du signe de son premier terme (i.e. du signe de $(-1)^{n+1}$) et $|R_n| \leq u_{n+1}$.

Exemple 2.59

- ▷ La série harmonique alternée $\sum_{n \geq 1} \frac{(-1)^n}{n}$ converge.
- ▷ La série $\sum_{n \geq 2} \frac{(-1)^n}{\ln n}$ converge.

Remarque 2.60

- ▷ Si $\sum_{n \geq n_0} (-1)^n u_n$ est une série alternée convergente, sa somme a le signe du premier terme de la série (ici le signe de $(-1)^{n_0} u_{n_0}$).

- ▷ La condition de décroissance de la suite u est essentielle ! Contre-exemple : la série $\sum_{n \geq 2} \frac{(-1)^n}{(-1)^n + \sqrt{n}}$

est une série alternée divergente.

De plus, cela fournit un contre-exemple au théorème de comparaison par équivalents si on ne tient pas compte de la condition sur le signe, qui doit être constant.

Exercice 2.61

Soit $\alpha > 1$. Pour $n \in \mathbb{N}^*$, on pose $u_n = \sum_{k=n}^{+\infty} \frac{(-1)^k}{k^\alpha + k}$.

Justifiez l'existence de u_n . Montrez que la série $\sum_{n \geq 1} u_n$ converge.

Correction 2.62

La suite $\left(\frac{1}{k^\alpha + k}\right)_{k \in \mathbb{N}^*}$ est positive, décroissante et converge vers 0 donc d'après le critère spécial des séries alternées, $\sum \frac{(-1)^k}{k^\alpha + k}$ converge.

Ainsi, $u_n = \sum_{k=n}^{+\infty} \frac{(-1)^k}{k^\alpha + k}$ existe pour $n \in \mathbb{N}^*$.

De plus, on a

$$|u_n| = \left| \sum_{k=n}^{+\infty} \frac{(-1)^k}{k^\alpha + k} \right| \leq \frac{1}{n^\alpha + n} \leq \frac{1}{n^\alpha}.$$

Or $\alpha > 1$ donc d'après le théorème de comparaison des séries à termes positifs, $\sum |u_n|$ converge et donc $\sum u_n$ converge.

Chapitre 3

Familles sommables

Sommaire

3.1	Sommes finies114
3.1.1	Définition	114
3.1.2	Propriétés	117
3.2	Conventions de calcul dans $\mathbb{R}_+ \cup \{+\infty\}$118
3.3	Somme d'une famille de réels positifs119
3.3.1	Propriétés	120
3.3.2	Théorème de sommation par paquets	121
3.3.3	Théorème de Fubini	121
3.4	Familles sommables dans un espace vectoriel normé de dimension finie122
3.4.1	Définitions	122
3.4.1.1	Cas réel	122
3.4.1.2	Cas complexe	122
3.4.1.3	Cas général	124
3.4.2	Propriétés	124
3.4.3	Théorème de sommation par paquets	125
3.4.4	Théorème de Fubini	127
3.4.5	Produit de Cauchy de deux séries	129

Dans ce chapitre, E désigne un espace vectoriel normé de dimension finie (qui peut être \mathbb{R} ou \mathbb{C}) et $\|\cdot\|$ la norme associée (qui est dans ce cas la valeur absolue ou le module).

Si A, B sont deux ensembles, alors on note $A \subseteq_f B$ pour indiquer que A est un sous-ensemble fini de B .

3.1 Sommes finies

3.1.1 Définition

D'abord un rappel : on définit par récurrence la somme de n éléments de E notés x_1, \dots, x_n par :

- ▷ si $n = 0$, alors $\sum_{k=1}^n x_k = 0$ (une somme vide a pour valeur 0 par convention) ;
- ▷ pour tout $n \in \mathbb{N}$, $\sum_{k=1}^{n+1} x_k = x_{n+1} + \sum_{k=1}^n x_k$.

On définit de même par récurrence les sommes de la forme $\sum_{k=p}^q x_k$ quand $p - 1 \leq q$ (si $q = p - 1$, la somme est vide donc vaut 0).

Proposition 3.1

Soient $n \in \mathbb{N}^*$ et $x_1, \dots, x_n \in E$.

Alors :

- (1) pour tout $(p, q) \in \llbracket 1 ; n \rrbracket^2$ tel que $p \leq q$, $\sum_{k=p}^n x_k = \sum_{k=p}^q x_k + \sum_{k=q+1}^n x_k$;
- (2) pour tout $\varphi \in \mathfrak{S}_n$, $\sum_{k=1}^n x_k = \sum_{k=1}^n x_{\varphi(k)}$.

Démonstration 3.2 (1)

On pose $\mathcal{P}(n)$ la proposition « $\forall (p, q) \in \llbracket 1 ; n \rrbracket, p \leq q \implies \sum_{k=p}^n x_k = \sum_{k=p}^q x_k + \sum_{k=q+1}^n x_k$ ».

Si $n = 1$, alors pour tout $(p, q) \in \llbracket 1 ; n \rrbracket^2$, $p = q = 1$ donc $\sum_{k=p}^n x_k = x_1 + 0 = \sum_{k=p}^q x_k + \sum_{k=q+1}^n x_k$ donc $\mathcal{P}(1)$ est vraie.

Si $\mathcal{P}(n)$ est vraie, alors soient $(x_1, \dots, x_{n+1}) \in E^{n+1}$ et $(p, q) \in \llbracket 1 ; n+1 \rrbracket^2$ tel que $p \leq q$:

- ▷ si $q \leq n$, alors par définition, $\sum_{k=p}^{n+1} x_k = \sum_{k=p}^n x_k + x_{n+1}$, donc d'après l'hypothèse de récurrence,
$$\sum_{k=p}^{n+1} x_k = \sum_{k=p}^q x_k + \sum_{k=q+1}^n x_k + x_{n+1} = \sum_{k=p}^q x_k + \sum_{k=q+1}^{n+1} x_k ;$$
- ▷ si $q = n+1$, alors $\sum_{k=p}^{n+1} x_k = \sum_{k=p}^q x_k + 0 = \sum_{k=p}^q x_k + \sum_{k=q+1}^{n+1} x_k$.

Dans les deux cas, on a montré $\sum_{k=p}^{n+1} x_k = \sum_{k=p}^q x_k + \sum_{k=q+1}^{n+1} x_k$. Autrement dit, $\mathcal{P}(n+1)$ est vraie.

D'après le principe de récurrence, pour tout $n \in \mathbb{N}^*$, $\mathcal{P}(n)$ est vraie. ■

Démonstration 3.3 (2)

On pose $\mathcal{P}(n)$ la proposition « $\forall \varphi \in \mathfrak{S}_n, \sum_{k=1}^n x_k = \sum_{k=1}^n x_{\varphi(k)}$ ».

$\mathcal{P}(1)$ est vraie car le seul élément de \mathfrak{S}_1 est l'application $1 \mapsto 1$.

Si $\mathcal{P}(n)$ est vraie, alors soient $(x_1, \dots, x_{n+1}) \in E^{n+1}$ et $\varphi \in \mathfrak{S}_{n+1}$:

- si $\varphi(n+1) = n+1$ alors φ induit une bijection de $\llbracket 1 ; n \rrbracket$ dans lui-même donc d'après l'hypothèse de récurrence, $\sum_{k=1}^n x_k = \sum_{k=1}^n x_{\varphi(k)}$ donc $\sum_{k=1}^{n+1} x_{\varphi(k)} = \sum_{k=1}^n x_{\varphi(k)} + x_{\varphi(n+1)} = \sum_{k=1}^n x_k + x_{n+1} = \sum_{k=1}^{n+1} x_k$;
- si $\varphi(n+1) = m \neq n+1$, alors on pose $\psi = (m \ n+1)\varphi$ et $a = \varphi^{-1}(n+1)$. On a alors $\psi(n+1) = n+1$, $\psi(a) = m$ et pour tout $k \in \llbracket 1 ; n+1 \rrbracket \setminus \{a, n+1\}$, $\psi(k) = \varphi(k)$. D'après le cas précédent, $\sum_{k=1}^{n+1} x_k = \sum_{k=1}^{n+1} x_{\psi(k)}$, donc en utilisant le résultat précédent :

$$\begin{aligned}
 \sum_{k=1}^{n+1} x_k &= \sum_{k=1}^{a-1} x_{\psi(k)} + x_{\psi(a)} + \sum_{k=a+1}^n x_{\psi(k)} + x_{\psi(n+1)} \\
 &= \sum_{k=1}^{a-1} x_{\varphi(k)} + x_m + \sum_{k=a+1}^n x_{\varphi(k)} + x_{n+1} \\
 &= \sum_{k=1}^{a-1} x_{\varphi(k)} + x_{\varphi(n+1)} + \sum_{k=a+1}^n x_{\varphi(k)} + x_{\varphi(a)} \\
 &= \sum_{k=1}^{n+1} x_{\varphi(k)}.
 \end{aligned}$$

Dans les deux cas, on a montré $\sum_{k=1}^{n+1} x_{\varphi(k)} = \sum_{k=1}^{n+1} x_k$. Autrement dit, $\mathcal{P}(n+1)$ est vraie.

D'après le principe de récurrence, pour tout $n \in \mathbb{N}^*$, $\mathcal{P}(n)$ est vraie. ■

Proposition 3.4

Soient I un ensemble fini et non-vidé d'indices, n son cardinal et f, g deux bijections de $\llbracket 1 ; n \rrbracket$ dans I (des énumérations de I).

Alors pour tout $(x_i)_{i \in I} \in E^I$, $\sum_{k=1}^n x_{f(k)} = \sum_{k=1}^n x_{g(k)}$.

Démonstration 3.5

On remarque que $g^{-1} \circ f$ est une bijection de $\llbracket 1 ; n \rrbracket$ dans lui-même donc d'après la Proposition 3.1 :

$$\sum_{k=1}^n x_{g(k)} = \sum_{k=1}^n x_{g(g^{-1} \circ f(k))} = \sum_{k=1}^n x_{f(k)}.$$
■

Autrement dit, quel que soit l'ordre dans lequel on numérote les éléments de la famille $(x_i)_{i \in I}$, on obtient toujours la même somme en les additionnant.

Définition 3.6

Si I est un ensemble fini d'indices et $(x_i)_{i \in I}$ une famille d'éléments de E , alors on pose $\sum_{i \in I} x_i$ la valeur d'une somme $\sum_{k=1}^n x_{f(k)}$, où f est une bijection de $\llbracket 1 ; n \rrbracket$ dans I quelconque.

Cette définition est cohérente, puisque la valeur de la somme $\sum_{k=1}^n x_{f(k)}$ ne dépend pas du choix de f d'après la proposition précédente. Autrement dit, il est inutile de connaître l'énumération choisie pour additionner les éléments de la famille, on peut considérer cette somme comme une somme « en vrac » de tous les éléments.

3.1.2 Propriétés

Proposition 3.7

Soient I un ensemble fini d'indices de cardinal n et $(x_i)_{i \in I}$ une famille d'éléments de E .

Alors :

- (1) pour toute bijection f d'un ensemble J dans I , $\sum_{i \in I} x_i = \sum_{j \in J} x_{f(j)}$ (changement d'indice dans une somme) ;
- (2) pour toute bijection f de I dans lui-même, $\sum_{i \in I} x_i = \sum_{i \in I} x_{f(i)}$ (propriété de commutativité) ;
- (3) pour tout couple (J, J') de parties de I disjointes et de réunion I , $\sum_{i \in I} x_i = \sum_{i \in J} x_i + \sum_{i \in J'} x_i$ (propriété d'associativité) ;
- (4) plus généralement, pour toute partition $(I_k)_{k \in K}$ de l'ensemble I , $\sum_{i \in I} x_i = \sum_{k \in K} \sum_{i \in I_k} x_i$.

Démonstration 3.8 (1)

Soit f une bijection de J dans I . On choisit une énumération ψ de J . Alors $f \circ \psi$ est une énumération de I .

Alors, par définition, $\sum_{i \in I} x_i = \sum_{k=1}^n x_{f \circ \psi(k)}$ et $\sum_{j \in J} x_{f(j)} = \sum_{k=1}^n x_{f \circ \psi(k)}$ donc $\sum_{i \in I} x_i = \sum_{j \in J} x_{f(j)}$. ■

Démonstration 3.9 (2)

Cas particulier $I = J$ du point précédent. ■

Démonstration 3.10 (3)

Soit (J, J') un couple de parties de I disjointes et de réunion I . On note q le cardinal de J , de sorte que $n - q$ est le cardinal de J' .

On choisit une énumération φ de J et une énumération ψ de J' . Alors l'application

$$\begin{aligned} \theta : \llbracket 1 ; n \rrbracket &\longrightarrow I \\ k &\longmapsto \begin{cases} \varphi(k) & \text{si } k \leq q \\ \psi(k - q) & \text{sinon} \end{cases} \end{aligned}$$

est une énumération de I .

Donc

$$\begin{aligned} \sum_{i \in I} x_i &= \sum_{k=1}^n x_{\theta(k)} \\ &= \sum_{k=1}^q x_{\theta(k)} + \sum_{k=q+1}^n x_{\theta(k)} \\ &= \sum_{k=1}^q x_{\varphi(k)} + \sum_{k=q+1}^n x_{\psi(k-q)} \\ &= \sum_{k=1}^q x_{\varphi(k)} + \sum_{k=1}^{n-q} x_{\psi(k)} \\ &= \sum_{i \in J} x_i + \sum_{i \in J'} x_i. \end{aligned}$$
■

Démonstration 3.11 (4)

Si $(I_k)_{k \in K}$ est une partition de l'ensemble I , l'ensemble K est fini donc par récurrence sur le cardinal b de K , on montre $\sum_{i \in I} x_i = \sum_{k \in K} \sum_{i \in I_k} x_i$ en utilisant le cas $b = 2$ démontré précédemment (il suffit de choisir un élément a de K , poser $J = I_a$ et $J' = \bigsqcup_{k \in K \setminus \{a\}} I_k$ et remarquer que la famille $(I_k)_{k \in K \setminus \{a\}}$ est une partition de l'ensemble J' et que le cardinal de $K \setminus \{a\}$ est $b - 1$). ■

3.2 Conventions de calcul dans $\mathbb{R}_+ \cup \{+\infty\}$

L'ensemble $\mathbb{R}_+ \cup \{+\infty\}$ est muni d'une addition : pour tout $(x, y) \in (\mathbb{R}_+ \cup \{+\infty\})^2$,

- ▷ si x et y sont réels, $x + y$ est la somme habituelle de deux réels positifs ;
- ▷ si $x = +\infty$ ou $y = +\infty$ alors on pose $x + y = +\infty$

et d'une multiplication :

- ▷ si x et y sont réels, xy est le produit habituel de deux réels positifs ;
- ▷ si $x = 0$ ou $y = 0$ alors on pose $xy = 0$;
- ▷ si $x = y = +\infty$ alors on pose $xy = +\infty$.

Il est aussi muni d'une relation d'ordre :

- ▷ si x et y sont deux réels, alors $x \leq y$ ou $x < y$ désignent les relations habituelles ;
- ▷ si x est réel et $y = +\infty$, alors on pose $x \leq +\infty$ et $x < +\infty$;
- ▷ si $x = y = +\infty$ alors $+\infty \leq +\infty$.

Proposition 3.12

L'addition dans $\mathbb{R}_+ \cup \{+\infty\}$ est associative, commutative et admet pour neutre 0.

La relation \leq est une relation d'ordre total dans $\mathbb{R}_+ \cup \{+\infty\}$.

De plus, l'addition et la multiplication sont compatibles avec la relation d'ordre : on peut additionner ou multiplier deux inégalités membre à membre.

Définition 3.13

Soit A une partie non-vide de $\mathbb{R}_+ \cup \{+\infty\}$.

Si A ne contient pas $+\infty$, alors :

- ▷ si A est majorée, elle possède une borne supérieure dans \mathbb{R} ;
- ▷ sinon on pose $\sup A = +\infty$.

Si A contient $+\infty$, on pose $\sup A = +\infty$.

Cette définition prolonge la notion de borne supérieure à toutes les parties de $\mathbb{R}_+ \cup \{+\infty\}$, au sens où pour toute partie A de $\mathbb{R}_+ \cup \{+\infty\}$, $\sup A$ est le plus petit majorant dans $\mathbb{R}_+ \cup \{+\infty\}$ de la partie A .

3.3 Somme d'une famille de réels positifs

Définition 3.14

Soit $(x_i)_{i \in I}$ une famille d'éléments de $\mathbb{R}_+ \cup \{+\infty\}$.

On pose $\sum_{i \in I} x_i = \sup \left\{ \sum_{i \in J} x_i \mid J \subseteq_f I \right\}$.

Remarque 3.15

Cette définition est sensée, car l'ensemble $\left\{ \sum_{i \in J} x_i \mid J \subseteq_f I \right\}$ est une partie de $\mathbb{R}_+ \cup \{+\infty\}$, donc possède toujours une borne supérieure dans $\mathbb{R}_+ \cup \{+\infty\}$.

Définition 3.16

Soit $(x_i)_{i \in I}$ une famille d'éléments de $\mathbb{R}_+ \cup \{+\infty\}$.

On dit que la famille $(x_i)_{i \in I}$ est sommable quand $\sum_{i \in I} x_i < +\infty$.

Évidemment, une famille sommable positive ne peut pas prendre la valeur $+\infty$, autrement dit, une famille sommable est nécessairement une famille de réels positifs.

3.3.1 Propriétés

Proposition 3.17

La somme d'une famille $(x_i)_{i \in I}$ d'éléments de $\mathbb{R}_+ \cup \{+\infty\}$ est invariante par permutation : si σ est une permutation de I , alors $\sum_{i \in I} x_i = \sum_{i \in I} x_{\sigma(i)}$.

En particulier, si $(x_i)_{i \in I}$ est une famille sommable, toute permutation de la famille est encore une famille sommable de même somme.

En particulier, dans le cas où $I = \mathbb{N}$, si une série à termes positifs $\sum u_n$ est convergente, alors on dit qu'elle est commutativement convergente : changer l'ordre des termes change bien sûr les valeurs des sommes partielles mais ne change pas la valeur de la limite de ces sommes partielles.

Proposition 3.18

Soient $(x_i)_{i \in I}, (y_i)_{i \in I}$ deux familles d'éléments de $\mathbb{R}_+ \cup \{+\infty\}$ et λ un réel positif.

Alors $\sum_{i \in I} (x_i + y_i) = \sum_{i \in I} x_i + \sum_{i \in I} y_i$ et $\sum_{i \in I} \lambda x_i = \lambda \sum_{i \in I} x_i$.

Corollaire 3.19

La somme de deux familles positives est sommable ssi les deux familles sont sommables.

Le produit par un réel strictement positif d'une famille positive est sommable ssi la famille est sommable.

Proposition 3.20

Soient $(x_i)_{i \in I}$, $(y_i)_{i \in I}$ deux familles d'éléments de $\mathbb{R}_+ \cup \{+\infty\}$.

Si pour tout $i \in I$, $0 \leq x_i \leq y_i$ et si la famille $(y_i)_{i \in I}$ est sommable, alors la famille $(x_i)_{i \in I}$ l'est aussi et $\sum_{i \in I} x_i \leq \sum_{i \in I} y_i$.

3.3.2 Théorème de sommation par paquets

Théorème 3.21

Soit $(x_i)_{i \in I}$ une famille de réels positifs.

Si I est partitionné en une famille $(I_p)_{p \in P}$ de parties, alors

$$\sum_{p \in P} \sum_{i \in I_p} x_i = \sum_{i \in I} x_i.$$

3.3.3 Théorème de Fubini

Théorème 3.22

Soit $(x_{ij})_{(i,j) \in I \times J}$ une famille de réels positifs. Alors

$$\sum_{(i,j) \in I \times J} x_{ij} = \sum_{j \in J} \sum_{i \in I} x_{ij} = \sum_{i \in I} \sum_{j \in J} x_{ij}.$$

Ce résultat se généralise par récurrence dans le cas d'un produit cartésien $I_1 \times \cdots \times I_k$.

Un cas particulier courant.

Proposition 3.23

Soient $(a_i)_{i \in I}$, $(b_j)_{j \in J}$ deux familles de réels positifs.

Alors la famille $(a_i b_j)_{(i,j) \in I \times J}$ est sommable ssi les familles $(a_i)_{i \in I}$ et $(b_j)_{j \in J}$ sont sommables et dans ce cas, on a

$$\sum_{(i,j) \in I \times J} a_i b_j = \sum_{i \in I} a_i \times \sum_{j \in J} b_j.$$

3.4 Familles sommables dans un espace vectoriel normé de dimension finie

3.4.1 Définitions

Définition 3.24

Soit $(x_i)_{i \in I}$ une famille de vecteurs de E .

On dit que la famille $(x_i)_{i \in I}$ est sommable quand la famille $(\|x_i\|)_{i \in I}$ est sommable, c'est-à-dire quand $\sum_{i \in I} \|x_i\| < +\infty$.

Cette définition est indépendante du choix de la norme, car en dimension finie, toutes les normes sont équivalentes.

3.4.1.1 Cas réel

Définition 3.25

Soit x un réel.

On appelle partie positive de x le réel $x^+ = \max(0, x)$ et partie négative de x le réel $x^- = -\min(x, 0)$.

On remarque les égalités suivantes : $|x| = x^+ + x^-$ et $x = x^+ - x^-$.

Proposition 3.26

Soit $(x_i)_{i \in I}$ une famille sommable de nombre réels.

Alors les familles positives $(x_i^+)_{i \in I}$ et $(x_i^-)_{i \in I}$ sont sommables et on a bien sûr $\sum_{i \in I} |x_i| = \sum_{i \in I} x_i^+ + \sum_{i \in I} x_i^-$.

On pose alors $\sum_{i \in I} x_i = \sum_{i \in I} x_i^+ - \sum_{i \in I} x_i^-$, qui est un réel tel que $\left| \sum_{i \in I} x_i \right| \leq \sum_{i \in I} |x_i|$.

3.4.1.2 Cas complexe

Proposition 3.27

Soit $(a_k)_{k \in I}$ une famille sommable de nombres complexes.

Alors les deux familles réelles $(\operatorname{Re} a_k)_{k \in I}$ et $(\operatorname{Im} a_k)_{k \in I}$ sont sommables.

On pose alors $\sum_{k \in I} a_k = \sum_{k \in I} \operatorname{Re} a_k + i \sum_{k \in I} \operatorname{Im} a_k$ qui est un complexe tel que $\left| \sum_{k \in I} a_k \right| \leq \sum_{k \in I} |a_k|$.

Exemple 3.28

- Toute famille finie est sommable et sa somme au sens des familles sommables est sa somme habituelle.
- Une suite $(a_n)_{n \in \mathbb{N}}$ est sommable ssi la série $\sum_n a_n$ est absolument convergente.

Exercice 3.29

Soit $\theta \in]0 ; 2\pi[$. Montrez que la famille $\left(\frac{e^{i\ell\theta}}{(k+\ell)^3} \right)_{(k,\ell) \in (\mathbb{N}^*)^2}$ est sommable.

Correction 3.30

On a

$$\begin{aligned} \sum_{(k,\ell) \in (\mathbb{N}^*)^2} \left| \frac{e^{i\ell\theta}}{(k+\ell)^3} \right| &= \sum_{(k,\ell) \in (\mathbb{N}^*)^2} \frac{1}{(k+\ell)^3} \\ &= \sum_{k \in \mathbb{N}^*} \sum_{\ell \in \mathbb{N}^*} \frac{1}{(k+\ell)^3} \\ &= \sum_{k=1}^{+\infty} \sum_{\ell=1}^{+\infty} \frac{1}{(k+\ell)^3}. \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{Fubini positif}$$

On veut donc montrer que pour $k \in \mathbb{N}^*$, la série $\sum_{\ell \geq 1} \frac{1}{(k+\ell)^3}$ converge et que la série $\sum_{k \geq 1} \sum_{\ell=1}^{+\infty} \frac{1}{(k+\ell)^3}$ converge.

Pour $k \geq 1$, on a $\frac{1}{(k+\ell)^3} \underset{\ell \rightarrow +\infty}{\sim} \frac{1}{\ell^3}$ donc par théorème de comparaison des séries à termes positifs, $\sum_{\ell \geq 1} \frac{1}{(k+\ell)^3}$ converge.

Or $\sum_{\ell=1}^{+\infty} \frac{1}{(k+\ell)^3} = \sum_{\ell=k+1}^{+\infty} \frac{1}{\ell^3} \underset{k \rightarrow +\infty}{\sim} \frac{1}{2k^2}$ (cf. DS1).

Donc $\sum_{k \geq 1} \sum_{\ell=1}^{+\infty} \frac{1}{(k+\ell)^3}$ converge par théorème de comparaison des séries à termes positifs.

Finalement, $\left(\frac{e^{i\ell\theta}}{(k+\ell)^3} \right)_{(k,\ell) \in (\mathbb{N}^*)^2}$ est sommable.

3.4.1.3 Cas général

Comme E est de dimension finie, on en choisit une base $\mathcal{B} = (e_1, \dots, e_p)$.

Pour toute famille sommable $(x_i)_{i \in I} \in E^I$, on note (x_{i1}, \dots, x_{ip}) les coordonnées de x_i dans la base \mathcal{B} .

Alors pour tout $k \in \llbracket 1 ; p \rrbracket$, la famille de réels ou de complexes $(x_{ik})_{i \in I}$ est sommable.

On pose alors $\sum_{i \in I} x_i$ le vecteur qui a pour coordonnées $\left(\sum_{i \in I} x_{i1}, \dots, \sum_{i \in I} x_{ip} \right)$ dans la base \mathcal{B} .

On note $\ell^1(I, E)$ l'ensemble des familles sommables de E indicées par I .

3.4.2 Propriétés

Proposition 3.31

Toute sous-famille d'une famille sommable de E est elle-même sommable.

Toute permutation d'une famille sommable de E est encore sommable et de même somme.

En particulier, les séries absolument convergentes sont commutativement convergentes.

Les familles sommables sont celles qui sont approchables par des familles finies à ε près au sens de la proposition suivante.

Comme pour les séries, on dispose d'un théorème de comparaison entre familles sommables.

Proposition 3.32

Soient $(a_i)_{i \in I}, (b_i)_{i \in I}$ deux familles d'éléments indicées par I .

Si pour tout $i \in I$, $0 \leq \|a_i\| \leq b_i$ et si la famille $(b_i)_{i \in I}$ est une famille sommable de réels positifs, alors la famille $(a_i)_{i \in I}$ est sommable et on a $\left\| \sum_{i \in I} a_i \right\| \leq \sum_{i \in I} \|a_i\| \leq \sum_{i \in I} b_i$.

La linéarité est encore vérifiée, mais n'est pas évidente au regard des définitions.

Proposition 3.33

L'ensemble $\ell^1(I, E)$ est un espace vectoriel et l'application $(a_i)_{i \in I} \mapsto \sum_{i \in I} a_i$ est une forme linéaire :

si $(a_i), (b_i) \in \ell^1(I, E)$ et λ est un scalaire, alors $\sum_{i \in I} (a_i + b_i) = \sum_{i \in I} a_i + \sum_{i \in I} b_i$ et $\sum_{i \in I} \lambda a_i = \lambda \sum_{i \in I} a_i$.

3.4.3 Théorème de sommation par paquets

Théorème 3.34

Soit $(a_i)_{i \in I}$ une famille sommable de E .

Si I est partitionné en une famille $(I_p)_{p \in P}$ de parties, alors pour tout $p \in P$, $(a_i)_{i \in I_p}$ est sommable et

$$\sum_{i \in I} a_i = \sum_{p \in P} \sum_{i \in I_p} a_i.$$

Exercice 3.35

Montrez que pour tout complexe z tel que $0 < |z| < 1$, la famille $(z^{|n|})_{n \in \mathbb{Z}}$ est sommable et calculez sa somme.

Correction 3.36

On a

$$\begin{aligned} \sum_{n \in \mathbb{Z}} |z^{|n|}| &= \sum_{n \in \mathbb{Z}} |z|^{|n|} \\ &= \sum_{n \in \mathbb{N}} |z|^n + \sum_{n \in \mathbb{Z}_-^*} |z|^{-n} \\ &= \sum_{n=0}^{+\infty} |z|^n + \sum_{n=1}^{+\infty} |z|^n \quad \left. \vphantom{\sum_{n=1}^{+\infty} |z|^n} \right) |z| < 1 \text{ donc } \sum |z|^n \text{ converge} \\ &< +\infty. \end{aligned}$$

Donc $(z^{|n|})_{n \in \mathbb{Z}}$ est sommable et on a

$$\begin{aligned} \sum_{n \in \mathbb{Z}} z^{|n|} &= \sum_{n \in \mathbb{N}} z^n + \sum_{n \in \mathbb{Z}_-^*} z^{-n} \\ &= \sum_{n=0}^{+\infty} z^n + \sum_{n=1}^{+\infty} z^n \\ &= \frac{1}{1-z} + \frac{z}{1-z} \\ &= \frac{1+z}{1-z}. \end{aligned}$$

Exercice 3.37

Montrez que la famille $\left(\frac{(-1)^n}{\max(m, n)^3} \right)_{m, n \geq 1}$ est sommable et calculez sa somme en fonction de $\zeta(2)$ et $\zeta(3)$.

On a
$$\sum_{(m,n) \in (\mathbb{N}^*)^2} \left| \frac{(-1)^n}{\max(m,n)^3} \right| = \sum_{(m,n) \in (\mathbb{N}^*)^2} \frac{1}{\max(m,n)^3}.$$

On pose $I_k = \{(m, k) \mid m \in \llbracket 1 ; k \rrbracket\} \cup \{(k, n) \mid n \in \llbracket 1 ; k \rrbracket\}$.

On a bien $(\mathbb{N}^*)^2 = \bigsqcup_{k \in \mathbb{N}^*} I_k$.

On a donc

$$\begin{aligned} \sum_{(m,n) \in (\mathbb{N}^*)^2} \frac{1}{\max(m,n)^3} &= \sum_{k \in \mathbb{N}^*} \sum_{(m,n) \in I_k} \frac{1}{\max(m,n)^3} \\ &= \sum_{k \in \mathbb{N}^*} \sum_{(m,n) \in I_k} \frac{1}{k^3} \\ &= \sum_{k \in \mathbb{N}^*} \frac{2k-1}{k^3} \\ &= \sum_{k=1}^{+\infty} \frac{2k-1}{k^3} \\ &< +\infty \end{aligned}$$

car $\frac{2k-1}{k^3} \underset{k \rightarrow +\infty}{\sim} \frac{2}{k^2}$ donc par théorème de comparaison des séries à termes positifs, $\sum \frac{2k-1}{k^3}$ converge.

Donc $\left(\frac{(-1)^n}{\max(m,n)^3} \right)_{(m,n) \in (\mathbb{N}^*)^2}$ est sommable.

Donc
$$S = \sum_{(m,n) \in (\mathbb{N}^*)^2} \frac{(-1)^n}{\max(m,n)^3} = \sum_{k \in \mathbb{N}^*} \frac{1}{k^3} \sum_{(m,n) \in I_k} (-1)^n.$$

Or

$$\begin{aligned} \sum_{(m,n) \in I_k} (-1)^n &= k (-1)^k + \sum_{n=1}^{k-1} (-1)^n \\ &= k (-1)^k + (-1) \frac{1 - (-1)^{k-1}}{1 - (-1)} \\ &= k (-1)^k - \frac{1}{2} (1 + (-1)^k). \end{aligned}$$

Donc

$$\begin{aligned}
 S &= \sum_{k=1}^{+\infty} \frac{k(-1)^k - \frac{1}{2}(1 + (-1)^k)}{k^3} \\
 &= \sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2} - \sum_{\substack{k=1 \\ k \text{ pair}}}^{+\infty} \frac{1}{k^3} \\
 &= \sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2} - \underbrace{\sum_{\ell=1}^{+\infty} \frac{1}{(2\ell)^3}}_{=\frac{1}{8}\zeta(3)}.
 \end{aligned}$$

Or

$$\begin{aligned}
 \sum_{k=1}^{+\infty} \frac{(-1)^k}{k^2} &= \sum_{k=1}^{+\infty} \frac{(-1)^k + 1 - 1}{k^2} \\
 &= \sum_{k=1}^{+\infty} \frac{(-1)^k + 1}{k^2} - \sum_{k=1}^{+\infty} \frac{1}{k^2} \\
 &= \sum_{\substack{k=1 \\ k \text{ pair}}} \frac{2}{k^2} - \zeta(2) \\
 &= \sum_{\ell=1}^{+\infty} \frac{2}{(2\ell)^2} - \zeta(2) \\
 &= \frac{-1}{2} \zeta(2).
 \end{aligned}$$

D'où $S = -\frac{1}{8}\zeta(3) - \frac{1}{2}\zeta(2)$.

3.4.4 Théorème de Fubini

Théorème 3.39

Soit $(a_{ij})_{(i,j) \in I \times J}$ une famille sommable de E .

Alors pour tout $i \in I$, la famille $(a_{ij})_{j \in J}$ est sommable; pour tout $j \in J$, la famille $(a_{ij})_{i \in I}$ est sommable et

$$\sum_{(i,j) \in I \times J} a_{ij} = \sum_{j \in J} \sum_{i \in I} a_{ij} = \sum_{i \in I} \sum_{j \in J} a_{ij}.$$

Ce résultat se généralise par récurrence dans le cas d'un produit cartésien $I_1 \times \cdots \times I_k$.

Exercice 3.40

Montrez que la famille $\left(\frac{(-1)^p}{q^p}\right)_{p,q \geq 2}$ est sommable et calculez sa somme.

Correction 3.41

On a

$$\begin{aligned}\sum_{p,q \geq 2} \left| \frac{(-1)^p}{q^p} \right| &= \sum_{p,q \geq 2} \frac{1}{q^p} \\ &= \sum_{q \geq 2} \sum_{p \geq 2} \frac{1}{q^p} \\ &= \sum_{q \geq 2} \frac{1/q^2}{1 - 1/q} \\ &= \sum_{q \geq 2} \frac{1}{q^2 - q} \\ &= \sum_{q \geq 2} \frac{1}{q(q-1)} \\ &= \sum_{q \geq 2} \left(\frac{1}{q-1} - \frac{1}{q} \right).\end{aligned}$$

Donc $\left(\frac{(-1)^p}{q^p}\right)_{p,q \geq 2}$ est sommable.

De plus, on a

$$\begin{aligned}\sum_{p,q \geq 2} \frac{(-1)^p}{q^p} &= \sum_{p,q \geq 2} \left(\frac{-1}{q} \right)^p \\ &= \sum_{q=2}^{+\infty} \sum_{p=2}^{+\infty} \left(\frac{-1}{q} \right)^p \\ &= \sum_{q=2}^{+\infty} \frac{1/q^2}{1 + 1/q} \\ &= \sum_{q=2}^{+\infty} \frac{1}{q(q+1)} \\ &= \sum_{q=2}^{+\infty} \left(\frac{1}{q} - \frac{1}{q+1} \right) \\ &= \frac{1}{2}.\end{aligned}$$

Un cas particulier courant.

Proposition 3.42

Soient $(a_i) \in \ell^1(I, E)$ et $(b_j) \in \ell^1(J, E)$.

Alors la famille $(a_i b_j)_{(i,j) \in I \times J}$ est sommable et

$$\sum_{(i,j) \in I \times J} a_i b_j = \sum_{i \in I} a_i \times \sum_{j \in J} b_j.$$

3.4.5 Produit de Cauchy de deux séries

Définition 3.43

Soient $\sum_{n \geq 0} a_n$ et $\sum_{n \geq 0} b_n$ deux séries à termes dans E .

On appelle produit de Cauchy des deux séries la série $\sum_{n \geq 0} c_n$ où pour tout $n \in \mathbb{N}$, $c_n = \sum_{k=0}^n a_k b_{n-k}$.

Remarque 3.44

Quand les séries ne commencent pas à partir du rang 0, il faut se méfier ! Une idée simple est de se ramener au cas précédent en décalant les indices.

Exemple très courant : les séries commençant au rang 1. Dans ce cas, le produit de Cauchy des séries

$\sum_{n \geq 1} a_n$ et $\sum_{n \geq 1} b_n$ est la série $\sum_{n \geq 1} c_n$ où pour tout $n \in \mathbb{N}^*$, $c_n = \sum_{k=1}^n a_k b_{n-k+1}$.

Théorème 3.45

Si les séries $\sum_{n \geq 0} a_n$ et $\sum_{n \geq 0} b_n$ sont absolument convergentes, alors leur produit de Cauchy est aussi absolument convergent et

$$\sum_{n=0}^{+\infty} c_n = \sum_{n=0}^{+\infty} a_n \times \sum_{n=0}^{+\infty} b_n.$$

Un exemple fondamental.

Proposition 3.46

Soit $z \in \mathbb{C}$. La série de terme général $\frac{z^n}{n!}$ est absolument convergente et

$$\sum_{n=0}^{+\infty} \frac{z^n}{n!} = e^z.$$

Remarque 3.47

L'absolue convergence des séries est indispensable ! Si on ne suppose que la convergence des séries, alors le produit de Cauchy peut très bien être une série divergente (voir exercice suivant).

Exercice 3.48

Soit $x > 0$. On pose b_n la somme partielle de la série $\sum_{n \geq 0} \frac{x^n}{n!}$ et $c_n = \frac{b_n}{x^n}$.

Donnez une CNS sur x pour que la série $\sum_{n \geq 0} c_n$ converge. Dans le cas où elle converge, donnez la valeur de sa somme en fonction de x .

Exercice 3.49

Pour tout $n \in \mathbb{N}^*$, on pose $a_n = \frac{(-1)^n}{\sqrt{n}}$.

Montrez que la série $\sum_{n \geq 1} a_n$ converge, mais que son produit de Cauchy avec elle-même diverge

(indication : pour tout $b > 0$, pour tout $x \in [0 ; b]$, $x(b-x) \leq \frac{b^2}{4}$).

Chapitre 4

Rappels et compléments d'algèbre linéaire

Sommaire

4.1	Sommes de sous-espaces vectoriels	132
4.1.1	Généralités	132
4.1.2	Sommes directes	132
4.1.3	Sous-espaces supplémentaires	134
4.1.4	Cas particulier de deux sous-espaces	138
4.1.5	Applications linéaires et sommes directes	138
4.2	Somme de sous-espaces vectoriels en dimension finie	139
4.2.1	Base adaptée à un sous-espace	139
4.2.2	Sommes directes et bases	140
4.2.3	Dimension d'une somme de sous-espaces vectoriels	140
4.2.4	Sous-espaces supplémentaires	142
4.2.5	Dimension d'une somme de deux sous-espaces vectoriels	144
4.3	Polynômes d'endomorphismes et de matrices	144
4.3.1	\mathbb{K} -algèbres	144
4.3.1.1	Définition	144
4.3.1.2	Polynômes d'éléments dans une \mathbb{K} -algèbre	145
4.3.2	Cas particulier des algèbres $\mathcal{L}(E)$ ou $\mathcal{M}_n(\mathbb{K})$	146
4.3.3	Polynôme annulateur d'une matrice ou d'un endomorphisme	146
4.3.4	Utilisation pratique d'un polynôme annulateur	149
4.3.4.1	Calcul de l'inverse	149
4.3.4.2	Calcul de puissances	152
4.4	Matrices semblables, trace	156
4.4.1	Trace d'une matrice	156
4.4.2	Matrices semblables	156
4.4.3	Trace d'un endomorphisme	157
4.5	Opérations par blocs	158
4.5.1	Cas général	158
4.5.2	Cas particuliers des matrices carrées	159
4.5.3	Interprétation des blocs	159

Dans tout ce chapitre, \mathbb{K} désigne un sous-corps de \mathbb{C} , en général \mathbb{R} ou \mathbb{C} .

4.1 Sommes de sous-espaces vectoriels

4.1.1 Généralités

Définition 4.1

Soient E un \mathbb{K} -espace vectoriel et F_1, \dots, F_n des sous-espaces vectoriels de E .

On appelle somme de F_1, \dots, F_n l'ensemble noté $F_1 + \dots + F_n$:

$$F_1 + \dots + F_n = \left\{ \sum_{i=1}^n x_i \mid (x_1, \dots, x_n) \in F_1 \times \dots \times F_n \right\}.$$

Proposition 4.2

Soit $(u_1, \dots, u_p, u_{p+1}, \dots, u_n) \in E^n$.

Alors $\text{Vect}(u_1, \dots, u_p, u_{p+1}, \dots, u_n) = \text{Vect}(u_1, \dots, u_p) + \text{Vect}(u_{p+1}, \dots, u_n)$.

Proposition 4.3

Avec les mêmes notations : $F_1 + \dots + F_n$ est un sous-espace vectoriel de E .

De plus, c'est le plus petit sous-espace vectoriel qui contient F_1, \dots, F_n .

Si on connaît des familles génératrices de chacun des sous-espaces vectoriels F_1, \dots, F_n , alors en concaténant ces familles, on obtient une famille génératrice de $F_1 + \dots + F_n$.

Conséquence : en fractionnant une famille génératrice de E en sous-familles, on décompose l'espace E en une somme de sous-espaces vectoriels.

4.1.2 Sommes directes

Définition 4.4

Soient E un \mathbb{K} -espace vectoriel et F_1, \dots, F_n des sous-espaces vectoriels de E .

On dit que la somme $F_1 + \dots + F_n$ est directe quand tout vecteur de $F_1 + \dots + F_n$ a une unique écriture $\sum_{i=1}^n x_i$ où $(x_1, \dots, x_n) \in F_1 \times \dots \times F_n$.

On dit aussi que les sous-espaces sont en somme directe. Dans ce cas, quand on veut insister sur cette propriété, on note la somme sous la forme $F_1 \oplus \dots \oplus F_n = \bigoplus_{i=1}^n F_i$.

Proposition 4.5

Avec les mêmes hypothèses.

La somme $F_1 + \cdots + F_n$ est directe ssi le vecteur nul a une unique décomposition $\sum_{i=1}^n x_i$ où $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$, qui est la décomposition triviale.

Autrement dit, la somme $F_1 + \cdots + F_n$ est directe ssi la seule solution de l'équation $\sum_{i=1}^n x_i = 0$ d'inconnue $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$ est le n -uplet nul.

Démonstration 4.6

\Rightarrow

Immédiat : le vecteur nul appartient à $F_1 + \cdots + F_n$ et a une unique écriture sous la forme $\sum_{i=1}^n x_i$ où $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$.

Or on en connaît une : $0 = 0 + \cdots + 0$.

Donc la seule solution à l'équation $\sum_{i=1}^n x_i = 0$ d'inconnue $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$ est la solution triviale $(0, \dots, 0)$.

\Leftarrow

Soit $z \in F_1 + \cdots + F_n$.

On veut montrer que z a une unique écriture $z = \sum_{i=1}^n x_i$ où $(x_1, \dots, x_n) \in F_1 \times \cdots \times F_n$.

Si $z = \sum_{i=1}^n x_i = \sum_{i=1}^n x'_i$ où $(x_1, \dots, x_n), (x'_1, \dots, x'_n) \in F_1 \times \cdots \times F_n$, alors $\sum_{i=1}^n (x_i - x'_i) = 0$.

Or pour tout $i \in \llbracket 1 ; n \rrbracket$, $x_i - x'_i \in F_i$ car F_i est un sous-espace vectoriel.

Ainsi, pour tout $i \in \llbracket 1 ; n \rrbracket$, $x_i - x'_i = 0$ i.e. $(x_1, \dots, x_n) = (x'_1, \dots, x'_n)$.

D'où l'unicité voulue. ■

Un exemple fondamental : si (v_1, \dots, v_n) est une famille libre, alors les droites vectorielles $\text{Vect}(v_i)$ sont en somme directe.

Proposition 4.7

Avec les mêmes hypothèses.

Si la somme $F_1 + \cdots + F_n$ est directe, alors en concaténant des familles libres de chacun des sous-espaces vectoriels, on obtient une famille libre.

Démonstration 4.8

Si $F_1 + \dots + F_n$ est une somme directe alors soient L_1, \dots, L_n des familles libres de F_1, \dots, F_n respectivement.

On note $L_i = (e_{i1}, \dots, e_{i\ell_i})$ où $i \in \llbracket 1 ; n \rrbracket$.

Soit $(\lambda_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq \ell_i}}$ une famille de scalaires telle que $\sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq \ell_i}} \lambda_{ij} e_{ij} = 0$.

Alors $\sum_{i=1}^n \sum_{j=1}^{\ell_i} \lambda_{ij} e_{ij} = 0$.

Or $F_1 + \dots + F_n$ est directe donc pour tout $i \in \llbracket 1 ; n \rrbracket$, $\sum_{j=1}^{\ell_i} \lambda_{ij} e_{ij} = 0$.

Or L_i est libre donc $\lambda_{i1} = \dots = \lambda_{i\ell_i} = 0$.

Donc la concaténation de L_1, \dots, L_n est libre. ■

4.1.3 Sous-espaces supplémentaires

Définition 4.9

Soient E un \mathbb{K} -espace vectoriel et F_1, \dots, F_n des sous-espaces vectoriels de E .

On dit que les sous-espaces F_1, \dots, F_n sont supplémentaires (dans E) quand $E = \bigoplus_{i=1}^n F_i$.

On déduit des deux parties précédentes le résultat de la décomposition d'un vecteur.

Proposition 4.10

Soient E un \mathbb{K} -espace vectoriel et F_1, \dots, F_n des sous-espaces vectoriels de E .

Il y a équivalence entre :

- ▷ les sous-espaces F_1, \dots, F_n sont supplémentaires
- ▷ tout vecteur de E peut s'écrire de façon unique comme somme de vecteurs des sous-espaces vectoriels F_1, \dots, F_n :

$$\forall v \in E, \exists! (v_1, \dots, v_n) \in F_1 \times \dots \times F_n, \quad v = \sum_{i=1}^n v_i.$$

Dans ce cas, soit v un vecteur de E . Il existe un unique n -uplet $(v_1, \dots, v_n) \in F_1 \times \dots \times F_n$ tel que $v = \sum_{i=1}^n v_i$.

Pour tout $j \in \llbracket 1 ; n \rrbracket$, on définit $p_j : E \longrightarrow E$ en posant $p_j(v) = v_j$.

Alors les applications p_j sont des projecteurs qui vérifient les propriétés :

$$\triangleright \sum_{i=1}^n p_i = \text{id}_E$$

\triangleright pour tout $(i, j) \in \llbracket 1 ; n \rrbracket^2$, $p_i \circ p_j = \delta_{ij} p_i$ où δ_{xy} est le symbole de Kronecker.

La réciproque est vraie : si (p_1, \dots, p_n) sont n projecteurs vérifiant les deux propriétés précédentes, alors les sous-espaces $(\text{Im } p_i)_{1 \leq i \leq n}$ sont supplémentaires.

Démonstration 4.11

Soient F_1, \dots, F_n supplémentaires dans E .

\triangleright Pour tout $j \in \llbracket 1 ; n \rrbracket$, p_j est un projecteur :

Soit $v \in E$. On écrit $v = v_1 + \dots + v_j + \dots + v_n$.

Par définition, on a $p_j(v) = v_j = 0 + \dots + 0 + v_j + 0 + \dots + 0$.

Donc, par définition, on a $p_j(p_j(v)) = p_j(v_j) = v_j$.

Donc $p_j^2 = p_j$.

Donc p est un projecteur (la linéarité est évidente).

\triangleright Pour tout $v \in E$, $v = v_1 + \dots + v_n = p_1(v) + \dots + p_n(v)$ donc $\sum_{i=1}^n p_i = \text{id}_E$.

\triangleright Montrons que pour tout $(i, j) \in \llbracket 1 ; n \rrbracket^2$, $p_i p_j = \delta_{ij} p_i$.

Si $i = j$, c'est vrai car p_i est un projecteur.

Si $i \neq j$, soit $v \in E$. On écrit $v = v_1 + \dots + v_n$.

Par définition, $p_j(v) = v_j = 0 + \dots + 0 + v_j + 0 + \dots + 0$.

D'où $p_i p_j(v) = 0$.

Réciproquement, montrons que si p_1, \dots, p_n sont des projecteurs tels que $\sum_{i=1}^n p_i = \text{id}_E$ et $\forall (i, j) \in \llbracket 1 ; n \rrbracket^2$, $p_i p_j = \delta_{ij} p_i$, alors $E = \bigoplus_{i=1}^n \text{Im } p_i$.

\triangleright Montrons que $\text{Im } p_1 + \dots + \text{Im } p_n$ est directe.

Soit $(v_1, \dots, v_n) \in \text{Im } p_1 \times \dots \times \text{Im } p_n$ telle que $v_1 + \dots + v_n = 0$.

Pour $i \in \llbracket 1 ; n \rrbracket$, on a

$$\begin{aligned} p_i(0) &= 0 = p_i(v_1) + \dots + p_i(v_n) \\ &= p_i p_1(v_1) + \dots + p_i p_i(v_i) + \dots + p_i p_n(v_n) \\ &= v_i. \end{aligned}$$

Donc $v_1 = \dots = v_n = 0$.

Donc la somme est directe.

► On a $\sum_{i=1}^n p_i = \text{id}_E$ donc pour tout $v \in E$, $v = \sum_{i=1}^n p_i(v)$.

Donc $v \in \text{Im } p_1 + \dots + \text{Im } p_n$.

Finalement, $E = \bigoplus_{i=1}^n \text{Im } p_i$. ■

Exercice 4.12

Soit $E = \mathcal{F}(\mathbb{C}, \mathbb{C})$. Pour $k \in \{0, 1, 2\}$, on pose $E_k = \{f \in E \mid \forall x \in \mathbb{C}, f(jx) = j^k f(x)\}$.

Montrez que E_0, E_1, E_2 sont trois sous-espaces vectoriels de E supplémentaires.

Correction 4.13

► Soit $k \in \{0, 1, 2\}$. E_k contient l'application nulle.

Soient $(f, g) \in E_k^2$, $\lambda \in \mathbb{C}$ et $x \in \mathbb{C}$.

On a

$$\begin{aligned} (\lambda f + g)(jx) &= \lambda f(jx) + g(jx) \\ &= \lambda j^k f(x) + j^k g(x) \\ &= j^k (\lambda f(x) + g(x)) \\ &= j^k (\lambda f + g)(x). \end{aligned}$$

Donc $\lambda f + g \in E_k$ i.e. E_k est un sous-espace vectoriel de E .

► Montrons que $\forall f \in E, \exists! (f_0, f_1, f_2) \in E_0 \times E_1 \times E_2, f = f_0 + f_1 + f_2$.

analyse

Soient $f \in E$ et $(f_0, f_1, f_2) \in E_0 \times E_1 \times E_2$ telles que $f = f_0 + f_1 + f_2$.

Pour tout $x \in \mathbb{C}$, on a

$$f(jx) = f_0(jx) + f_1(jx) + f_2(jx) = f_0(x) + j f_1(x) + j^2 f_2(x)$$

et

$$f(j^2x) = f_0(x) + j^2 f_1(x) + j f_2(x).$$

D'où

$$\begin{cases} f_0(x) + f_1(x) + f_2(x) = f(x) \\ f_0(x) + j f_1(x) + j^2 f_2(x) = f(jx) \\ f_0(x) + j^2 f_1(x) + j f_2(x) = f(j^2x) \end{cases}$$

Or $\begin{vmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j^4 \end{vmatrix}$ est un déterminant de Vandermonde et est donc non-nul.

En faisant $L_1 + L_2 + L_3$, on obtient

$$f_0(x) = \frac{1}{3} (f(x) + f(jx) + f(j^2x)).$$

En faisant $L_1 + j^2 L_2 + j L_3$, on obtient

$$f_1(x) = \frac{1}{3} (f(x) + j^2 f(jx) + j f(j^2x)).$$

En faisant $L_1 + j L_2 + j^2 L_3$, on obtient

$$f_2(x) = \frac{1}{3} (f(x) + j f(jx) + j^2 f(j^2x)).$$

Ceci prouve l'unicité de la décomposition (si elle existe).

synthèse

On pose $f_k : x \mapsto \frac{1}{3} (f(x) + j^{2k} f(jx) + j^k f(j^2x))$ pour $k \in \{0, 1, 2\}$.

On a $\forall x \in \mathbb{C}, f(x) = f_0(x) + f_1(x) + f_2(x)$ i.e. $f = f_0 + f_1 + f_2$.

De plus, on a $f_0(jx) = \frac{1}{3} (f(jx) + f(j^2x) + f(j^3x)) = f_0(x)$ car $j^3 = 1$ donc $f_0 \in E_0$.

On montre de même que $f_1 \in E_1$ et $f_2 \in E_2$.

conclusion On a bien $E = \bigoplus_{k=0}^2 E_k$.

Proposition 4.14

Avec les mêmes hypothèses.

Si les sous-espaces F_1, \dots, F_n sont supplémentaires, alors en concaténant des bases de chacun des sous-espaces vectoriels, on obtient une base de E .

4.1.4 Cas particulier de deux sous-espaces

Proposition 4.15

Soient E un \mathbb{K} -espace vectoriel et F, G deux sous-espaces vectoriels de E .

La somme $F + G$ est directe ssi $F \cap G = \{0\}$.

Attention ! Il ne faut pas généraliser à trois ou plus sous-espaces. Même si $F_1 \cap F_2 \cap F_3$ est le sous-espace nul, on ne peut pas conclure que la somme $F_1 + F_2 + F_3$ est directe.

4.1.5 Applications linéaires et sommes directes

Proposition 4.16

Soient E, F deux \mathbb{K} -espaces vectoriels, E_1, \dots, E_n des sous-espaces vectoriels supplémentaires dans E et f_1, \dots, f_n des applications linéaires de E_1, \dots, E_n dans F respectivement.

Alors il existe une unique application linéaire f de E dans F telle que pour tout $i \in \llbracket 1 ; n \rrbracket$, $f|_{E_i} = f_i$.

Autrement dit, pour définir une application linéaire sur une somme directe de sous-espaces vectoriels, il suffit de la définir sur chacun des sous-espaces vectoriels.

Démonstration 4.17

analyse

On suppose que f existe telle que $\forall i \in \llbracket 1 ; n \rrbracket$, $f|_{E_i} = f_i$.

Soit $x \in E$.

Comme $E = \bigoplus_{i=1}^n E_i$, il existe un unique n -uplet $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ tel que $x = \sum_{i=1}^n x_i$.

f étant linéaire, on a $f(x) = \sum_{i=1}^n f(x_i)$.

Or pour tout $i \in \llbracket 1 ; n \rrbracket$, $x_i \in E_i$ et $f|_{E_i} = f_i$ donc $f(x_i) = f_i(x_i)$.

Donc $f(x) = \sum_{i=1}^n f_i(x_i)$.

L'analyse prouve l'unicité de f (si elle existe).

synthèse

Pour $x \in E$ qu'on écrit $\sum_{i=1}^n x_i$ où $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$, on pose $f(x) = \sum_{i=1}^n f_i(x_i)$.

► Montrons que pour tout $i \in \llbracket 1 ; n \rrbracket$, $f|_{E_i} = f_i$.

Pour $i \in \llbracket 1 ; n \rrbracket$, pour $x \in E_i$, on a $x = 0 + \cdots + 0 + x + 0 + \cdots + 0$.

Donc $f(x) = f_1(0) + \cdots + f_{i-1}(0) + f_i(x) + f_{i+1}(0) + \cdots + f_n(0) = f_i(x)$.

Donc $f|_{E_i} = f_i$.

► Montrons que f est linéaire.

Soient $(x, y) \in E^2$ et $\lambda \in \mathbb{K}$.

On écrit $x = \sum_{i=1}^n x_i$ et $y = \sum_{i=1}^n y_i$ où $(x_1, \dots, x_n), (y_1, \dots, y_n) \in E_1 \times \cdots \times E_n$.

Alors

$$\begin{aligned} f(\lambda x + y) &= \sum_{i=1}^n f_i(\lambda x_i + y_i) \\ &= \sum_{i=1}^n (\lambda f_i(x_i) + f_i(y_i)) \\ &= \lambda \sum_{i=1}^n f_i(x_i) + \sum_{i=1}^n f_i(y_i) \\ &= \lambda f(x) + f(y). \end{aligned}$$

■

Exercice 4.18

Soient E_1, \dots, E_n des sous-espaces vectoriels supplémentaires. Quelles sont les applications qui induisent l'application identité sur un des E_i et l'application nulle sur les autres E_j ?

Correction 4.19

On pose $f = \text{id}_{E_i}$ sur E_i et $f = 0$ sur E_j pour $j \neq i$.

f est le projecteur sur E_i parallèlement à $\bigoplus_{j \neq i} E_j$.

4.2 Somme de sous-espaces vectoriels en dimension finie

4.2.1 Base adaptée à un sous-espace

Définition 4.20

Soient E un \mathbb{K} -espace vectoriel de dimension n et F un sous-espace vectoriel de E de dimension p .

On appelle base de E adaptée à F toute base de E qui contient une base de F . Quitte à changer l'ordre des vecteurs, on peut supposer dans une base adaptée à F que les p premiers vecteurs de la base forment une base de F .

Définition 4.21

Soient E un \mathbb{K} -espace vectoriel de dimension n et (F_1, \dots, F_p) une famille de sous-espaces vectoriels supplémentaires dans E .

On appelle base adaptée à la somme $E = \bigoplus_{i=1}^p F_i$ la concaténation de bases de chacun des sous-espaces F_1, \dots, F_p (dans cet ordre).

Si \mathcal{B} est une base de E , alors en fractionnant la base en sous-familles, les sous-espaces vectoriels engendrés par chacune de ces sous-familles sont supplémentaires et la base est alors adaptée à la somme des sous-espaces.

4.2.2 Sommes directes et bases

On donne un moyen simple de vérifier qu'une somme est directe, voire plus.

Proposition 4.22

Soient E un \mathbb{K} -espace vectoriel de dimension n et (F_1, \dots, F_p) une famille de sous-espaces vectoriels de E .

Si en concaténant des bases de chacun des sous-espaces vectoriels F_1, \dots, F_p on obtient une famille libre, alors les sous-espaces vectoriels sont en somme directe.

Si en concaténant des bases de chacun des sous-espaces vectoriels F_1, \dots, F_p on obtient une base de E , alors les sous-espaces vectoriels sont supplémentaires.

4.2.3 Dimension d'une somme de sous-espaces vectoriels

Proposition 4.23

Soient E un \mathbb{K} -espace vectoriel de dimension finie et F_1, \dots, F_n des sous-espaces vectoriels de E .

Alors $\dim \sum_{i=1}^n F_i \leq \sum_{i=1}^n \dim F_i$.

Démonstration 4.24

On note $p_i = \dim F_i$ pour $i \in \llbracket 1 ; n \rrbracket$.

Pour tout $i \in \llbracket 1 ; n \rrbracket$, on choisit une base \mathcal{B}_i de F_i et on a $F_i = \text{Vect}(\mathcal{B}_i)$.

Donc $F_1 + \dots + F_n = \text{Vect}(\mathcal{B}_1 \dots \mathcal{B}_n)$.

Donc

$$\begin{aligned} \dim(F_1 + \dots + F_n) &\leq \text{Card}(\mathcal{B}_1 \dots \mathcal{B}_n) \\ &= \sum_{i=1}^n \dim F_i. \end{aligned}$$

■

Démonstration 4.25 (Autre méthode)

On pose l'application linéaire

$$\begin{aligned} \varphi : F_1 \times \dots \times F_n &\longrightarrow E \\ (x_1, \dots, x_n) &\longmapsto \sum_{i=1}^n x_i \end{aligned}$$

On a $\text{Im } \varphi = F_1 + \dots + F_n$.

Donc

$$\begin{aligned} \dim(F_1 + \dots + F_n) &= \dim \text{Im } \varphi \\ &= \text{rg } \varphi \\ &= \dim(F_1 \times \dots \times F_n) - \dim \ker \varphi \\ &\leq \dim(F_1 \times \dots \times F_n) \\ &= \sum_{i=1}^n \dim F_i. \end{aligned}$$

■

Il y a égalité quand la somme est directe.

Théorème 4.26

Soient E un \mathbb{K} -espace vectoriel de dimension finie et F_1, \dots, F_n des sous-espaces vectoriels de E .

Alors F_1, \dots, F_n sont en somme directe ssi $\dim \sum_{i=1}^n F_i = \sum_{i=1}^n \dim F_i$.

Démonstration 4.27

On reprend la Démonstration 4.25.

Si $\dim \sum_{i=1}^n F_i = \sum_{i=1}^n \dim F_i$, alors $\dim \ker \varphi = 0$.

Donc $\ker \varphi = \left\{ (x_1, \dots, x_n) \in F_1 \times \dots \times F_n \left| \sum_{i=1}^n x_i = 0 \right. \right\} = \{(0, \dots, 0)\}$.

Ainsi, $F_1 + \dots + F_n$ est directe.

Et réciproquement. ■

4.2.4 Sous-espaces supplémentaires

En dimension finie, on a une façon plus simple de prouver que des sous-espaces vectoriels sont supplémentaires.

Proposition 4.28 (Trois pour le prix de deux)

Soient E un \mathbb{K} -espace vectoriel de dimension finie et F_1, \dots, F_n des sous-espaces vectoriels de E .

Quand deux propriétés parmi les trois suivantes sont vraies, alors la troisième l'est aussi :

- ▷ $E = \sum_{i=1}^n F_i$
- ▷ $\dim \sum_{i=1}^n F_i = \sum_{i=1}^n \dim F_i$
- ▷ $\dim E = \sum_{i=1}^n \dim F_i$

Donc dans ce cas, les sous-espaces vectoriels F_1, \dots, F_n sont supplémentaires.

En pratique, le cas le plus utile est le suivant : si $\dim \sum_{i=1}^n F_i = \sum_{i=1}^n \dim F_i = \dim E$, alors les sous-espaces vectoriels F_1, \dots, F_n sont supplémentaires.

Démonstration 4.29

Si $\dim \sum_{i=1}^n F_i = \sum_{i=1}^n \dim F_i$ et $\dim \sum_{i=1}^n F_i = \dim E$.

Alors $F_1 + \dots + F_n$ est directe et $\sum_{i=1}^n F_i = E$.

Alors $E = \bigoplus_{i=1}^n F_i$. ■

Exercice 4.30

Dans \mathbb{R}^4 , soient $H = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + z + t = 0 \text{ et } x - y - z + t = 0\}$, $F = \text{Vect}((-1, 0, 1, 1))$ et $G = \text{Vect}((2, -1, 1, 0))$.

F , G et H sont-ils supplémentaires ?

Correction 4.31

H est l'ensemble des solutions d'un système linéaire homogène :

$$S : \begin{cases} x + z + t = 0 \\ x - y - z + t = 0 \end{cases}$$

i.e. $H = \ker f$ où $f : (x, y, z, t) \mapsto (x + z + t, x - y - z + t)$.

D'après le théorème du rang, on a $\dim H = \dim \mathbb{R}^4 - \text{rg } S = 4 - 2 = 2$.

F et G sont des droites vectorielles donc $\dim F = \dim G = 1$.

On a donc $\dim \mathbb{R}^4 = \dim F + \dim G + \dim H$.

On a

$$\begin{aligned} v = (x, y, z, t) \in H &\iff \begin{cases} x + z + t = 0 \\ x - y - z + t = 0 \end{cases} \\ &\iff \begin{cases} x = -z - t \\ y = -2z \end{cases} \\ &\iff v = (-z - t, -2z, z, t) \\ &\iff v = z \underbrace{(-1, -2, 1, 0)}_{h_1} + t \underbrace{(-1, 0, 0, 1)}_{h_2}. \end{aligned}$$

Donc $H = \text{Vect}(h_1, h_2)$.

On note $f = (-1, 0, 1, 1)$ et $g = (2, -1, 1, 0)$.

On a alors $F + G + H = \text{Vect}(f, g, h_1, h_2)$.

Donc $\dim(F + G + H) = \text{rg}(f, g, h_1, h_2)$.

On note \mathcal{B}_0 la base canonique de \mathbb{R}^4 .

On a

$$A = \text{Mat}_{\mathcal{B}_0}(f, g, h_1, h_2) = \begin{pmatrix} -1 & 2 & -1 & -1 \\ 0 & -1 & -2 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{matrix} L_3 \leftarrow \tilde{L}_3 + L_1 \\ L_4 \leftarrow \tilde{L}_4 + L_1 \end{matrix} \begin{pmatrix} -1 & 2 & -1 & -1 \\ 0 & -1 & -2 & 0 \\ 0 & 3 & 0 & -1 \\ 0 & 2 & -1 & 0 \end{pmatrix}$$

$$\begin{matrix} L_3 \leftarrow \tilde{L}_3 + 3L_2 \\ L_4 \leftarrow \tilde{L}_4 + 2L_2 \end{matrix} \begin{pmatrix} -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & -6 & -1 \\ 0 & 0 & -5 & 0 \end{pmatrix}.$$

Donc $\text{rg } A = 4$.

Donc $\dim(F + G + H) = \dim \mathbb{R}^4 = \dim F + \dim G + \dim H$.

Finalement, on a $\mathbb{R}^4 = F \oplus G \oplus H$.

4.2.5 Dimension d'une somme de deux sous-espaces vectoriels

Rappel : la formule de Grassmann.

Proposition 4.32

Soient E un \mathbb{K} -espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E .

Alors $\dim(F + G) = \dim F + \dim G - \dim(F \cap G)$.

4.3 Polynômes d'endomorphismes et de matrices

4.3.1 \mathbb{K} -algèbres

4.3.1.1 Définition

Définition 4.33

Un ensemble A est appelé \mathbb{K} -algèbre quand A est à la fois un anneau et un \mathbb{K} -espace vectoriel, dont les multiplications sont compatibles.

Il y a donc trois lois dans une \mathbb{K} -algèbre :

- une addition classique $+$;

- une multiplication externe . ;
- une multiplication interne, compatible avec la précédente :

$$\forall (\lambda, a, b) \in \mathbb{K} \times A^2, \quad \lambda.(ab) = (\lambda.a)b = a(\lambda.b).$$

On qualifie les \mathbb{K} -algèbres par du vocabulaire des anneaux (algèbres intègres, algèbres principales, etc) ou des espaces vectoriels (algèbres de dimension finie, etc).

Exemple 4.34

- \mathbb{K} est lui-même une \mathbb{K} -algèbre, où les deux multiplications sont confondues ; \mathbb{C} est aussi une \mathbb{R} -algèbre de dimension 2.
- $\mathbb{K}[X]$ est une \mathbb{K} -algèbre intègre, commutative et de dimension finie.
- Si I est un intervalle, $\mathcal{F}(I, \mathbb{K})$ est une \mathbb{K} -algèbre commutative, non-intègre et de dimension finie.
- Si $n \in \mathbb{N}^*$, $\mathcal{M}_n(\mathbb{K})$ est une \mathbb{K} -algèbre de dimension n^2 qui n'est ni intègre ni commutative.
- Si E est un \mathbb{K} -espace vectoriel, alors $\mathcal{L}(E)$ est une \mathbb{K} -algèbre de dimension finie ssi E l'est aussi qui n'est ni intègre ni commutative.

4.3.1.2 Polynômes d'éléments dans une \mathbb{K} -algèbre

Proposition 4.35

Soient A une \mathbb{K} -algèbre et $a \in A$.

Pour $P = \sum_{i=0}^n c_i X^i \in \mathbb{K}[X]$, on pose $P(a) = \sum_{i=0}^n c_i a^i$.

L'application $\begin{array}{ccc} \mathbb{K}[X] & \longrightarrow & A \\ P & \longmapsto & P(a) \end{array}$ est alors un morphisme d'algèbres :

$$\forall (P, Q) \in \mathbb{K}[X]^2, \quad \forall \lambda \in \mathbb{K}, \quad \begin{cases} (P+Q)(a) = P(a) + Q(a) \\ (PQ)(a) = P(a)Q(a) \\ (\lambda P)(a) = \lambda P(a) \end{cases}$$

De plus, on note $\mathbb{K}[a]$ l'ensemble $\{P(a) \mid P \in \mathbb{K}[X]\}$: cet ensemble est stable par les lois de A , on dit que c'est une sous-algèbre de A .

4.3.2 Cas particulier des algèbres $\mathcal{L}(E)$ ou $\mathcal{M}_n(\mathbb{K})$

E étant un \mathbb{K} -espace vectoriel, l'ensemble $\mathcal{L}(E)$ est une \mathbb{K} -algèbre. De même, n étant un entier non-nul, $\mathcal{M}_n(\mathbb{K})$ est une \mathbb{K} -algèbre. Dans ces algèbres, on définit naturellement la notion de polynôme d'endomorphisme ou de matrice. Bien sûr, ces notions sont liées par choix d'une base de l'espace.

Proposition 4.36

Soient E un \mathbb{K} -espace vectoriel de dimension n , \mathcal{B} une base de E et $f \in \mathcal{L}(E)$ tel que $\text{Mat}_{\mathcal{B}}(f) = A$.

Alors pour tout polynôme $P \in \mathbb{K}[X]$, $P(f)$ a pour matrice $P(A)$ dans la base \mathcal{B} .

Démonstration 4.37

Comme $A = \text{Mat}_{\mathcal{B}}(f)$, pour tout $k \in \mathbb{N}$, $A^k = \text{Mat}_{\mathcal{B}}(f^k)$.

Donc $\forall (a_0, \dots, a_p) \in \mathbb{K}^p$, $\text{Mat}_{\mathcal{B}}\left(\sum_{i=0}^p a_i f^i\right) = \sum_{i=0}^p a_i A^i$ c'est-à-dire

$$\forall P \in \mathbb{K}[X], \text{Mat}_{\mathcal{B}}(P(f)) = P(A).$$

■

Remarque 4.38

- La « multiplication » dans $\mathcal{L}(E)$ est la composition \circ . Donc la deuxième propriété de la Proposition 4.35 doit être comprise comme suit : si $f \in \mathcal{L}(E)$, alors $(PQ)(f) = P(f) \circ Q(f)$.
- Même si les multiplications dans $\mathcal{M}_n(\mathbb{K})$ ou $\mathcal{L}(E)$ ne sont pas commutatives en général, on peut intervertir l'ordre des polynômes car la multiplication dans $\mathbb{K}[X]$ est commutative : si $A \in \mathcal{M}_n(\mathbb{K})$, $(PQ)(A) = P(A)Q(A) = Q(A)P(A)$; si $u \in \mathcal{L}(E)$, $(PQ)(u) = P(u) \circ Q(u) = Q(u) \circ P(u)$.
- Attention aux notations ! Si $f \in \mathcal{L}(E)$, $x \in E$ et $P \in \mathbb{K}[X]$, alors l'application de $P(f)$ au vecteur x se note $P(f)(x)$ et pas $P(f(x))$, notation qui n'a aucun sens.

4.3.3 Polynôme annulateur d'une matrice ou d'un endomorphisme

Définition 4.39

Soient $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{K})$. On appelle polynôme annulateur de A tout polynôme non-nul $P \in \mathbb{K}[X]$ tel que $P(A) = 0$.

Soient E un \mathbb{K} -espace vectoriel et $u \in \mathcal{L}(E)$. On appelle polynôme annulateur de u tout polynôme non-nul $P \in \mathbb{K}[X]$ tel que $P(u) = 0$.

Remarque 4.40

Attention à ne pas confondre les notions : si P est un polynôme annulateur de la matrice A (on dit aussi que P annule A par abus de langage), on ne dit pas que A est une racine de P !

Une racine d'un polynôme est un nombre...

De même, si $P(u) = 0$, on ne dit pas que u est une racine de P , ça n'a aucun sens.

Définition 4.41

Si $A \in \mathcal{M}_n(\mathbb{K})$, alors l'ensemble $\text{Ann}(A) = \{P \in \mathbb{K}[X] \mid P(A) = 0\}$ est appelé idéal annulateur de A .

Si u est un endomorphisme d'un espace vectoriel E , alors l'ensemble $\text{Ann}(u) = \{P \in \mathbb{K}[X] \mid P(u) = 0\}$ est appelé idéal annulateur de u .

Théorème 4.42

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors $\text{Ann}(A)$ est un sous-espace vectoriel de $\mathbb{K}[X]$ stable par \times . De plus, il existe un unique polynôme unitaire μ_A tel que $\text{Ann}(A) = \mu_A \mathbb{K}[X]$.

Soient E un \mathbb{K} -espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$. Alors $\text{Ann}(u)$ est un sous-espace vectoriel de $\mathbb{K}[X]$ stable par \times . De plus, il existe un unique polynôme unitaire μ_u tel que $\text{Ann}(u) = \mu_u \mathbb{K}[X]$.

Démonstration 4.43

► Soient $(P, Q) \in \text{Ann}(A)^2$ et $\lambda \in \mathbb{K}$.

On a

$$(P + Q)(A) = P(A) + Q(A) = 0 + 0 = 0$$

et

$$(PQ)(A) = P(A)Q(A) = 0 \times 0 = 0$$

et

$$(\lambda P)(A) = \lambda P(A) = \lambda \times 0 = 0.$$

Donc $P + Q \in \text{Ann}(A)$, $PQ \in \text{Ann}(A)$ et $\lambda P \in \text{Ann}(A)$.

Remarque : on a même montré que $\forall P \in \mathbb{K}[X], \forall Q \in \text{Ann}(A), PQ \in \text{Ann}(A)$.

► On montre d'abord l'existence d'un polynôme annulateur de A non-nul.

On considère la famille (I_n, A, \dots, A^{n^2}) qui contient $n^2 + 1$ éléments de $\mathcal{M}_n(\mathbb{K})$, espace de dimension n^2 , et est donc liée : il existe $(a_0, \dots, a_{n^2}) \in \mathbb{K}^{n^2+1}$ tel que

$$(a_0, \dots, a_{n^2}) \neq (0, \dots, 0) \quad \text{et} \quad \sum_{i=0}^{n^2} a_i A^i = 0.$$

Alors le polynôme $\sum_{i=0}^{n^2} a_i X^i$ est un polynôme annulateur de A .

► On pose $D = \{\deg P \mid P \in \text{Ann}(A) \setminus \{0\}\}$.

D'après ce qui précède, D est une partie non-vide de \mathbb{N} .

Donc D admet un minimum $d \geq 1$ d'après le principe fondamental de \mathbb{N} , associé à un polynôme $P \neq 0$ tel que $P(A) = 0$.

En notant λ le coefficient dominant de P , le polynôme $\mu = \frac{1}{\lambda}P$ est annulateur de A et unitaire.

► Montrons que $\text{Ann}(A) = \mu\mathbb{K}[X]$.

L'inclusion $\mu\mathbb{K}[X] \subseteq \text{Ann}(A)$ est triviale.

Soit $P \in \text{Ann}(A)$. On effectue la division euclidienne de P par μ : il existe $(Q, R) \in \mathbb{K}[X]^2$ tel que

$$P = Q\mu + R \quad \text{et} \quad \deg R < \deg \mu = d.$$

Donc $P(A) = Q(A)\mu(A) + R(A)$.

Or $P(A) = \mu(A) = 0$ donc $R(A) = 0$ i.e. $R \in \text{Ann}(A)$.

Par définition de d , on a $R = 0$.

Donc P est un multiple de μ .

D'où $\text{Ann}(A) \subseteq \mu\mathbb{K}[X]$.

Donc $\text{Ann}(A) = \mu\mathbb{K}[X]$.

► Montrons maintenant l'unicité de μ .

Si ν est un polynôme annulateur de A de degré d et unitaire, alors $\text{Ann}(A) = \nu\mathbb{K}[X] = \mu\mathbb{K}[X]$.

Donc $\nu \mid \mu$ et $\mu \mid \nu$.

Or μ et ν sont unitaires donc $\mu = \nu$. ■

Remarque 4.44

Ce dernier résultat est faux en dimension infinie. Contre-exemple : l'endomorphisme $u : \mathbb{K}[X] \longrightarrow \mathbb{K}[X]$ défini par $u(P) = XP$.

Définition 4.45

On appelle polynôme minimal d'une matrice carrée A le polynôme μ_A précédent (noté aussi parfois π_A). C'est le polynôme unitaire de degré minimal qui annule A .

On appelle polynôme minimal d'un endomorphisme u en dimension finie le polynôme μ_u précédent (noté aussi parfois π_u). C'est le polynôme unitaire de degré minimal qui annule u .

Autrement dit, on a l'équivalence : pour tout $P \in \mathbb{K}[X]$,

$$P(u) = 0 \iff \mu_u \mid P.$$

De même, on a l'équivalence : pour tout $P \in \mathbb{K}[X]$,

$$P(A) = 0 \iff \mu_A \mid P.$$

Les polynômes annulateurs sont donc les multiples des polynômes minimaux.

On verra plus tard qu'on peut trouver des polynômes annulateurs de plus petits degrés que ceux donnés par le théorème précédent.

En général, il est souvent pénible de calculer le polynôme minimal. En pratique, on se contente de trouver des polynômes annulateurs de degrés pas trop grands (et souvent, il s'agit du polynôme minimal).

Remarque 4.46

Si $A \in \mathcal{M}_n(\mathbb{K})$ et \mathbb{K}' est un sous-corps de \mathbb{C} qui contient \mathbb{K} (on dit que \mathbb{K}' est une extension de \mathbb{K}), alors le polynôme minimal de A , vue comme matrice de $\mathcal{M}_n(\mathbb{K}')$, est a priori différent de celui de A vue comme matrice de $\mathcal{M}_n(\mathbb{K})$. On peut seulement affirmer pour l'instant que $\mu_{A\mathbb{K}'}$ divise $\mu_{A\mathbb{K}}$.

En fait, on montre plus loin que le polynôme minimal ne dépend pas du corps \mathbb{K} .

4.3.4 Utilisation pratique d'un polynôme annulateur

4.3.4.1 Calcul de l'inverse

Proposition 4.47

Soient $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{K})$.

Alors A est inversible ssi A possède un polynôme annulateur P tel que 0 ne soit pas racine de P . Dans ce cas, A^{-1} est un polynôme en A .

De même, soient E un \mathbb{K} -espace vectoriel et $f \in \mathcal{L}(E)$.

Alors si f possède un polynôme annulateur P tel que 0 ne soit pas racine de P , f est un automorphisme de E . Dans ce cas, f^{-1} est un polynôme en f . La réciproque est vraie si E est de dimension finie.

Démonstration 4.48



On suppose que A possède un polynôme annulateur P tel que $P(0) \neq 0$.

En notant $P = \sum_{i=0}^p a_i X^i$, on a $\sum_{i=0}^p a_i A^i = 0$ et $a_0 \neq 0$.

Donc $a_0 I_n + a_1 A + \cdots + a_p A^p = 0$.

Donc $a_0 I_n = -a_1 A - \cdots - a_p A^p$.

$$\text{Donc } I_n = A \underbrace{\left(\frac{-a_1}{a_0} I_n - \cdots - \frac{a_p}{a_0} A^{p-1} \right)}_B.$$

On a trouvé une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = I_n = BA$.

Donc A est inversible et $A^{-1} = B = Q(A)$ où $Q = \sum_{i=0}^{p-1} \frac{-a_{i+1}}{a_0} X^i$.

\Rightarrow

On suppose A inversible. On veut montrer que $\mu_A(0) \neq 0$.

Par l'absurde, on suppose $\mu_A(0) = 0$.

On a $\mu_A = a_1 X + \cdots + X^p$ et $\mu_A(A) = 0$.

Donc $a_1 A + \cdots + A^p = 0$.

A étant inversible, on multiplie par A^{-1} et on obtient $a_1 I_n + a_2 A + \cdots + A^{p-1} = 0$.

Donc $a_1 + a_2 X + \cdots + X^{p-1}$ est annulateur de A et de degré strictement inférieur à celui de μ_A : contradiction.

Donc $\mu_A(0) \neq 0$. ■

Exercice 4.49

On pose $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. Déterminez un polynôme annulateur de A , montrez que A est inversible et calculez A^{-1} .

Correction 4.50

$$\text{On a } A^2 = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix} \text{ et } A^3 = \begin{pmatrix} 1 & 3 & 11 \\ 0 & 1 & 7 \\ 0 & 0 & 8 \end{pmatrix}.$$

Donc on a

$$A^3 = xA^2 + yA + zI_3 \iff \begin{cases} 1 = x + y + z \\ 8 = 4x + 2y + z \\ 7 = 3x + y \\ 11 = 4x + y \\ 3 = 2x + y \end{cases}$$

$$\iff \begin{cases} x = 4 \\ y = -5 \\ z = 2 \end{cases}$$

Donc $A^3 = 4A^2 - 5A + 2I_3$.

Donc $P = X^3 - 4X^2 + 5X - 2$ est un polynôme annulateur de A .

On a $P(0) \neq 0$ donc A est inversible.

On obtient

$$A^{-1} = \frac{1}{2} (A^2 - 4A + 5I_3) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 1/2 \end{pmatrix}.$$

Exercice 4.51

Soient E un \mathbb{K} -espace vectoriel et $p \in \mathcal{L}(E)$ un projecteur. Déterminez un polynôme annulateur de p .

Soit $\lambda \in \mathbb{K}$. On pose $f = p - \lambda \text{id}_E$. Déterminez un polynôme annulateur de f et vérifiez que f est un automorphisme pour presque toutes les valeurs de λ ; dans ce cas, calculez son inverse.

Correction 4.52

p est un projecteur donc $p^2 = p$ donc $X^2 - X$ est un polynôme annulateur de p .

Donc le polynôme annulateur minimal de p est un diviseur unitaire de $X^2 - X : X^2 - X, X$ (auquel cas $p = 0$) ou $X - 1$ (auquel cas $p = \text{id}_E$).

On a $f = p - \lambda \text{id}_E$ donc

$$\begin{aligned} f^2 &= p^2 - 2\lambda p + \lambda^2 \text{id}_E \\ &= (1 - 2\lambda) p + \lambda^2 \text{id}_E \\ &= (1 - 2\lambda) f + \lambda(1 - 2\lambda) \text{id}_E + \lambda^2 \text{id}_E \\ &= (1 - 2\lambda) f + (-\lambda^2 + \lambda) \text{id}_E. \end{aligned}$$

Donc $P = X^2 + (2\lambda - 1)X + (\lambda^2 - \lambda)$ est annulateur de f .

Si $\lambda \neq 0$ et $\lambda \neq 1$ alors $\lambda^2 - \lambda = P(0) = 0$.

Donc d'après la Proposition 4.47, $f \in \text{GL}(E)$.

De plus, $f^{-1} = \frac{1}{\lambda - \lambda^2} (f + (2\lambda - 1) \text{id}_E)$.

Si $\lambda = 0$ alors $f = p$ est un automorphisme ssi $p = \text{id}_E$.

Si $\lambda = 1$ alors $f = p - \text{id}_E$ est un automorphisme ssi $p = 0$.

4.3.4.2 Calcul de puissances

Proposition 4.53

Soient $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{K})$. On choisit un polynôme annulateur P de la matrice A .

Alors pour tout $p \in \mathbb{N}$, $A^p = R_p(A)$ où R_p est le reste de la division euclidienne de X^p par P .

De même, soient E un \mathbb{K} -espace vectoriel et $f \in \mathcal{L}(E)$ qui possède un polynôme annulateur P .

Alors pour tout $p \in \mathbb{N}$, $f^p = R_p(f)$ où R_p est le reste de la division euclidienne de X^p par P .

Démonstration 4.54

On effectue la division euclidienne de X^p par P : $X^p = PQ + R_p$ et $\deg R_p < \deg P$.

Alors $A^p = P(A)Q(A) + R_p(A) = R_p(A)$ car $P(A) = 0$. ■

Conséquence :

- si A possède un polynôme annulateur de degré a , alors $\mathbb{K}[A] = \{P(A) \mid P \in \mathbb{K}_{a-1}[X]\}$;
- si f possède un polynôme annulateur de degré a , alors $\mathbb{K}[f] = \{P(f) \mid P \in \mathbb{K}_{a-1}[X]\}$.

Proposition 4.55

Si p est le degré du polynôme minimal d'une matrice A , alors $\dim \mathbb{K}[A] = p$ et (I_n, A, \dots, A^{p-1}) est une base de $\mathbb{K}[A]$.

Si p est le degré du polynôme minimal d'un endomorphisme f d'un espace vectoriel E , alors $\dim \mathbb{K}[f] = p$ et $(\text{id}_E, f, \dots, f^{p-1})$ est une base de $\mathbb{K}[f]$.

Démonstration 4.56

On note $p = \deg \mu_A$.

Soit $B \in \mathbb{K}[A]$.

Il existe $P \in \mathbb{K}[X]$ tel que $B = P(A)$.

On effectue la division euclidienne de P par $\mu_A : P = Q\mu_A + R$ avec $\deg R < p$.

Donc $P(A) = Q(A)\mu_A(A) + R(A) = R(A) \in \text{Vect}(I_n, A, \dots, A^{p-1})$.

Donc $\mathbb{K}[A] \subseteq \text{Vect}(I_n, A, \dots, A^{p-1})$.

L'inclusion réciproque est immédiate.

Donc $\mathbb{K}[A] = \text{Vect}(I_n, A, \dots, A^{p-1})$.

Si (I_n, A, \dots, A^{p-1}) est liée, alors il existe $(a_0, \dots, a_{p-1}) \neq (0, \dots, 0)$ tel que $\sum_{i=0}^{p-1} a_i A^i = 0$.

En posant $P = \sum_{i=0}^{p-1} a_i X^i$, on a un polynôme annulateur de A de degré strictement inférieur à $p = \deg \mu_A$: contradiction.

Donc (I_n, A, \dots, A^{p-1}) est libre et est donc une base de $\mathbb{K}[A]$ et $\dim \mathbb{K}[A] = p$. ■

Exercice 4.57

Soient E un espace vectoriel et $(p, q) \in \mathcal{L}(E)^2$ deux projecteurs tels que $p + q = \text{id}_E$. Vérifiez que $p \circ q = q \circ p = 0$. Déterminez un polynôme annulateur de $f = 2p + 3q$. Donnez une expression générale de f^k en fonction de f et k .

Correction 4.58

On a $p + q = \text{id}_E$ donc $p \circ p + p \circ q = p$ donc $p \circ q = 0$ car $p \circ p = p$.

De même, $q \circ p = 0$.

On a donc $f^2 = 4p^2 + 6pq + 6qp + 9q^2 = 4p + 9q$.

Donc

$$\begin{aligned} f^2 = xf + y\text{id}_E &\iff \begin{cases} 4 = 2x + y \\ 9 = 3x + y \end{cases} \\ &\iff \begin{cases} 5 = x \\ -6 = y \end{cases} \end{aligned}$$

Donc $P = X^2 - 5X + 6$ est un polynôme annulateur de f dont les racines sont 2 et 3.

Soit $k \in \mathbb{N}$.

On effectue la division euclidienne de X^k par P :

$$X^k = (X^2 - 5X + 6)Q + a_k X + b_k.$$

En appliquant en 2 et en 3, on obtient

$$\begin{cases} 2a_k + b_k = 2^k \\ 3a_k + b_k = 3^k \end{cases} \quad \text{donc} \quad \begin{cases} a_k = 3^k - 2^k \\ b_k = 3 \times 2^k - 2 \times 3^k \end{cases}$$

Donc $f^k = a_k f + b_k \text{id}_E = (3^k - 2^k) f + (3 \times 2^k - 2 \times 3^k) \text{id}_E$.

Remarque 4.59

Plus généralement, si P est annulateur de f et de degré N à N racines simples $\omega_1, \dots, \omega_N$.

Alors $X^k = PQ + c_{N-1}X^{N-1} + \dots + c_1X + c_0$.

Alors on a un système linéaire à N équations et N inconnues

$$\begin{cases} c_0 + c_1\omega_1 + c_2\omega_1^2 + \dots + c_{N-1}\omega_1^{N-1} = \omega_1^k \\ \vdots \\ c_0 + c_1\omega_N + c_2\omega_N^2 + \dots + c_{N-1}\omega_N^{N-1} = \omega_N^k \end{cases}$$

Ce système a pour déterminant

$$\begin{vmatrix} 1 & \omega_1 & \omega_1^2 & \dots & \omega_1^{N-1} \\ & & \vdots & & \vdots \\ 1 & \omega_N & \omega_N^2 & \dots & \omega_N^{N-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (\omega_j - \omega_i) \neq 0$$

et admet donc une unique solution.

Exercice 4.60

On pose $A = \begin{pmatrix} 2 & -4 & -5 \\ -1 & 2 & 2 \\ 1 & -2 & -2 \end{pmatrix}$. Vérifiez que $P = X^3 - 2X^2 + X$ est un polynôme annulateur de A . Donnez une expression générale de A^p en fonction de A et p .

Correction 4.61

On a $A^2 = \begin{pmatrix} 3 & -6 & -8 \\ -2 & 4 & 5 \\ 2 & -4 & -5 \end{pmatrix}$ et $A^3 = \begin{pmatrix} 4 & -8 & -11 \\ -3 & 6 & 8 \\ 3 & -6 & -8 \end{pmatrix}$ donc $P = X^3 - 2X^2 + X = X(X-1)^2$ est annulateur de A .

Soit $k \in \mathbb{N}$.

On effectue la division euclidienne de X^k par P :

$$X^k = X(X-1)^2 Q + a_k X^2 + b_k X + c_k.$$

En évaluant en 0 et 1, on obtient

$$\begin{cases} c_k = 0^k \\ a_k + b_k + c_k = 1^k \\ 2a_k + b_k = k1^k \end{cases}$$

Or 1 est racine double donc $P'(1) = 0$ donc $2a_k + b_k = k$.

Donc, si $k \neq 0$, on a

$$\begin{cases} c_k = 0 \\ a_k + b_k = 1 \\ 2a_k + b_k = k \end{cases} \quad \text{donc} \quad \begin{cases} a_k = k - 1 \\ b_k = 2 - k \\ c_k = 0 \end{cases}$$

Donc pour tout $k \in \mathbb{N}^*$, $A^k = (k - 1) A^2 + (2 - k) A$.

Corollaire 4.62

Si \mathbb{K}' est une extension de \mathbb{K} , alors pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, ses polynômes minimaux relativement à \mathbb{K} et \mathbb{K}' sont égaux : $\mu_{A\mathbb{K}} = \mu_{A\mathbb{K}'}$.

Autrement dit, le polynôme minimal ne dépend pas du corps \mathbb{K} .

Démonstration 4.63

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

On note $\text{Ann}_{\mathbb{K}}(A) = \{P \in \mathbb{K}[X] \mid P(A) = 0\}$ et $\text{Ann}_{\mathbb{K}'}(A) = \{P \in \mathbb{K}'[X] \mid P(A) = 0\}$.

Soit $P \in \text{Ann}_{\mathbb{K}}(A)$.

Comme $\mathbb{K} \subseteq \mathbb{K}'$, on a $P \in \text{Ann}_{\mathbb{K}'}(A)$.

Donc $\mu_{A\mathbb{K}}[X] = \text{Ann}_{\mathbb{K}}(A) \subseteq \text{Ann}_{\mathbb{K}'}(A) = \mu_{A\mathbb{K}'}[X]$.

Donc $\mu_{A\mathbb{K}} \in \text{Ann}_{\mathbb{K}'}(A)$ donc $\mu_{A\mathbb{K}'} \mid \mu_{A\mathbb{K}}$.

Donc $p' = \deg \mu_{A\mathbb{K}'} \leq \deg \mu_{A\mathbb{K}} = p$.

Or (I_n, A, \dots, A^{p-1}) est libre dans le \mathbb{K} -espace vectoriel $\mathbb{K}[A]$ et $(I_n, A, \dots, A^{p'-1})$ est libre dans le \mathbb{K}' -espace vectoriel $\mathbb{K}'[A]$.

On veut montrer que (I_n, A, \dots, A^{p-1}) est libre dans le \mathbb{K}' -espace vectoriel $\mathbb{K}'[A]$.

Par l'absurde, si cette famille est liée, alors il existe $(\lambda_0, \dots, \lambda_{p-1}) \in (\mathbb{K}')^p$ tel que

$$\begin{cases} (\lambda_0, \dots, \lambda_{p-1}) \neq (0, \dots, 0) \\ \sum_{i=0}^{p-1} \lambda_i A^i = 0 \end{cases}$$

On pose $q = \max \{i \in \llbracket 0 ; p-1 \rrbracket \mid \lambda_i \neq 0\}$.

On a alors $A^q = - \sum_{i=0}^{q-1} \frac{\lambda_i}{\lambda_q} A^i$.

Cette égalité est interprétable en un système linéaire à n^2 équations dont les « inconnues » sont $\frac{-\lambda_i}{\lambda_q}$ et les coefficients sont les coefficients de A^q, \dots, I_n qui sont tous dans \mathbb{K} .

En résolvant ce système linéaire, on obtient des solutions qui sont dans \mathbb{K} .

Donc on obtient une relation de dépendance linéaire entre les matrices $(I_n, \dots, A^q, \dots, A^{p-1})$ à coefficients dans \mathbb{K} : contradiction car (I_n, \dots, A^{p-1}) est libre.

Donc $p \leq \dim_{\mathbb{K}'} \mathbb{K}'[A] = p'$.

D'où $p = p'$.

Donc, comme $\mu_{A\mathbb{K}'} \mid \mu_{A\mathbb{K}}$ et que ces polynômes sont unitaires, on a

$$\mu_{A\mathbb{K}'} = \mu_{A\mathbb{K}}. \quad \blacksquare$$

4.4 Matrices semblables, trace

4.4.1 Trace d'une matrice

Définition 4.64

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On appelle trace de A la somme de ses coefficients diagonaux :

$$\text{tr } A = \sum_{i=1}^n a_{ii}.$$

L'application trace vérifie de remarquables propriétés.

Proposition 4.65

- La trace est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$.
- Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, $\text{tr}(A^\top) = \text{tr } A$.
- Pour tout $(A, B) \in \mathcal{M}_n(\mathbb{K})^2$, $\text{tr}(AB) = \text{tr}(BA)$.

4.4.2 Matrices semblables

Définition 4.66

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$.

On dit que A et B sont semblables quand il existe $P \in \text{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Proposition 4.67

La relation de similitude entre matrices de $\mathcal{M}_n(\mathbb{K})$ est une relation d'équivalence.

La relation de similitude est une relation très contraignante. Il n'existe pas de caractérisation simple de la similitude entre deux matrices carrées : savoir si deux matrices sont semblables est un problème difficile.

D'après la formule de changement de base, on a immédiatement le résultat suivant.

Proposition 4.68

Deux matrices de $\mathcal{M}_n(\mathbb{K})$ sont semblables ssi elle représentent un même endomorphisme dans des bases différentes. La matrice P est la matrice de passage d'une base à l'autre.

Corollaire 4.69

Deux matrices semblables ont même rang, même trace et même déterminant.

Mais c'est loin d'être suffisant pour être semblables.

Proposition 4.70

Si A et B sont deux matrices semblables, alors pour tout polynôme $P \in \mathbb{K}[X]$, $P(A)$ et $P(B)$ sont semblables avec la même matrice de passage.

Démonstration 4.71

Si A et B sont semblables, il existe $Q \in \text{GL}_n(\mathbb{K})$ telle que $B = Q^{-1}AQ$.

Alors, par récurrence sur $k \in \mathbb{N}$, on a $B^k = Q^{-1}A^kQ$.

Puis, par combinaison linéaire, pour tout $P \in \mathbb{K}[X]$, $P(B) = Q^{-1}P(A)Q$. ■

4.4.3 Trace d'un endomorphisme

Proposition 4.72

Soient E un \mathbb{K} -espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$.

Toutes les matrices carrées représentant f ont la même trace. Cette trace ne dépend donc pas du choix de la base dans laquelle on écrit la matrice de f , elle ne dépend que de f : on l'appelle trace de f et on la note $\text{tr } f$.

On peut alors reformuler les résultats sur la trace d'une matrice.

Proposition 4.73

► *La trace est une forme linéaire sur $\mathcal{L}(E)$.*

► *Pour tout $(u, v) \in \mathcal{L}(E)^2$, $\text{tr}(u \circ v) = \text{tr}(v \circ u)$.*

4.5 Opérations par blocs

4.5.1 Cas général

Soit $(n, p) \in (\mathbb{N}^*)^2$. On fixe deux entiers k, ℓ tels que $1 \leq k \leq n-1$ et $1 \leq \ell \leq p-1$.

À toute matrice $M \in \mathcal{M}_{np}(\mathbb{K})$, on associe quatre matrices obtenues en découpant la matrice en blocs :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

où $A = (m_{ij})_{\substack{1 \leq i \leq k \\ 1 \leq j \leq \ell}}$, $B = (m_{ij})_{\substack{1 \leq i \leq k \\ \ell+1 \leq j \leq p}}$, $C = (m_{ij})_{\substack{k+1 \leq i \leq n \\ 1 \leq j \leq \ell}}$ et $D = (m_{ij})_{\substack{k+1 \leq i \leq n \\ \ell+1 \leq j \leq p}}$.

Cette décomposition par blocs permet de faire des calculs formellement comme s'il s'agissait de nombres.

Proposition 4.74

Soient $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ et $M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$ deux matrices de même taille décomposées de la même façon en blocs et $\lambda \in \mathbb{K}$.

Alors $M + M' = \begin{pmatrix} A + A' & B + B' \\ C + C' & D + D' \end{pmatrix}$ et $\lambda M = \begin{pmatrix} \lambda A & \lambda B \\ \lambda C & \lambda D \end{pmatrix}$.

Proposition 4.75

Soient $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ et $M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$ deux matrices telles que le produit MM' existe et décomposées en blocs.

Alors, sous réserve que les blocs soient de tailles compatibles pour la multiplication, on a

$$MM' = \begin{pmatrix} AA' + BC' & AB' + BD' \\ CA' + DC' & CB' + DD' \end{pmatrix}.$$

Remarque 4.76

- Comme les symboles mis en jeu ne sont pas des nombres mais des matrices, il est indispensable de respecter l'ordre dans les produits.
- On peut généraliser à un nombre quelconque de blocs, pas forcément deux en ligne ou en colonne.

4.5.2 Cas particuliers des matrices carrées

Si M est une matrice de $\mathcal{M}_n(\mathbb{K})$, alors avec les mêmes notations, on choisit toujours A et D carrées elles aussi. Dans ce paragraphe, on suppose que c'est le cas.

Définition 4.77

On dit que M est triangulaire supérieure par blocs quand il existe des matrices carrées A_1, \dots, A_k telles que M soit de la forme

$$M = \begin{pmatrix} A_1 & ? & ? & \dots & ? \\ 0 & A_2 & ? & \dots & ? \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & A_{k-1} & ? \\ 0 & \dots & 0 & 0 & A_k \end{pmatrix}.$$

On définit de même la notion de matrice triangulaire inférieure par blocs.

Définition 4.78

On dit que M est diagonale par blocs quand il existe des matrices carrées A_1, \dots, A_k telles que M soit de la forme

$$M = \begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & A_{k-1} & 0 \\ 0 & \dots & 0 & 0 & A_k \end{pmatrix}.$$

Les résultats sur les matrices triangulaires ou diagonales restent valables par blocs : la somme et le produit de deux matrices triangulaires supérieures par blocs de mêmes tailles l'est encore, et de même pour les matrices triangulaires inférieures par blocs et les matrices diagonales par blocs.

Une conséquence est qu'une matrice M triangulaire par blocs est inversible ssi tous les blocs diagonaux sont inversibles.

Dans ce cas, l'inverse de M est triangulaire par blocs et ses blocs diagonaux sont les inverses des blocs diagonaux de M .

En particulier, l'inverse d'une matrice M diagonale par blocs est la matrice diagonale par blocs dont les blocs diagonaux sont les inverses de ceux de M .

De plus, le déterminant d'une matrice triangulaire par blocs est le produit des déterminants des blocs diagonaux.

4.5.3 Interprétation des blocs

Définition 4.79

Soient E un \mathbb{K} -espace vectoriel, $f \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E .

On dit que F est stable par f quand $f(F) \subseteq F$, *i.e.* pour tout $x \in F$, $f(x) \in F$.

Dans ce cas, on peut définir une application φ de F dans F en posant pour tout $x \in F$, $\varphi(x) = f(x)$.

Il est facile de vérifier que φ est un endomorphisme de F , appelé endomorphisme induit par f dans F .

Exemple 4.80

Si g est un endomorphisme de E qui commute avec f (i.e. $fg = gf$), alors $\ker g$ et $\text{Im } g$ sont stables par f .

Proposition 4.81

Soient E un \mathbb{K} -espace vectoriel de dimension n , $f \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E de dimension p .

Si F est stable par f , alors il existe une base de E dans laquelle la matrice de f est triangulaire supérieure par bloc, le premier bloc étant de taille (p, p) :

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \text{ et } A \in \mathcal{M}_p(\mathbb{K}).$$

Réciproquement, si f possède une matrice de cette forme, alors le sous-espace vectoriel engendré par les p premiers vecteurs est stable par f .

Démonstration 4.82

On choisit une base de E adaptée à F : $\mathcal{B} = \left(\underbrace{e_1, \dots, e_p}_{\text{base de } F}, e_{p+1}, \dots, e_n \right)$.

Pour tout $j \in \llbracket 1 ; p \rrbracket$, en notant $f(e_j) = (a_{1j}, \dots, a_{nj})_{\mathcal{B}}$, on a

$$\begin{aligned} f(e_j) \in F &\iff f(e_j) \in \text{Vect}(e_1, \dots, e_p) \\ &\iff \forall i \geq p+1, a_{ij} = 0 \\ &\iff f(e_j) = (a_{1j}, \dots, a_{pj}, 0, \dots, 0)_{\mathcal{B}}. \end{aligned}$$

Alors, si F est stable par f , on a

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} a_{11} & \dots & a_{1p} & ? & \dots & ? \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p1} & \dots & a_{pp} & \vdots & & \vdots \\ 0 & \dots & 0 & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & ? & \dots & ? \end{pmatrix}.$$

Et réciproquement. ■

Proposition 4.83

Soient E un \mathbb{K} -espace vectoriel de dimension n et $f \in \mathcal{L}(E)$.

Si F_1, \dots, F_k sont des sous-espaces vectoriels supplémentaires stables par f de dimensions respectives p_1, \dots, p_k , alors il existe une base de E dans laquelle la matrice de f est diagonale par blocs, la taille du i -ème bloc étant (p_i, p_i) :

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & A_{k-1} & 0 \\ 0 & \dots & 0 & 0 & A_k \end{pmatrix}.$$

Réciproquement, si f possède une matrice dans une certaine base qui est diagonale par blocs et contenant k blocs carrés, alors il existe k sous-espaces vectoriels F_1, \dots, F_k stables par f et supplémentaires dans E .

Chapitre 5

Réduction des endomorphismes

Sommaire

5.1	Éléments propres d'un endomorphisme	163
5.1.1	Valeurs propres et vecteurs propres	163
5.1.2	Lien avec les polynômes annulateurs	164
5.1.3	Sous-espaces propres	166
5.2	Polynôme caractéristique d'un endomorphisme	168
5.2.1	Caractérisation des valeurs propres en dimension finie	168
5.2.2	Définition et lien avec les valeurs propres	168
5.2.3	Ordre de multiplicité et dimension du sous-espace propre	171
5.2.4	Endomorphisme scindé	172
5.3	Éléments propres d'une matrice carrée.	173
5.3.1	Valeurs propres et vecteurs propres	173
5.3.2	Lien avec les polynômes annulateurs	174
5.3.3	Sous-espaces propres	174
5.4	Polynôme caractéristique d'une matrice carrée	175
5.4.1	Définition et lien avec les valeurs propres	175
5.4.2	Ordre de multiplicité et dimension du sous-espace propre	177
5.4.3	Matrice scindée	177
5.5	Endomorphismes diagonalisables, matrices diagonalisables	178
5.5.1	Définition	178
5.5.2	Caractérisations équivalentes	179
5.5.3	Lien avec le polynôme caractéristique	181
5.6	Lien entre diagonalisabilité et polynômes annulateurs	182
5.6.1	Racines du polynôme minimal	182
5.6.2	Lemme des noyaux	183
5.6.3	Application à la diagonalisabilité	185
5.6.4	Diagonalisabilité d'un endomorphisme induit	188
5.7	Quelques applications de la diagonalisation	189
5.7.1	Puissances d'une matrice, suites récurrentes linéairement	189
5.7.2	Systèmes d'équations différentielles	190
5.8	Endomorphismes trigonalisables, matrices trigonalisables	190
5.8.1	Définition et propriétés	190
5.8.2	Caractérisation équivalente	191
5.8.3	Théorème de Cayley-Hamilton	194
5.8.4	Sous-espaces caractéristiques	195

5.9	Endomorphismes nilpotents, matrices nilpotentes	197
5.9.1	Généralités	197
5.9.2	Éléments propres d'un nilpotent	199
5.9.3	Application aux sous-espaces caractéristiques d'un endomorphisme	199

Dans ce chapitre, \mathbb{K} désigne un sous-corps de \mathbb{C} , en général \mathbb{R} ou \mathbb{C} .

5.1 Éléments propres d'un endomorphisme

Dans cette section, E est un \mathbb{K} -espace vectoriel de dimension quelconque, finie ou non.

5.1.1 Valeurs propres et vecteurs propres

Définition 5.1

Soient $f \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$.

On dit que λ est une valeur propre de f quand il existe un vecteur v non-nul tel que $f(v) = \lambda v$.

Si λ est une valeur propre de f , alors tout vecteur non-nul v tel que $f(v) = \lambda v$ est appelé vecteur propre associé à la valeur propre λ .

Remarque 5.2

Si $f(v) = \lambda v$ et $v \neq 0$ alors pour tout $\alpha \neq 0$, $f(\alpha v) = \alpha f(v) = \alpha(\lambda v) = \lambda(\alpha v)$. Donc αv est un vecteur propre de f pour la valeur propre λ .

Exemple 5.3

- Pour tout $\alpha \in \mathbb{K}$, αid_E a pour unique valeur propre α et tout vecteur non-nul de E est un vecteur propre associé.
- Si p est un projecteur non-trivial (*i.e.* $p \neq 0$ et $p \neq \text{id}_E$), alors p a pour seules valeurs propres 0 et 1.
- De même, si s est une symétrie non-triviale (*i.e.* $s \neq \text{id}_E$ et $s \neq -\text{id}_E$), alors les valeurs propres de s sont 1 et -1 .
- L'endomorphisme de $\mathbb{K}[X]$ $P \mapsto XP$ n'a pas de valeur propre.

L'ensemble des valeurs propres d'un endomorphisme f est appelé le spectre de f et est noté $\text{Sp}_{\mathbb{K}}(f)$ ou plus simplement $\text{Sp}(f)$ (en toute rigueur, cette définition est fautive en dimension infinie, mais à notre niveau, cette approximation est acceptable).

Définition 5.4

On appelle droite propre d'un endomorphisme toute droite dirigée par un vecteur propre.

Proposition 5.5

Les droites propres d'un endomorphisme sont exactement les droites stables par cet endomorphisme.

Exercice 5.6

Soit $f \in \mathcal{L}(\mathbb{R}^{\mathbb{N}})$ défini par : si $(u_n) \in \mathbb{R}^{\mathbb{N}}$, on pose $f(u) = (u_{n+1})$. Quelles sont les valeurs propres de f et les vecteurs propres associés ?

Exercice 5.7

Même question avec d l'opérateur de dérivation dans $\mathcal{C}^{\infty}(\mathbb{R}, \mathbb{R})$.

Exercice 5.8

Même question avec D l'opérateur de dérivation dans $\mathbb{R}[X]$.

5.1.2 Lien avec les polynômes annulateurs

En dimension quelconque, il est souvent difficile de trouver les valeurs propres d'un endomorphisme. La connaissance d'un polynôme annulateur peut aider.

Lemme 5.9

Soient $f \in \mathcal{L}(E)$ et $P \in \mathbb{K}[X]$. Si λ est une valeur propre de f et v un vecteur propre associé, alors $P(f)(v) = P(\lambda)v$.

Démonstration 5.10

On montre par récurrence la propriété $\mathcal{P}(k)$: « $f^k(v) = \lambda^k v$ ».

On a $f^0(v) = v = \lambda^0 v$.

Si $\mathcal{P}(k)$ est vraie, alors

$$\begin{aligned} f^{k+1}(v) &= f f^k(v) \\ &= f(\lambda^k v) \\ &= \lambda^k f(v) \\ &= \lambda^k \lambda v \\ &= \lambda^{k+1} v. \end{aligned}$$

D'où $\mathcal{P}(k+1)$.

Par récurrence, pour tout $k \in \mathbb{N}$, $\mathcal{P}(k)$ est vraie.

On écrit $P = \sum_{i=0}^n a_i X^i$.

Alors $P(f) = \sum_{i=0}^n a_i f^i$.

Donc

$$\begin{aligned} P(f)(v) &= \sum_{i=0}^n a_i f^i(v) \\ &= \sum_{i=0}^n a_i (\lambda^i v) \\ &= v \sum_{i=0}^n a_i \lambda^i \\ &= P(\lambda) v. \end{aligned}$$

■

Si $P \in \mathbb{K}[X]$, on note $Z_{\mathbb{K}}(P)$ l'ensemble des racines de P dans \mathbb{K} .

Proposition 5.11

Soit $f \in \mathcal{L}(E)$.

Si P est un polynôme annulateur de f , alors $\text{Sp}(f) \subseteq Z_{\mathbb{K}}(P)$.

Démonstration 5.12

Il existe $v \neq 0$ tel que $f(v) = \lambda v$.

D'après le lemme précédent, $P(f)(v) = P(\lambda) v$.

Or $P(f) = 0$ donc $P(\lambda) v = 0$.

Or $v \neq 0$ donc $P(\lambda) = 0$.

Donc $\lambda \in Z_{\mathbb{K}}(P)$.

■

Remarque 5.13

Attention ! La réciproque est fautive. Contre-exemple : le polynôme $P = X^2 - 1$ est annulateur de id_E et pourtant -1 , qui est racine de P , n'est pas valeur propre de id_E .

Exercice 5.14

Soit $n \geq 2$. Pour $M \in \mathcal{M}_n(\mathbb{K})$, on pose $f(M) = M + M^T + \text{tr}(M) I_n$: f est clairement un endomorphisme de $\mathcal{M}_n(\mathbb{K})$.

Déterminez un polynôme annulateur de f de degré 3 et déduisez-en les valeurs propres de f .

5.1.3 Sous-espaces propres

Proposition 5.15

Soient $f \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$.

Alors λ est valeur propre de f ssi $\ker(f - \lambda \text{id}_E) \neq \{0\}$, autrement dit ssi $f - \lambda \text{id}_E$ n'est pas injectif.

Démonstration 5.16

On a

$$\begin{aligned}\lambda \in \text{Sp}(f) &\iff \exists v \in E, v \neq 0 \text{ et } f(v) = \lambda v \\ &\iff \exists v \in E, v \neq 0 \text{ et } f(v) - \lambda v = 0 \\ &\iff \exists v \in E, v \neq 0 \text{ et } (f - \lambda \text{id}_E)(v) = 0 \\ &\iff \exists v \in E, v \neq 0 \text{ et } v \in \ker(f - \lambda \text{id}_E) \\ &\iff \ker(f - \lambda \text{id}_E) \neq \{0\} \\ &\iff f - \lambda \text{id}_E \text{ non-injective.} \quad \blacksquare\end{aligned}$$

Définition 5.17

Soit $f \in \mathcal{L}(E)$.

Si $\lambda \in \text{Sp}(f)$, le noyau $\ker(f - \lambda \text{id}_E)$ est appelé le sous-espace propre associé à la valeur propre λ . Il est souvent noté $\text{sep}(f, \lambda)$.

Par conséquent, $\text{sep}(f, \lambda)$ est l'ensemble des vecteurs propres associés à la valeur propre λ auquel on ajoute le vecteur nul.

Remarque 5.18

Un cas particulier important : 0 est valeur propre ssi f n'est pas injective.

Exercice 5.19

Soit u un endomorphisme ayant pour matrice $M = \begin{pmatrix} -3 & 4 & -4 \\ 4 & -3 & 3 \\ 4 & -4 & 4 \end{pmatrix}$ dans une certaine base \mathcal{B} .

Calculez $M^3 + 2M^2 - 3M$. Déduisez-en les valeurs propres de u puis déterminez les sous-espaces propres associés.

Proposition 5.20

Tout sous-espace propre d'un endomorphisme est stable par cet endomorphisme. L'endomorphisme induit sur un sous-espace propre est alors une homothétie.

Démonstration 5.21

Soit $v \in \text{sep}(f, \lambda)$.

On a $f(v) = \lambda v$.

Donc $f(f(v)) = \lambda f(v)$.

Donc $f(v) \in \text{sep}(f, \lambda)$.

Donc le sous-espace propre $\text{sep}(f, \lambda)$ est stable par f .

De plus, l'endomorphisme induit par f sur ce sous-espace est

$$\begin{array}{ccc} \text{sep}(f, \lambda) & \longrightarrow & \text{sep}(f, \lambda) \\ v & \longmapsto & f(v) = \lambda v \end{array}$$

i.e. l'homothétie de rapport λ . ■

Théorème 5.22

Soient $f \in \mathcal{L}(E)$ et $\lambda_1, \dots, \lambda_p$ des valeurs propres distinctes de f .

Alors les sous-espaces propres $(\text{sep}(f, \lambda_i))_{1 \leq i \leq p}$ sont en somme directe.

Autrement dit, toute famille de vecteurs propres associés à des valeurs propres distinctes est libre.

Démonstration 5.23

Soit $(v_1, \dots, v_p) \in \prod_{i=1}^p \text{sep}(f, \lambda_i)$ tel que $v_1 + \dots + v_p = 0$ (1).

On veut montrer que $v_1 = \dots = v_p = 0$.

On applique f à (1) : $f(v_1) + \dots + f(v_p) = 0$ i.e. $\lambda_1 v_1 + \dots + \lambda_p v_p = 0$.

On réitère $p - 2$ fois et on obtient le système suivant :

$$\left\{ \begin{array}{l} v_1 + \dots + v_p = 0 \\ \lambda_1 v_1 + \dots + \lambda_p v_p = 0 \\ \lambda_1^2 v_1 + \dots + \lambda_p^2 v_p = 0 \\ \vdots \\ \lambda_1^{p-1} v_1 + \dots + \lambda_p^{p-1} v_p = 0 \end{array} \right.$$

La matrice de ce système linéaire est

$$\begin{pmatrix} 1 & \dots & \dots & 1 \\ \lambda_1 & \dots & \dots & \lambda_p \\ \lambda_1^2 & \dots & \dots & \lambda_p^2 \\ \vdots & & & \vdots \\ \lambda_1^{p-1} & \dots & \dots & \lambda_p^{p-1} \end{pmatrix}$$

i.e. une matrice de Vandermonde inversible car les λ_i sont distincts donc le système a une unique solution $(v_1, \dots, v_p) = (0, \dots, 0)$. ■

Remarque 5.24

Quand on demande de déterminer les éléments propres d'un endomorphisme, on demande de déterminer les valeurs propres et les vecteurs propres associés, *i.e.* les sous-espaces propres.

À partir de maintenant, il est toujours supposé que E est de dimension finie n

5.2 Polynôme caractéristique d'un endomorphisme

5.2.1 Caractérisation des valeurs propres en dimension finie

Proposition 5.25

Soient $f \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. Alors

$$\lambda \in \text{Sp}(f) \iff \text{rg}(f - \lambda \text{id}_E) < n.$$

Dans ce cas, $\dim \text{sep}(f, \lambda) = n - \text{rg}(f - \lambda \text{id}_E)$.

Démonstration 5.26

D'après le théorème du rang, on a

$$n = \underbrace{\dim \ker(f - \lambda \text{id}_E)}_{=\dim \text{sep}(f, \lambda)} + \text{rg}(f - \lambda \text{id}_E).$$

Donc $\dim \text{sep}(f, \lambda) = n - \text{rg}(f - \lambda \text{id}_E)$.

On obtient l'inégalité voulue grâce à la Proposition 5.15. ■

5.2.2 Définition et lien avec les valeurs propres

Définition 5.27

Soit $f \in \mathcal{L}(E)$.

On appelle polynôme caractéristique de f le polynôme $\chi_f = \det(X \text{id}_E - f)$.

La notation χ_f est très courante : elle est à connaître.

Théorème 5.28

Soit $f \in \mathcal{L}(E)$.

Alors χ_f est un polynôme unitaire de degré n de $\mathbb{K}[X]$ et les valeurs propres de f sont exactement les racines dans \mathbb{K} de $\chi_f : Z_{\mathbb{K}}(\chi_f) = \text{Sp}(f)$.

Par conséquent, un endomorphisme d'un espace de dimension n a au plus n valeurs propres distinctes.

Démonstration 5.29

On choisit une base \mathcal{B} de E et on pose $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ la matrice de f dans \mathcal{B} .

On a

$$\begin{aligned}
\chi_f &= \det(\text{Id}_E - f) \\
&= \det(XI_n - A) \\
&= \begin{vmatrix} X - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & X - a_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & -a_{n-1n} \\ -a_{n1} & \dots & -a_{nn-1} & X - a_{nn} \end{vmatrix}
\end{aligned}$$

On pose $c_{ij} = \begin{cases} -a_{ij} & \text{si } i \neq j \\ X - a_{ii} & \text{sinon} \end{cases}$

Alors

$$\begin{aligned}
\chi_f &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^n c_{i\sigma(i)} \\
&= \varepsilon(\text{id}) \prod_{i=1}^n c_{ii} + \sum_{\sigma \in \mathfrak{S}_n \setminus \{\text{id}\}} \varepsilon(\sigma) \prod_{i=1}^n c_{i\sigma(i)}
\end{aligned}$$

On remarque que si $\sigma \in \mathfrak{S}_n$ alors σ a n points fixes si $\sigma = \text{id}$ et σ a moins de $n - 2$ points fixes sinon donc si $\sigma \neq \text{id}$, il existe au moins deux entiers $i, j \in \llbracket 1 ; n \rrbracket$ tels que $\sigma(i) \neq i$ et $\sigma(j) \neq j$.

Donc pour toute permutation $\sigma \neq \text{id}$, parmi les facteurs du produit $\prod_{i=1}^n a_{i\sigma(i)}$, il en existe au moins

deux qui sont de la forme $a_{??}$ donc $\prod_{i=1}^n c_{i\sigma(i)}$ est un polynôme de degré au plus $n - 2$.

Donc $\deg \chi_f = n$ et χ_f est unitaire.

De plus, on a

$$\begin{aligned}
\lambda \in \text{Sp}(f) &\iff f - \lambda \text{id}_E \text{ n'est pas injectif} \\
&\iff f - \lambda \text{id}_E \text{ n'est pas bijectif} \\
&\iff \det(f - \lambda \text{id}_E) = 0 \\
&\iff \det(\lambda \text{id}_E - f) = 0 \\
&\iff \chi_f(\lambda) = 0.
\end{aligned}$$

} dimension finie

■

Exercice 5.30

Montrez que si $\dim E = 2$, alors pour tout $f \in \mathcal{L}(E)$, $\chi_f = X^2 - \operatorname{tr}(f)X + \det f$.

Exercice 5.31

Calculez le polynôme caractéristique d'un endomorphisme de matrice $\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$ et donnez ses valeurs propres.

Exercice 5.32

Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , $s = \sum_{i=1}^n e_i$ et $f \in \mathcal{L}(E)$ tel que pour tout $j \in \llbracket 1 ; n \rrbracket$, $f(e_j) = e_j + s$.

Calculez son polynôme caractéristique et ses éléments propres.

On peut noter un lien avec la trace et le déterminant.

Proposition 5.33

Soit $f \in \mathcal{L}(E)$.

Alors $\chi_f = X^n - \operatorname{tr}(f)X^{n-1} + \dots + (-1)^n \det f$.

Démonstration 5.34

► On a $\chi_f = \det(X\operatorname{Id}_E - f)$ donc

$$\chi_f(0) = \det(-f) = (-1)^n \det f$$

est le coefficient constant de χ_f .

► On avait $\chi_f = \prod_{i=1}^n (X - a_{ii}) + Q$ avec $\deg Q \leq n - 2$ (cf. Démonstration 5.29).

Donc le coefficient d'indice $n - 1$ est celui du produit $\prod_{i=1}^n (X - a_{ii})$.

Or on a

$$\begin{aligned} (X - a_{11}) \dots (X - a_{nn}) &= X^n + (-a_{11} - \dots - a_{nn})X^{n-1} + \dots \\ &= X^n - \operatorname{tr}(f)X^{n-1} + \dots \end{aligned}$$

■

5.2.3 Ordre de multiplicité et dimension du sous-espace propre

Définition 5.35

Soient $f \in \mathcal{L}(E)$ et $\lambda \in \text{Sp}(f)$.

On appelle ordre de multiplicité de la valeur propre λ son ordre de multiplicité en tant que racine de χ_f .

Proposition 5.36

Soient $f \in \mathcal{L}(E)$, F un sous-espace vectoriel de E stable par f et g l'endomorphisme induit par f dans F .

Alors χ_g divise χ_f .

Démonstration 5.37

On choisit une base $\mathcal{B} = (e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ de E adaptée à F : (e_1, \dots, e_p) est une base de F .

Alors

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} A & ? \\ 0 & B \end{pmatrix}$$

où $A \in \mathcal{M}_p(\mathbb{K})$ et $B \in \mathcal{M}_{n-p}(\mathbb{K})$.

On remarque que $A = \text{Mat}_{(e_1, \dots, e_p)}(g)$.

Alors

$$\begin{aligned} \chi_f &= \begin{vmatrix} XI_p - A & -? \\ 0 & XI_{n-p} - B \end{vmatrix} \\ &= \underbrace{\det(XI_p - A)}_{=\chi_g} \det(XI_{n-p} - B). \end{aligned}$$

Donc $\chi_g \mid \chi_f$. ■

Une conséquence très importante de ce résultat est le théorème suivant.

Théorème 5.38

Soient $f \in \mathcal{L}(E)$ et $\lambda \in \text{Sp}(f)$.

Si λ est une valeur propre d'ordre α , alors $1 \leq \dim \text{sep}(f, \lambda) \leq \alpha$.

Démonstration 5.39

Si $\lambda \in \text{Sp}(f)$ alors $\text{sep}(f, \lambda)$ est stable par f et l'endomorphisme induit par f dans $\text{sep}(f, \lambda)$ est l'homothétie de rapport $\lambda : g = \lambda \text{id}$.

On note $p = \dim \text{sep}(f, \lambda)$.

On a

$$\chi_g = \begin{vmatrix} X - \lambda & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & X - \lambda \end{vmatrix}_{[p]} = (X - \lambda)^p.$$

D'après la Proposition 5.36, $(X - \lambda)^p \mid \chi_f$ donc $p \leq \alpha$.

De plus, on a $1 \leq p$ car $\text{sep}(f, \lambda) \neq \{0\}$. ■

Exercice 5.40

Soit f un endomorphisme de matrice $\begin{pmatrix} 3 & -4 & -5 \\ -1 & 3 & 2 \\ 1 & -2 & -1 \end{pmatrix}$. Déterminez les valeurs propres de f , leur multiplicité et la dimension des sous-espaces propres associés.

5.2.4 Endomorphisme scindé

Définition 5.41

On dit qu'un endomorphisme de E est scindé quand son polynôme caractéristique est scindé dans $\mathbb{K}[X]$.

Dans le cas d'un endomorphisme scindé, on connaît alors la somme et le produit des valeurs propres.

Proposition 5.42

Si $f \in \mathcal{L}(E)$ est scindé et a pour valeurs propres $\lambda_1, \dots, \lambda_p$ avec les ordres de multiplicité $\alpha_1, \dots, \alpha_p$, alors

$$\text{tr } f = \sum_{k=1}^p \alpha_k \lambda_k \quad \text{et} \quad \det f = \prod_{k=1}^p \lambda_k^{\alpha_k}.$$

Démonstration 5.43

Relations coefficients/racines. ■

Si $\mathbb{K} = \mathbb{C}$ alors on est dans ce cas, car tous les polynômes de $\mathbb{C}[X]$ sont scindés dans $\mathbb{C}[X]$ d'après le théorème de d'Alembert-Gauss.

Mais si $\mathbb{K} = \mathbb{R}$, alors il faut se méfier des raisonnements hâtifs : comme un \mathbb{R} -endomorphisme peut ne pas avoir de valeurs propres réelles, la trace et le déterminant peuvent ne pas avoir de rapport avec les valeurs propres.

Exercice 5.44

Soit f un endomorphisme d'un \mathbb{C} -espace vectoriel de dimension $n \geq 2$ dont la matrice dans une base est remplie par ligne de 1, ligne de 2, etc. Sans calculer le polynôme caractéristique, déterminez les valeurs propres complexes de f , leur multiplicité et la dimension des sous-espaces propres associés.

Remarque 5.45

Dans le langage courant, on dit souvent que la trace est la somme des valeurs propres. Cette phrase est correcte seulement si l'on sous-entend que l'on parle de la somme des valeurs propres comptées chacune avec son ordre de multiplicité.

On rencontre en fait deux types de résultats à propos des valeurs propres :

- ceux où l'on parle des valeurs propres distinctes (comme le Théorème 5.28) ;
- ceux où l'on parle des valeurs propres comptées selon leur multiplicité (comme la Proposition 5.42).

Il faut donc être très attentif à la façon dont on considère les valeurs propres.

5.3 Éléments propres d'une matrice carrée

Soit $n \in \mathbb{N}^*$. Les matrices-colonnes d'ordre n sont les matrices de $\mathcal{M}_{n1}(\mathbb{K})$, espace souvent identifié avec \mathbb{K}^n .

5.3.1 Valeurs propres et vecteurs propres

Définition 5.46

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$.

On dit que λ est valeur propre de A quand il existe une matrice-colonne X non-nulle telle que $AX = \lambda X$.

Si λ est une valeur propre de A , alors toute matrice-colonne non-nulle X telle que $AX = \lambda X$ est appelée vecteur propre associé à la valeur propre λ .

Exemple 5.47

- Pour tout $\alpha \in \mathbb{K}$, αI_n a pour unique valeur propre α et toute matrice-colonne non-nulle est un vecteur propre associé.
- Si A est une matrice diagonale, alors ses valeurs propres sont les coefficients diagonaux et des vecteurs propres associés sont les colonnes remplies de 0 sauf un seul coefficient égal à 1.

L'ensemble des valeurs propres d'une matrice A est appelé le spectre de A et est noté $\text{Sp}_{\mathbb{K}}(A)$ ou plus simplement $\text{Sp}(A)$.

Mais comme une matrice à coefficients réels est aussi une matrice à coefficients complexes, il vaut mieux savoir si on parle des valeurs propres réelles ou complexes. Il est donc préférable d'indiquer clairement le corps de base, comme le montre le résultat suivant.

Proposition 5.48

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et \mathbb{K}' une extension de \mathbb{K} dans \mathbb{C} .

Alors $\text{Sp}_{\mathbb{K}}(A) \subseteq \text{Sp}_{\mathbb{K}'}(A)$.

Proposition 5.49

Soient $A \in \mathcal{M}_n(\mathbb{K})$, $f \in \mathcal{L}(E)$ et \mathcal{B} une base de E .

Si $A = \underset{\mathcal{B}}{\text{Mat}}(f)$, alors $\text{Sp}_{\mathbb{K}}(A) = \text{Sp}(f)$.

Par conséquent, deux matrices semblables ont les mêmes valeurs propres (mais attention, pas forcément les mêmes vecteurs propres).

5.3.2 Lien avec les polynômes annulateurs

Proposition 5.50

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

Si P est un polynôme annulateur de A , alors $\text{Sp}_{\mathbb{K}}(A) \subseteq Z_{\mathbb{K}}(P)$.

Attention ! La réciproque est fausse. Contre-exemple : le polynôme $P = X^2 - 1$ est annulateur de I_n et pourtant -1 , qui est racine de P , n'est pas valeur propre de I_n .

5.3.3 Sous-espaces propres

Proposition 5.51

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$.

Alors λ est valeur propre de A ssi $A - \lambda I_n$ n'est pas inversible, autrement dit ssi $\text{rg}(A - \lambda I_n) < n$ ou $\det(A - \lambda I_n) = 0$.

Si $\lambda \in \text{Sp}_{\mathbb{K}}(A)$, le sous-espace propre associé à la valeur propre λ est l'ensemble des vecteurs propres associés à la valeur propre λ auquel on ajoute le vecteur nul. Il est souvent noté $\text{sep}_{\mathbb{K}}(A, \lambda)$:

$$\text{sep}_{\mathbb{K}}(A, \lambda) = \{X \in \mathcal{M}_{n,1}(\mathbb{K}) \mid AX = \lambda X\}.$$

Proposition 5.52

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. Alors

$$\lambda \in \text{Sp}_{\mathbb{K}}(A) \iff \text{rg}(A - \lambda I_n) < n.$$

Dans ce cas, $\dim \text{sep}_{\mathbb{K}}(A, \lambda) = n - \text{rg}(A - \lambda I_n)$.

Attention ! Dans la relation $\dim \text{sep}_{\mathbb{K}}(A, \lambda) = n - \text{rg}(A - \lambda I_n)$, c'est n , pas n^2 ! Il s'agit de la dimension de $\mathcal{M}_{n,1}(\mathbb{K})$, pas celle de $\mathcal{M}_n(\mathbb{K})$.

Remarque 5.53

Un cas particulier important : 0 est valeur propre ssi A n'est pas inversible, c'est-à-dire ssi $\text{rg } A < n$.

Théorème 5.54

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda_1, \dots, \lambda_p$ des valeurs propres distinctes de A .

Alors les sous-espaces propres $(\text{sep}_{\mathbb{K}}(A, \lambda_i))_{1 \leq i \leq p}$ sont en somme directe.

Autrement dit, toute famille de vecteurs propres associés à des valeurs propres distinctes est libre.

Remarque 5.55

Quand on demande de déterminer les éléments propres d'une matrice, on demande de déterminer les valeurs propres et les vecteurs propres associés, *i.e.* les sous-espaces propres.

5.4 Polynôme caractéristique d'une matrice carrée

5.4.1 Définition et lien avec les valeurs propres

Définition 5.56

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

On appelle polynôme caractéristique de A le polynôme $\chi_A = \det(XI_n - A)$.

Proposition 5.57

Soient $A \in \mathcal{M}_n(\mathbb{K})$, $f \in \mathcal{L}(E)$ et \mathcal{B} une base de E .

Si $A = \text{Mat}_{\mathcal{B}}(f)$, alors $\chi_A = \chi_f$.

Par conséquent, deux matrices semblables ont le même polynôme caractéristique.

Théorème 5.58

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

Alors χ_A est un polynôme unitaire de degré n de $\mathbb{K}[X]$ et les valeurs propres de A sont exactement les racines de χ_A dans \mathbb{K} .

Par conséquent, une matrice carrée de taille (n, n) a au plus n valeurs propres distinctes.

Corollaire 5.59

L'ensemble $\text{GL}_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})$.

Démonstration 5.60

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

On veut montrer qu'il existe une suite de matrices inversibles qui converge vers A .

Considérons la suite $\left(A + \frac{1}{k}I_n\right)_{k \in \mathbb{N}^*}$.

On a $\lim_{k \rightarrow +\infty} \left(A + \frac{1}{k}I_n\right) = A$.

Montrons qu'à partir d'un certain rang, cette suite est constituée de matrices inversibles.

Pour tout $k \in \mathbb{N}^*$, $A + \frac{1}{k}I_n$ n'est pas inversible $\iff \frac{-1}{k}$ est valeur propre de A .

- ▷ Si A n'a que des valeurs propres positives ou nulles, alors comme pour tout $k \in \mathbb{N}^*$, $\frac{-1}{k} < 0$, $\frac{-1}{k}$ n'est pas valeur propre.
- ▷ Si A possède au moins une valeur propre strictement négative, on pose $r = \min \{|\lambda| \mid \lambda \in \text{Sp}(A) \cap \mathbb{R}_-^*\} > 0$.
Dès que $\frac{1}{k} < r$, il est certain que $\frac{-1}{k}$ n'est pas valeur propre. ■

On peut noter un lien avec la trace et le déterminant.

Proposition 5.61

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

Alors $\chi_A = X^n - \text{tr}(A)X^{n-1} + \dots + (-1)^n \det A$.

5.4.2 Ordre de multiplicité et dimension du sous-espace propre

Définition 5.62

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \text{Sp}_{\mathbb{K}}(A)$.

On appelle ordre de multiplicité de la valeur propre λ son ordre de multiplicité en tant que racine de χ_A .

Théorème 5.63

Soient $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \text{Sp}_{\mathbb{K}}(A)$.

Si λ est une valeur propre d'ordre α , alors $1 \leq \dim \text{sep}_{\mathbb{K}}(A, \lambda) \leq \alpha$.

Proposition 5.64

Soient $A \in \mathcal{M}_n(\mathbb{K})$, $f \in \mathcal{L}(E)$ et \mathcal{B} une base de E .

Si $A = \underset{\mathcal{B}}{\text{Mat}}(f)$, alors $\dim \text{sep}_{\mathbb{K}}(A, \lambda) = \dim \text{sep}(f, \lambda)$.

Par conséquent, deux matrices semblables ont des sous-espaces propres de même dimension (mais pas les mêmes vecteurs propres).

5.4.3 Matrice scindée

Définition 5.65

On dit qu'une matrice de $\mathcal{M}_n(\mathbb{K})$ est scindée quand son polynôme caractéristique est scindé dans $\mathbb{K}[X]$.

Dans le cas d'une matrice scindée, on connaît alors la somme et le produit des valeurs propres.

Proposition 5.66

Si $A \in \mathcal{M}_n(\mathbb{K})$ est scindée et a pour valeurs propres $\lambda_1, \dots, \lambda_p$ avec les ordres de multiplicité $\alpha_1, \dots, \alpha_p$, alors

$$\text{tr } A = \sum_{k=1}^p \alpha_k \lambda_k \quad \text{et} \quad \det A = \prod_{k=1}^p \lambda_k^{\alpha_k}.$$

Si $\mathbb{K} = \mathbb{C}$, alors on est dans ce cas, car tous les polynômes de $\mathbb{C}[X]$ sont scindés dans $\mathbb{C}[X]$ d'après le théorème de d'Alembert-Gauss.

Mais si $\mathbb{K} = \mathbb{R}$, alors il faut se méfier des raisonnements hâtifs : comme un polynôme à coefficients réels peut ne pas avoir de racines réelles, la trace et le déterminant peuvent ne pas avoir de rapport avec les valeurs propres.

5.5 Endomorphismes diagonalisables, matrices diagonalisables

5.5.1 Définition

Définition 5.67

Soient $f \in \mathcal{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{K})$.

On dit que f est diagonalisable quand il existe une base de E constituée de vecteurs propres de f .

On dit que A est diagonalisable dans $\mathcal{M}_n(\mathbb{K})$ (ou \mathbb{K} -diagonalisable) quand il existe une base de $\mathcal{M}_{n1}(\mathbb{K})$ constituée de vecteurs propres de A .

D'après le lien entre les endomorphismes et les matrices carrées, un endomorphisme est diagonalisable ssi sa matrice dans n'importe quelle base est diagonalisable.

Exercice 5.68

La matrice $\begin{pmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}$ est-elle \mathbb{R} -diagonalisable ? \mathbb{C} -diagonalisable ?

Exercice 5.69

Montrez que la matrice $A = \begin{pmatrix} 5 & -8 & -4 \\ 8 & -15 & -8 \\ -10 & 20 & 11 \end{pmatrix}$ est diagonalisable.

Exercice 5.70

Même exercice avec $B = \begin{pmatrix} 0 & 1 & -1 \\ 2 & 1 & 1 \\ 4 & -2 & 4 \end{pmatrix}$.

Exercice 5.71

La matrice $C = \begin{pmatrix} 11 & 7 & -3 \\ 11 & 7 & -3 \\ 66 & 42 & -18 \end{pmatrix}$ est-elle diagonalisable ?

Proposition 5.72

Si un endomorphisme (une matrice) est diagonalisable, alors il (elle) est scindé(e).

Mais la réciproque est fausse.

5.5.2 Caractérisations équivalentes

On note $\mathcal{D}_n(\mathbb{K})$ l'ensemble des matrices diagonales de $\mathcal{M}_n(\mathbb{K})$.

Proposition 5.73

Soient $f \in \mathcal{L}(E)$ et $A \in \mathcal{M}_n(\mathbb{K})$.

f est diagonalisable ssi il existe une base \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(f) \in \mathcal{D}_n(\mathbb{K})$. Dans ce cas, les valeurs propres de f sont les éléments diagonaux de cette matrice.

A est \mathbb{K} -diagonalisable ssi elle est \mathbb{K} -semblable à une matrice diagonale, i.e. il existe $P \in \text{GL}_n(\mathbb{K})$ et $D \in \mathcal{D}_n(\mathbb{K})$ tel que $A = PDP^{-1}$. Dans ce cas, les valeurs propres de A sont les éléments diagonaux de D .

Démonstration 5.74

Si f est diagonalisable, il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E constituée de vecteurs propres, i.e.

$$\text{pour tout } j \in \llbracket 1 ; n \rrbracket, f(e_j) = \lambda_j e_j$$

où λ_j est la valeur propre associée à e_j .

Donc

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \in \mathcal{D}_n(\mathbb{K}).$$

Et réciproquement. ■

Exemple 5.75

- Toute matrice diagonale est diagonalisable, car elle est semblable à elle-même.
- Les projecteurs et les symétries sont diagonalisables.

Remarque 5.76

Quitte à changer l'ordre des vecteurs dans la base, on peut ranger les valeurs propres sur la diagonale dans l'ordre qu'on veut.

Exemple 5.77

Si $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 2 & -1 & 3 \end{pmatrix}$ et $D = P^{-1}AP$, alors la colonne 1 de P est un vecteur propre de A pour la valeur propre 1 et les deux autres sont des vecteurs propres pour la valeur propre 3, donc en posant $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 3 & -1 & 2 \end{pmatrix}$, on a $Q^{-1}AQ = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Lemme 5.78

Soit $f \in \mathcal{L}(E)$ diagonalisable : il existe une base de E dans laquelle la matrice D de f est diagonale.

Les valeurs propres de f sont les éléments diagonaux de D et si λ est un tel nombre, alors la dimension de $\text{sep}(f, \lambda)$ est le nombre d'occurrences de λ dans la diagonale de D .

On en déduit les théorèmes suivants.

Théorème 5.79

Soit $f \in \mathcal{L}(E)$.

Il y a équivalence entre les propositions suivantes :

- ▷ f est diagonalisable
- ▷ les sous-espaces propres de f sont supplémentaires dans E
- ▷ $\sum_{\lambda \in \text{Sp}(f)} \dim \text{sep}(f, \lambda) = n$

Démonstration 5.80

Les sous-espaces propres d'un endomorphisme sont en somme directe donc ils sont supplémentaires ssi la somme de leurs dimensions est celle de l'espace E . ■

Et sa version matricielle.

Théorème 5.81

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

Il y a équivalence entre les propositions suivantes :

- ▷ A est diagonalisable dans $\mathcal{M}_n(\mathbb{K})$
- ▷ les sous-espaces propres de A dans $\mathcal{M}_{n1}(\mathbb{K})$ sont supplémentaires dans $\mathcal{M}_{n1}(\mathbb{K})$
- ▷ $\sum_{\lambda \in \text{Sp}_{\mathbb{K}}(A)} \dim \text{sep}_{\mathbb{K}}(A, \lambda) = n$

Exercice 5.82

On pose $A = \begin{pmatrix} 0 & 1 & -1 \\ 2 & 1 & 1 \\ 4 & -2 & 4 \end{pmatrix}$. On a vu à l'Exercice 5.70 que A est diagonalisable. Diagonalisez A .

5.5.3 Lien avec le polynôme caractéristique

Théorème 5.83

Soit $f \in \mathcal{L}(E)$.

Il y a équivalence entre les propositions suivantes :

- ▷ f est diagonalisable
- ▷ f est scindé et pour tout $\lambda \in \text{Sp}(f)$, la dimension de $\text{sep}(f, \lambda)$ est égale à l'ordre de multiplicité de λ

Démonstration 5.84

Si $\lambda \in \text{Sp}(f)$, on note $\omega(\lambda)$ l'ordre de multiplicité de la valeur propre λ .



Si f est diagonalisable alors $\sum_{\lambda \in \text{Sp}(f)} \dim \text{sep}(f, \lambda) = n = \sum_{\lambda \in \text{Sp}(f)} \omega(\lambda)$.

Donc $\sum_{\lambda \in \text{Sp}(f)} \underbrace{(\omega(\lambda) - \dim \text{sep}(f, \lambda))}_{\geq 0 \text{ d'après le Théorème 5.38}} = 0$.

Or une somme de réels positifs est nulle ssi tous ces réels sont nuls donc

$$\forall \lambda \in \text{Sp}(f), \quad \omega(\lambda) = \dim \text{sep}(f, \lambda).$$



Si f est scindé et $\forall \lambda \in \text{Sp}(f), \quad \omega(\lambda) = \dim \text{sep}(f, \lambda)$, alors χ_f est scindé.

Donc $\sum_{\lambda \in \text{Sp}(f)} \omega(\lambda) = \deg \chi_f = n$.

Donc $\sum_{\lambda \in \text{Sp}(f)} \dim \text{sep}(f, \lambda) = n$.

Donc f est diagonalisable d'après le Théorème 5.79. ■

Et sa version matricielle.

Théorème 5.85

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

Il y a équivalence entre les propositions suivantes :

- ▷ A est diagonalisable dans $\mathcal{M}_n(\mathbb{K})$

- A est scindée et pour tout $\lambda \in \text{Sp}_{\mathbb{K}}(A)$, la dimension de $\text{sep}_{\mathbb{K}}(A, \lambda)$ est égale à l'ordre de multiplicité de λ

Dans le cas où $\mathbb{K} = \mathbb{C}$, la condition « être scindé » est automatiquement satisfaite.

Un cas particulier très courant.

Proposition 5.86

Si un endomorphisme de E possède exactement n valeurs propres distinctes, alors il est diagonalisable.

Si une matrice de $\mathcal{M}_n(\mathbb{K})$ possède exactement n valeurs propres distinctes dans \mathbb{K} , alors elle est diagonalisable dans $\mathcal{M}_n(\mathbb{K})$.

Exercice 5.87

Montrez que la matrice $\begin{pmatrix} -4 & 8 & 22 \\ -2 & 3 & 4 \\ -1 & 2 & 7 \end{pmatrix}$ est diagonalisable.

Théorème 5.88 (Théorème spectral)

Si A est une matrice réelle symétrique, alors A est diagonalisable.

Démonstration 5.89

★★ Admis, sera démontré plus tard (cf. Démonstration 9.96) ★★

■

5.6 Lien entre diagonalisabilité et polynômes annulateurs

5.6.1 Racines du polynôme minimal

Proposition 5.90

Soit $f \in \mathcal{L}(E)$. Les racines de μ_f sont exactement les valeurs propres de $f : Z_{\mathbb{K}}(\mu_f) = \text{Sp}(f)$.

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les racines dans \mathbb{K} de μ_A sont exactement les valeurs propres dans \mathbb{K} de $A : Z_{\mathbb{K}}(\mu_A) = \text{Sp}_{\mathbb{K}}(A)$.

Démonstration 5.91

\supseteq Cf. Proposition 5.11 car μ_f est un polynôme annulateur de f .

\subseteq

Soit $\lambda \in Z_{\mathbb{K}}(\mu_f)$.

Alors $X - \lambda \mid \mu_f$, i.e. il existe $Q \in \mathbb{K}[X]$ tel que $\mu_f = (X - \lambda)Q$.

Alors $\mu_f(f) = 0 = (f - \lambda \text{id}_E) \circ Q(f)$.

Donc pour tout $x \in E$, $(f - \lambda \text{id}_E)(Q(f)(x)) = 0$.

Donc $\text{Im } Q(f) \subseteq \ker(f - \lambda \text{id}_E)$.

Or $\deg Q < \deg \mu_f$ donc Q n'est pas annulateur de f , i.e. $Q(f) \neq 0$, i.e. $\text{Im } Q(f) \neq \{0\}$.

Donc $\ker(f - \lambda \text{id}_E) \neq \{0\}$, i.e. $\lambda \in \text{Sp}(f)$.

Donc $Z_{\mathbb{K}}(\mu_f) \subseteq \text{Sp}(f)$. ■

5.6.2 Lemme des noyaux

Proposition 5.92

Soient $f \in \mathcal{L}(E)$ et $P, Q \in \mathbb{K}[X]$ tels que $P \wedge Q = 1$.

Alors $\ker(PQ)(f) = \ker P(f) \oplus \ker Q(f)$.

Démonstration 5.93

D'après le théorème de Bézout, il existe $(U, V) \in \mathbb{K}[X]^2$ tel que $UP + VQ = 1$.

Donc $(UP)(f) + (VQ)(f) = \text{id}_E$, i.e.

$$U(f) \circ P(f) + V(f) \circ Q(f) = \text{id}_E \quad (1)$$

Soit $x \in \ker P(f) \cap \ker Q(f)$.

On a $P(f)(x) = 0$ et $Q(f)(x) = 0$.

Donc, en appliquant (1) sur le vecteur x , on obtient

$$\begin{aligned} x &= U(f)(P(f)(x)) + V(f)(Q(f)(x)) \\ &= U(f)(0) + V(f)(0) \\ &= 0. \end{aligned}$$

Donc $\ker P(f)$ et $\ker Q(f)$ sont en somme directe.

□

On a $(PQ)(f) = P(f) \circ Q(f) = Q(f) \circ P(f)$.

Donc $\ker P(f) \subseteq \ker (PQ)(f)$ et $\ker Q(f) \subseteq \ker (PQ)(f)$.

Donc $\ker P(f) \oplus \ker Q(f) \subseteq \ker (PQ)(f)$.

□

Soit $x \in \ker (PQ)(f)$.

On veut montrer qu'il existe $(a, b) \in \ker P(f) \times \ker Q(f)$ tel que $x = a + b$.

On applique (1) sur x :

$$x = U(f) \circ P(f)(x) + V(f) \circ Q(f)(x).$$

On pose $a = V(f) \circ Q(f)(x)$.

On a

$$P(f)(a) = P(f)(V(f) \circ Q(f)(x)) = (P(f) \circ V(f) \circ Q(f))(x).$$

Or $\mathbb{K}[f]$ est une algèbre commutative donc

$$\begin{aligned} P(f)(a) &= (V(f) \circ P(f) \circ Q(f))(x) \\ &= V(f)(P(f) \circ Q(f)(x)) \\ &= V(f)((PQ)(f)(x)) \\ &= V(f)(0) \\ &= 0. \end{aligned}$$

Donc $a \in \ker P(f)$.

De même, $b = U(f) \circ P(f)(x) \in \ker Q(f)$.

Finalement, on a

$$\ker (PQ)(f) = \ker P(f) \oplus \ker Q(f). \quad \blacksquare$$

Proposition 5.94

Soient $f \in \mathcal{L}(E)$ et $P_1, \dots, P_k \in \mathbb{K}[X]$ premiers entre eux deux à deux. On pose $P = \prod_{i=1}^k P_i$.

Alors $\ker P(f) = \bigoplus_{i=1}^k \ker P_i(f)$.

Démonstration 5.95

On note $\mathcal{P}(k)$ le prédicat énoncé.

- On a clairement $\mathcal{P}(1)$ et $\mathcal{P}(2)$ est vraie (cf. Proposition 5.92).
- Soit $k \in \mathbb{N}^*$ tel que $\mathcal{P}(k)$ soit vraie.

Soient $P_1, \dots, P_{k+1} \in \mathbb{K}[X]$ premiers entre eux deux à deux.

On a $P_1 \dots P_k \wedge P_{k+1} = 1$.

D'après $\mathcal{P}(2)$, on a

$$\ker(P_1 \dots P_{k+1})(f) = \ker(P_1 \dots P_k)(f) \oplus \ker P_{k+1}(f).$$

Puis, par hypothèse de récurrence, on a

$$\ker(P_1 \dots P_k)(f) = \bigoplus_{i=1}^k \ker P_i(f).$$

Finalement, on a

$$\ker(P_1 \dots P_{k+1})(f) = \bigoplus_{i=1}^{k+1} \ker P_i(f)$$

d'où $\mathcal{P}(k+1)$.

- D'après le principe de récurrence, pour tout $k \in \mathbb{N}^*$, $\mathcal{P}(k)$ est vraie. ■

5.6.3 Application à la diagonalisabilité

Définition 5.96

Un polynôme est dit simplement scindé quand il est scindé et à racines simples.

Théorème 5.97

Soit $f \in \mathcal{L}(E)$.

Il y a équivalence entre les propositions suivantes :

(α) f est diagonalisable

(β) μ_f est simplement scindé

(γ) il existe un polynôme annulateur de f simplement scindé

(δ) le polynôme $\prod_{\lambda \in \text{Sp}(f)} (X - \lambda)$ est un polynôme annulateur de f

Démonstration 5.98 ((β) \implies (γ))

Immédiat car μ_f est annulateur de f . ■

Démonstration 5.99 ((γ) \implies (β))

Si P est simplement scindé et $P(f) = 0$ alors $\mu_f \mid P$ donc μ_f est simplement scindé. ■

Démonstration 5.100 ((β) \implies (δ))

On sait que $Z_{\mathbb{K}}(\mu_f) = \text{Sp}(f)$ donc si μ_f est simplement scindé, alors $\mu_f = \prod_{\lambda \in \text{Sp}(f)} (X - \lambda)$. Or μ_f est annulateur de f . ■

Démonstration 5.101 ((δ) \implies (γ))

Immédiat. ■

Démonstration 5.102 ((α) \implies (δ))

Supposons f diagonalisable, i.e. il existe une base \mathcal{B} de E telle que $\text{Mat}_{\mathcal{B}}(f)$ soit diagonale :

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \Lambda_1 & 0 & \dots & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & \dots & 0 & \Lambda_k \end{pmatrix} = D$$

où $\lambda_1, \dots, \lambda_k$ sont les valeurs propres distinctes de f et pour tout $i \in \llbracket 1 ; k \rrbracket$, $\Lambda_i = \begin{pmatrix} \lambda_i & 0 & \dots & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & \dots & 0 & \lambda_i \end{pmatrix}$.

On pose $P = \prod_{i=1}^k (X - \lambda_i)$.

Or on a pour tout $Q \in \mathbb{K}[X]$, $Q(D) = \begin{pmatrix} Q(\Lambda_1) & 0 & \dots & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & \dots & 0 & Q(\Lambda_k) \end{pmatrix}$

et pour tout $Q \in \mathbb{K}[X]$, pour tout $i \in \llbracket 1 ; k \rrbracket$, $Q(\Lambda_i) = \begin{pmatrix} Q(\lambda_i) & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & 0 \\ & & & Q(\lambda_i) \end{pmatrix}$.

En particulier, $P(D) = 0$ car $\{\lambda_1, \dots, \lambda_k\} = Z_{\mathbb{K}}(P)$. ■

Démonstration 5.103 ((δ) \implies (α))

On pose $\text{Sp}(f) = \{\lambda_1, \dots, \lambda_k\}$ et pour tout $i \in \llbracket 1 ; k \rrbracket$, $P_i = X - \lambda_i$.

Les polynômes P_1, \dots, P_k sont premiers entre eux deux à deux donc d'après le lemme des noyaux, on a

$$\underbrace{\ker(P_1 \dots P_k)(f)}_{=0} = \bigoplus_{i=1}^k \underbrace{\ker P_i(f)}_{=\text{sep}(f, \lambda_i)}.$$

D'où $E = \bigoplus_{i=1}^k \text{sep}(f, \lambda_i)$.

Donc d'après le Théorème 5.79, f est diagonalisable. ■

Et sa version matricielle.

Théorème 5.104

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

Il y a équivalence entre les propositions suivantes :

- ▷ A est diagonalisable dans $\mathcal{M}_n(\mathbb{K})$
- ▷ μ_A est simplement scindé
- ▷ il existe un polynôme annulateur de A simplement scindé dans $\mathbb{K}[X]$
- ▷ le polynôme $\prod_{\lambda \in \text{Sp}_{\mathbb{K}}(A)} (X - \lambda)$ est un polynôme annulateur de A

Exercice 5.105

On pose $A = \begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$. Calculez $(A + I_3)^3$. A est-elle diagonalisable ?

Exercice 5.106

Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $A^3 = I_n$. Selon que \mathbb{K} soit égal à \mathbb{C} ou \mathbb{R} , à quelle condition A est-elle \mathbb{K} -diagonalisable ?

5.6.4 Diagonalisabilité d'un endomorphisme induit

Proposition 5.107

Soient $f \in \mathcal{L}(E)$, F un sous-espace vectoriel de E stable par f et g l'endomorphisme induit par f dans F .

Alors μ_g divise μ_f .

Démonstration 5.108

On a $g : F \longrightarrow F$
 $x \longmapsto f(x)$

Pour tout $x \in F$, pour tout $P \in \mathbb{K}[X]$, $P(g)(x) = P(f)(x)$.

Or $\mu_f(f) = 0$ donc pour tout $x \in F$, $\mu_f(g)(x) = 0$, i.e. μ_f est annulateur de g .

Donc $\mu_g \mid \mu_f$. ■

Corollaire 5.109

Soient $f \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E stable par f .

Si f est diagonalisable, alors l'endomorphisme induit par f dans F est aussi diagonalisable.

Démonstration 5.110

Si f est diagonalisable, d'après le Théorème 5.97, μ_f est simplement scindé.

Or $\mu_g \mid \mu_f$ donc μ_g est simplement scindé.

Donc g est diagonalisable d'après le Théorème 5.97. ■

Remarque 5.111

On a également :

► $\text{Sp}(g) = Z_{\mathbb{K}}(\mu_g) \subseteq Z_{\mathbb{K}}(\mu_f) = \text{Sp}(f)$

► si x est un vecteur propre de g pour la valeur propre λ , i.e. $\begin{cases} x \in F \\ x \neq 0 \\ g(x) = f(x) = \lambda x \end{cases}$ alors x est un vecteur propre de f dans F , et réciproquement.

On a donc $\text{sep}(g, \lambda) = \text{sep}(f, \lambda) \cap F$.

Exercice 5.112

Soit f un endomorphisme de matrice $\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ dans une base $\mathcal{B} = (e_1, e_2, e_3)$. Déterminez les sous-espaces vectoriels de E stables par f .

Exercice 5.113 (Codiagonalisation ou diagonalisation simultanée)

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ diagonalisables et qui commutent.

Montez qu'il existe $P \in \text{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ et $P^{-1}BP$ sont diagonales.

5.7 Quelques applications de la diagonalisation

5.7.1 Puissances d'une matrice, suites récurrentes linéairement

Un petit lemme déjà rencontré.

Lemme 5.114

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ et $P \in \text{GL}_n(\mathbb{K})$ telles que $A = PBP^{-1}$.

Alors pour tout $k \in \mathbb{N}$, $A^k = PB^kP^{-1}$.

Le lemme précédent est particulièrement utile si A est diagonalisable et si on choisit $B = D$, matrice diagonale semblable à A , car calculer les puissances d'une matrice diagonale est très facile.

Grâce à la diagonalisation de A , on peut espérer exprimer la forme générale des suites récurrentes linéaires (voir le chapitre précédent, section sur les polynômes annulateurs).

Exercice 5.115

Soient u, v, w les trois suites réelles telles que $u_0 = v_0 = w_0 = 1$ et

$$\text{pour tout } n \in \mathbb{N}, \quad \begin{cases} u_{n+1} = u_n - v_n \\ v_{n+1} = -4u_n + 4v_n - 6w_n \\ w_{n+1} = -3u_n + 3v_n - 4w_n \end{cases}$$

Déterminez des expressions de u_n, v_n, w_n en fonction de n .

Cette technique s'applique en particulier aux suites u vérifiant une relation de récurrence linéaire de la forme : pour tout $n \in \mathbb{N}$, $u_{n+d} = a_{d-1}u_{n+d-1} + \cdots + a_2u_{n+2} + a_1u_{n+1} + a_0u_n$.

On pose alors $X_n = \begin{pmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{n+d-1} \end{pmatrix}$ et $A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & \dots & a_{d-1} \end{pmatrix} \in \mathcal{M}_d(\mathbb{K})$.

Alors pour tout $n \in \mathbb{N}$, $X_{n+1} = AX_n$ et on est ramené au cas précédent.

La matrice A s'appelle la matrice-compagnon du polynôme $P = X^d - a_{d-1}X^{d-1} - \dots - a_1X - a_0$: elle a la propriété remarquable que son polynôme caractéristique est P , son polynôme minimal est aussi P et donc que ses valeurs propres sont les racines de P . C'est pourquoi le polynôme P est appelé polynôme caractéristique associé à la suite u (cas déjà étudié en première année : $d = 2$).

On en déduit que A est diagonalisable ssi P est simplement scindé et dans ce cas, A possède d valeurs propres distinctes. Dans ce cas, en notant $\lambda_1, \dots, \lambda_p$ les valeurs propres distinctes, la suite u est combinaison linéaire des suites géométriques $(\lambda_1^n), \dots, (\lambda_d^n)$.

Exercice 5.116

Explicitez l'unique suite (u_n) vérifiant

$$u_0 = 0, u_1 = 1, u_2 = 5 \quad \text{et} \quad \forall n \in \mathbb{N}, \quad u_{n+3} = 6u_{n+2} - 11u_{n+1} + 6u_n.$$

5.7.2 Systèmes d'équations différentielles

Ce point sera traité dans le chapitre sur les équations différentielles linéaires.

5.8 Endomorphismes trigonalisables, matrices trigonalisables

5.8.1 Définition et propriétés

Définition 5.117

Un endomorphisme est dit trigonalisable quand il existe une base dans laquelle sa matrice est triangulaire supérieure.

Une matrice carrée de $\mathcal{M}_n(\mathbb{K})$ est dite trigonalisable dans $\mathcal{M}_n(\mathbb{K})$ quand elle est semblable à une matrice triangulaire dans $\mathcal{M}_n(\mathbb{K})$.

Remarque 5.118

- Si un endomorphisme (une matrice) est diagonalisable, alors il (elle) est trigonalisable.
- Si une matrice est trigonalisable, ses valeurs propres sont les nombres sur la diagonale de toute matrice triangulaire semblable.

Exercice 5.119

On considère la matrice $M = \begin{pmatrix} -2 & -1 & 7 \\ 5 & 4 & -8 \\ 1 & 1 & 1 \end{pmatrix}$ et f un endomorphisme de matrice M . Déterminez les éléments propres de M . Est-elle diagonalisable? En complétant une famille libre de vecteurs propres, déterminez une base \mathcal{B} de l'espace où la matrice de f est triangulaire supérieure, puis trigonalisez M .

Exercice 5.120

Soit f un endomorphisme de matrice $A = \begin{pmatrix} 2 & -4 & -5 \\ -1 & 2 & 2 \\ 1 & -2 & -2 \end{pmatrix}$. Montrez que f n'est pas diagonalisable mais est trigonalisable et donnez une base de trigonalisation de f . Donnez une forme générale pour A^n .

Quand un endomorphisme ou une matrice n'est pas diagonalisable, on peut espérer qu'il ou elle est trigonalisable : faute de grives, on se contente de merles !

Remarque 5.121

On ne confondra pas la trigonalisation d'une matrice carrée et la transformation par lignes (ou colonnes) des matrices vue en première année ! Seule la trigonalisation fournit des matrices semblables ! La transformation par lignes ne conserve que le rang !

5.8.2 Caractérisation équivalente

La trigonalisabilité est beaucoup plus courante que la diagonalisabilité, comme on le voit grâce aux résultats suivants.

Proposition 5.122

Un endomorphisme (une matrice) est trigonalisable ssi il (elle) est scindé(e).

Démonstration 5.123

On pose $\mathcal{P}(n)$: « si f est un endomorphisme d'un espace de dimension n et si χ_f est scindé, alors f est trigonalisable ».

► $\mathcal{P}(1)$ est vraie car tout endomorphisme en dimension 1 est trigonalisable.

► Supposons $\mathcal{P}(n-1)$.

Soient E un espace de dimension n et $f \in \mathcal{L}(E)$ tel que χ_f soit scindé.

Comme χ_f est scindé, il existe $\lambda \in \mathbb{K}$ tel que λ soit racine de χ_f et donc une valeur propre de f à laquelle on associe un vecteur propre u_1 .

Comme $u_1 \neq 0$, d'après le théorème de la base incomplète, il existe $(u_2, \dots, u_n) \in E^{n-1}$ tel que $\mathcal{B}_0 = (u_1, \dots, u_n)$ soit une base de E .

On a

$$\text{Mat}_{\mathcal{B}_0}(f) = \left(\begin{array}{c|c} \lambda & L \\ \hline 0 & B \\ \vdots & \\ 0 & \end{array} \right).$$

Donc

$$\chi_f = \left| \begin{array}{c|c} X - \lambda & -L \\ \hline 0 & XI_{n-1} - B \\ \vdots & \\ 0 & \end{array} \right| = (X - \lambda) \chi_B.$$

On pose $F = \text{Vect}(u_2, \dots, u_n)$.

Soit $g \in \mathcal{L}(F)$ tel que $\text{Mat}_{(u_2, \dots, u_n)}(g) = B$.

On a $\chi_g = \chi_B$ scindé donc par hypothèse de récurrence, g est trigonalisable : il existe une base

$$(u'_2, \dots, u'_n) \text{ de } F \text{ telle que } \text{Mat}_{(u'_2, \dots, u'_n)}(g) = \begin{pmatrix} t_{22} & t_{23} & \dots & t_{2n} \\ & \ddots & \ddots & \vdots \\ 0 & & & t_{n-1n} \\ \vdots & & & \vdots \\ 0 & \dots & 0 & t_{nn} \end{pmatrix} = T.$$

La famille $\mathcal{B} = (u_1, u'_2, \dots, u'_n)$ est une base de E .

On veut montrer que

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda & \alpha_2 & \dots & \alpha_n \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & 0 & & & t_{n-1n} \\ 0 & 0 & \dots & 0 & t_{nn} \end{pmatrix}.$$

On a $g = p \circ f|_F$ où p est le projecteur sur F parallèlement à $\text{Vect}(u_1)$.

Donc pour tout $x \in F$, $f(x) = \underbrace{g(x)}_{\in F} + \alpha u_1$ où $\alpha \in \mathbb{K}$.

De plus,

$$\text{pour tout } j \in \llbracket 2 ; n \rrbracket, \quad f(u'_j) = g(u'_j) + \alpha_j u_1$$

$$= \sum_{i=2}^j t_{ij} u'_i + \alpha_j u_1.$$

D'où $\mathcal{P}(n)$.

► Par récurrence, pour tout $n \in \mathbb{N}^*$, $\mathcal{P}(n)$ est vraie. ■

En particulier, quand $\mathbb{K} = \mathbb{C}$, tous les endomorphismes sont trigonalisables, toutes les matrices de $\mathcal{M}_n(\mathbb{C})$ sont trigonalisables dans $\mathcal{M}_n(\mathbb{C})$.

En pratique, quand on cherche à trigonaliser un endomorphisme, on peut chercher une base dans laquelle la matrice est triangulaire supérieure avec des 1 ou des 0 sur la sur-diagonale et des 0 sur les diagonales partielles encore au-dessus (c'est démontrable, mais c'est difficile à démontrer, cela s'appelle le théorème de Jordan – hors-programme –).

Théorème 5.124

Soit $f \in \mathcal{L}(E)$.

Il y a équivalence entre les propositions suivantes :

- ▷ f est trigonalisable
- ▷ χ_f est scindé
- ▷ μ_f est scindé
- ▷ il existe un polynôme annulateur de f scindé

Et sa version matricielle.

Théorème 5.125

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

Il y a équivalence entre les propositions suivantes :

- ▷ A est trigonalisable dans $\mathcal{M}_n(\mathbb{K})$
- ▷ χ_A est scindé
- ▷ μ_A est scindé
- ▷ il existe un polynôme annulateur de A qui est scindé dans $\mathbb{K}[X]$

Exercice 5.126

Soit $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Calculez A^2 , puis A^3 . La matrice A est-elle diagonalisable ? trigonalisable ? Dans l'affirmative, diagonalisez ou trigonalisez la.

5.8.3 Théorème de Cayley-Hamilton

Théorème 5.127

Le polynôme caractéristique d'un endomorphisme (d'une matrice carrée) est un polynôme annulateur.

Démonstration 5.128

On pose $\mathcal{P}(n)$: « si $A \in \mathcal{M}_n(\mathbb{C})$, alors $\chi_A(A) = 0$ ».

► Si $n = 1$: on pose $A = (a)$.

On a $\chi_A = X - a$ donc $\chi_A(A) = A - aI_1 = (a) - (a) = 0$.

D'où $\mathcal{P}(1)$.

► Supposons $\mathcal{P}(n-1)$.

Soit $A \in \mathcal{M}_n(\mathbb{C})$.

Le polynôme χ_A est scindé donc A est trigonalisable dans $\mathcal{M}_n(\mathbb{C})$, i.e. il existe $P \in \text{GL}_n(\mathbb{C})$ et

$$T \in \mathcal{T}_n^+(\mathbb{C}) \text{ telles que } A = PTP^{-1}, \text{ avec } T = \begin{pmatrix} \lambda_1 & ? & \dots & ? \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & ? \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

$$\text{On a } \chi_A = \chi_T = \prod_{i=1}^n (X - \lambda_i).$$

On peut écrire

$$T = \begin{pmatrix} \lambda_1 & ? & \dots & ? \\ 0 & \lambda_2 & ? & ? \\ \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & ? \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & ? & \dots & ? \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{pmatrix} \begin{matrix} \\ U \\ \\ \end{matrix}$$

où $U \in \mathcal{M}_{n-1}(\mathbb{C})$.

$$\text{On a } \chi_U = \prod_{i=2}^n (X - \lambda_i) \text{ donc } \chi_A = \chi_T = (X - \lambda_1) \chi_U.$$

Donc

$$\begin{aligned} \chi_A(A) &= (A - \lambda_1 I_n) \chi_U(A) \\ &= (PTP^{-1} - \lambda_1 I_n) \chi_U(PTP^{-1}) \\ &= P(T - \lambda_1 I_n) \chi_U(T) P^{-1}. \end{aligned} \quad \left. \vphantom{\begin{aligned} \chi_A(A) &= (A - \lambda_1 I_n) \chi_U(A) \\ &= (PTP^{-1} - \lambda_1 I_n) \chi_U(PTP^{-1}) \\ &= P(T - \lambda_1 I_n) \chi_U(T) P^{-1}. \end{aligned}} \right) (PTP^{-1})^k = PT^k P^{-1}$$

Or on a

$$\underbrace{\begin{pmatrix} 0 & ? & \dots & ? \\ 0 & \lambda_2 - \lambda_1 & ? & ? \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}}_{T - \lambda_1 I_n} \underbrace{\begin{pmatrix} \chi_U(\lambda_1) & ? & \dots & ? \\ 0 & & & \\ \vdots & & \chi_U(U) & \\ 0 & & & \end{pmatrix}}_{\chi_U(T)} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & 0 & \\ 0 & & & \end{pmatrix}$$

car $\chi_U(U) = 0$.

Donc $\chi_A(A) = 0$, d'où $\mathcal{P}(n)$.

► Par récurrence, pour tout $n \in \mathbb{N}^*$, $\mathcal{P}(n)$ est vraie. ■

Corollaire 5.129

Le polynôme minimal divise le polynôme caractéristique. Donc en dimension n , le polynôme minimal est de degré au plus n .

Les polynômes minimal et caractéristique partagent les mêmes racines dans \mathbb{C} (en fait dans tout corps \mathbb{K}) mais pas avec les mêmes ordres de multiplicité : si f est scindé, alors en notant $\lambda_1, \dots, \lambda_k$ les k valeurs propres distinctes de f , on peut écrire

$$\chi_f = \prod_{i=1}^k (X - \lambda_i)^{\alpha_i} \quad \text{et} \quad \mu_f = \prod_{i=1}^k (X - \lambda_i)^{\beta_i}$$

où pour tout $i \in \llbracket 1 ; k \rrbracket$, $1 \leq \beta_i \leq \alpha_i$.

5.8.4 Sous-espaces caractéristiques

Définition 5.130

Soit $f \in \mathcal{L}(E)$ un endomorphisme scindé. On écrit $\chi_f = \prod_{i=1}^k (X - \lambda_i)^{\alpha_i}$ où $\lambda_1, \dots, \lambda_k$ sont les k valeurs propres distinctes de f .

Les sous-espaces caractéristiques de f sont les noyaux $\ker (f - \lambda_i \text{id}_E)^{\alpha_i}$.

Proposition 5.131

Les sous-espaces caractéristiques d'un endomorphisme scindé sont supplémentaires et stables par f .

Démonstration 5.132

► Soient $\lambda \in \text{Sp}(f)$ et α l'ordre de multiplicité de λ .

Soit $x \in \ker(f - \lambda \text{id}_E)^\alpha$.

On a

$$\begin{aligned}(f - \lambda \text{id}_E)^\alpha(f(x)) &= ((f - \lambda \text{id}_E)^\alpha \circ f)(x) && \left. \begin{array}{l} \text{composée de deux} \\ \text{polynômes en } f \text{ donc} \\ \text{commutative} \end{array} \right\} \\ &= (f \circ (f - \lambda \text{id}_E)^\alpha)(x) \\ &= f(0) \\ &= 0.\end{aligned}$$

Donc $f(x) \in \ker(f - \lambda \text{id}_E)^\alpha$.

Donc $\ker(f - \lambda \text{id}_E)^\alpha$ est stable par f .

► On a $\chi_f = \prod_{i=1}^k (X - \lambda_i)^{\alpha_i}$: produit de polynômes premiers entre eux deux à deux.

D'après le lemme des noyaux, on a

$$\ker \chi_f(f) = \bigoplus_{i=1}^k \ker(f - \lambda_i \text{id}_E)^{\alpha_i}.$$

Or $\chi_f(f) = 0$ d'après le théorème de Cayley-Hamilton donc

$$E = \bigoplus_{i=1}^k \ker(f - \lambda_i \text{id}_E)^{\alpha_i}.$$

■

Théorème 5.133

Tout endomorphisme scindé possède une base dans laquelle sa matrice est diagonale par blocs telle que :

- *il y a autant de blocs que de valeurs propres : à chaque valeur propre, on associe un unique bloc ;*
- *chaque bloc est de la forme $\lambda I_r + U$ où λ est la valeur propre associée au bloc, r est l'ordre de multiplicité de λ et U est une matrice strictement triangulaire supérieure de $\mathcal{M}_r(\mathbb{K})$*

Toute matrice scindée est semblable à une matrice diagonale par blocs vérifiant les conditions précédentes.

Démonstration 5.134

Soient $\lambda \in \text{Sp}(f)$ et α son ordre de multiplicité.

Sur $F = \ker(f - \lambda \text{id}_E)^\alpha$, f induit un endomorphisme \tilde{f} tel que $(\tilde{f} - \lambda \text{id}_F)^\alpha = 0$.

Donc $(X - \lambda)^\alpha$ est un polynôme annulateur de \tilde{f} qui est scindé donc \tilde{f} a pour unique valeur propre λ et est trigonalisable.

Donc il existe une base \mathcal{B}_λ de F telle que

$$\text{Mat}_{\mathcal{B}_\lambda}(\tilde{f}) = \begin{pmatrix} \lambda & ? & \dots & ? \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & ? \\ 0 & \dots & 0 & \lambda \end{pmatrix} = \lambda I_\alpha + U.$$

Comme $E = \bigoplus_{i=1}^k \ker(f - \lambda_i \text{id}_E)^{\alpha_i}$, en concaténant de telles bases, on obtient une base de E dans laquelle la matrice de f est

$$\begin{pmatrix} \lambda_1 I_{\alpha_1} + U_1 & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ & 0 & \ddots & 0 \\ & & & \lambda_k I_{\alpha_k} + U_k \end{pmatrix}.$$

■

Corollaire 5.135

La dimension d'un sous-espace caractéristique est l'ordre de multiplicité de la valeur propre associée.

5.9 Endomorphismes nilpotents, matrices nilpotentes

5.9.1 Généralités

Définition 5.136

Soit $u \in \mathcal{L}(E)$. On dit que u est nilpotent quand il existe $p \in \mathbb{N}$ tel que $u^p = 0$.

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On dit que A est nilpotente quand il existe $p \in \mathbb{N}$ tel que $A^p = 0$.

Le plus petit indice p satisfaisant à la condition précédente s'appelle l'indice de nilpotence de u (de A).

Proposition 5.137

Toute matrice strictement triangulaire (supérieure ou inférieure) est nilpotente. Par conséquent, les matrices semblables à une matrice strictement triangulaire sont nilpotentes.

Démonstration 5.138

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice strictement triangulaire :

$$A = \begin{pmatrix} 0 & ? & \dots & ? \\ & 0 & \ddots & \vdots \\ & \vdots & \ddots & 0 \\ 0 & \dots & 0 & 0 \end{pmatrix}.$$

On a $\chi_A = X^n$ et $\chi_A(A) = 0$ donc $A^n = 0$. ■

Dans la décomposition en sous-espaces caractéristiques, on a vu apparaître des matrices $\lambda I_r + U$: les matrices U sont nilpotentes.

L'ensemble des matrices nilpotentes n'a pas de structure particulière : en général, la somme et le produit de deux matrices nilpotentes ne sont pas nilpotents. Néanmoins, avec une condition de commutation supplémentaire, on a quelques résultats.

Proposition 5.139

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ deux matrices nilpotentes.

Si A et B commutent, alors $A + B$ et AB sont nilpotentes.

Démonstration 5.140

Soit $(k, \ell) \in \mathbb{N}^2$ tel que $A^k = 0$ et $B^\ell = 0$.

Supposons $AB = BA$.

On a

$$\begin{aligned} (AB)^{\min(k, \ell)} &= A^{\min(k, \ell)} B^{\min(k, \ell)} \\ &= 0 \end{aligned}$$

et

$$\begin{aligned} (A + B)^{k+\ell} &= \sum_{i=0}^{k+\ell} \binom{k+\ell}{i} A^i B^{k+\ell-i} \\ &= \sum_{i=0}^k \binom{k+\ell}{i} A^i \underbrace{B^{k+\ell-i}}_{=0} + \sum_{i=k+1}^{k+\ell} \binom{k+\ell}{i} \underbrace{A^i}_{=0} B^{k+\ell-i} \\ &= 0. \end{aligned}$$
■

On a bien sûr les mêmes résultats concernant les endomorphismes nilpotents.

5.9.2 Éléments propres d'un nilpotent

Proposition 5.141

Un endomorphisme en dimension n est nilpotent ssi son polynôme caractéristique est X^n , i.e. s'il est scindé et admet 0 comme unique valeur propre.

Une matrice de $\mathcal{M}_n(\mathbb{K})$ est nilpotente ssi son polynôme caractéristique est X^n , i.e. si elle est scindée et admet 0 comme unique valeur propre.

L'indice de nilpotence dans ces deux cas est alors le degré du polynôme minimal ; il est donc inférieur ou égal à n .

Démonstration 5.142

Si f est nilpotent alors il existe $k \in \mathbb{N}$ tel que $f^k = 0$ donc X^k est annulateur de f donc $\text{Sp}(f) = \{0\}$ donc $\chi_f = X^n$.

Réciproquement, si $\chi_f = X^n$, d'après le théorème de Cayley-Hamilton, $f^n = 0$ donc f est nilpotent.

Or $\mu_f \mid \chi_f$ donc μ_f est de la forme X^ℓ avec $\ell \leq n$ et par définition de μ_f , ℓ est l'indice de nilpotence de f . ■

Mis à part la matrice nulle, aucune matrice nilpotente n'est diagonalisable : c'est une idée parfois utile pour prouver qu'une matrice est nulle (diagonalisable et nilpotente implique nulle).

Proposition 5.143

Tout endomorphisme nilpotent est trigonalisable : il existe une base dans laquelle sa matrice est triangulaire supérieure stricte. Réciproquement, si un endomorphisme est trigonalisable et n'a que 0 pour valeur propre, alors il est nilpotent.

Toute matrice nilpotente est trigonalisable : elle est semblable à une matrice triangulaire supérieure stricte. La réciproque est vraie.

5.9.3 Application aux sous-espaces caractéristiques d'un endomorphisme

Proposition 5.144

Soit $f \in \mathcal{L}(E)$.

Pour toute valeur propre λ de f , si α est l'ordre de multiplicité de λ dans le polynôme minimal de f , le sous-espace caractéristique associé est aussi le noyau $\ker(f - \lambda \text{id}_E)^\alpha$.

Lemme 5.145

Si $F_1, \dots, F_k, G_1, \dots, G_k$ vérifient $\bigoplus_{i=1}^k F_i = \bigoplus_{i=1}^k G_i$ et pour tout $i \in \llbracket 1 ; k \rrbracket$, $F_i \subseteq G_i$, alors pour tout $i \in \llbracket 1 ; k \rrbracket$, $F_i = G_i$.

Démonstration 5.146

Soient $i \in \llbracket 1 ; k \rrbracket$ et $x \in G_i$.

On a $x \in \bigoplus_{j=1}^k G_j = \bigoplus_{j=1}^k F_j$.

Donc il existe $(y_1, \dots, y_k) \in F_1 \times \dots \times F_k$ tel que

$$\underbrace{x}_{\in G_i} = \underbrace{y_1}_{\in F_1 \subseteq G_1} + \dots + \underbrace{y_k}_{\in F_k \subseteq G_k}.$$

Or la somme $\bigoplus_{j=1}^k G_j$ est directe donc par unicité

$$\left\{ \begin{array}{l} y_1 = 0 \\ \vdots \\ y_{i-1} = 0 \\ y_i = x \\ y_{i+1} = 0 \\ \vdots \\ y_k = 0 \end{array} \right.$$

Donc $x = y_i \in F_i$.

Donc $F_i \subseteq G_i$.

Donc $F_i = G_i$. ■

Démonstration 5.147 (de la Proposition 5.144)

On note $\chi_f = \prod_{i=1}^k (X - \lambda_i)^{\alpha_i}$ et $\mu_f = \prod_{i=1}^k (X - \lambda_i)^{\beta_i}$ où pour tout $i \in \llbracket 1 ; k \rrbracket$, $\alpha_i \geq \beta_i \geq 1$.

On veut montrer que pour tout $i \in \llbracket 1 ; k \rrbracket$, $\ker(f - \lambda_i \text{id}_E)^{\alpha_i} = \ker(f - \lambda_i \text{id}_E)^{\beta_i}$.

Comme $\beta_i \leq \alpha_i$, on a immédiatement pour tout $i \in \llbracket 1 ; k \rrbracket$, $\ker(f - \lambda_i \text{id}_E)^{\beta_i} \subseteq \ker(f - \lambda_i \text{id}_E)^{\alpha_i}$.

Comme $\mu_f(f) = \chi_f(f) = 0$, d'après le lemme des noyaux :

$$E = \bigoplus_{i=1}^k \ker(f - \lambda_i \text{id}_E)^{\beta_i} = \bigoplus_{i=1}^k \ker(f - \lambda_i \text{id}_E)^{\alpha_i}.$$

D'après le Lemme 5.145, on en déduit

$$\forall i \in \llbracket 1 ; k \rrbracket, \ker(f - \lambda_i \text{id}_E)^{\alpha_i} = \ker(f - \lambda_i \text{id}_E)^{\beta_i}. \quad \blacksquare$$

On peut même démontrer mieux.

Proposition 5.148

Soient $f \in \mathcal{L}(E)$, $\lambda \in \text{Sp}(f)$ et α l'ordre de multiplicité de λ dans le polynôme minimal de f .

Alors la suite des noyaux $\left(\ker(f - \lambda \text{id}_E)^k\right)_{k \in \mathbb{N}}$ est strictement croissante jusqu'au rang α , puis constante à partir du rang α :

$$\{0\} \subsetneq \ker(f - \lambda \text{id}_E) \subsetneq \ker(f - \lambda \text{id}_E)^2 \subsetneq \dots \subsetneq \ker(f - \lambda \text{id}_E)^\alpha = \ker(f - \lambda \text{id}_E)^{\alpha+1} = \dots$$

Chapitre 6

Intégrales généralisées

Sommaire

Fonctions continues par morceaux sur un intervalle	203
6.1 Intégrales généralisées sur $[a ; +\infty[$	203
6.1.1 Définition et exemples fondamentaux	203
6.1.2 Propriétés	205
6.1.3 Cas des fonctions réelles positives	207
6.1.4 Théorème de comparaison entre fonctions positives	208
6.1.5 Lien avec les séries	211
6.2 Intégrales généralisées sur d'autres types d'intervalles	215
6.2.1 Intégrales généralisées sur $[a ; b[$	215
6.2.2 Intégrales généralisées sur $]a ; b]$	215
6.2.3 Intégrales généralisées sur $]a ; b[$	217
6.2.4 Propriétés communes à toutes ces intégrales	219
6.2.4.1 Changement de variable	219
6.2.4.2 Intégration par parties	222
6.2.4.3 Primitives	224
6.3 Résumé pour étudier la convergence d'une intégrale	225
6.4 Fonctions intégrables sur un intervalle	227
6.4.1 Intégrales absolument convergentes	227
6.4.2 Fonctions intégrables	229
6.4.3 Théorème de comparaison des fonctions intégrables	230
6.5 Intégration des relations de comparaison	231
6.5.1 Théorème de comparaison par domination	232
6.5.2 Théorème de comparaison par équivalence	233

Dans ce chapitre, \mathbb{K} désigne l'ensemble \mathbb{R} ou \mathbb{C} . Les fonctions considérées dans ce chapitre sont à valeurs dans \mathbb{K} .

On suppose connue la notion d'intégrale sur un segment d'une fonction continue par morceaux à valeurs réelles ou complexes (*cf.* cours de première année).

Si f est une fonction continue sur un segment $[a ; b]$ (ou $[b ; a]$), on note $\int_a^b f = \int_a^b f(t) dt$ (ou toute autre lettre à la place de t) l'intégrale de f entre a et b : quand on a pas besoin de nommer la variable d'intégration, on ne la note pas, mais si on la note alors on n'oublie pas l'élément différentiel.

En préambule, on généralise la notion de fonction continue par morceaux sur un intervalle quelconque.

Fonctions continues par morceaux sur un intervalle

Définition 6.1

Soit I un intervalle quelconque.

On dit qu'une fonction est continue par morceaux sur I quand elle est continue par morceaux sur tout segment inclus dans I .

Exemple 6.2

- La fonction $t \mapsto \begin{cases} -\ln t & \text{si } t \in]0 ; 1] \\ e^{-t} & \text{si } t \in [1 ; +\infty[\end{cases}$ est continue par morceaux sur $]0 ; +\infty]$.
- La fonction $t \mapsto \left\lfloor \frac{1}{t} \right\rfloor$ est continue par morceaux sur $]0 ; +\infty]$.

Dans toute la suite, on note $\mathcal{C}_m^0(I, \mathbb{K})$ l'ensemble des fonctions continues par morceaux sur I et à valeurs dans \mathbb{K} .

Proposition 6.3

L'ensemble $\mathcal{C}_m^0(I, \mathbb{K})$ est une \mathbb{K} -algèbre.

6.1 Intégrales généralisées sur $[a ; +\infty[$

Dans cette section, a est un réel.

6.1.1 Définition et exemples fondamentaux

Définition 6.4

Soit $f \in \mathcal{C}_m^0([a ; +\infty[, \mathbb{K})$.

On dit que l'intégrale $\int_a^{+\infty} f$ converge (ou qu'elle est convergente, ou simplement qu'elle existe) quand $\int_a^x f$ a une limite finie quand x tend vers $+\infty$.

Dans ce cas, on pose $\int_a^{+\infty} f = \lim_{x \rightarrow +\infty} \int_a^x f$.

Dans le cas contraire, on dit que l'intégrale $\int_a^{+\infty}$ diverge (ou qu'elle est divergente).

Remarque 6.5

Une intégrale généralisée $\int_a^{+\infty}$ est une limite et une limite n'existe pas toujours. Avant d'utiliser une telle intégrale dans un raisonnement ou un calcul, on doit donc toujours justifier son existence !

Les résultats ci-dessous sont à connaître.

Exemple 6.6

- Soit α un réel. L'intégrale $\int_0^{+\infty} e^{-\alpha t} dt$ converge ssi $\alpha > 0$.
- Soit α un réel. L'intégrale $\int_1^{+\infty} \frac{1}{t^\alpha} dt$ converge ssi $\alpha > 1$ (intégrale dite de Riemann).

Démonstration 6.7

- La fonction $t \mapsto e^{-\alpha t}$ est continue sur $[0 ; +\infty[$.

Pour $x \geq 0$, si $\alpha \neq 0$, on a

$$\int_0^x e^{-\alpha t} dt = \left[\frac{1}{-\alpha} e^{-\alpha t} \right]_0^x = \frac{1}{\alpha} (1 - e^{-\alpha x}).$$

Donc si $\alpha > 0$, on a

$$\int_0^x e^{-\alpha t} dt \xrightarrow{x \rightarrow +\infty} \frac{1}{\alpha}$$

et si $\alpha < 0$, on a

$$\int_0^x e^{-\alpha t} dt \xrightarrow{x \rightarrow +\infty} +\infty.$$

De plus, si $\alpha = 0$, on a

$$\int_0^x e^{-\alpha t} dt = x \xrightarrow{x \rightarrow +\infty} +\infty.$$

Donc $\int_0^{+\infty} e^{-\alpha t} dt$ converge ssi $\alpha > 0$ et, dans ce cas, $\int_0^{+\infty} e^{-\alpha t} dt = \frac{1}{\alpha}$.

- La fonction $t \mapsto \frac{1}{t^\alpha}$ est continue sur $[1 ; +\infty[$.

Pour $x \geq 1$, si $\alpha \neq 1$, on a

$$\int_1^x \frac{1}{t^\alpha} dt = \int_1^x t^{-\alpha} dt = \left[\frac{1}{-\alpha + 1} t^{-\alpha+1} \right]_1^x = \frac{1}{\alpha - 1} (1 - x^{-\alpha+1}).$$

Donc si $\alpha > 1$, on a

$$\int_1^x \frac{1}{t^\alpha} dt \xrightarrow{x \rightarrow +\infty} \frac{1}{\alpha - 1}$$

et si $\alpha < 1$, on a

$$\int_1^x \frac{1}{t^\alpha} dt \xrightarrow{x \rightarrow +\infty} +\infty.$$

De plus, si $\alpha = 1$ alors

$$\int_1^x \frac{1}{t^\alpha} dt = \ln x \xrightarrow{x \rightarrow +\infty} +\infty.$$

Donc $\int_1^{+\infty} \frac{1}{t^\alpha} dt$ converge ssi $\alpha > 1$ et, dans ce cas, $\int_1^{+\infty} \frac{1}{t^\alpha} dt = \frac{1}{\alpha - 1}$. ■

Exercice 6.8

Quelle est la nature de l'intégrale $\int_0^{+\infty} \frac{1}{1+t^2} dt$?

Correction 6.9

La fonction $t \mapsto \frac{1}{1+t^2}$ est continue sur $[0 ; +\infty[$.

Pour $x \geq 0$, on a

$$\int_0^x \frac{1}{1+t^2} dt = [\operatorname{Arctan} t]_0^x = \operatorname{Arctan} x \xrightarrow{x \rightarrow +\infty} \frac{\pi}{2}$$

.

Donc $\int_0^{+\infty} \frac{1}{1+t^2} dt$ converge et vaut $\frac{\pi}{2}$.

Exercice 6.10

Quelle est la nature de l'intégrale $\int_1^{+\infty} \frac{\ln t}{t} dt$.

Correction 6.11

La fonction $t \mapsto \frac{\ln t}{t}$ est continue sur $[1 ; +\infty[$.

Pour $x \geq 1$, on a

$$\int_1^x \frac{\ln t}{t} dt = \left[\frac{\ln^2 t}{2} \right]_1^x = \frac{\ln^2 x}{2} \xrightarrow{x \rightarrow +\infty} +\infty.$$

Donc $\int_1^{+\infty} \frac{\ln t}{t} dt$ diverge.

6.1.2 Propriétés

La convergence de l'intégrale ne dépend pas de la borne a , ce qui généralise la relation de Chasles.

Proposition 6.12

Soient $f \in \mathcal{C}_m^0([a ; +\infty[, \mathbb{K})$ et $b \in [a ; +\infty[$.

Alors l'intégrale $\int_a^{+\infty} f$ converge ssi $\int_b^{+\infty} f$ converge.

Dans ce cas, on a $\int_a^{+\infty} f = \int_a^b f + \int_b^{+\infty} f$.

Démonstration 6.13

Soit $x \geq a$.

Par relation de Chasles sur les intégrales classiques, on a

$$\int_a^x f = \underbrace{\int_a^b f}_{\in \mathbb{K}} + \int_b^x f.$$

Donc $\int_a^x f$ a une limite finie quand $x \rightarrow +\infty$ ssi $\int_b^x f$ a une limite finie quand $x \rightarrow +\infty$ par opération sur les limites.

Dans ce cas, on a

$$\begin{aligned} \int_a^{+\infty} f &= \lim_{x \rightarrow +\infty} \int_a^x f \\ &= \lim_{x \rightarrow +\infty} \left(\int_a^b f + \int_b^x f \right) \\ &= \int_a^b f + \lim_{x \rightarrow +\infty} \int_b^x f \\ &= \int_a^b f + \int_b^{+\infty} f. \end{aligned}$$

■

Dans le cas convergent, on retrouve la linéarité.

Proposition 6.14

Soient $(f, g) \in \mathcal{C}_m^0([a; +\infty[, \mathbb{K})^2$ et $(\lambda, \mu) \in \mathbb{K}^2$.

Si les intégrales $\int_a^{+\infty} f$ et $\int_a^{+\infty} g$ convergent, alors $\int_a^{+\infty} (\lambda f + \mu g)$ converge.

Dans ce cas, on a $\int_a^{+\infty} (\lambda f + \mu g) = \lambda \int_a^{+\infty} f + \mu \int_a^{+\infty} g$.

Démonstration 6.15

Pour $x \geq 0$, on a

$$\int_a^x (\lambda f + \mu g) = \lambda \int_a^x f + \mu \int_a^x g$$

par linéarité de l'intégrale classique.

Puis, d'après les théorèmes d'opérations sur les limites, si $\int_a^{+\infty} f$ et $\int_a^{+\infty} g$ convergent, les deux fonctions $x \mapsto \int_a^x f$ et $x \mapsto \int_a^x g$ ont des limites finies en $+\infty$ donc $x \mapsto \lambda \int_a^x f + \mu \int_a^x g$ en a une aussi, i.e. $\int_a^{+\infty} (\lambda f + \mu g)$ converge.

Dans ce cas, $\int_a^{+\infty} (\lambda f + \mu g) = \lambda \int_a^{+\infty} f + \mu \int_a^{+\infty} g$.

■

Remarque 6.16

- ▷ « La somme d'une intégrale convergente et d'une divergente est divergente ».
- ▷ Il n'y a rien à dire *a priori* sur la « somme de deux intégrales divergentes ».

Démonstration 6.17

Montrons que si $\int_a^{+\infty} f$ converge et $\int_a^{+\infty} g$ diverge, alors $\int_a^{+\infty} (f + g)$ diverge.

Par l'absurde, si $\int_a^{+\infty} (f + g)$ converge alors comme $g = (f + g) - f$, d'après la proposition précédente, $\int_a^{+\infty} g$ converge : contradiction. ■

6.1.3 Cas des fonctions réelles positives

Quand une fonction f est positive et continue par morceaux, dans le cas où l'intégrale $\int_a^{+\infty} f$ diverge, on pose par convention $\int_a^{+\infty} f = +\infty$, ce qui permet de donner un sens à toutes les intégrales de fonctions positives.

Proposition 6.18

Soit $f \in \mathcal{C}_m^0([a; +\infty[, \mathbb{R})$.

Si $f \geq 0$, alors $\int_a^{+\infty} f \geq 0$.

Si, de plus, f est continue et prend au moins une valeur strictement positive, alors $\int_a^{+\infty} f > 0$.

Ceci est vrai en particulier quand f est continue et strictement positive sur $[a; +\infty[$.

Démonstration 6.19

- ▷ Si $f \geq 0$ sur $[a; +\infty[$, alors pour tout $x \in [a; +\infty[$, $f \geq 0$ sur $[a; x]$.

Donc par positivité de l'intégrale classique, $\int_a^x f \geq 0$.

Donc si $\int_a^{+\infty} f$ diverge, alors

$$\int_a^{+\infty} f = +\infty \geq 0$$

et si $\int_a^{+\infty} f$ converge, on a

$$\int_a^{+\infty} f \geq 0$$

par passage à la limite quand $x \rightarrow +\infty$.

- Si f est continue, positive et prend au moins une valeur strictement positive en x_0 , pour $x \geq x_0 + 1$, on a

$$\int_a^x f = \int_a^{x_0+1} f + \int_{x_0+1}^x f.$$

D'après le théorème de stricte positivité de l'intégrale classique, comme f est continue et positive sur $[a ; x_0 + 1]$, $f(x_0) > 0$ et $x_0 \in [a ; x_0 + 1]$, on a $\int_a^{x_0+1} f > 0$.

Or $\int_{x_0+1}^x f \geq 0$ donc $\int_a^x f \geq \int_a^{x_0+1} f$.

Par passage à la limite quand $x \rightarrow +\infty$, on obtient

$$\int_a^{+\infty} f \geq \int_a^{x_0+1} f > 0. \quad \blacksquare$$

Remarque 6.20 (Contraposée)

Si f est continue et positive sur $[a ; +\infty[$ et $\int_a^{+\infty} f = 0$, alors $f = 0$ sur $[a ; +\infty[$.

On en déduit la propriété de croissance des intégrales.

Proposition 6.21

Soit $(f, g) \in \mathcal{C}_m^0([a ; +\infty[, \mathbb{R})^2$ tel que f et g soient positives.

Si $f \leq g$, alors $\int_a^{+\infty} f \leq \int_a^{+\infty} g$.

6.1.4 Théorème de comparaison entre fonctions positives

D'abord une condition nécessaire et suffisante de convergence dans le cas d'une fonction positive.

Proposition 6.22

Soit $f \in \mathcal{C}_m^0([a ; +\infty[, \mathbb{R})$.

Si $f \geq 0$, alors l'intégrale $\int_a^{+\infty} f$ converge ssi la fonction $x \mapsto \int_a^x f$ est majorée.

Démonstration 6.23

Si $f \geq 0$ sur $[a ; +\infty[$, alors $x \mapsto \int_a^x f$ est croissante sur $[a ; +\infty[$.

Donc pour $(x, y) \in [a ; +\infty[^2$ tel que $x \geq y$, on a

$$\int_a^x f = \int_a^y f + \int_y^x f \geq \int_a^y f.$$

Donc $\int_a^{+\infty} f$ converge ssi la fonction croissante $x \mapsto \int_a^x f$ a une limite finie en $+\infty$ ssi la fonction croissante $x \mapsto \int_a^x f$ est majorée, par théorème de la limite monotone. ■

On en déduit un théorème de comparaison du même type que celui sur les séries.

Théorème 6.24

Soit $(f, g) \in \mathcal{C}_m^0([a ; +\infty[, \mathbb{R})^2$ tel que f et g soient positives.

- Si $f \leq g$ et $\int_a^{+\infty} g$ converge, alors $\int_a^{+\infty} f$ converge.
- Si $f \leq g$ et $\int_a^{+\infty} f$ diverge, alors $\int_a^{+\infty} g$ diverge.
- Si $f \underset{+\infty}{\sim} g$ alors les intégrales $\int_a^{+\infty} f$ et $\int_a^{+\infty} g$ sont de même nature : l'une converge ssi l'autre converge.

Démonstration 6.25

- Si $0 \leq f \leq g$ sur $[a ; +\infty[$, alors pour tout $x \in [a ; +\infty[$, $\int_a^x f \leq \int_a^x g$.

Donc si $\int_a^{+\infty} g$ converge, alors $x \mapsto \int_a^x g$ est majorée d'après la Proposition 6.22.

Donc $x \mapsto \int_a^x f$ est majorée.

Donc $\int_a^{+\infty} f$ converge d'après la Proposition 6.22.

- Par contraposée, si $\int_a^{+\infty} f$ diverge, alors $\int_a^{+\infty} g$ diverge.
- Si $f \underset{+\infty}{\sim} g$, il existe une fonction φ définie au voisinage de $+\infty$ telle que $f = g\varphi$ et $\varphi \xrightarrow{+\infty} 1$.

Comme $\varphi \xrightarrow{+\infty} 1$, au voisinage de $+\infty$, on a $\frac{1}{2} \leq \varphi \leq \frac{3}{2}$.

Donc, au voisinage de $+\infty$, on a $0 \leq \frac{1}{2}g \leq \varphi g = f \leq \frac{3}{2}g$.

D'après ce qui précède, si $\int_a^{+\infty} f$ converge, alors $\int_a^{+\infty} \frac{1}{2}g$ converge donc $\int_a^{+\infty} g = 2 \int_a^{+\infty} \frac{1}{2}g$ converge et si $\int_a^{+\infty} f$ diverge, alors $\int_a^{+\infty} \frac{3}{2}g$ diverge donc $\int_a^{+\infty} g = \frac{2}{3} \int_a^{+\infty} \frac{3}{2}g$ diverge. ■

Remarque 6.26

- ▷ Dans ce théorème, il suffit que les inégalités soient vraies au voisinage de $+\infty$ seulement.
- ▷ Si les fonctions sont à valeurs négatives, on se ramène à ce théorème en travaillant avec les fonctions opposées. Ce qui compte est donc qu'elles soient de signe constant.
- ▷ Avec des fonctions dont le signe n'est pas constant, ce théorème est faux. Il faut donc bien s'assurer et mettre en valeur que les fonctions sont positives (ou négatives).
- ▷ On compare les fonctions, pas les intégrales ! N'écrivez pas des symboles $\int_a^{+\infty} \dots$ partout.

Exercice 6.27

Montrez que $\int_1^{+\infty} \frac{e^{-t}}{t} dt$ converge.

Correction 6.28

Les fonctions $t \mapsto \frac{e^{-t}}{t}$ et $t \mapsto e^{-t}$ sont continues sur $[1 ; +\infty[$.

Pour tout $t \geq 1$, $0 \leq \frac{e^{-t}}{t} \leq e^{-t}$ or $\int_1^{+\infty} e^{-t} dt$ converge donc par comparaison de fonctions positives, $\int_1^{+\infty} \frac{e^{-t}}{t} dt$ converge.

Exercice 6.29

Montrez que $\int_2^{+\infty} \frac{1}{\ln t} dt$ diverge.

Correction 6.30

La fonction $t \mapsto \frac{1}{\ln t}$ est continue sur $[2 ; +\infty[$.

Pour $t \geq 2$, on a $\frac{1}{\ln t} \geq \frac{1}{t} \geq 0$ car $\forall t > 0$, $\ln t \leq t - 1 \leq t$ car \ln est concave.

Or $\int_2^{+\infty} \frac{1}{t} dt$ diverge donc par comparaison de fonctions positives, $\int_2^{+\infty} \frac{1}{\ln t} dt$ diverge.

Exercice 6.31

Montrez que pour tout $\alpha > 1$, $\int_1^{+\infty} \frac{\ln t}{t^\alpha} dt$ converge.

Correction 6.32

Pour tout $\alpha > 1$, $t \mapsto \frac{\ln t}{t^\alpha}$ est continue sur $[1; +\infty[$.

- Si $\alpha > 2$, comme on a $0 \leq \frac{\ln t}{t^\alpha} \leq \frac{t}{t^\alpha} = \frac{1}{t^{\alpha-1}}$ pour $t \in [1; +\infty[$ et $\int_1^{+\infty} \frac{1}{t^{\alpha-1}} dt$ converge (car $\alpha - 1 > 1$), on en déduit que $\int_1^{+\infty} \frac{\ln t}{t^\alpha} dt$ converge.
- Il existe $a > 1$ tel que pour tout $t \in [a; +\infty[$, $0 \leq \ln t \leq t^{\alpha-1/2}$ car $\ln t \underset{t \rightarrow +\infty}{=} o(t^{\alpha-1/2})$ car $\frac{\alpha-1}{2} > 0$.

Donc pour tout $t \in [a; +\infty[$, $0 \leq \frac{\ln t}{t^\alpha} \leq \frac{t^{\alpha-1/2}}{t^\alpha} = \frac{1}{t^{\alpha+1/2}}$.

Comme $\alpha > 1$, $\frac{\alpha+1}{2} > 1$ donc $\int_a^{+\infty} \frac{1}{t^{\alpha+1/2}} dt$ converge.

Donc $\int_a^{+\infty} \frac{\ln t}{t^\alpha} dt$ converge *i.e.* $\int_1^{+\infty} \frac{\ln t}{t^\alpha} dt$ converge.

6.1.5 Lien avec les séries

Le théorème de comparaison série-intégrale peut se réécrire comme suit.

Proposition 6.33

Soit $f \in \mathcal{C}_m^0(\mathbb{R}_+, \mathbb{R})$ positive et décroissante.

La série $\sum f(n)$ et l'intégrale $\int_0^{+\infty} f$ sont de même nature.

Dans le même genre, on peut étudier la convergence d'une intégrale d'une fonction positive par l'intermédiaire d'une série.

Proposition 6.34

Soient $f \in \mathcal{C}_m^0(\mathbb{R}_+, \mathbb{R})$ positive et u une suite positive et strictement croissante qui diverge vers $+\infty$.

L'intégrale $\int_0^{+\infty} f$ converge ssi la série $\sum_{n \geq 0} \int_{u_n}^{u_{n+1}} f$ converge.

Démonstration 6.35

\Rightarrow

Si $f \geq 0$ sur $[0 ; +\infty[$, alors $F : x \mapsto \int_0^x f$ est croissante.

Donc si $\int_0^{+\infty} f$ converge, alors $\int_0^x f$ a une limite réelle ℓ quand $x \rightarrow +\infty$.

Donc comme $u_n \xrightarrow[n \rightarrow +\infty]{} +\infty$, on a $\int_0^{u_n} f \xrightarrow[n \rightarrow +\infty]{} \ell$.

Or $\int_0^{u_n} f = \int_0^{u_0} f + \int_{u_0}^{u_1} f + \cdots + \int_{u_{n-1}}^{u_n} f$ donc

$$\sum_{k=0}^{n-1} \int_{u_k}^{u_{k+1}} f \xrightarrow[n \rightarrow +\infty]{} \ell - \int_0^{u_0} f$$

i.e. la série $\sum_n \int_{u_n}^{u_{n+1}} f$ converge.

\Leftarrow

Si la série $\sum_n \int_{u_n}^{u_{n+1}} f$ converge, $\sum_{k=0}^{n-1} \int_{u_k}^{u_{k+1}} f = F(u_n) - F(u_0)$ a une limite réelle quand $n \rightarrow +\infty$ i.e. $(F(u_n))$ a une limite réelle.

Or F est croissante donc d'après le théorème de la limite monotone, F a une limite $L \in \mathbb{R} \cup \{+\infty\}$ en $+\infty$.

D'après la caractérisation séquentielle de la limite, on a $F(u_n) \xrightarrow[n \rightarrow +\infty]{} L$ car $u_n \xrightarrow[n \rightarrow +\infty]{} +\infty$.

Donc $L \in \mathbb{R}$. Donc $\int_0^{+\infty} f$ converge. ■

Remarque 6.36

L'hypothèse de positivité de f est essentielle.

Contre-exemple si on ne suppose rien sur le signe de f :

Pour tout $n \in \mathbb{N}$, $\int_{2n\pi}^{2(n+1)\pi} \sin = 0$ donc $\sum \int_{2n\pi}^{2(n+1)\pi} \sin$ converge mais $\int_0^x \sin = 1 - \cos x$ n'a pas de limite quand $x \rightarrow +\infty$ i.e. $\int_0^{+\infty} \sin$ diverge.

Exercice 6.37

Montrez que l'intégrale $\int_1^{+\infty} \frac{|\sin t|}{t} dt$ diverge.

Correction 6.38

La fonction $t \mapsto \frac{|\sin t|}{t}$ est continue sur $[1 ; +\infty[$.

D'après la Proposition 6.34, $\int_1^{+\infty} \frac{|\sin t|}{t} dt$ converge ssi la série $\sum_n \int_{u_n}^{u_{n+1}} \frac{|\sin t|}{t} dt$ converge, où (u_n) est une suite positive et strictement croissante.

On prend $u_n = n\pi$ pour $n \geq 1$.

Pour $n \in \mathbb{N}^*$, on a $\int_{n\pi}^{(n+1)\pi} \frac{|\sin t|}{t} dt \geq \int_{n\pi}^{(n+1)\pi} \frac{|\sin t|}{(n+1)\pi} dt$ car

$$\begin{aligned} \forall t \in [n\pi ; (n+1)\pi], \quad n\pi \leq t \leq (n+1)\pi \text{ donc } \frac{1}{(n+1)\pi} \leq \frac{1}{t} \leq \frac{1}{n\pi} \\ \text{donc } \frac{|\sin t|}{(n+1)\pi} \leq \frac{|\sin t|}{t} \leq \frac{|\sin t|}{n\pi}. \end{aligned}$$

Donc on a

$$\begin{aligned} \int_{n\pi}^{(n+1)\pi} \frac{|\sin t|}{t} dt &\geq \frac{1}{(n+1)\pi} \int_{n\pi}^{(n+1)\pi} |\sin t| dt \\ &= \frac{1}{(n+1)\pi} \int_0^\pi |\sin t| dt \quad \left. \begin{array}{l} \text{ } \\ \text{ } \end{array} \right\} t \mapsto |\sin t| \text{ est } \pi\text{-périodique} \\ &= \frac{2}{(n+1)\pi}. \end{aligned}$$

Or $\sum_n \frac{2}{(n+1)\pi}$ diverge donc $\int_1^{+\infty} \frac{|\sin t|}{t} dt$ diverge.

Exercice 6.39

En utilisant l'inégalité $\sin t \geq \frac{2}{\pi}t$, valable pour tout $t \in [0 ; \frac{\pi}{2}]$, montrez que l'intégrale $\int_0^{+\infty} \frac{1}{1 + e^x |\sin x|} dx$ converge.

Correction 6.40

La fonction $x \mapsto \frac{1}{1 + e^x |\sin x|}$ est continue et positive sur $[0 ; +\infty[$ donc $\int_0^{+\infty} \frac{1}{1 + e^x |\sin x|} dx$ converge

ssi la série $\sum_n \int_{n\pi}^{(n+1)\pi} \frac{1}{1 + e^x |\sin x|} dx$ converge.

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

$$\begin{aligned} 0 \leq \int_{n\pi}^{(n+1)\pi} \frac{1}{1 + e^x |\sin x|} dx &\quad \left. \begin{array}{l} \text{ } \\ \text{ } \end{array} \right\} \begin{cases} x = t + n\pi \\ dx = dt \end{cases} \\ = \int_0^\pi \frac{1}{1 + e^{t+n\pi} |\sin t|} dt. \end{aligned}$$

Or, comme on a $\forall t \in \left[0 ; \frac{\pi}{2}\right]$, $\frac{2}{\pi}t \leq \sin t$, on a

$$1 + e^{n\pi} \frac{2}{\pi} t \leq 1 + e^{t+n\pi} \frac{2}{\pi} t \leq 1 + e^{t+n\pi} \sin t$$

donc

$$\begin{aligned} \int_0^{\pi/2} \frac{1}{1 + e^{t+n\pi} |\sin t|} dt &\leq \int_0^{\pi/2} \frac{1}{1 + e^{t+n\pi} \frac{2}{\pi} t} dt \\ &\leq \int_0^{\pi/2} \frac{1}{1 + e^{n\pi} \frac{2}{\pi} t} dt \\ &= \left[\frac{\ln \left(1 + e^{n\pi} \frac{2}{\pi} t \right)}{e^{n\pi} \frac{2}{\pi}} \right]_0^{\pi/2} \\ &= \frac{\pi}{2} \frac{\ln (1 + e^{n\pi})}{e^{n\pi}}. \end{aligned}$$

De même, on a $\forall t \in \left[\frac{\pi}{2} ; \pi\right]$, $\frac{2}{\pi}(\pi - t) \leq \sin t$ donc

$$\begin{aligned} \int_{\pi/2}^{\pi} \frac{1}{1 + e^{t+n\pi} \sin t} dt &\leq \int_{\pi/2}^{\pi} \frac{1}{1 + e^{n\pi} \frac{2}{\pi} (\pi - t)} dt \\ &= \left[\frac{\ln \left(1 + e^{n\pi} \frac{2}{\pi} (\pi - t) \right)}{-e^{n\pi} \frac{2}{\pi}} \right]_{\pi/2}^{\pi} \\ &= \frac{\pi}{2} \frac{\ln (1 + e^{n\pi})}{e^{n\pi}}. \end{aligned}$$

Ainsi, pour tout $n \in \mathbb{N}$, on a

$$0 \leq \int_{n\pi}^{(n+1)\pi} \frac{1}{1 + e^x |\sin x|} dx \leq \pi \frac{\ln (1 + e^{n\pi})}{e^{n\pi}} = a_n.$$

Or $1 + e^{n\pi} \underset{n \rightarrow +\infty}{\sim} e^{n\pi} \xrightarrow[n \rightarrow +\infty]{} +\infty$ donc $\ln (1 + e^{n\pi}) \underset{n \rightarrow +\infty}{\sim} \ln (e^{n\pi}) = n\pi$.

Donc $a_n \underset{n \rightarrow +\infty}{\sim} \pi^2 \frac{n}{e^{n\pi}} \underset{n \rightarrow +\infty}{=} o\left(\frac{1}{n^2}\right)$.

Or $\sum \frac{1}{n^2}$ converge donc $\sum a_n$ converge.

Donc $\sum \int_{n\pi}^{(n+1)\pi} \frac{1}{1 + e^x |\sin x|} dx$ converge.

Donc $\int_0^{+\infty} \frac{1}{1 + e^x |\sin x|} dx$ converge.

6.2 Intégrales généralisées sur d'autres types d'intervalles

6.2.1 Intégrales généralisées sur $[a ; b[$

Dans cette partie, a est un réel et b est un réel ou $+\infty$, de sorte que $a < b$.

Définition 6.41

Soit $f \in \mathcal{C}_m^0([a ; b[, \mathbb{K})$.

On dit que l'intégrale $\int_a^b f$ converge quand $\int_a^x f$ a une limite finie quand x tend vers b^- .

Dans ce cas, on pose $\int_a^b f = \lim_{x \rightarrow b^-} \int_a^x f$.

Dans le cas contraire, on dit que l'intégrale $\int_a^b f$ diverge (ou qu'elle est divergente).

6.2.2 Intégrales généralisées sur $]a ; b]$

Dans cette partie, a est un réel ou $-\infty$ et b un réel, de sorte que $a < b$.

Définition 6.42

Soit $f \in \mathcal{C}_m^0(]a ; b], \mathbb{K})$.

On dit que l'intégrale $\int_a^b f$ converge quand $\int_x^b f$ a une limite finie quand x tend vers a^+ .

Dans ce cas, on pose $\int_a^b f = \lim_{x \rightarrow a^+} \int_x^b f$.

Dans le cas contraire, on dit que l'intégrale $\int_a^b f$ diverge (ou qu'elle est divergente).

Les résultats suivants sont à connaître.

Exemple 6.43

▷ $\int_0^1 \ln(t) dt$ converge.

▷ Soit α un réel. L'intégrale $\int_0^1 \frac{1}{t^\alpha} dt$ converge ssi $\alpha < 1$ (intégrale dite de Riemann).

Démonstration 6.44

► La fonction \ln est continue sur $]0 ; 1]$.

Pour $x \in]0 ; 1]$, on a

$$\int_x^1 \ln(t) dt = [t \ln t - t]_x^1 = -1 - x \ln x + x.$$

Or $\lim_{x \rightarrow 0} x \ln x = 0$ donc $\int_x^1 \ln(t) dt \xrightarrow{x \rightarrow 0} -1$.

Donc $\int_0^1 \ln(t) dt$ converge et vaut -1 .

► La fonction $t \mapsto \frac{1}{t^\alpha}$ est continue sur $]0 ; 1]$.

Pour $x \in]0 ; 1]$, on a

$$\int_x^1 \frac{1}{t^\alpha} dt = \int_x^1 t^{-\alpha} dt = \left[\frac{1}{1-\alpha} t^{1-\alpha} \right]_x^1 = \frac{1}{1-\alpha} (1 - x^{1-\alpha}).$$

Donc

$$\begin{cases} \int_x^1 \frac{1}{t^\alpha} dt \xrightarrow{x \rightarrow 0} \frac{1}{1-\alpha} & \text{si } \alpha < 1 \\ \int_x^1 \frac{1}{t^\alpha} dt \xrightarrow{x \rightarrow 0} +\infty & \text{si } \alpha > 1 \\ \int_x^1 \frac{1}{t^\alpha} dt = -\ln x \xrightarrow{x \rightarrow 0} +\infty & \text{si } \alpha = 1 \end{cases}$$

Donc $\int_0^1 \frac{1}{t^\alpha} dt$ converge ssi $\alpha < 1$ et vaut $\frac{1}{1-\alpha}$. ■

Exercice 6.45

Montrez que l'intégrale $\int_1^2 \frac{1}{\ln t} dt$ diverge.

À quelle condition sur α l'intégrale $\int_1^2 \frac{(t-1)^\alpha}{\ln t} dt$ converge-t-elle ?

Correction 6.46

► La fonction $t \mapsto \frac{1}{\ln t}$ est continue sur $]1 ; 2]$.

On veut montrer que $\int_1^2 \frac{1}{\ln t} dt$ diverge *i.e.* $\int_x^2 \frac{1}{\ln t} dt$ n'a pas de limite réelle quand $x \rightarrow 1^+$.

On effectue le changement de variable $\begin{cases} s = \ln t \\ ds = \frac{dt}{t} \end{cases} \iff \begin{cases} t = e^s \\ dt = e^s ds \end{cases}$

On a ramené le problème à la non-existence d'une limite réelle de $\int_X^{\ln 2} \frac{e^s}{s} ds$ quand $X \rightarrow 0$

i.e. à la divergence de $\int_0^{\ln 2} \frac{e^s}{s} ds$.

Or pour tout $s > 0$, $\frac{e^s}{s} \geq \frac{1}{s} \geq 0$ et $\int_0^{\ln 2} \frac{1}{s} ds$ diverge.

Donc par comparaison de fonctions positives, $\int_0^{\ln 2} \frac{e^s}{s} ds$ diverge *i.e.* $\int_1^2 \frac{1}{\ln t} dt$ diverge.

► Autre méthode :

Pour tout $t > 1$, $\ln t \leq t - 1$ donc $0 \leq \frac{1}{t-1} \leq \frac{1}{\ln t}$.

Or $\int_1^2 \frac{1}{t-1} dt = \int_0^1 \frac{1}{t} dt$ diverge.

Donc par comparaison de fonctions positives, $\int_1^2 \frac{1}{\ln t} dt$ diverge.

► La fonction $t \mapsto \frac{(t-1)^\alpha}{\ln t}$ est continue et positive sur $]1; 2]$.

On a $\frac{(t-1)^\alpha}{\ln t} \underset{t \rightarrow 1}{\sim} (t-1)^{\alpha-1} = \frac{1}{(t-1)^{1-\alpha}}$ car $\ln t \underset{t \rightarrow 1}{\sim} t-1$.

Donc $\int_1^2 \frac{(t-1)^\alpha}{\ln t} dt$ converge ssi $1-\alpha < 1$ *i.e.* $\alpha > 0$.

On peut remarquer que par changement de variable $x \mapsto -x$, l'intégrale $\int_a^b f(t) dt$ est de même nature (et, dans le cas convergent, a la même valeur) que $\int_{-b}^{-a} f(-u) du$. Les résultats valables en un point réel ne dépendent donc pas du côté du point où on se place.

Exemple 6.47

► Si a est un réel, alors $\int_a^{a+1} \frac{1}{(t-a)^\alpha} dt$ converge ssi $\alpha < 1$.

► Si a est un réel, alors $\int_{a-1}^a \frac{1}{(a-t)^\alpha} dt$ converge ssi $\alpha < 1$.

► $\int_0^1 \ln(1-t) dt$ converge.

6.2.3 Intégrales généralisées sur $]a; b[$

Dans cette partie, a et b sont des réels ou des infinis tels que $a < b$.

Définition 6.48

Soit $f \in \mathcal{C}_m^0(]a; b[, \mathbb{K})$.

On dit que l'intégrale $\int_a^b f$ converge quand il existe $c \in]a; b[$ tel que $\int_a^c f$ et $\int_c^b f$ convergent.

Dans ce cas, on pose $\int_a^b f = \int_a^c f + \int_c^b f = \lim_{x \rightarrow a^+} \int_x^c f + \lim_{y \rightarrow b^-} \int_c^y f$.

Dans le cas contraire, on dit que l'intégrale $\int_a^b f$ diverge (ou qu'elle est divergente).

Remarque 6.49

Grâce à la relation de Chasles, on constate que la valeur de c n'est finalement pas importante : si ça marche pour un certain réel $c \in]a ; b[$, alors ça marche pour toute autre valeur prise dans $]a ; b[$.

Exercice 6.50

Montrez que l'intégrale de Gauss $\int_{-\infty}^{+\infty} e^{-t^2} dt$ converge.

Correction 6.51

La fonction $t \mapsto e^{-t^2}$ est continue et positive sur $]-\infty ; +\infty[$.

$\int_1^{+\infty} e^{-t^2} dt$ converge car pour tout $t \in [1 ; +\infty[$, $t^2 \geq t$ donc $0 \leq e^{-t^2} \leq e^{-t}$ et $\int_1^{+\infty} e^{-t} dt$ converge.

Or $t \mapsto e^{-t^2}$ est paire donc $\int_{-\infty}^{-1} e^{-t^2} dt$ est de même nature que $\int_1^{+\infty} e^{-t^2} dt$ donc elle converge.

Finalement, $\int_{-\infty}^{+\infty} e^{-t^2} dt$ converge.

Exercice 6.52

Même chose avec l'intégrale $\int_0^1 \frac{1}{\sqrt{t(1-t)}} dt$.

Correction 6.53

La fonction $t \mapsto \frac{1}{\sqrt{t(1-t)}}$ est continue et positive sur $]0 ; 1[$.

Pour tout $t \in \left]0 ; \frac{1}{2}\right]$, $0 \leq \frac{1}{\sqrt{t(1-t)}} \leq \frac{2}{\sqrt{t}} = \frac{2}{t^{1/2}}$.

Or $\int_0^{1/2} \frac{1}{t^{1/2}} dt$ converge donc $\int_0^{1/2} \frac{1}{\sqrt{t(1-t)}} dt$ converge.

Pour tout $t \in \left[\frac{1}{2} ; 1\right[$, $0 \leq \frac{1}{\sqrt{t(1-t)}} \leq \frac{2}{\sqrt{1-t}} = \frac{2}{(1-t)^{1/2}}$.

Or $\int_{1/2}^1 \frac{1}{(1-t)^{1/2}} dt$ converge donc $\int_{1/2}^1 \frac{1}{\sqrt{t(1-t)}} dt$ converge.

Finalement, $\int_0^1 \frac{1}{\sqrt{t(1-t)}} dt$ converge.

Exercice 6.54

Même chose avec l'intégrale $\int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}} dt$.

Correction 6.55

La fonction $t \mapsto \frac{e^{-t}}{\sqrt{t}}$ est continue et positive sur $]0 ; +\infty[$.

Pour tout $t \geq 1$, $0 \leq \frac{e^{-t}}{\sqrt{t}} \leq e^{-t}$ et $\int_1^{+\infty} e^{-t} dt$ converge donc $\int_1^{+\infty} \frac{e^{-t}}{\sqrt{t}} dt$ converge.

On a $\frac{e^{-t}}{\sqrt{t}} \underset{t \rightarrow 0}{\sim} \frac{1}{\sqrt{t}}$ et $\int_0^1 \frac{1}{\sqrt{t}} dt$ converge donc $\int_0^1 \frac{e^{-t}}{\sqrt{t}} dt$ converge.

Finalement, $\int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}} dt$ converge.

Exercice 6.56

Montrez que pour tout $\alpha > 0$, $\int_0^{+\infty} \frac{1}{t^\alpha} dt$ diverge.

Correction 6.57

Clair.

6.2.4 Propriétés communes à toutes ces intégrales

Toutes les propriétés vues dans la première section sont préservées : linéarité, positivité, croissance, relation de Chasles.

En voici trois autres.

6.2.4.1 Changement de variable

Proposition 6.58

Soient a, b, α, β des réels ou des infinis tels que $a < b$ et $\alpha < \beta$ et $f \in \mathcal{C}_m^0(]a ; b[, \mathbb{K})$.

Si φ est une bijection de classe \mathcal{C}^1 et strictement croissante de $]a ; \beta[$ dans $]a ; b[$, alors les intégrales $\int_a^b f(t) dt$ et $\int_\alpha^\beta f \circ \varphi(u) \varphi'(u) du$ sont de même nature et, si elles sont convergentes, sont égales.

Démonstration 6.59

Soit $(x, y) \in \mathbb{R}^2$ tel que $a < x \leq y < b$.

D'après la formule classique du changement de variable, on a

$$\int_x^y f(t) dt = \int_{\varphi^{-1}(x)}^{\varphi^{-1}(y)} f(\varphi(u)) \varphi'(u) du.$$

On a $\varphi^{-1}(x) \xrightarrow{x \rightarrow a} \alpha$ et $\varphi^{-1}(y) \xrightarrow{y \rightarrow b} \beta$.

Donc

$$\begin{aligned} \int_a^b f \text{ converge} &\iff \int_x^y f \text{ a une limite finie quand } x \rightarrow a^+ \text{ et } y \rightarrow b^- \\ &\iff \int_{\varphi^{-1}(x)}^{\varphi^{-1}(y)} f(\varphi(u)) \varphi'(u) du \text{ a une limite finie quand } x \rightarrow a^+ \text{ et } y \rightarrow b^- \\ &\iff \int_\lambda^\mu f(\varphi(u)) \varphi'(u) du \text{ a une limite finie quand } \lambda \rightarrow \alpha^+ \text{ et } \mu \rightarrow \beta^- \\ &\iff \int_\alpha^\beta f(\varphi(u)) \varphi'(u) du \text{ converge.} \end{aligned}$$

■

Remarque 6.60

On a évidemment un résultat analogue avec un changement de variable strictement décroissant et des bornes inversées.

Comme une bijection de classe \mathcal{C}^1 entre deux intervalles est forcément strictement monotone, l'hypothèse de stricte monotonie est redondante ; mais comme elle est explicitement dans le programme de MPI, il vaut mieux la préciser (de toute façon, elle sera évidente dans les cas pratiques et ne nécessitera pas de longues preuves).

Exercice 6.61

Montrez que $\int_0^{\pi/2} \sqrt{\tan t} dt$ converge et qu'on a l'égalité $\int_0^{\pi/2} \sqrt{\tan t} dt = \int_0^{+\infty} \frac{\sqrt{u}}{1+u^2} du$.

Correction 6.62

On effectue le changement de variable $u = \tan t$.

Sous réserve de convergence, on a

$$\int_0^{\pi/2} \sqrt{\tan t} dt = \int_0^{+\infty} \frac{\sqrt{u}}{1+u^2} du.$$

Or, comme $u \mapsto \frac{\sqrt{u}}{1+u^2}$ est continue et positive sur $[0; +\infty[$ et $\frac{\sqrt{u}}{1+u^2} \underset{u \rightarrow +\infty}{\sim} \frac{1}{u^{3/2}}$ et $\frac{3}{2} > 1$, $\int_0^{+\infty} \frac{\sqrt{u}}{1+u^2} du$ converge.

Exercice 6.63

Montrez que $\int_0^{+\infty} \frac{\ln t}{1+t^2} dt$ converge et vaut 0.

Déduisez-en la valeur de $\int_0^{+\infty} \frac{\ln t}{a^2+t^2} dt$ où $a > 0$.

Correction 6.64

► La fonction $\varphi : t \mapsto \frac{\ln t}{1+t^2}$ est continue sur $]0 ; +\infty[$.

Sur $]0 ; 1]$, φ est négative et sur $[1 ; +\infty[$, φ est positive.

On a $\varphi(t) \underset{t \rightarrow 0}{\sim} \ln t$ et $\int_0^1 \ln(t) dt$ converge donc $\int_0^1 \varphi(t) dt$ converge par comparaison de fonctions à signe constant.

On a $\ln t \underset{t \rightarrow +\infty}{=} o(\sqrt{t})$ donc il existe $A > 0$ tel que pour tout $t \geq A$, $0 \leq \ln t \leq \sqrt{t}$ donc

$$0 \leq \frac{\ln t}{1+t^2} \leq \frac{\sqrt{t}}{1+t^2} \leq \frac{\sqrt{t}}{t^2} = \frac{1}{t^{3/2}}$$

.

Or $\frac{3}{2} > 1$ donc $\int_A^{+\infty} \frac{1}{t^{3/2}} dt$ converge donc $\int_A^{+\infty} \frac{\ln t}{1+t^2} dt$ converge et donc $\int_1^{+\infty} \frac{\ln t}{1+t^2} dt$ converge.

On veut montrer que $\int_0^1 \frac{\ln t}{1+t^2} dt = - \int_1^{+\infty} \frac{\ln t}{1+t^2} dt$ de sorte que $\int_0^{+\infty} \frac{\ln t}{1+t^2} dt = 0$.

En effectuant le changement de variable $\begin{cases} t = \frac{1}{u} \\ dt = -\frac{1}{u^2} du \end{cases}$ (où $t \mapsto \frac{1}{t}$ est une bijection \mathcal{C}^1 strictement décroissante de $]0 ; 1]$ dans $[1 ; +\infty[$), on obtient

$$\int_0^1 \frac{\ln t}{1+t^2} dt = \int_{+\infty}^1 \frac{\ln 1/u}{1+(1/u)^2} \left(\frac{-1}{u^2} \right) du = - \int_1^{+\infty} \frac{\ln u}{1+u^2} du.$$

D'où l'égalité.

$$\text{Donc } \int_0^{+\infty} \frac{\ln t}{1+t^2} dt = 0.$$

► Soit $a > 0$.

$$\int_0^{+\infty} \frac{\ln t}{a^2+t^2} dt \text{ converge (idem).}$$

On a

$$\begin{aligned}
 \int_0^{+\infty} \frac{\ln t}{a^2 + t^2} dt &= \frac{1}{a^2} \int_0^{+\infty} \frac{\ln t}{1 + (t/a)^2} dt \\
 &= \frac{1}{a^2} \int_0^{+\infty} \frac{a \ln(au)}{1 + u^2} du \quad \left. \begin{array}{l} t = au \\ dt = a du \end{array} \right\} \\
 &= \frac{1}{a} \int_0^{+\infty} \frac{\ln a}{1 + u^2} du + \underbrace{\frac{1}{a} \int_0^{+\infty} \frac{\ln u}{1 + u^2} du}_{=0} \\
 &= \frac{\pi \ln a}{2a}.
 \end{aligned}$$

6.2.4.2 Intégration par parties

Si f est une fonction définie sur $]a ; b[$ et a des limites finies en a^+ et en b^- , on note $[f]_a^b = \lim_{b^-} f - \lim_{a^+} f$.

Proposition 6.65

Soient a, b des réels ou des infinis tels que $a < b$ et $(f, g) \in \mathcal{C}^1(]a ; b[, \mathbb{K})$.

Si parmi les trois quantités suivantes

$$\int_a^b f'g, \quad \int_a^b fg', \quad [fg]_a^b$$

deux existent, alors la troisième existe aussi et, dans ce cas, on a l'égalité habituelle

$$\int_a^b f'g = [fg]_a^b - \int_a^b fg'$$

.

Démonstration 6.66

Soit $(x, y) \in \mathbb{R}^2$ tel que $a < x \leq y < b$.

On a $\int_x^y fg' = [fg]_x^y - \int_x^y f'g$.

Si parmi les quantités $\int_x^y fg'$, $\int_x^y f'g$ et $[fg]_x^y = f(y)g(y) - f(x)g(x)$ deux ont des limites finies quand $x \rightarrow a$ et $y \rightarrow b$, alors la troisième aussi (par opération sur les limites).

Dans ce cas, les limites sont égales :

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

■

En pratique, pour éviter d'écrire des choses qui n'ont pas de sens, il vaut mieux revenir à une vraie intégration par parties sur un segment $[x ; y] \subseteq]a ; b[$, s'assurer qu'on peut faire tendre x vers a et y vers b , puis le faire effectivement pour obtenir la relation entre les intégrales.

Exercice 6.67

Pour $n \in \mathbb{N}$, on pose $u_n = \int_0^{+\infty} t^n e^{-t} dt$.

Montrez que les intégrales u_n convergent, donnez une relation de récurrence simple entre u_n et u_{n+1} , puis donnez la valeur de u_n en fonction de n .

Correction 6.68

Pour $n \in \mathbb{N}$, la fonction $t \mapsto t^n e^{-t}$ est continue sur $[0 ; +\infty[$.

Sous réserve de convergence, on pose $u_n = \int_0^{+\infty} t^n e^{-t} dt$ pour $n \in \mathbb{N}$.

On a $t^{n+2} \underset{t \rightarrow +\infty}{=} o(e^t)$ donc $t^n e^{-t} \underset{t \rightarrow +\infty}{=} o\left(\frac{1}{t^2}\right)$.

Donc il existe $a > 0$ tel que $\forall t \geq a, 0 \leq t^n e^{-t} \leq \frac{1}{t^2}$.

Or $\int_a^{+\infty} \frac{1}{t^2} dt$ converge donc $\int_a^{+\infty} t^n e^{-t} dt$ converge et donc $\int_0^{+\infty} t^n e^{-t} dt$ converge.

Sous réserve de convergence, par intégration par parties, on a

$$\int_0^{+\infty} t^{n+1} e^{-t} dt = [-t^{n+1} e^{-t}]_0^{+\infty} + \int_0^{+\infty} (n+1) t^n e^{-t} dt.$$

Les deux intégrales convergent d'après ce qui précède donc l'intégration par parties est licite.

Ainsi, comme $[-t^{n+1} e^{-t}]_0^{+\infty} = 0$, on a $u_{n+1} = (n+1) u_n$.

Or $u_0 = \int_0^{+\infty} e^{-t} dt = 1$.

Donc $\forall n \in \mathbb{N}, u_n = n!$ (récurrence immédiate).

Exercice 6.69

Montrez que l'intégrale $\int_0^{+\infty} \frac{1 - \cos t}{t^2} dt$ converge, puis déduisez-en que l'intégrale $\int_0^{+\infty} \frac{\sin t}{t} dt$ converge.

Correction 6.70

► La fonction $t \mapsto \frac{1 - \cos t}{t^2}$ est continue sur $]0 ; +\infty[$.

Pour tout $t > 0$, $-1 \leq \cos t \leq 1$ donc $0 \leq 1 - \cos t \leq 2$ donc $0 \leq \frac{1 - \cos t}{t^2} \leq \frac{2}{t^2}$.

Or $\int_1^{+\infty} \frac{1}{t^2} dt$ converge donc $\int_1^{+\infty} \frac{1 - \cos t}{t^2} dt$ converge.

On a $\cos t \underset{t \rightarrow 0}{=} 1 - \frac{t^2}{2} + o(t^2)$ donc $1 - \cos t \underset{t \rightarrow 0}{\sim} \frac{t^2}{2}$ donc $\frac{1 - \cos t}{t^2} \underset{t \rightarrow 0}{\sim} \frac{1}{2}$.

Or $\int_0^1 \frac{1}{2} dt$ converge donc $\int_0^1 \frac{1 - \cos t}{t^2} dt$ converge.

Finalement, $\int_0^{+\infty} \frac{1 - \cos t}{t^2} dt$ converge.

► Soit $(x, y) \in \mathbb{R}^2$ tel que $0 < x \leq y$.

On a

$$\int_x^y \frac{1 - \cos t}{t^2} dt = \left[-\frac{1 - \cos t}{t} \right]_x^y + \int_x^y \frac{\sin t}{t} dt.$$

On sait déjà que $\int_x^y \frac{1 - \cos t}{t^2} dt$ a une limite réelle quand $x \rightarrow 0$ et $y \rightarrow +\infty$.

De plus, on a $\left[-\frac{1 - \cos t}{t} \right]_x^y = \frac{1 - \cos x}{x} - \frac{1 - \cos y}{y}$.

Or $\frac{1 - \cos x}{x} \underset{x \rightarrow 0}{\sim} \frac{x}{2}$ donc $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x} = 0$ et $\frac{1 - \cos y}{y} \underset{y \rightarrow +\infty}{=} \mathcal{O}\left(\frac{1}{y}\right)$ donc $\lim_{y \rightarrow +\infty} \frac{1 - \cos y}{y} = 0$ par encadrement.

Donc $\int_x^y \frac{\sin t}{t} dt$ a une limite réelle quand $x \rightarrow 0$ et $y \rightarrow +\infty$ i.e. $\int_0^{+\infty} \frac{\sin t}{t} dt$ converge (NB : elle se nomme intégrale de Dirichlet).

► On peut aussi montrer que $\int_0^{+\infty} \frac{\sin t}{t} dt = \int_0^{+\infty} \frac{\sin^2 t}{t^2} dt$.

En effet, on a

$$\begin{aligned} \int_0^{+\infty} \frac{\sin t}{t} dt &= \int_0^{+\infty} \frac{1 - \cos t}{t^2} dt \\ &= \int_0^{+\infty} \frac{2 \sin^2 t/2}{t^2} dt \\ &= \int_0^{+\infty} \frac{2 \sin^2 u}{4u^2} du \quad \left\{ \begin{array}{l} t = 2u \\ dt = 2 du \end{array} \right. \\ &= \int_0^{+\infty} \frac{\sin^2 u}{u^2} du. \end{aligned}$$

6.2.4.3 Primitives

Proposition 6.71

Soient $f \in \mathcal{C}_m^0(]a ; b[, \mathbb{K})$ et $c \in]a ; b[$ tel que l'intégrale $\int_a^c f$ converge.

Alors la fonction $x \mapsto \int_a^x f$ est l'unique primitive de f sur $]a ; b[$ qui a pour limite 0 en a^+ .

Démonstration 6.72

La fonction f est continue sur $]a ; b[$ donc elle possède une primitive F sur $]a ; b[$. Par exemple $F : x \mapsto \int_c^x f$ où $x \in]a ; b[$.

$\int_a^c f$ converge donc F a une limite finie en a^+ .

Donc

$$\begin{aligned}\forall x \in]a ; b[, \quad \int_a^x f &= \lim_{y \rightarrow a^+} \int_y^x f \\ &= \lim_{y \rightarrow a^+} (F(x) - F(y)) \\ &= F(x) - \underbrace{\lim_{y \rightarrow a^+} F(y)}_{\text{constante}}.\end{aligned}$$

Donc $x \mapsto \int_a^x f$ est une primitive de f sur $]a ; b[$.

De plus, $\lim_{x \rightarrow a^+} \int_a^x f = \lim_{x \rightarrow a^+} F(x) - \lim_{y \rightarrow a^+} F(y) = 0$. ■

Proposition 6.73

Soient $f \in \mathcal{C}_m^0(]a ; b[, \mathbb{K})$ et $c \in]a ; b[$ tel que $\int_c^b f$ converge.

Alors la fonction $x \mapsto \int_x^b f$ est l'opposée de l'unique primitive de f sur $]a ; b[$ qui a pour limite 0 en b^- .

Exemple 6.74

► La fonction $x \mapsto \int_0^x \ln(t) dt$ est la primitive de \ln qui a pour limite 0 en 0.

► La fonction $x \mapsto \int_x^{+\infty} \frac{e^{it}}{t^2} dt$ est définie sur $]0 ; +\infty[$, de classe \mathcal{C}^1 sur cet intervalle, et sa dérivée est la fonction $x \mapsto -\frac{e^{ix}}{x^2}$.

6.3 Résumé pour étudier la convergence d'une intégrale

On veut savoir si une intégrale $\int_a^b f$ existe, où a et b sont des réels ou des infinis tels que $a < b$.

D'abord, on détermine le plus grand sous-ensemble de $[a ; b]$ sur lequel f est continue par morceaux :

- si c'est $[a ; b]$, alors il n'y a aucun problème d'existence de l'intégrale : c'est une bête intégrale classique ;
- si c'est $[a ; b[$ (avec a réel) ou $]a ; b]$ (avec b réel), alors il faut étudier le comportement de f au voisinage du point ouvert ;
- si c'est $]a ; b[$, alors on choisit arbitrairement un point $c \in]a ; b[$ et on se ramène deux fois au cas précédent.

Un petit résultat qui supprime parfois le problème en un point ouvert réel : pensez à étudier la limite de la fonction : si elle est réelle, c'est réglé. On dit qu'on a une fausse singularité en ce point réel.

Proposition 6.75

Soient a et b deux réels tels que $a < b$ et $f \in \mathcal{C}_m^0(]a ; b], \mathbb{K})$.

Si f a une limite réelle en a par valeurs supérieures, alors on peut prolonger f par continuité en a , le prolongement \bar{f} est une fonction continue par morceaux sur $[a ; b]$ et l'intégrale $\int_a^b f$ converge et vaut $\int_a^b \bar{f}$.

Remarque 6.76

Si f est continue par morceaux sur $[0 ; +\infty[$ et si $f(x) \xrightarrow{x \rightarrow +\infty} \ell$, alors si $\ell \neq 0$, $\int_0^{+\infty} f$ diverge (grossièrement).

Exercice 6.77

Justifiez que l'intégrale $\int_0^{+\infty} \frac{\sin^2 t}{t^2} dt$ converge.

Correction 6.78

La fonction $t \mapsto \frac{\sin^2 t}{t^2}$ est continue sur $]0 ; +\infty[$.

On a $\sin t \underset{t \rightarrow 0}{\sim} t$ donc $\frac{\sin^2 t}{t^2} \underset{t \rightarrow 0}{\rightarrow} 1$ donc $\int_0^1 \frac{\sin^2 t}{t^2} dt$ converge.

Pour tout $t \geq 1$, $0 \leq \frac{\sin^2 t}{t^2} \leq \frac{1}{t^2}$ or $\int_1^{+\infty} \frac{1}{t^2} dt$ converge donc $\int_1^{+\infty} \frac{\sin^2 t}{t^2} dt$ converge.

Finalement, $\int_0^{+\infty} \frac{\sin^2 t}{t^2} dt$ converge.

Exercice 6.79

Montrez que l'intégrale $\int_0^{+\infty} \frac{t \ln t}{1+t^3} dt$ converge.

6.4 Fonctions intégrables sur un intervalle

Dans cette section, a et b sont des réels ou des infinis tels que $a < b$. On note $I =]a ; b[$.

6.4.1 Intégrales absolument convergentes

Définition 6.80

Soit $f \in \mathcal{C}_m^0(I, \mathbb{K})$.

On dit que l'intégrale $\int_a^b f$ converge absolument (ou est absolument convergente) quand l'intégrale $\int_a^b |f|$ converge.

Le théorème suivant est primordial pour la suite du cours.

Théorème 6.81

Soit $f \in \mathcal{C}_m^0(I, \mathbb{K})$.

Si l'intégrale $\int_a^b f$ converge absolument, alors l'intégrale $\int_a^b f$ converge.

Dans ce cas, on a $\left| \int_a^b f \right| \leq \int_a^b |f|$.

Démonstration 6.82

► Si f est à valeurs réelles, on a $0 \leq f + |f| \leq 2|f|$.

Donc si $\int_a^b |f|$ converge, alors par comparaison de fonctions positives, $\int_a^b (f + |f|)$ converge.

Donc $\int_a^b f = \int_a^b (f + |f|) - \int_a^b |f|$ converge.

► Si f est à valeurs complexes, on a $|\operatorname{Re} f| \leq |f|$ et $|\operatorname{Im} f| \leq |f|$.

Donc si $\int_a^b |f|$ converge, $\int_a^b |\operatorname{Re} f|$ et $\int_a^b |\operatorname{Im} f|$ convergent par comparaison de fonctions positives.

Donc d'après ce qui précède, $\int_a^b \operatorname{Re} f$ et $\int_a^b \operatorname{Im} f$ convergent.

Donc $\int_a^b f = \int_a^b \operatorname{Re} f + i \int_a^b \operatorname{Im} f$ converge. ■

Exercice 6.83

Montrez que l'intégrale $\int_0^{+\infty} \frac{\sin^3 t}{t^3} dt$ est absolument convergente et donc convergente.

Correction 6.84

La fonction $t \mapsto \frac{\sin^3 t}{t^3}$ est continue sur $]0 ; +\infty[$.

On a $\left| \frac{\sin^3 t}{t^3} \right| \underset{t \rightarrow 0}{\sim} \left| \frac{t^3}{t^3} \right| \xrightarrow{t \rightarrow 0} 1$: fausse singularité en 0, donc $\int_0^1 \left| \frac{\sin^3 t}{t^3} \right| dt$ converge.

Pour tout $t \geq 1$, $\left| \frac{\sin^3 t}{t^3} \right| \leq \frac{1}{t^3}$ or $\int_1^{+\infty} \frac{1}{t^3} dt$ converge donc $\int_1^{+\infty} \frac{\sin^3 t}{t^3} dt$ converge absolument par comparaison de fonctions positives.

Finalement, $\int_0^{+\infty} \frac{\sin^3 t}{t^3} dt$ converge absolument et donc converge.

Exercice 6.85

Montrez que si m est un complexe de partie réelle strictement positive, alors l'intégrale $\int_0^{+\infty} e^{-mt} dt$ converge et donnez sa valeur.

Déduisez-en l'existence et la valeur des intégrales $\int_0^{+\infty} \sin(t) e^{-t} dt$ et $\int_0^{+\infty} \cos(t) e^{-t} dt$.

Correction 6.86

► On note $m = a + ib$ où $(a, b) \in \mathbb{R}_+^* \times \mathbb{R}$.

Pour $t \in [0 ; +\infty[$, on a $e^{-mt} = e^{-at} e^{-ibt}$ donc $|e^{-mt}| = e^{-at}$.

Or $a > 0$ donc $\int_0^{+\infty} |e^{-mt}| dt$ converge donc $\int_0^{+\infty} e^{-mt} dt$ converge.

► Pour $X \geq 0$, on a $\int_0^X e^{-mt} dt = \left[\frac{-e^{-mt}}{m} \right]_0^X = \frac{1}{m} (1 - e^{-mX})$.

Or $|e^{-mX}| = e^{-aX} \xrightarrow{X \rightarrow +\infty} 0$.

Donc $\int_0^{+\infty} e^{-mt} dt = \frac{1}{m}$.

► Pour $t \geq 0$, on a $\begin{cases} |\sin(t) e^{-t}| \leq e^{-t} \\ |\cos(t) e^{-t}| \leq e^{-t} \end{cases}$ et $\int_0^{+\infty} e^{-t} dt$ converge donc $A = \int_0^{+\infty} \cos(t) e^{-t} dt$ et

$B = \int_0^{+\infty} \sin(t) e^{-t} dt$ convergent (absolument).

On a

$$\begin{aligned} A + iB &= \int_0^{+\infty} (\cos t + i \sin t) e^{-t} dt \\ &= \int_0^{+\infty} e^{it} e^{-t} dt \\ &= \int_0^{+\infty} e^{-(1-i)t} dt \\ &= \frac{1}{1-i} \\ &= \frac{1+i}{2}. \end{aligned}$$

$$\text{Donc } A = \operatorname{Re} \frac{1+i}{2} = \frac{1}{2} \text{ et } B = \operatorname{Im} \frac{1+i}{2} = \frac{1}{2}.$$

Remarque 6.87

La réciproque est fautive ! On a montré que $\int_0^{+\infty} \frac{\sin t}{t} dt$ converge et qu'elle ne converge pas absolument.

6.4.2 Fonctions intégrables

Définition 6.88

Soit $f \in \mathcal{C}_m^0(I, \mathbb{K})$.

On dit que f est intégrable sur I quand l'intégrale $\int_a^b f$ converge absolument.

On note alors aussi $\int_I f = \int_I f(t) dt = \int_a^b f$.

L'ensemble des fonctions intégrables sur I est souvent noté $\mathcal{L}^1(I, \mathbb{K})$, L comme Lebesgue, mathématicien français de la fin du 19ème et début du 20ème siècle. Par abus de notation, on écrit parfois « f est \mathcal{L}^1 » pour « f est intégrable sur I ».

Exemple 6.89

Soient $a \in \mathbb{R}$ et $\varepsilon \in \mathbb{R}_+^*$.

- La fonction $t \mapsto e^{-t}$ est intégrable sur $[\varepsilon ; +\infty[$.
Plus généralement, la fonction $t \mapsto t^a e^{-t}$ est intégrable sur $[\varepsilon ; +\infty[$ (voire $[0 ; +\infty[$ si $a > -1$).
- La fonction $t \mapsto \frac{1}{t^\alpha}$ est intégrable sur $[\varepsilon ; +\infty[$ ssi $\alpha > 1$ et sur $]0 ; \varepsilon]$ ssi $\alpha < 1$.
- La fonction \ln est intégrable sur $]0 ; \varepsilon]$.

Proposition 6.90

$\mathcal{L}^1(I, \mathbb{K})$ est un \mathbb{K} -espace vectoriel.

Proposition 6.91 (Stricte positivité de l'intégrale)

Si f est continue, intégrable sur I et $\int_I |f| = 0$, alors $f = 0$.

Par contraposée, si f est continue, intégrable sur I et $f \neq 0$, alors $\int_I |f| > 0$.

6.4.3 Théorème de comparaison des fonctions intégrables

Rappel 6.92

Soient f et g deux fonctions définies au voisinage d'un point $p \in \mathbb{R} \cup \{-\infty, +\infty\}$.

- $f = \mathcal{O}(g)$ au voisinage de p signifie qu'il existe $K > 0$ et V un voisinage de p tels que

$$\forall x \in V, |f(x)| \leq K |g(x)|.$$

- $f = o(g)$ au voisinage de p signifie qu'il existe une fonction $\varepsilon > 0$ et V un voisinage de p tels que

$$\forall x \in V, |f(x)| \leq \varepsilon(x) |g(x)| \quad \text{et} \quad \lim_{x \rightarrow p} \varepsilon(x) = 0.$$

Dans le cas où g ne s'annule pas (ce qui, en pratique, est toujours le cas) :

- $f = \mathcal{O}(g)$ au voisinage de p signifie que $\frac{f}{g}$ est bornée au voisinage de p .
- $f = o(g)$ au voisinage de p signifie que $\frac{f}{g}$ a pour limite 0 en p .

Théorème 6.93

Soit $(f, g) \in \mathcal{C}_m^0(I, \mathbb{K})^2$.

- Si $|f| \leq |g|$ sur I et si g est intégrable, alors f est intégrable.
- Si $f = \mathcal{O}(g)$ au voisinage des bornes ouvertes de I et g est intégrable, alors f est intégrable. C'est vrai en particulier si $f = o(g)$.
- Si $f \sim g$ au voisinage des bornes ouvertes de I , alors il y a équivalence entre l'intégrabilité de f et l'intégrabilité de g .

Exercice 6.94

Montrez que la fonction $t \mapsto \frac{\cos t}{\sqrt{t}} e^{-t}$ est intégrable sur $]0 ; +\infty[$.

Correction 6.95

La fonction $f : t \mapsto \frac{\cos t}{\sqrt{t}} e^{-t}$ est continue sur $]0 ; +\infty[$.

On a $\frac{\cos t}{\sqrt{t}} e^{-t} \underset{t \rightarrow 0}{\sim} \frac{1}{\sqrt{t}}$ et $t \mapsto \frac{1}{\sqrt{t}}$ est intégrable sur $]0 ; 1]$ donc f aussi.

Pour tout $t \geq 1$, $\left| \frac{\cos t}{\sqrt{t}} e^{-t} \right| \leq e^{-t}$ et $t \mapsto e^{-t}$ est intégrable sur $[1 ; +\infty[$ donc f aussi.

Finalement, f est intégrable sur $]0 ; +\infty[$.

Exercice 6.96

La fonction $t \mapsto \frac{\ln t}{\cos t + t^2}$ est-elle intégrable sur $]0 ; +\infty[$?

Correction 6.97

Pour tout $t > 1$, $t^2 > 1 \geq -\cos t$ donc $t^2 + \cos t > 0$.

Pour tout $t \in]0 ; 1]$, $t^2 > 0$ et $\cos t > 0$ donc $t^2 + \cos t > 0$.

Donc $f : t \mapsto \frac{\ln t}{\cos t + t^2}$ est continue sur $]0 ; +\infty[$.

On a $\frac{\ln t}{\cos t + t^2} \underset{t \rightarrow 0}{\sim} \ln t$ et \ln est intégrable sur $]0 ; 1]$ donc f aussi.

On a $\cos t + t^2 \underset{t \rightarrow +\infty}{\sim} t^2$ donc $f(t) \underset{t \rightarrow +\infty}{\sim} \frac{\ln t}{t^2} \underset{t \rightarrow +\infty}{=} o\left(\frac{1}{t^{3/2}}\right)$.

Or $t \mapsto \frac{1}{t^{3/2}}$ est intégrable sur $[1 ; +\infty[$ donc f aussi.

Finalement, f est intégrable sur $]0 ; +\infty[$.

6.5 Intégration des relations de comparaison

Les résultats présentés portent sur des fonctions intégrables sur $[a ; +\infty[$. On obtient évidemment des résultats analogues sur les autres types d'intervalles.

6.5.1 Théorème de comparaison par domination

Dans le cas convergent d'abord, les « restes partiels » suivent la même relation de comparaison.

Théorème 6.98

Soient f, g deux fonctions définies sur $[a ; +\infty[$ avec g à valeurs dans \mathbb{R}_+ .

Si $f = \mathcal{O}(g)$ et g est intégrable sur $[a ; +\infty[$, alors f est intégrable sur $[a ; +\infty[$.

De plus, $\int_x^{+\infty} f = \mathcal{O}\left(\int_x^{+\infty} g\right)$ quand $x \rightarrow +\infty$.

Si $f = o(g)$ et g est intégrable sur $[a ; +\infty[$, alors f est intégrable sur $[a ; +\infty[$.

De plus, $\int_x^{+\infty} f = o\left(\int_x^{+\infty} g\right)$ quand $x \rightarrow +\infty$.

Démonstration 6.99

► Si $f = \mathcal{O}(g)$, alors il existe $K > 0$ et $A \geq a$ tels que $\forall x \geq A, |f(x)| \leq Kg(x)$.

Donc si g est \mathcal{L}^1 alors $|f|$ est \mathcal{L}^1 et donc f est \mathcal{L}^1 .

De plus,

$$\forall x \geq A, \forall t \geq x, |f(t)| \leq Kg(t)$$

$$\text{donc } \int_x^{+\infty} |f(t)| dt \leq K \int_x^{+\infty} g(t) dt$$

$$\text{donc } \left| \int_x^{+\infty} f(t) dt \right| \leq \int_x^{+\infty} |f(t)| dt \leq K \int_x^{+\infty} g(t) dt$$

$$\text{i.e. } \int_x^{+\infty} f(t) dt = \mathcal{O}\left(\int_x^{+\infty} g(t) dt\right).$$

► Idem avec un « pour tout » devant K et un o à la fin pour le cas $f = o(g)$. ■

Dans le cas divergent ensuite, les « intégrales partielles » suivent aussi la même relation de comparaison.

Théorème 6.100

Soient f, g deux fonctions définies sur $[a ; +\infty[$ avec g à valeurs dans \mathbb{R}_+ .

Si $f = \mathcal{O}(g)$ et f n'est pas intégrable sur $[a ; +\infty[$, alors g n'est pas intégrable sur $[a ; +\infty[$.

De plus, $\int_a^x f = \mathcal{O}\left(\int_a^x g\right)$ quand $x \rightarrow +\infty$.

Si $f = o(g)$ et f n'est pas intégrable sur $[a ; +\infty[$, alors g n'est pas intégrable sur $[a ; +\infty[$.

De plus, $\int_a^x f = o\left(\int_a^x g\right)$ quand $x \rightarrow +\infty$.

Démonstration 6.101

► Si $f = \mathcal{O}(g)$ alors il existe $K > 0$ et $A \geq a$ tels que $\forall x \geq A, |f(x)| \leq Kg(x)$.

Donc si f n'est pas \mathcal{L}^1 , $\int_a^{+\infty} |f|$ diverge donc $\int_a^{+\infty} g$ diverge.

Pour $x \geq A$, pour $t \in [A; x]$, on a $\int_A^x |f| \leq K \int_A^x g$ donc

$$\int_a^x |f| = \int_a^A |f| + \int_A^x |f| \leq \int_a^A |f| + K \int_A^x g.$$

Or $\int_A^x g \xrightarrow{x \rightarrow +\infty} +\infty$.

Donc il existe $B \geq a$ tel que $\forall x \geq B, \int_a^A |f| \leq K \int_A^x g$.

Pour $x \geq \max(A, B)$, on a $\int_a^x |f| \leq 2K \int_A^x g \leq 2K \int_a^x g$.

Donc $\left| \int_a^x f \right| \leq \int_a^x |f| \leq 2K \int_a^x g$.

Donc $\int_a^x f \underset{x \rightarrow +\infty}{=} \mathcal{O} \left(\int_a^x g \right)$.

► Idem avec un « pour tout » devant le K et un o à la fin dans le cas $f = o(g)$. ■

6.5.2 Théorème de comparaison par équivalence

Théorème 6.102

Soient f, g deux fonctions définies sur $[a; +\infty[$ avec g à valeurs dans \mathbb{R}_+ .

Si $f \sim g$, alors l'intégrabilité de f est équivalente à l'intégrabilité de g .

De plus :

► si les fonctions sont intégrables, alors les restes partiels sont équivalents :

$$\int_x^{+\infty} f \underset{x \rightarrow +\infty}{\sim} \int_x^{+\infty} g$$

► si les fonctions ne sont pas intégrables, alors les intégrales partielles divergent et sont équivalentes :

$$\int_a^x f \underset{x \rightarrow +\infty}{\sim} \int_a^x g.$$

Démonstration 6.103

On a $f \underset{+\infty}{\sim} g \iff f - g \underset{+\infty}{=} o(g)$.

Si g est \mathcal{L}^1 alors d'après le Théorème 6.98, on a

$$\begin{aligned}\int_x^{+\infty} (f - g) &\underset{x \rightarrow +\infty}{=} o\left(\int_x^{+\infty} g\right) \\ \int_x^{+\infty} f - \int_x^{+\infty} g &\underset{x \rightarrow +\infty}{=} o\left(\int_x^{+\infty} g\right) \\ \int_x^{+\infty} f &\underset{x \rightarrow +\infty}{\sim} \int_x^{+\infty} g.\end{aligned}$$

Si g n'est pas \mathcal{L}^1 , d'après le Théorème 6.100, on a

$$\begin{aligned}\int_a^x (f - g) &\underset{x \rightarrow +\infty}{=} o\left(\int_a^x g\right) \\ \int_a^x f - \int_a^x g &\underset{x \rightarrow +\infty}{=} o\left(\int_a^x g\right) \\ \int_a^x f &\underset{x \rightarrow +\infty}{\sim} \int_a^x g.\end{aligned}$$

■

Chapitre 7

Intégrales à paramètre

Sommaire

7.1	Introduction	235
7.2	Convergence simple	236
7.2.1	Convergence simple d'une suite de fonctions	236
7.2.2	Convergence simple d'une série de fonctions	238
7.3	Suites et séries de fonctions intégrables	239
7.3.1	Théorème de convergence dominée	239
7.3.2	Théorème d'intégration terme à terme	242
7.4	Fonctions définies par une intégrale à paramètre	244
7.4.1	Continuité	244
7.4.2	Dérivabilité	245
7.5	Domination sur des sous-intervalles	249
7.6	Complément : la fonction Γ d'Euler	252

Dans ce chapitre, \mathbb{K} désigne l'ensemble \mathbb{R} ou \mathbb{C} . Les fonctions dans ce chapitre sont à valeurs dans \mathbb{K} .

On considère dans ce chapitre des intégrales de la forme $\int_a^b f(p, t) dt$ où $f(p, t)$ est une expression qui dépend de deux variables p et t , p pouvant être de n'importe quel type mais t bien sûr réelle. Par habitude, on distingue dans le vocabulaire ces deux variables : t est appelée la variable d'intégration (notez le dt qui le signale) et p est appelée le paramètre.

L'intégrale $\int_a^b f(p, t) dt$ est donc une intégrale qui dépend du paramètre p (mais qui ne dépend bien entendu pas de t) et l'objet de ce chapitre est d'étudier des résultats concernant cette dépendance vis-à-vis de p ; en somme, d'étudier des propriétés de l'application $p \mapsto \int_a^b f(p, t) dt$.

Les sections 2 et 3 étudient surtout le cas où p est un paramètre entier naturel, les suivantes le cas où p est un paramètre réel.

7.1 Introduction

Pour commencer, un exercice d'intervention de symboles, qui marque le début de l'étude de ce problème général et qui va nous occuper durant quelques chapitres.

Exercice 7.1

Pour $n \in \mathbb{N}^*$, on pose $f_n(x) = n^2 e^{-nx} (1 - e^{-x})$.

Pour $x \geq 0$, que vaut $\lim_{n \rightarrow +\infty} f_n(x)$?

Montrez la convergence de l'intégrale $\int_0^{+\infty} f_n(x) dx$ et donnez sa valeur en fonction de n .

Comparez $\lim_{n \rightarrow +\infty} \int_0^{+\infty} f_n(x) dx$ et $\int_0^{+\infty} \left(\lim_{n \rightarrow +\infty} f_n(x) \right) dx$.

Quelle conclusion peut-on en tirer ?

Correction 7.2

► Si $x = 0$, alors $f_n(x) = 0 \xrightarrow{n \rightarrow +\infty} 0$.

Si $x > 0$, alors $0 < e^{-x} < 1$ donc $e^{-nx} = (e^{-x})^n \xrightarrow{n \rightarrow +\infty} 0$.

Donc $f_n(x) \xrightarrow{n \rightarrow +\infty} 0$ car $n^2 \underset{n \rightarrow +\infty}{=} o((e^x)^n)$.

Finalement, $\forall x \geq 0, f_n(x) \xrightarrow{n \rightarrow +\infty} 0$.

► f_n est \mathcal{C}^0 sur $[0 ; +\infty[$.

Pour $x \geq 0$, $0 \leq f_n(x) \leq n^2 e^{-nx}$ or $x \mapsto e^{-nx}$ est \mathcal{L}^1 sur $[0 ; +\infty[$ donc f_n aussi.

On a

$$\begin{aligned} \int_0^{+\infty} n^2 e^{-nx} (1 - e^{-x}) dx &= n^2 \left(\int_0^{+\infty} e^{-nx} dx - \int_0^{+\infty} e^{-(n+1)x} dx \right) \\ &= n^2 \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= \frac{n}{n+1}. \end{aligned}$$

► On a $\lim_{n \rightarrow +\infty} \int_0^{+\infty} f_n(x) dx = 1$ et $\int_0^{+\infty} \left(\lim_{n \rightarrow +\infty} f_n(x) \right) dx = \int_0^{+\infty} 0 dx = 0$.

Or $1 \neq 0$.

► Conclusion : on ne peut pas intervertir les symboles \lim et \int *a priori*.

7.2 Convergence simple

7.2.1 Convergence simple d'une suite de fonctions

Définition 7.3

Soient A une partie de \mathbb{R} et (f_n) une suite de fonctions définies sur A .

On dit que la suite (f_n) converge simplement sur A quand pour tout $x \in A$, la suite numérique $(f_n(x))$ converge.

Dans ce cas, on peut définir une fonction f sur A en posant, pour tout $x \in A$, $f(x) = \lim_{n \rightarrow +\infty} f_n(x)$.

La fonction f est alors appelée limite simple sur A de la suite (f_n) et on dit que la suite (f_n) converge simplement vers f sur A .

Exercice 7.4

Étudiez, pour $n \in \mathbb{N}^*$, la convergence simple de la suite de fonctions $f_n : x \mapsto \frac{ne^{-x} + x^2}{n + x}$ sur $[0 ; +\infty[$.

Correction 7.5

Les fonctions f_n sont toutes définies sur $[0 ; +\infty[$.

Pour $x \geq 0$, on a $f_n(x) = \frac{ne^{-x} + x^2}{n + x} \underset{n \rightarrow +\infty}{\sim} \frac{ne^{-x}}{n} = e^{-x} \xrightarrow{n \rightarrow +\infty} e^{-x}$.

Donc (f_n) converge simplement vers $x \mapsto e^{-x}$ sur $[0 ; +\infty[$.

Exercice 7.6

Même question avec la suite de fonctions $f_n : x \mapsto \frac{x^n}{1 + x^n}$ sur $[0 ; +\infty[$.

Correction 7.7

Les fonctions f_n sont toutes définies sur $[0 ; +\infty[$.

Pour $x \geq 0$, on a $f_n(x) = \frac{x^n}{1 + x^n}$.

Si $x \in [0 ; 1[$, on a $f_n(x) \xrightarrow{n \rightarrow +\infty} 0$.

Si $x = 1$, on a $f_n(x) = \frac{1}{2} \xrightarrow{n \rightarrow +\infty} \frac{1}{2}$.

Si $x \in]1 ; +\infty[$, on a $f_n(x) \underset{n \rightarrow +\infty}{\sim} 1 \xrightarrow{n \rightarrow +\infty} 1$.

Donc (f_n) converge simplement vers $x \mapsto \begin{cases} 0 & \text{si } 0 \leq x < 1 \\ 1/2 & \text{si } x = 1 \\ 1 & \text{sinon} \end{cases}$ sur $[0 ; +\infty[$.

Exercice 7.8

Même question avec la suite de fonctions $f_n : x \mapsto n^\alpha x^n (1 - x)$ où α est un réel strictement positif.

Correction 7.9

Si $x = 0$ ou $x = 1$, on a $f_n(x) = 0 \xrightarrow{n \rightarrow +\infty} 0$.

Si $x \in]0 ; 1[$, on a $n^\alpha \underset{n \rightarrow +\infty}{=} o\left(\left(\frac{1}{x}\right)^n\right)$ donc $n^\alpha x^n \xrightarrow{n \rightarrow +\infty} 0$ donc $f_n(x) \xrightarrow{n \rightarrow +\infty} 0$.

Donc la suite (f_n) converge simplement vers la fonction nulle sur $[0 ; 1]$.

7.2.2 Convergence simple d'une série de fonctions

Définition 7.10

Soient A une partie de \mathbb{R} et (f_n) une suite de fonctions définies sur A .

On dit que la série $\sum_{n \geq 0} f_n$ converge simplement sur A quand pour tout $x \in A$, la série numérique $\sum_{n \geq 0} f_n(x)$ converge.

Autrement dit, la série de fonctions $\sum_{n \geq 0} f_n$ converge simplement sur A quand la suite des sommes partielles $\left(\sum_{k=0}^n f_k\right)$ converge simplement sur A .

Dans ce cas, on peut définir une fonction f sur A en posant, pour tout $x \in A$, $f(x) = \sum_{n=0}^{+\infty} f_n(x)$.

La fonction f est alors appelée (fonction) somme sur A de la série $\sum_{n \geq 0} f_n$.

Exercice 7.11

Étudiez, pour $n \in \mathbb{N}^*$, la convergence simple de la série de fonctions $f_n : x \mapsto \frac{nx^2}{n^3 + x^2}$ sur $[0 ; +\infty[$.

Correction 7.12

Si $x = 0$, on a $f_n(x) = 0$ donc $\sum f_n(0)$ converge.

Si $x > 0$, on a $f_n(x) \underset{n \rightarrow +\infty}{\sim} \frac{nx^2}{n^3} = \frac{x^2}{n^2}$.

Donc par théorème de comparaison des séries à termes positifs, $\sum_{n \geq 1} f_n$ converge simplement sur $[0 ; +\infty[$.

Exercice 7.13

Même question avec la série de fonctions $f_n : x \mapsto \frac{x^n}{1+x^n}$ sur $[0 ; +\infty[$.

Correction 7.14

Si $x \geq 1$, la suite $(f_n(x))_n$ converge vers $\frac{1}{2}$ ou 1.

Si $0 \leq x < 1$, $f_n(x) \underset{n \rightarrow +\infty}{\sim} x^n$ or $\sum x^n$ converge donc par théorème de comparaison des séries à termes positifs, $\sum_{n \geq 1} f_n(x)$ converge.

Donc $\sum_{n \geq 1} f_n$ converge simplement sur $[0 ; 1[$.

Exercice 7.15

Même question avec la série de fonctions $f_n : x \mapsto \frac{\sin(nx)}{n^3 + x^3}$ sur $[0 ; +\infty[$.

Correction 7.16

Pour $x \geq 0$, on a $|f_n(x)| \leq \frac{1}{n^3}$ donc $\sum_{n \geq 1} f_n(x)$ converge (absolument) sur $[0 ; +\infty[$.

7.3 Suites et séries de fonctions intégrables

Dans cette section, tous les théorèmes sont admis (démonstrations très difficiles!).

7.3.1 Théorème de convergence dominée

Théorème 7.17

Soient I un intervalle et (f_n) une suite de fonctions continues par morceaux sur I .

Si

- ▷ la suite (f_n) converge simplement sur I vers une fonction f qui est continue par morceaux sur I
- ▷ il existe une fonction φ , intégrable sur I et à valeurs positives, telle que
pour tout $n \in \mathbb{N}$, $|f_n| \leq \varphi$ sur I (hypothèse de domination)

alors les fonctions f et f_n sont toutes intégrables sur I et $\int_I f_n \xrightarrow{n \rightarrow +\infty} \int_I f$.

L'hypothèse de domination est essentielle! Il s'agit donc de trouver une fonction φ (dont on dit qu'elle domine la suite (f_n)) intégrable et, surtout, qui ne dépend pas de n !

Exercice 7.18

Montrez que la suite d'intégrales $\left(\int_0^{+\infty} \cos(t) e^{-nt^2} dt \right)_{n \geq 1}$ est bien définie et qu'elle converge vers 0.

Correction 7.19

On pose $f_n : t \mapsto \cos(t) e^{-nt^2}$ sur $[0 ; +\infty[$.

Pour $n \in \mathbb{N}^*$, f_n est \mathcal{C}^0 sur $[0 ; +\infty[$.

Si $t = 0$, $f_n(t) = 1 \xrightarrow{n \rightarrow +\infty} 1$.

Sinon, $f_n(t) \xrightarrow{n \rightarrow +\infty} 0$.

Donc (f_n) converge simplement vers $f : t \mapsto \begin{cases} 1 & \text{si } t = 0 \\ 0 & \text{sinon} \end{cases}$ sur $[0 ; +\infty[$.

Pour $n \in \mathbb{N}^*$ et $t \in [0 ; +\infty[$, on a

$$|f_n(t)| \leq e^{-t^2}.$$

Or $t \mapsto e^{-t^2}$ est \mathcal{L}^1 sur $[0 ; +\infty[$ donc d'après le théorème de convergence dominée, pour tout $n \in \mathbb{N}^*$, f_n est \mathcal{L}^1 sur $[0 ; +\infty[$ et

$$\int_0^{+\infty} f_n \xrightarrow{n \rightarrow +\infty} \int_0^{+\infty} f = 0.$$

Exercice 7.20

Montrez que la suite d'intégrales $\left(\int_0^{\pi/2} \sin^n t dt \right)$ converge vers 0.

Correction 7.21

Pour $n \in \mathbb{N}$, on pose $f_n : t \mapsto \sin^n t$ sur $[0 ; \pi/2]$.

(f_n) est une suite de fonctions \mathcal{C}^0 sur le segment $[0 ; \pi/2]$ donc \mathcal{L}^1 sur $[0 ; \pi/2]$.

Pour $t \in [0 ; \pi/2]$, $f_n(t) \xrightarrow{n \rightarrow +\infty} 0$ car $0 \leq \sin^n t < 1$.

Si $t = \frac{\pi}{2}$, $f_n(t) \xrightarrow{n \rightarrow +\infty} 1$.

Donc (f_n) converge simplement vers $f : t \mapsto \begin{cases} 0 & \text{si } t \in [0 ; \pi/2[\\ 1 & \text{sinon} \end{cases}$ sur $[0 ; \pi/2]$.

Pour $n \in \mathbb{N}$ et $t \in [0 ; \pi/2]$, on a

$$|f_n(t)| \leq 1.$$

Or $t \mapsto 1$ est \mathcal{L}^1 sur $[0 ; \pi/2]$.

Donc d'après le théorème de convergence dominée, on a

$$\int_0^{\pi/2} f_n \xrightarrow{n \rightarrow +\infty} \int_0^{\pi/2} f = 0.$$

Exercice 7.22

Montrez que pour tout $n \geq 2$, $t \mapsto \frac{1}{1+t^n}$ est intégrable sur $[0; +\infty[$, puis donnez la limite des intégrales quand $n \rightarrow +\infty$.

Correction 7.23

Pour $n \in \mathbb{N}^* \setminus \{1\}$, on pose $f_n : t \mapsto \frac{1}{1+t^n}$ sur $[0; +\infty[$.

Pour $0 \leq t < 1$, $f_n(t) \xrightarrow{n \rightarrow +\infty} 1$.

Pour $t = 1$, $f_n(t) = \frac{1}{2} \xrightarrow{n \rightarrow +\infty} \frac{1}{2}$.

Pour $t > 1$, $f_n(t) \xrightarrow{n \rightarrow +\infty} 0$.

Donc (f_n) converge simplement vers $f : t \mapsto \begin{cases} 1 & \text{si } 0 \leq t < 1 \\ 1/2 & \text{si } t = 1 \\ 0 & \text{sinon} \end{cases}$ sur $[0; +\infty[$.

Pour $n \geq 2$ et $t \in [0; +\infty[$, on a

$$|f_n(t)| \leq \begin{cases} 1 & \text{si } t \leq 1 \\ 1/t^2 & \text{sinon} \end{cases}$$

Or $t \mapsto \begin{cases} 1 & \text{si } t \leq 1 \\ 1/t^2 & \text{sinon} \end{cases}$ est \mathcal{L}^1 sur $[0; +\infty[$ donc par théorème de convergence dominée, $(f_n)_{n \geq 2}$ est une suite de fonctions \mathcal{L}^1 sur $[0; +\infty[$ et

$$\int_0^{+\infty} f_n \xrightarrow{n \rightarrow +\infty} \int_0^{+\infty} f = 1.$$

Quitte à utiliser la caractérisation séquentielle de la limite, on peut étendre le théorème précédent à des fonctions paramétrées par un réel.

Théorème 7.24

Soient I, A deux intervalles, $\alpha \in \overline{A}$ et $(f_a)_{a \in A}$ une famille de fonctions continues par morceaux sur I .

Si

► pour tout $x \in I$, $f_a(x) \xrightarrow{a \rightarrow \alpha} f(x)$ où f est une fonction continue par morceaux sur I

► il existe une fonction φ intégrable sur I et à valeurs positives, telle que

pour tout $a \in A$, $|f_a| \leq \varphi$ sur I (hypothèse de domination)

alors les fonctions f et f_a sont toutes intégrables sur I et $\int_I f_a \xrightarrow{a \rightarrow \alpha} \int_I f$.

Démonstration 7.25

Si pour $x \in I$, $f_a(x) \xrightarrow{a \rightarrow \alpha} f(x)$, d'après la caractérisation séquentielle de la limite, pour toute suite $(a_n) \in A^{\mathbb{N}}$ convergeant vers α , on a $f_{a_n}(x) \xrightarrow{n \rightarrow +\infty} f(x)$.

La suite de fonctions $(f_{a_n})_{n \in \mathbb{N}}$ converge simplement vers f .

Or, pour $x \in I$ et $n \in \mathbb{N}$, $|f_{a_n}(x)| \leq \varphi(x)$ et φ est \mathcal{L}^1 sur I .

Donc d'après le théorème de convergence dominée, $\int_I f_{a_n} \xrightarrow{n \rightarrow +\infty} \int_I f$.

Ceci est vrai pour toute suite $(a_n) \in A^{\mathbb{N}}$ qui converge vers α donc d'après la caractérisation séquentielle de la limite, $\int_I f_a \xrightarrow{a \rightarrow \alpha} \int_I f$. ■

7.3.2 Théorème d'intégration terme à terme

Théorème 7.26

Soient I un intervalle et (f_n) une suite de fonctions continues par morceaux sur I .

Si

- ▷ la série $\sum_{n \geq 0} f_n$ converge simplement sur I vers une fonction continue par morceaux sur I
- ▷ pour tout $n \in \mathbb{N}$, f_n est intégrable sur I
- ▷ la série numérique $\sum_{n \geq 0} \int_I |f_n|$ converge

alors la fonction somme $\sum_{n=0}^{+\infty} f_n$ est intégrable sur I et $\int_I \sum_{n=0}^{+\infty} f_n = \sum_{n=0}^{+\infty} \int_I f_n$.

L'hypothèse de convergence de la série des intégrales est essentielle, mais hélas très contraignante. Il arrive souvent qu'il soit plus facile d'utiliser le théorème de convergence dominée sur les sommes partielles de la série de fonctions.

Exercice 7.27

Justifiez l'existence et calculez $\int_0^{+\infty} \frac{t}{e^t - 1} dt$.

Correction 7.28

On pose $f : t \mapsto \frac{t}{e^t - 1}$ sur $]0 ; +\infty[$.

On a $e^t - 1 \underset{t \rightarrow 0}{\sim} t$ donc $f(t) \underset{t \rightarrow 0}{\sim} 1 \xrightarrow[t \rightarrow 0]{} 1$: fausse singularité, donc $\int_0^1 f$ converge.

On a $f(t) \underset{t \rightarrow +\infty}{\sim} \frac{t}{e^t} = o\left(\frac{1}{t^2}\right)$ donc $\int_1^{+\infty} f$ converge.

Donc $\int_0^{+\infty} f$ converge.

Pour $t > 0$, $f(t) = \sum_{n=0}^{+\infty} ?$

On a $f(t) = \frac{te^{-t}}{e^{-t}(e^t - 1)} = \frac{te^{-t}}{1 - e^{-t}}$.

Or $0 < e^{-t} < 1$ donc $f(t) = te^{-t} \sum_{n=0}^{+\infty} (e^{-t})^n = \sum_{n=0}^{+\infty} te^{-(n+1)t}$.

On pose $f_n : t \mapsto te^{-(n+1)t}$ pour $n \in \mathbb{N}$.

La série de fonctions $\sum f_n$ converge simplement sur $]0 ; +\infty[$ et $\sum_{n=0}^{+\infty} f_n = f$.

Donc $\int_0^{+\infty} f = \int_0^{+\infty} \sum_{n=0}^{+\infty} f_n$.

Pour $n \in \mathbb{N}$, f_n est \mathcal{C}^0 sur $[0 ; +\infty[$ et $f_n(t) \underset{t \rightarrow +\infty}{=} o\left(\frac{1}{t^2}\right)$.

Donc pour $n \in \mathbb{N}$, f_n est \mathcal{L}^1 sur $[0 ; +\infty[$ donc sur $]0 ; +\infty[$.

Étudions la convergence de la série $\sum_{n \geq 0} \int_0^{+\infty} |f_n|$.

On a

$$\int_0^{+\infty} |f_n| = \int_0^{+\infty} te^{-(n+1)t} dt = \left[\frac{-1}{n+1} te^{-(n+1)t} \right]_0^{+\infty} - \frac{-1}{n+1} \int_0^{+\infty} e^{-(n+1)t} dt = \frac{1}{(n+1)^2}.$$

Donc la série $\sum_{n \geq 0} \int_0^{+\infty} |f_n|$ converge.

D'après le théorème d'intégration terme à terme, on a donc

$$\int_0^{+\infty} f = \sum_{n=0}^{+\infty} \int_0^{+\infty} f_n = \sum_{n=0}^{+\infty} \frac{1}{(n+1)^2} = \frac{\pi^2}{6}.$$

Remarque 7.29

Si on ajoute l'hypothèse « les fonctions f_n sont positives » alors f est intégrable sur I ssi $\sum \int_I f_n$ converge et, dans ce cas, $\int_I f = \sum_{n=0}^{+\infty} \int_I f_n$.

7.4 Fonctions définies par une intégrale à paramètre

On s'intéresse aux propriétés des fonctions définies par des intégrales du type $x \mapsto \int_I f(x, t) dt$. On dit que x est un paramètre de l'intégrale $\int_I f(x, t) dt$.

7.4.1 Continuité

Théorème 7.30 (Théorème de continuité sous le signe intégrale)

Soient A, I deux intervalles de \mathbb{R} et $f : (x, t) \mapsto f(x, t)$ une fonction définie sur $A \times I$.

Si

- ▷ pour tout $x \in A$, $t \mapsto f(x, t)$ est continue par morceaux sur I
- ▷ pour tout $t \in I$, $x \mapsto f(x, t)$ est continue sur A
- ▷ il existe une fonction φ intégrable sur I et à valeurs positives, telle que
pour tout $(x, t) \in A \times I$, $|f(x, t)| \leq \varphi(t)$ (hypothèse de domination)

alors pour tout $x \in A$, $t \mapsto f(x, t)$ est intégrable sur I et la fonction $x \mapsto \int_I f(x, t) dt$ est continue sur A .

Démonstration 7.31

Soit $x_0 \in A$.

On a $\lim_{x \rightarrow x_0} f(x, t) = f(x_0, t)$ pour tout $t \in I$.

De plus, on a $|f(x, t)| \leq \varphi(t)$ pour tout $t \in I$, $x \in A$.

Donc d'après le Théorème 7.24, on a

$$\lim_{x \rightarrow x_0} \int_I f(x, t) dt = \int_I \left(\lim_{x \rightarrow x_0} f(x, t) \right) dt = \int_I f(x_0, t) dt.$$

Donc $x \mapsto \int_I f(x, t) dt$ est continue en x_0 . ■

Exercice 7.32

Montrez que la fonction $g : x \mapsto \int_0^{+\infty} \frac{\cos(xt^2)}{1+t^2} dt$ est définie et continue sur \mathbb{R} .

Correction 7.33

On pose $f : (x, t) \mapsto \frac{\cos(xt^2)}{1+t^2}$ sur $\mathbb{R} \times [0 ; +\infty[$.

f est continue sur $\mathbb{R} \times [0 ; +\infty[$ donc pour tout $x \in \mathbb{R}$, $t \mapsto f(x, t)$ est \mathcal{C}^0 sur $[0 ; +\infty[$ et pour tout $t \in [0 ; +\infty[$, $x \mapsto f(x, t)$ est \mathcal{C}^0 sur \mathbb{R} .

Pour $x \in \mathbb{R}$, $t \in [0 ; +\infty[$, on a $|f(x, t)| \leq \frac{1}{1+t^2}$.

Or $t \mapsto \frac{1}{1+t^2}$ est \mathcal{L}^1 sur $[0 ; +\infty[$.

Donc d'après le théorème de continuité sous le signe intégrale, g est continue sur \mathbb{R} et $t \mapsto f(x, t)$ est \mathcal{L}^1 sur $[0 ; +\infty[$.

Exercice 7.34

Montrez que la fonction $h : u \mapsto \int_0^1 \text{Arctan}(u + x \ln x) dx$ est définie et continue sur \mathbb{R} .

Correction 7.35

On pose $f : (u, x) \mapsto \text{Arctan}(u + x \ln x)$ sur $\mathbb{R} \times]0 ; 1]$.

f est continue sur $\mathbb{R} \times]0 ; 1]$.

Pour $u \in \mathbb{R}$, $x \in]0 ; 1]$, on a $|f(u, x)| \leq \frac{\pi}{2}$.

Or $x \mapsto \frac{\pi}{2}$ est \mathcal{L}^1 sur $]0 ; 1]$.

Donc d'après le théorème de continuité sous le signe intégrale, $x \mapsto f(u, x)$ est \mathcal{L}^1 sur $]0 ; 1]$ et h est continue sur \mathbb{R} .

7.4.2 Dérivabilité

Théorème 7.36 (Théorème de dérivation sous le signe intégrale)

Soient A, I deux intervalles de \mathbb{R} et $f : (x, t) \mapsto f(x, t)$ une fonction définie sur $A \times I$.

Si

- ▷ pour tout $x \in A$, $t \mapsto f(x, t)$ est continue par morceaux et intégrable sur I

- ▷ pour tout $t \in I$, $x \mapsto f(x, t)$ est de classe \mathcal{C}^1 sur A
- ▷ pour tout $x \in A$, $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue par morceaux sur I
- ▷ il existe une fonction φ intégrable et à valeurs positives, telle que

$$\text{pour tout } (x, t) \in A \times I, \left| \frac{\partial f}{\partial x}(x, t) \right| \leq \varphi(t) \text{ (hypothèse de domination)}$$

alors la fonction $g : x \mapsto \int_I f(x, t) dt$ est de classe \mathcal{C}^1 sur A et pour tout $x \in A$, $g'(x) = \int_I \frac{\partial f}{\partial x}(x, t) dt$.

Démonstration 7.37

La fonction $g : x \mapsto \int_I f(x, t) dt$ est bien définie sur A d'après la première hypothèse.

Soient $x_0 \in A$ et $x \neq x_0$.

On veut montrer que g est dérivable en x_0 et $g'(x_0) = \int_I \frac{\partial f}{\partial x}(x_0, t) dt$, i.e.

$$\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = \int_I \frac{\partial f}{\partial x}(x_0, t) dt.$$

L'hypothèse de domination justifie la convergence de cette intégrale.

$$\text{On a } \frac{g(x) - g(x_0)}{x - x_0} = \int_I \frac{f(x, t) - f(x_0, t)}{x - x_0} dt.$$

On veut intervertir $\lim_{x \rightarrow x_0}$ et \int_I .

Pour $t \in I$, $x \mapsto f(x, t)$ est \mathcal{C}^1 donc d'après le théorème des accroissements finis, il existe $c_{x,t} \in]x_0; x[$ tel que

$$\frac{f(x, t) - f(x_0, t)}{x - x_0} = \frac{\partial f}{\partial x}(c_{x,t}, t).$$

Par théorème d'encadrement, pour $t \in I$, $\lim_{x \rightarrow x_0} c_{x,t} = x_0$.

Or $x \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue sur A donc par composition des limites,

$$\lim_{x \rightarrow x_0} \frac{\partial f}{\partial x}(c_{x,t}, t) = \frac{\partial f}{\partial x}(x_0, t).$$

Or par hypothèse de domination, $\left| \frac{\partial f}{\partial x}(c_{x,t}, t) \right| \leq \varphi(t)$.

Donc d'après le théorème de convergence dominée, on a

$$\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \int_I \frac{\partial f}{\partial x}(c_{x,t}, t) dt = \int_I \left(\lim_{x \rightarrow x_0} \frac{\partial f}{\partial x}(c_{x,t}, t) \right) dt = \int_I \frac{\partial f}{\partial x}(x_0, t) dt.$$

On a donc montré que g est dérivable sur A et

$$\forall x \in A, \quad g'(x) = \int_I \frac{\partial f}{\partial x}(x, t) dt.$$

Avec les hypothèses $x \mapsto \frac{\partial f}{\partial x}(x, t)$ continue sur A et $\left| \frac{\partial f}{\partial x}(c_{x,t}, t) \right| \leq \varphi(t)$ et le Théorème 7.30, g' est en plus continue sur A , i.e. g est de classe \mathcal{C}^1 sur A . ■

Exercice 7.38

Montrez que la fonction $g : x \mapsto \int_0^\pi \cos(x \sin t) dt$ est définie et de classe \mathcal{C}^1 sur \mathbb{R} .

Correction 7.39

On pose $f : (x, t) \mapsto \cos(x \sin t)$ sur $\mathbb{R} \times [0 ; \pi]$.

f est de classe \mathcal{C}^1 sur $\mathbb{R} \times [0 ; \pi]$ (par opérations et compositions de fonctions \mathcal{C}^1).

Pour $x \in \mathbb{R}$, $t \mapsto \cos(x \sin t)$ est \mathcal{C}^0 sur le segment $[0 ; \pi]$ donc y est intégrable.

On a $\frac{\partial f}{\partial x}(x, t) = -\sin(t) \sin(x \sin t)$.

Pour $x \in \mathbb{R}$, $t \in [0 ; \pi]$, on a $\left| \frac{\partial f}{\partial x}(x, t) \right| \leq 1$.

Or $t \mapsto 1$ est \mathcal{L}^1 sur $[0 ; \pi]$.

D'après le théorème de dérivation sous le signe intégrale, g est de classe \mathcal{C}^1 sur \mathbb{R} et, pour $x \in \mathbb{R}$, $g'(x) = -\int_0^\pi \sin(t) \sin(x \sin t) dt$.

Exercice 7.40

Montrez que la fonction $g : x \mapsto \int_0^{+\infty} \cos(xt^2) e^{-t} dt$ est définie et de classe \mathcal{C}^1 sur \mathbb{R} .

Correction 7.41

On pose $f : (x, t) \mapsto \cos(xt^2) e^{-t}$ sur $\mathbb{R} \times [0 ; +\infty[$.

f est de classe \mathcal{C}^1 sur $\mathbb{R} \times [0 ; +\infty[$.

Pour $x \in \mathbb{R}$, $t \in [0 ; +\infty[$, $|f(x, t)| \leq e^{-t}$ or $t \mapsto e^{-t}$ est \mathcal{L}^1 sur $[0 ; +\infty[$ donc $t \mapsto f(x, t)$ l'est aussi.

On a $\frac{\partial f}{\partial x}(x, t) = -t^2 e^{-t} \sin(xt^2)$.

Donc $\left| \frac{\partial f}{\partial x}(x, t) \right| \leq t^2 e^{-t}$.

Or $t \mapsto t^2 e^{-t}$ est \mathcal{L}^1 sur $[0 ; +\infty[$ (car $t^2 e^{-t} \underset{t \rightarrow +\infty}{=} o\left(\frac{1}{t^{7/2}}\right)$).

D'après le théorème de dérivation sous le signe intégrale, g est de classe \mathcal{C}^1 sur \mathbb{R} et, pour $x \in \mathbb{R}$,
 $g'(x) = - \int_0^{+\infty} t^2 e^{-t} \sin(xt^2) dt$.

Ce théorème est généralisable pour des dérivations d'ordre plus élevé.

Théorème 7.42

Soient A, I deux intervalles de \mathbb{R} et $f : (x, t) \mapsto f(x, t)$ une fonction définie sur $A \times I$.

Si

- ▷ pour tout $x \in A$, $t \mapsto f(x, t)$ est continue par morceaux et intégrable sur I
- ▷ pour tout $t \in I$, $x \mapsto f(x, t)$ est de classe \mathcal{C}^k sur A
- ▷ pour tout $x \in A$, pour tout $j \in \llbracket 1 ; k-1 \rrbracket$, $t \mapsto \frac{\partial^j f}{\partial x^j}(x, t)$ est continue par morceaux et intégrable sur I
- ▷ il existe une fonction φ intégrable sur I et à valeurs positives, telle que

$$\text{pour tout } (x, t) \in A \times I, \quad \left| \frac{\partial^k f}{\partial x^k}(x, t) \right| \leq \varphi(t) \quad (\text{hypothèse de domination})$$

alors la fonction $g : x \mapsto \int_I f(x, t) dt$ est de classe \mathcal{C}^k sur A et pour tout $x \in A$, pour tout $j \in \llbracket 1 ; k \rrbracket$, $g^{(j)}(x) = \int_I \frac{\partial^j f}{\partial x^j}(x, t) dt$.

Exercice 7.43

Montrez que la fonction $g : x \mapsto \int_0^{+\infty} \cos(xt^2) e^{-t} dt$ est définie et de classe \mathcal{C}^∞ sur \mathbb{R} .

Correction 7.44

On reprend les notations de la Correction 7.41.

f est \mathcal{C}^∞ sur $\mathbb{R} \times [0 ; +\infty[$.

Pour montrer que g est \mathcal{C}^∞ sur \mathbb{R} , il suffit de donner une domination pour toutes les dérivées partielles :
 pour $k \in \mathbb{N}$, $x \in \mathbb{R}$, $t \in [0 ; +\infty[$,

$$\left| \frac{\partial^k f}{\partial x^k}(x, t) \right| \leq \varphi_k(t)$$

où φ_k est \mathcal{L}^1 sur $[0 ; +\infty[$.

On a $e^{ix} = \cos x + i \sin x$ donc

$$\frac{d^k}{dx^k} e^{ix} = \cos^{(k)} x + i \sin^{(k)} x = i^k e^{ix}$$

donc

$$\cos^{(k)} x = \operatorname{Re} \left(e^{ik\pi/2} e^{ix} \right) = \cos \left(x + k \frac{\pi}{2} \right)$$

donc

$$\frac{\partial^k f}{\partial x^k} (x, t) = t^{2k} e^{-t} \cos \left(xt^2 + k \frac{\pi}{2} \right).$$

Donc $\left| \frac{\partial^k f}{\partial x^k} (x, t) \right| \leq t^{2k} e^{-t}$ et $t \mapsto t^{2k} e^{-t}$ est \mathcal{L}^1 sur $[0; +\infty[$.

Donc d'après le théorème de dérivation sous le signe intégrale, g est \mathcal{C}^∞ sur \mathbb{R} et, pour $x \in \mathbb{R}$, $k \in \mathbb{N}$,
 $g^{(k)}(x) = \int_0^{+\infty} t^{2k} e^{-t} \cos \left(xt^2 + k \frac{\pi}{2} \right) dt.$

7.5 Domination sur des sous-intervalles

La continuité étant une propriété locale, il est souvent inutile d'avoir une domination globale sur A pour conclure. En général, on peut se contenter de domination sur des parties plus petites que A , en général les segments inclus dans A , ou toute famille recouvrante de parties de A .

Définition 7.45

Soit A un intervalle.

Une famille \mathcal{F} de parties de A est dite recouvrante quand sa réunion est A : $A = \bigcup_{X \in \mathcal{F}} X$.

Proposition 7.46

Soit A un intervalle.

La famille des segments inclus dans A est recouvrante : $A = \bigcup_{(a,b) \in A^2} [a; b]$.

Démonstration 7.47

Tout intervalle est convexe.

Si $x \in A$, il existe $(a, b) \in A^2$ tel que $a \leq x \leq b$.

Donc $x \in [a; b]$.

Donc $x \in \bigcup_{(a,b) \in A^2} [a; b]$.

Donc $A \subseteq \bigcup_{(a,b) \in A^2} [a; b] \subseteq A$. ■

On en déduit alors le théorème suivant, dont il vaut mieux à mon avis, sur chaque exercice, présenter le détail des idées.

Théorème 7.48

Soient A, I deux intervalles de \mathbb{R} et $f : (x, t) \mapsto f(x, t)$ une fonction définie sur $A \times I$.

Soit \mathcal{F} une famille recouvrante de parties de A .

Si

- ▷ pour tout $x \in A$, $t \mapsto f(x, t)$ est continue par morceaux sur I
- ▷ pour tout $t \in I$, $x \mapsto f(x, t)$ est continue sur A
- ▷ pour toute partie F de \mathcal{F} , il existe une fonction φ_F intégrable sur I et à valeurs positives, telle que

$$\text{pour tout } (x, t) \in F \times I, |f(x, t)| \leq \varphi_F(t) \text{ (hypothèse de domination)}$$

alors pour tout $x \in A$, $t \mapsto f(x, t)$ est intégrable sur I et la fonction $x \mapsto \int_I f(x, t) dt$ est continue sur A .

On a de même une version locale des théorèmes de dérivation sous le signe intégrale.

Autrement dit, au lieu de chercher à appliquer les théorèmes précédents directement sur A , on trouve une famille recouvrante de sous-intervalles sur chacun desquels on peut appliquer les théorèmes précédents, conclure à la continuité ou dérivabilité sur chaque sous-intervalle, puis signaler que par réunion, la propriété reste valable sur A .

Exercice 7.49

Montrez que la fonction $f : x \mapsto \int_0^{+\infty} \frac{e^{-xt}}{\sqrt{t}} dt$ est définie et continue sur $]0 ; +\infty[$.

Correction 7.50

On pose $\varphi : (x, t) \mapsto \frac{e^{-xt}}{\sqrt{t}}$ sur $]0 ; +\infty[\times]0 ; +\infty[$.

φ est continue sur $]0 ; +\infty[^2$.

Pour $t \in]0 ; +\infty[$, $x \in]0 ; +\infty[$, $|\varphi(x, t)| \leq \frac{1}{\sqrt{t}}$ est la meilleure majoration possible valable pour $x > 0$.

Soit $a > 0$.

Pour $t > 0$, $x \in [a ; +\infty[$, $|\varphi(x, t)| \leq \frac{e^{-at}}{\sqrt{t}}$.

Or $\frac{e^{-at}}{\sqrt{t}} \underset{t \rightarrow 0}{\sim} \frac{1}{t^{1/2}}$ donc $t \mapsto \frac{e^{-at}}{\sqrt{t}}$ est \mathcal{L}^1 sur $]0 ; 1]$.

De plus, $\frac{e^{-at}}{\sqrt{t}} \underset{t \rightarrow +\infty}{\sim} o\left(\frac{1}{t^2}\right)$ donc $t \mapsto \frac{e^{-at}}{\sqrt{t}}$ est \mathcal{L}^1 sur $[1 ; +\infty[$.

D'après le théorème de continuité sous le signe intégrale, f est continue sur $[a ; +\infty[$.

Donc f est continue sur $]0 ; +\infty[= \bigcup_{a>0} [a ; +\infty[$.

Exercice 7.51

Montrez que la fonction $f : x \mapsto \int_1^{+\infty} \frac{\ln(1+xt)}{t^2} dt$ est définie et continue sur $[0 ; +\infty[$ et qu'elle est de classe \mathcal{C}^1 sur $]0 ; +\infty[$. Donnez une expression simple de $f'(x)$ pour tout $x > 0$.

Correction 7.52

On pose $\varphi : (x, t) \mapsto \frac{\ln(1+xt)}{t^2}$ sur $[0 ; +\infty[\times [1 ; +\infty[$.

φ est de classe \mathcal{C}^∞ sur $[0 ; +\infty[\times [1 ; +\infty[$.

Soit $a > 0$.

Pour $x \in [0 ; a]$, $t \in [1 ; +\infty[$, $\left| \frac{\ln(1+xt)}{t^2} \right| \leq \frac{\ln(1+at)}{t^2}$.

Or $t \mapsto \frac{\ln(1+at)}{t^2}$ est continue sur $[1 ; +\infty[$ et $\frac{\ln(1+at)}{t^2} \underset{t \rightarrow +\infty}{\sim} \frac{\ln t}{t^2} \underset{t \rightarrow +\infty}{=} o\left(\frac{1}{t^{3/2}}\right)$.

Donc $t \mapsto \frac{\ln(1+at)}{t^2}$ est \mathcal{L}^1 sur $[1 ; +\infty[$.

Donc f est continue sur $[0 ; a]$ et donc sur $[0 ; +\infty[= \bigcup_{a>0} [0 ; a]$.

On a $\frac{\partial \varphi}{\partial x}(x, t) = \frac{t}{t^2(1+xt)} = \frac{1}{t(1+xt)}$.

Soit $a > 0$.

Pour $x \in [a ; +\infty[$, $t \in [1 ; +\infty[$, $\left| \frac{\partial \varphi}{\partial x}(x, t) \right| \leq \frac{1}{at^2}$.

Or $t \mapsto \frac{1}{at^2}$ est \mathcal{L}^1 sur $[1 ; +\infty[$.

Donc f est \mathcal{C}^1 sur $[a ; +\infty[$ et donc sur $]0 ; +\infty[= \bigcup_{a>0} [a ; +\infty[$.

Pour $x > 0$, calculons $f'(x) = \int_1^{+\infty} \frac{1}{t(1+xt)} dt$.

On a $\frac{1}{t(1+xt)} = \frac{1}{t} - \frac{x}{1+xt}$.

Donc

$$\begin{aligned}
 \int_1^X \frac{1}{t(1+xt)} dt &= \int_1^X \frac{1}{t} dt - x \int_1^X \frac{1}{1+xt} dt \\
 &= \ln X - \ln(1+xX) + \ln(1+x) \\
 &= \ln X - \ln X - \ln\left(x + \frac{1}{x}\right) + \ln(1+x) \\
 &= \ln \frac{1+x}{x}.
 \end{aligned}$$

Donc

$$f'(x) = \int_1^{+\infty} \frac{1}{t(1+xt)} dt = \lim_{X \rightarrow +\infty} \frac{1+x}{x} = \frac{1+x}{x}.$$

7.6 Complément : la fonction Γ d'Euler

Pour $x \in \mathbb{R}$, on pose, quand cela a un sens

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

Cette fonction très courante a les propriétés suivantes :

- Γ est définie sur $]0 ; +\infty[$
- Γ est de classe \mathcal{C}^∞ sur $]0 ; +\infty[$
- pour tout $x > 0$, $\Gamma(x+1) = x\Gamma(x)$
- pour tout $n \in \mathbb{N}^*$, $\Gamma(n) = (n-1)!$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- il existe un unique $\alpha \in]1 ; 2[$ tel que $\Gamma'(\alpha) = 0$ et Γ est strictement décroissante sur $]0 ; \alpha]$ et strictement croissante sur $]\alpha ; +\infty[$
- Γ est convexe sur $]0 ; +\infty[$
- Γ a des limites infinies en 0 et en $+\infty$.

Démonstration 7.53

On pose $f : (x, t) \mapsto e^{(x-1)\ln t} e^{-t}$ sur $]0 ; +\infty[\times]0 ; +\infty[$.

f est \mathcal{C}^∞ sur $]0 ; +\infty[^2$.

Soient $b > a > 0$.

Pour $n \in \mathbb{N}$, $x \in [a ; b]$, $t \in]0 ; +\infty[$, on a

$$\frac{\partial^n f}{\partial x^n}(x, t) = \ln^n(t) e^{(x-1)\ln t} e^{-t} = \ln^n(t) \frac{1}{t^{1-x}} e^{-t}.$$

$$\text{Donc } \left| \frac{\partial^n f}{\partial x^n}(x, t) \right| = |\ln^n t| e^{(x-1)\ln t} e^{-t}.$$

On a $a - 1 \leq x - 1 \leq b - 1$.

$$\text{Donc } \begin{cases} (x-1)\ln t \leq (b-1)\ln t & \text{si } t \geq 1 \\ (x-1)\ln t \leq (a-1)\ln t & \text{sinon} \end{cases}$$

$$\text{Donc } \left| \frac{\partial^n f}{\partial x^n}(x, t) \right| \leq \varphi(t) \text{ où } \varphi : t \mapsto \begin{cases} |\ln^n t| t^{b-1} e^{-t} & \text{si } t \geq 1 \\ |\ln^n t| t^{a-1} e^{-t} & \text{sinon} \end{cases}$$

Quand $t \rightarrow +\infty$, $|\ln^n t| t^{b-1} e^{-t} = o\left(\frac{1}{t^2}\right)$ donc φ est \mathcal{L}^1 sur $[1 ; +\infty[$.

Quand $t \rightarrow 0$, $\varphi(t) \sim \frac{|\ln^n t|}{t^{1-a}} = o\left(\frac{1}{t^\alpha}\right)$ où $1-a < \alpha < 1$ donc φ est \mathcal{L}^1 sur $]0 ; 1]$.

Donc φ est \mathcal{L}^1 sur $]0 ; +\infty[$.

Donc Γ est \mathcal{C}^∞ sur $[a ; b]$.

Donc Γ est \mathcal{C}^∞ sur $]0 ; +\infty[= \bigcup_{0 < a < b} [a ; b]$. ■

Chapitre 8

Espaces préhilbertiens réels

Sommaire

8.1	Généralités	254
8.1.1	Produit scalaire	254
8.1.2	Exemples fondamentaux	255
8.1.3	Norme euclidienne	256
8.1.4	Vecteurs orthogonaux	259
8.2	Bases orthonormées	260
8.2.1	Familles orthonormées	260
8.2.2	Existence de bases orthonormées	261
8.2.3	Calculs en base orthonormée	263
8.3	Sous-espaces orthogonaux.	263
8.3.1	Orthogonalité de deux sous-espaces vectoriels	263
8.3.2	Orthogonal d'un sous-espace vectoriel	264
8.4	Projection orthogonale sur un sous-espace vectoriel de dimension finie	266
8.4.1	Projection orthogonale	266
8.4.2	Distance à un sous-espace vectoriel	267

Dans tout ce chapitre, E désigne un \mathbb{R} -espace vectoriel.

8.1 Généralités

8.1.1 Produit scalaire

Définition 8.1

On appelle produit scalaire sur E toute application φ de E^2 dans \mathbb{R} qui est

- ▷ bilinéaire (linéaire par rapport à chacune de ses deux variables)
- ▷ symétrique : pour tout $(x, y) \in E^2$, $\varphi(x, y) = \varphi(y, x)$
- ▷ définie-positive : pour tout $x \in E$, $\varphi(x, x) \geq 0$ et $\varphi(x, x) = 0 \iff x = 0$.

Remarque 8.2

Pour montrer que φ est un produit scalaire, on montre en général d'abord que φ est symétrique, puis qu'elle est linéaire à gauche, la linéarité à droite découlant alors de la symétrie.

Définition 8.3

Quand E est muni d'un produit scalaire, on dit que E est un espace préhilbertien. Quand, de plus, E est de dimension finie, on dit que E est un espace euclidien.

En général, on note $\langle \cdot | \cdot \rangle$ les produits scalaires.

8.1.2 Exemples fondamentaux

(1) Le produit scalaire de la géométrie vérifie toutes ces propriétés.

(2) Si $E = \mathbb{R}^n$, soit $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$, on pose $\varphi(x, y) = \sum_{i=1}^n x_i y_i$: φ est appelé le produit scalaire canonique sur \mathbb{R}^n .

(3) Plus généralement, si E est un \mathbb{R} -espace vectoriel de dimension n , alors à toute base \mathcal{B} de E , on peut associer un produit scalaire : si x et y sont deux vecteurs de coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$

et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}}$, on pose $\varphi(x, y) = \sum_{i=1}^n x_i y_i$. L'expression matricielle du produit scalaire est alors $\varphi(x, y) = X^T Y$.

(4) Si a, b sont deux réels tels que $a < b$, $I = [a ; b]$ et $E = \mathcal{C}^0(I, \mathbb{R})$, alors pour f, g deux éléments de E , on pose $\varphi(f, g) = \int_a^b fg$: φ est un produit scalaire sur E .

(5) Si I est un intervalle et $E = \mathcal{C}^0(I, \mathbb{R}) \cap \mathcal{L}^2(I, \mathbb{R})$, ensemble des fonctions f à valeurs réelles, continues sur I et telles que f^2 soit intégrables sur I , alors pour f, g deux éléments de E , on pose $\varphi(f, g) = \int_I fg$: φ est un produit scalaire sur E .

(6) Dans $\mathcal{M}_n(\mathbb{R})$, l'application $(A, B) \mapsto \text{tr}(A^T B)$ est un produit scalaire, c'est même le produit scalaire canonique.

Démonstration 8.4 (5)

Montrons que si $f, g \in \mathcal{L}^2(I, \mathbb{R})$ alors $fg \in \mathcal{L}^1(I, \mathbb{R})$.

Pour tout $(a, b) \in \mathbb{R}^2$, on a $|ab| \leq \frac{a^2 + b^2}{2}$.

Donc pour $x \in I$, on a $0 \leq |f(x)g(x)| \leq \frac{f^2(x) + g^2(x)}{2}$.

Or f^2 et g^2 sont intégrables sur I donc par comparaison de fonctions positives, $|fg|$ est intégrable sur I et donc fg est intégrable sur I .

La fonction $\varphi : (f, g) \mapsto \int_I fg$ est donc bien définie sur E^2 .

La symétrie et la bilinéarité de φ sont évidentes.

Pour tout $f \in E$, $\varphi(f, f) = \int_I f^2 \geq 0$ et comme f^2 est positive et continue sur I , on a

$$\begin{aligned} \int_I f^2 = 0 &\iff f^2 = 0 \\ &\iff f = 0. \end{aligned}$$

Donc φ est un produit scalaire sur E . ■

Démonstration 8.5 (6)

Pour $(A, B) \in \mathcal{M}_n(\mathbb{R})^2$, on a

$$\text{tr}(B^T A) = \text{tr}\left((A^T B)^T\right) = \text{tr}(A^T B)$$

donc on a la symétrie.

La bilinéarité est évidente par linéarité de la trace et de la transposition et par la bilinéarité du produit matriciel.

Pour $A \in \mathcal{M}_n(\mathbb{R})$, on a $\text{tr}(A^T A) = \sum_{1 \leq i, j \leq n} a_{i,j}^2$ donc il est clair que $\text{tr}(A^T A) = 0 \iff A = 0_n$. ■

8.1.3 Norme euclidienne

Définition 8.6

Soit E un espace préhilbertien. On note $\langle \cdot | \cdot \rangle$ le produit scalaire sur E .

On appelle norme euclidienne associée au produit scalaire l'application de E dans \mathbb{R}_+ définie par

$$\forall x \in E, \|x\| = \sqrt{\langle x | x \rangle}.$$

Remarque 8.7

Cette définition a bien un sens, car d'après les propriétés d'un produit scalaire, pour tout $x \in E$, $\langle x | x \rangle \geq 0$ donc $\sqrt{\langle x | x \rangle}$ existe.

On vérifie alors les résultats suivants, inspirés par la géométrie habituelle dans un triangle ou un parallélogramme.

Proposition 8.8

Avec les mêmes notations, pour tout $(x, y) \in E^2$,

- ▷ $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2 \langle x | y \rangle$ (égalité d'Al-Kashi)
- ▷ $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2 \langle x | y \rangle$ (égalité d'Al-Kashi)
- ▷ $\|x + y\|^2 + \|x - y\|^2 = 2 \|x\|^2 + 2 \|y\|^2$ (identité du parallélogramme)
- ▷ $\|x + y\|^2 - \|x - y\|^2 = 4 \langle x | y \rangle$ (identité de polarisation).

Et encore

Proposition 8.9

Avec les mêmes notations,

- ▷ $|\langle x | y \rangle| \leq \|x\| \|y\|$ (inégalité de Cauchy-Schwarz)
- ▷ $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire)
- ▷ pour tout $\lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda| \|x\|$
- ▷ $\|x\| = 0 \iff x = 0$.

Remarque 8.10

Il y a égalité dans l'inégalité de Cauchy-Schwarz ssi x et y sont colinéaires.

Il y a égalité dans l'inégalité triangulaire ssi x et y sont colinéaires de même sens.

Démonstration 8.11 (Inégalité de Cauchy-Schwarz)

Soit $(x, y) \in E^2$.

Si l'un des deux vecteurs est nul, l'inégalité est vraie.

Supposons $x \neq 0$ et $y \neq 0$.

On pose $p : \mathbb{R} \longrightarrow \mathbb{R}$
 $t \longmapsto \|tx + y\|^2$

p est à valeurs positives et, pour $t \in \mathbb{R}$, on a

$$\begin{aligned} p(t) &= \|tx\|^2 + \|y\|^2 + 2 \langle tx | y \rangle \\ &= t^2 \langle x | x \rangle + \|y\|^2 + 2t \langle x | y \rangle. \end{aligned}$$

Or $\langle x | x \rangle > 0$ car $x \neq 0$.

Donc p est un polynôme du second degré de signe constant, donc son discriminant est négatif ou nul, *i.e.*

$$(2 \langle x | y \rangle)^2 - 4 \|x\|^2 \|y\|^2 \leq 0.$$

Donc $|\langle x | y \rangle| \leq \|x\| \|y\|$.

Il y a égalité ssi le trinôme p possède une unique racine réelle t_0 .

Dans ce cas, $p(t_0) = 0 = \|t_0 x + y\|^2 = \langle t_0 x + y | t_0 x + y \rangle$.

Par définie-positivité de $\langle \cdot | \cdot \rangle$, $t_0 x + y = 0$ donc $y = -t_0 x$ donc x et y sont colinéaires.

Et réciproquement. ■

Démonstration 8.12 (Inégalité triangulaire)

Pour $(x, y) \in E^2$, on a

$$\begin{aligned} \|x + y\| \leq \|x\| + \|y\| &\iff \|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2 \|x\| \|y\| \\ &\iff \|x\|^2 + \|y\|^2 + 2 \langle x | y \rangle \leq \|x\|^2 + \|y\|^2 + 2 \|x\| \|y\| \\ &\iff \langle x | y \rangle \leq \|x\| \|y\| \end{aligned} \quad \left. \vphantom{\|x + y\| \leq \|x\| + \|y\|} \right\} \text{Al-Kashi}$$

ce qui est vrai d'après l'inégalité de Cauchy-Schwarz.

Il y a égalité ssi $\langle x | y \rangle = \|x\| \|y\|$, ce qui implique le cas d'égalité de l'inégalité de Cauchy-Schwarz, donc x et y sont colinéaires.

Or $\langle x | y \rangle = \|x\| \|y\| \geq 0$ donc x et y sont positivement colinéaires.

Et réciproquement. ■

On dit qu'un vecteur de E est unitaire (ou normalisé) si sa norme vaut 1. À tout vecteur $x \in E \setminus \{0\}$, on associe deux vecteurs unitaires : $\frac{x}{\|x\|}$ et $-\frac{x}{\|x\|}$.

Exercice 8.13

Soit $(a_1, \dots, a_n, b_1, \dots, b_n) \in \mathbb{R}^{2n}$.

Donnez une inégalité liant $\sum_{k=1}^n a_k b_k$, $\sum_{k=1}^n a_k^2$ et $\sum_{k=1}^n b_k^2$.

Correction 8.14

On pose $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n) \in \mathbb{R}^n$.

On note $\langle \cdot | \cdot \rangle$ le produit scalaire canonique sur \mathbb{R}^n .

D'après l'inégalité de Cauchy-Schwarz, on a $|\langle a | b \rangle| \leq \|a\| \|b\|$ donc

$$\left| \sum_{i=1}^n a_i b_i \right| \leq \sqrt{\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2}.$$

Exercice 8.15

Soit $f \in \mathcal{C}^0([a; b], \mathbb{R}_+^*)$. Montrez que

$$(b-a)^2 \leq \left(\int_a^b f \right) \left(\int_a^b \frac{1}{f} \right).$$

Correction 8.16

On pose $\langle \cdot | \cdot \rangle : \mathcal{C}^0([a; b], \mathbb{R}_+^*)^2 \longrightarrow \mathbb{R}$
 $(f, g) \longmapsto \int_a^b fg$

D'après l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned} \left| \left\langle \sqrt{f} \mid \sqrt{\frac{1}{f}} \right\rangle \right| &\leq \left\| \sqrt{f} \right\| \left\| \sqrt{\frac{1}{f}} \right\| \\ |b-a| &\leq \sqrt{\int_a^b f} \sqrt{\int_a^b \frac{1}{f}} \\ (b-a)^2 &\leq \left(\int_a^b f \right) \left(\int_a^b \frac{1}{f} \right). \end{aligned}$$

8.1.4 Vecteurs orthogonaux

Définition 8.17

Soit E un espace préhilbertien. On note $\langle \cdot | \cdot \rangle$ le produit scalaire sur E .

On dit que deux vecteurs x, y sont orthogonaux (pour ce produit scalaire) quand $\langle x | y \rangle = 0$.

On peut alors noter $x \perp y$ pour signifier que x et y sont orthogonaux.

Plus généralement, si x_1, \dots, x_n sont n vecteurs de E , on dit que la famille (x_1, \dots, x_n) est une famille orthogonale quand

$$\text{pour tout } (i, j) \in \llbracket 1; n \rrbracket^2 \text{ tel que } i \neq j, \quad \langle x_i | x_j \rangle = 0.$$

On retrouve alors le célèbre théorème de Pythagore.

Proposition 8.18

Avec les mêmes notations,

$$x \perp y \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Exercice 8.19

Soient E un \mathbb{R} -espace vectoriel de dimension au moins 2 et u, v deux vecteurs non-colinéaires de E .

Montrez qu'il existe un produit scalaire sur E pour lequel u et v sont orthogonaux.

Correction 8.20

On pose $n \geq 2$ la dimension de E .

Comme u et v ne sont pas colinéaires, (u, v) est libre.

Donc d'après le théorème de la base incomplète, il existe une base (u, v, e_3, \dots, e_n) de E .

En posant φ le produit scalaire associé à cette base, on a

$$\varphi(u, v) = 1 \times 0 + 0 \times 1 + 0 \times 0 + \dots + 0 \times 0 = 0.$$

Donc $u \perp v$ pour le produit scalaire φ .

8.2 Bases orthonormées

8.2.1 Familles orthonormées

Définition 8.21

Soit E un espace préhilbertien.

Une famille de vecteurs de E est dite orthonormée (ou orthonormale) quand elle est orthogonale et ses vecteurs sont unitaires.

Proposition 8.22

Une famille orthogonale sans vecteur nul est libre. En particulier, une famille orthonormée est libre.

Une famille orthonormée génératrice de E est donc une base orthonormée de E .

Démonstration 8.23

Soit $(v_1, \dots, v_p) \in E^p$ une famille orthogonale sans vecteur nul.

Soit $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$ tel que $\lambda_1 v_1 + \dots + \lambda_p v_p = 0$.

Soit $j \in \llbracket 1 ; p \rrbracket$.

On a $\langle \lambda_1 v_1 + \dots + \lambda_p v_p \mid v_j \rangle = \langle 0 \mid v_j \rangle = 0$, or

$$\begin{aligned} \langle \lambda_1 v_1 + \dots + \lambda_p v_p \mid v_j \rangle &= \lambda_1 \langle v_1 \mid v_j \rangle + \dots + \lambda_p \langle v_p \mid v_j \rangle \\ &= \lambda_j \langle v_j \mid v_j \rangle. \end{aligned}$$

Or $v_j \neq 0$ donc $\langle v_j | v_j \rangle \neq 0$ donc $\lambda_j = 0$.

Donc $\lambda_1 = \dots = \lambda_p = 0$.

Donc (v_1, \dots, v_p) est libre. ■

Exercice 8.24

Généralisez l'exercice précédent.

Correction 8.25

Pour toute famille libre, il existe un produit scalaire tel que cette famille soit orthogonale.

8.2.2 Existence de bases orthonormées

Théorème 8.26

Soit E un espace euclidien.

Il existe dans E des bases orthonormées.

De plus, pour toute base (v_1, \dots, v_n) de E , il existe une base orthonormée (e_1, \dots, e_n) de E telle que

$$\text{pour tout } k \in \llbracket 1 ; n \rrbracket, \text{ Vect}(v_1, \dots, v_k) = \text{Vect}(e_1, \dots, e_k).$$

La démonstration repose sur l'algorithme d'orthogonalisation/orthonormalisation de Schmidt.

Démonstration 8.27

Soit (v_1, \dots, v_n) une base de E .

(1) On pose $u_1 = v_1$.

(2) On choisit $u_2 = v_2 - \lambda_1 u_1$ où λ_1 est bien choisi pour que

$$u_1 \perp u_2 \text{ i.e. } \langle u_1 | u_2 \rangle = 0$$

$$\text{i.e. } \langle u_1 | v_2 - \lambda_1 u_1 \rangle = 0$$

$$\text{i.e. } \langle u_1 | v_2 \rangle - \lambda_1 \langle u_1 | u_1 \rangle = 0$$

$$\text{i.e. } \lambda_1 = \frac{\langle u_1 | v_2 \rangle}{\langle u_1 | u_1 \rangle}.$$

Par suite, si on a construit (u_1, \dots, u_k) une famille orthogonale telle que

$$\text{Vect}(u_1, \dots, u_k) = \text{Vect}(v_1, \dots, v_k) :$$

On pose $u_{k+1} = v_{k+1} - \alpha_1 u_1 - \alpha_2 u_2 - \dots - \alpha_k u_k$ où $\alpha_1, \dots, \alpha_k$ sont bien choisis pour obtenir $u_{k+1} \perp u_i$ pour $i \in \llbracket 1 ; k \rrbracket$.

Alors

$$\begin{aligned}\langle u_{k+1} \mid u_i \rangle &= \langle v_{k+1} \mid u_i \rangle - \sum_{j=1}^k \alpha_j \langle u_j \mid u_i \rangle \\ &= \langle v_{k+1} \mid u_i \rangle - \alpha_i \langle u_i \mid u_i \rangle.\end{aligned}$$

Donc $\alpha_i = \frac{\langle v_{k+1} \mid u_i \rangle}{\langle u_i \mid u_i \rangle}$.

u_{k+1} ainsi construit est orthogonal à u_1, \dots, u_k donc (u_1, \dots, u_{k+1}) est orthogonale et $u_{k+1} \in \text{Vect}(u_{k+1}, u_1, \dots, u_k) = \text{Vect}(v_1, \dots, v_{k+1})$.

Ainsi, (u_1, \dots, u_n) est une base orthogonale de E . ■

Remarque 8.28

Si l'on souhaite obtenir une base orthonormée, on divise par les normes.

On en déduit le théorème de la base orthonormée incomplète.

Théorème 8.29

Soit E un espace euclidien.

Toute famille orthonormée de E peut être complétée en une base orthonormée de E .

Exercice 8.30

Dans \mathbb{R}^n muni du produit scalaire canonique, on pose $u = (1, \dots, n)$.

Complétez la famille u en une base orthonormée de \mathbb{R}^n .

Correction 8.31

$$\begin{array}{ccccc} 1 & 2 & 0 & \dots & 0 \\ 2 & -1 & 0 & & \vdots \\ \vdots & 0 & 4 & & \vdots \\ \vdots & \vdots & -3 & & \vdots \text{ puis Gram-Schmidt.} \\ \vdots & \vdots & 0 & & 0 \\ \vdots & \vdots & \vdots & & n \\ n & 0 & 0 & \dots & n-1 \end{array}$$

8.2.3 Calculs en base orthonormée

Soient E un espace euclidien et $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E .

Soient $x, y \in E$, de coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}}$.

Alors

$$\langle x \mid y \rangle = \sum_{i=1}^n x_i y_i = X^\top Y \quad \|x\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{X^\top X} \quad \forall i \in \llbracket 1 ; n \rrbracket, \quad x_i = \langle x \mid e_i \rangle.$$

Démonstration 8.32

On a

$$\begin{aligned} \langle x \mid y \rangle &= \sum_{i=1}^n x_i \langle e_i \mid y \rangle \\ &= \sum_{i=1}^n x_i \sum_{j=1}^n y_j \langle e_i \mid e_j \rangle \\ &= \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i \mid e_j \rangle \\ &= \sum_{i=1}^n x_i y_i. \end{aligned} \quad \left. \begin{array}{l} \searrow \\ \swarrow \end{array} \right\} \langle e_i \mid e_j \rangle = \delta_{i,j} \text{ car } \mathcal{B} \text{ est orthonormée}$$

De même, on a $\langle x \mid e_j \rangle = \sum_{i=1}^n x_i \langle e_i \mid e_j \rangle = x_j$. ■

8.3 Sous-espaces orthogonaux

8.3.1 Orthogonalité de deux sous-espaces vectoriels

Définition 8.33

Soient E un espace préhilbertien, F, G deux sous-espaces vectoriels de E et $u \in E$.

On dit que u est orthogonal (ou normal) à F quand u est orthogonal à tous les vecteurs de F .

On dit que F et G sont orthogonaux quand tout vecteur de F et tout vecteur de G sont orthogonaux, autrement dit quand

$$\text{pour tout } (x, y) \in F \times G, \quad \langle x \mid y \rangle = 0.$$

Proposition 8.34

Si F est de dimension finie et a pour famille génératrice (v_1, \dots, v_k) , alors u est orthogonal à F ssi pour tout $i \in \llbracket 1 ; k \rrbracket$, $\langle u \mid v_i \rangle = 0$.

Proposition 8.35

Si F et G sont orthogonaux, alors ils sont en somme directe : $F \cap G = \{0\}$.

Démonstration 8.36

Supposons $F \perp G$.

Soit $x \in F \cap G$.

Comme $x \in F$ et $x \in G$, on a $x \perp x$ i.e. $\langle x \mid x \rangle = 0$ i.e. $x = 0$.

D'où $F \cap G = \{0\}$.

On peut généraliser à k sous-espaces vectoriels deux à deux orthogonaux F_1, \dots, F_k .

Soit $(x_1, \dots, x_k) \in F_1 \times \dots \times F_k$ tel que $x_1 + \dots + x_k = 0$.

Pour $i \in \llbracket 1 ; k \rrbracket$, on a

$$\begin{aligned} \langle x_1 + \dots + x_k \mid x_i \rangle &= \langle x_1 \mid x_i \rangle + \dots + \langle x_k \mid x_i \rangle \\ &= \langle x_i \mid x_i \rangle \\ &= 0. \end{aligned}$$

Donc $x_i = 0$.

Donc F_1, \dots, F_k sont en somme directe orthogonale. ■

8.3.2 Orthogonal d'un sous-espace vectoriel

Définition 8.37

Soient E un espace préhilbertien et F un sous-espace vectoriel de E .

On note F^\perp l'ensemble des vecteurs normaux à F :

$$F^\perp = \{v \in E \mid \forall x \in F, \langle v \mid x \rangle = 0\}.$$

Avec cette notation, on a clairement l'équivalence :

$$F \text{ et } G \text{ sont orthogonaux} \iff F \subseteq G^\perp \text{ ou, ce qui revient au même, } G \subseteq F^\perp.$$

Théorème 8.38

Soient E un espace préhilbertien et F un sous-espace vectoriel de E .

Alors F^\perp est un sous-espace vectoriel de E , orthogonal à F et donc en somme directe avec F .

Démonstration 8.39

Pour $x \in E$, on pose $\varphi_x : E \longrightarrow \mathbb{R}$
 $y \longmapsto \langle x | y \rangle$

Pour $x \neq 0$, on a $\varphi_x \neq 0$ (car $\varphi_x(x) = \langle x | x \rangle = \|x\|^2 > 0$).

Alors

$$\begin{aligned} F^\perp &= \{v \in E \mid \forall x \in F, \langle x | v \rangle = 0\} \\ &= \{v \in E \mid \forall x \in F, v \in \ker \varphi_x\} \\ &= \bigcap_{x \in F} \ker \varphi_x. \end{aligned}$$

Donc F^\perp est un sous-espace vectoriel de E .

De plus, par définition de F^\perp , on a $\forall x \in F, \forall y \in F^\perp, \langle x | y \rangle = 0$.

Donc F et F^\perp sont orthogonaux (et donc en somme directe). ■

Proposition 8.40

Soient E un espace préhilbertien et F un sous-espace vectoriel de E .

Alors $F \subseteq (F^\perp)^\perp$.

Remarque 8.41

En général, F^\perp n'est pas supplémentaire à F et F n'est pas égal à $(F^\perp)^\perp$.

Remarque 8.42

Dans le cas où F est une droite vectorielle dirigée par un vecteur u , on note plutôt $G = u^\perp$ l'orthogonal de F . Dans ce cas, u^\perp est un hyperplan et on dit alors que u est un vecteur normal à G .

Exercice 8.43

Montrez que si F est un sous-espace vectoriel de E , alors \overline{F} est un sous-espace vectoriel de E , que $F^\perp = \overline{F}^\perp$ et que F^\perp est fermé.

8.4 Projection orthogonale sur un sous-espace vectoriel de dimension finie

8.4.1 Projection orthogonale

Définition 8.44

Soient E un espace préhilbertien et F un sous-espace vectoriel de E de dimension finie.

Alors F^\perp est un supplémentaire de F , appelé le supplémentaire orthogonal de F .

Le projecteur sur F parallèlement à F^\perp est appelé le projecteur orthogonal sur F .

La symétrie orthogonale par rapport à F est la symétrie par rapport à F parallèlement à F^\perp .

Si on connaît une base orthonormée (e_1, \dots, e_p) de F , alors il est facile de calculer la projection orthogonale de x sur F :

$$p_F(x) = \sum_{i=1}^p \langle x | e_i \rangle e_i.$$

Démonstration 8.45

On a $\dim F = p$.

F possède une base orthonormée $\mathcal{B} = (e_1, \dots, e_p)$.

Soit $x \in E$. On cherche $y \in F$ tel que $x - y \perp F$.

analyse

Si y existe alors $y \in F = \text{Vect}(e_1, \dots, e_p)$ donc $y = \sum_{i=1}^p y_i e_i$.

On a $x - y \perp F$ donc pour $j \in \llbracket 1 ; p \rrbracket$, $x - y \perp e_j$, i.e.

$$\begin{aligned} \langle x - y | e_j \rangle &= 0 \\ \langle x | e_j \rangle - \langle y | e_j \rangle &= 0 \\ \langle x | e_j \rangle - y_j &= 0 \\ y_j &= \langle x | e_j \rangle. \end{aligned}$$

L'analyse prouve l'unicité de y .

synthèse

On pose $y = \sum_{i=1}^p \underbrace{\langle x | e_i \rangle}_{y_i} e_i$.

On a $y \in \text{Vect}(e_1, \dots, e_p) = F$.

Pour $j \in \llbracket 1 ; p \rrbracket$, on a

$$\begin{aligned}\langle x - y \mid e_j \rangle &= \langle x \mid e_j \rangle - \langle y \mid e_j \rangle \\ &= \langle x \mid e_j \rangle - y_j \\ &= 0.\end{aligned}$$

Donc $x - y \perp F$.

conclusion

On a montré

$$\forall x \in E, \exists ! y \in F, x - y \in F^\perp$$

i.e.

$$\forall x \in E, \exists ! (y, z) \in F \times F^\perp, x = y + z$$

i.e.

$$E = F \oplus F^\perp.$$

■

On en déduit l'inégalité de Bessel.

Proposition 8.46

Soient E un espace préhilbertien et F un sous-espace vectoriel de E de dimension finie.

Si p est le projecteur orthogonal sur F , alors pour tout $x \in E$, $\|p(x)\| \leq \|x\|$.

Démonstration 8.47

Soit $x \in E$.

On pose $y = p_F(x)$.

On a $x - y \perp F$ donc $x - y \perp y$ donc

$$\begin{aligned}\|x\|^2 &= \|(x - y) + y\|^2 \\ &= \|x - y\|^2 + \|y\|^2 \\ &\geq \|y\|^2.\end{aligned} \quad \left. \vphantom{\begin{aligned}\|x\|^2 &= \|(x - y) + y\|^2 \\ &= \|x - y\|^2 + \|y\|^2 \\ &\geq \|y\|^2.\end{aligned}} \right\} \text{Pythagore}$$

D'où $\forall x \in E, \|p_F(x)\| \leq \|x\|$.

NB : on en déduit que p_F est continu.

■

8.4.2 Distance à un sous-espace vectoriel

Proposition 8.48

Soient E un espace préhilbertien, F un sous-espace vectoriel de E de dimension finie et $x \in E$.

Soit y la projection orthogonale de x sur F .

Pour tout $z \in F$, $\|x - y\| \leq \|x - z\|$, avec égalité ssi $z = y$.

Autrement dit, le projeté orthogonal de x sur F est l'unique vecteur de F qui minimise la distance entre x et un point de F .

$\|x - y\|$ est appelé la distance de x à F , c'est la plus petite des distances entre x et un élément de F , notée $d(x, F)$.

Démonstration 8.49

Pour $z \in F$, $z - y \in F$ et $y - x \in F^\perp$ donc $z - y \perp y - x$, donc

$$\begin{aligned} \|z - x\|^2 &= \|(z - y) + (y - x)\|^2 \\ &= \|z - y\|^2 + \|y - x\|^2 \quad \left. \vphantom{\|z - y\|^2} \right\} \text{Pythagore} \\ &\geq \|y - x\|^2. \end{aligned}$$

Si (e_1, \dots, e_p) est une base orthonormée de F , on a $y = \sum_{i=1}^p \langle x | e_i \rangle e_i$.

De plus, on a $\|x\|^2 = \|y\|^2 + \|x - y\|^2$ donc $\|x - y\|^2 = \|x\|^2 - \|y\|^2$.

Or $\|y\|^2 = \sum_{i=1}^p \langle x | e_i \rangle^2$.

Donc $d(x, F) = \|x - y\| = \sqrt{\|x\|^2 - \sum_{i=1}^p \langle x | e_i \rangle^2}$. ■

Remarque 8.50

Tout ce qui précède est évidemment valable si E est de dimension finie.

Dans ce cas, pour tout sous-espace vectoriel F de E , F^\perp est un supplémentaire de F dans E .

Par conséquent, $\dim F^\perp = \dim E - \dim F$.

Chapitre 9

Endomorphismes dans un espace euclidien

Sommaire

9.1	Adjoint d'un endomorphisme	269
9.1.1	Représentation des formes linéaires	269
9.1.2	Adjoint	270
9.1.3	Matrice de l'adjoint	274
9.1.4	Stabilité de sous-espaces vectoriels	276
9.2	Orientation d'un \mathbb{R}-espace vectoriel de dimension finie.	276
9.3	Isométries vectorielles	277
9.4	Matrices orthogonales	280
9.4.1	Déterminant d'une isométrie vectorielle	283
9.4.2	Changements de bases orthonormées	284
9.4.3	Produit mixte	285
9.4.4	Produit vectoriel en dimension 3	286
9.5	Étude en dimension 2	287
9.6	Réduction des isométries vectorielles ou des matrices orthogonales .	289
9.6.1	Réduction des isométries vectorielles	289
9.6.2	Réduction des matrices orthogonales	294
9.6.3	Étude en dimension 3	294
9.7	Endomorphismes auto-adjoints	300
9.7.1	Définition et propriétés	300
9.7.2	Théorème spectral	301
9.8	Endomorphismes auto-adjoints positifs, définis-positifs	304
9.8.1	Endomorphismes auto-adjoints positifs	304
9.8.2	Matrices symétriques positives	306

Dans tout ce chapitre, E désigne un espace euclidien de dimension n , muni du produit scalaire $\langle \cdot | \cdot \rangle$.

9.1 Adjoint d'un endomorphisme

9.1.1 Représentation des formes linéaires

Le théorème suivant est parfois appelé théorème de représentation de Riesz.

Proposition 9.1

Soit φ une forme linéaire sur E .

Il existe un unique vecteur $v \in E$ tel que pour tout $x \in E$, $\varphi(x) = \langle v | x \rangle$.

Démonstration 9.2

On choisit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E .

Soit $x \in E$ tel que $x = (x_1, \dots, x_n)_{\mathcal{B}}$.

$$\text{On a } \varphi(x) = \varphi\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i \varphi(e_i).$$

On pose alors $v = (\varphi(e_1), \dots, \varphi(e_n))_{\mathcal{B}}$ de sorte que $\varphi(x) = \langle v | x \rangle$.

On a unicité par unicité des coordonnées d'un vecteur. ■

9.1.2 Adjoint

Lemme 9.3

Pour $a, b \in E$, on a

$$a = b \iff \forall x \in E, \langle a | x \rangle = \langle b | x \rangle.$$

Démonstration 9.4



Pour $x \in E$, on a $\langle a - b | x \rangle = 0$ donc $\langle a - b | a - b \rangle = 0$ donc $a - b = 0$ donc $a = b$.



Clair. ■

Proposition 9.5

Soit $f \in \mathcal{L}(E)$.

Il existe un unique endomorphisme $g \in \mathcal{L}(E)$ tel que pour tout $(x, y) \in E^2$, $\langle f(x) | y \rangle = \langle x | g(y) \rangle$.

Démonstration 9.6

Pour tout $y \in E$, l'application $E \longrightarrow \mathbb{R}$ est linéaire.
$$x \longmapsto \langle f(x) \mid y \rangle$$

D'après la Proposition 9.1, il existe un unique vecteur $g(y)$ tel que

$$\forall x \in E, \quad \langle f(x) \mid y \rangle = \langle g(y) \mid x \rangle.$$

On a donc construit une application $g : E \longrightarrow E$.

Soient $(y, z) \in E^2$ et $\lambda \in \mathbb{R}$.

Pour tout $x \in E$, on a

$$\begin{aligned} \langle g(\lambda y + z) \mid x \rangle &= \langle f(x) \mid \lambda y + z \rangle \\ &= \lambda \langle f(x) \mid y \rangle + \langle f(x) \mid z \rangle \\ &= \lambda \langle g(y) \mid x \rangle + \langle g(z) \mid x \rangle \\ &= \langle \lambda g(y) + g(z) \mid x \rangle. \end{aligned}$$

On en déduit $g(\lambda y + z) = \lambda g(y) + g(z)$ d'après le Lemme 9.3. ■

Définition 9.7

L'endomorphisme g est appelé l'adjoint de f et est noté f^* .

Par bilinéarité et symétrie du produit scalaire, on en déduit les propriétés élémentaires de l'adjonction.

Proposition 9.8

- (1) L'application $f \longmapsto f^*$ est linéaire.
- (2) Pour tout $f \in \mathcal{L}(E)$, $f^{**} = f$.
- (3) Pour tout $(f, g) \in \mathcal{L}(E)^2$, $(f \circ g)^* = g^* \circ f^*$.

Démonstration 9.9 (2)

Pour $(x, y) \in E^2$, on a

$$\langle f(x) \mid y \rangle = \langle x \mid f^*(y) \rangle = \langle f^{**}(x) \mid y \rangle.$$

D'après le Lemme 9.3, on a $f(x) = f^{**}(x)$.

Donc $f = f^{**}$. ■

Démonstration 9.10 (3)

Pour tout $(x, y) \in E^2$, on a

$$\begin{aligned}\langle f \circ g(x) \mid y \rangle &= \langle f(g(x)) \mid y \rangle \\ &= \langle g(x) \mid f^*(y) \rangle \\ &= \langle x \mid g^*(f^*(y)) \rangle \\ &= \langle x \mid g^* \circ f^*(y) \rangle.\end{aligned}$$

Par unicité de l'adjoint, $(f \circ g)^* = g^* \circ f^*$. ■

Démonstration 9.11 (1)

Soient $(f, g) \in \mathcal{L}(E)^2$ et $\lambda \in \mathbb{R}$.

Pour tout $(x, y) \in E^2$, on a

$$\begin{aligned}\langle (\lambda f + g)(x) \mid y \rangle &= \langle \lambda f(x) + g(x) \mid y \rangle \\ &= \lambda \langle f(x) \mid y \rangle + \langle g(x) \mid y \rangle \\ &= \lambda \langle x \mid f^*(y) \rangle + \langle x \mid g^*(y) \rangle \\ &= \langle x \mid \lambda f^*(y) + g^*(y) \rangle \\ &= \langle x \mid (\lambda f^* + g^*)(y) \rangle.\end{aligned}$$

Par unicité de l'adjoint, $(\lambda f + g)^* = \lambda f^* + g^*$. ■

Exercice 9.12

Montrez que si f est un projecteur orthogonal, alors $f^* = f$.

Correction 9.13

Soit f un projecteur orthogonal.

Comme E est de dimension finie, on a $E = \text{Im } f \oplus \ker f$.

On choisit une base orthonormée de $\ker f$ et de $\text{Im } f$: en les concaténant, on obtient une base orthonormée de E

$$\mathcal{B} = (e_1, \dots, e_r, e_{r+1}, \dots, e_n).$$

Soit $(x, y) \in E^2$. On a

$$f(x) = \sum_{i=1}^r \langle x \mid e_i \rangle e_i \quad \text{et} \quad f(y) = \sum_{i=1}^r \langle y \mid e_i \rangle e_i.$$

Alors

$$\begin{aligned}\langle f(x) | y \rangle &= \left\langle \sum_{i=1}^r \langle x | e_i \rangle e_i \middle| y \right\rangle \\ &= \sum_{i=1}^r \langle x | e_i \rangle \langle e_i | y \rangle \\ &= \left\langle x \middle| \sum_{i=1}^r \langle e_i | y \rangle e_i \right\rangle \\ &= \langle x | f(y) \rangle.\end{aligned}$$

Donc $f = f^*$.

Exercice 9.14

Premièrement,

Soit $f \in \mathcal{L}(E)$.

Montrez que $\text{Im } f^* = (\ker f)^\perp$ et $\ker f^* = (\text{Im } f)^\perp$.

Comparez $\text{rg } f$ et $\text{rg } f^*$.

Correction 9.15

Soit $y \in \text{Im } f^*$. Il existe $a \in E$ tel que $f^*(a) = y$.

Alors pour $x \in \ker f$, on a

$$\langle x | y \rangle = \langle x | f^*(a) \rangle = \langle a | f(x) \rangle = 0.$$

Donc $x \perp y$ et donc $y \in (\ker f)^\perp$.

Donc $\text{Im } f^* \subseteq (\ker f)^\perp$.

Deuxièmement,

Soient $x \in \ker f^*$ et $y \in \text{Im } f$.

Il existe $a \in E$ tel que $f(a) = y$.

Donc

$$\langle x | y \rangle = \langle x | f(a) \rangle = \langle f^*(x) | a \rangle = 0.$$

Donc $x \perp y$, donc $x \in (\text{Im } f)^\perp$ et donc $\ker f^* \subseteq (\text{Im } f)^\perp$.

De plus, on a

$$\dim \text{Im } f^* \leq \dim (\ker f)^\perp = n - \dim \ker f = \dim \text{Im } f$$

et

$$\dim \ker f^* \leq \dim (\text{Im } f)^\perp = n - \dim \text{Im } f = \dim \ker f.$$

Or $\dim \ker f^* = n - \dim \operatorname{Im} f^* \leq n - \dim \operatorname{Im} f$ donc

$$\dim \operatorname{Im} f^* \geq \dim \operatorname{Im} f.$$

D'où

$$\dim \operatorname{Im} f^* = \dim \operatorname{Im} f = n - \dim \ker f = \dim (\ker f)^\perp$$

et

$$\dim \ker f^* = \dim \ker f = n - \dim \operatorname{Im} f = \dim (\operatorname{Im} f)^\perp.$$

Donc on a

$$\operatorname{Im} f^* = (\ker f)^\perp$$

et

$$\ker f^* = (\operatorname{Im} f)^\perp$$

et

$$\operatorname{rg} f = \operatorname{rg} f^*.$$

Exercice 9.16

Soit $f \in \mathcal{L}(E)$.

Montrez que $\operatorname{rg} f = \operatorname{rg} (f^* \circ f)$.

Correction 9.17

Pour tout $(f, g) \in \mathcal{L}(E)^2$, on a

$$\begin{cases} \operatorname{rg} (g \circ f) \leq \operatorname{rg} f & (1) \\ \operatorname{rg} (f \circ g) \leq \operatorname{rg} f & (2) \end{cases}$$

En effet, on a $\operatorname{Im} (f \circ g) \subseteq \operatorname{Im} f$ donc (2) et on a $\ker f \subseteq \ker (f \circ g)$ et le théorème du rang donc (1).

Pour avoir $\operatorname{rg} (f^* \circ f) = \operatorname{rg} f$, il suffit donc de montrer que $\ker f = \ker (f^* \circ f)$.

On a clairement $\ker f \subseteq \ker (f^* \circ f)$.

Soit $x \in \ker (f^* \circ f)$.

On a $f^* \circ f(x) = 0$ donc $\langle f^* \circ f(x) | x \rangle = 0$ donc $\langle f(x) | f(x) \rangle = 0$ donc $\|f(x)\| = 0$ donc $f(x) = 0$.

Donc $x \in \ker f$ et donc $\ker (f^* \circ f) \subseteq \ker f$.

9.1.3 Matrice de l'adjoint

Proposition 9.18

Soient \mathcal{B} une base orthonormée de E et $(f, g) \in \mathcal{L}(E)^2$.

$$\text{On a } g = f^* \iff \text{Mat}_{\mathcal{B}}(g) = \left(\text{Mat}_{\mathcal{B}}(f) \right)^{\top}.$$

Démonstration 9.19

Soit \mathcal{B} une base orthonormée de E .

Soient $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, et x, y de coordonnées X, Y dans \mathcal{B} .

On pose $A = \text{Mat}_{\mathcal{B}}(f)$.

$f(x)$ a pour coordonnées AX dans \mathcal{B} .

Donc

$$\begin{aligned} \langle f(x) | y \rangle &= (AX)^{\top} Y \\ &= X^{\top} A^{\top} Y \\ &= \langle x | f^*(y) \rangle \\ &= X^{\top} (BY) \end{aligned}$$

où $B = \text{Mat}_{\mathcal{B}}(f^*)$.

Donc $\forall (X, Y) \in \mathcal{M}_{n,1}(\mathbb{R})^2$, $X^{\top} A^{\top} Y = X^{\top} B Y$.

Donc $B = A^{\top}$.

Et réciproquement. ■

Remarque 9.20

Attention, ceci n'est valable qu'en base orthonormée. En base quelconque, c'est plus compliqué.

Exercice 9.21

Soit $f \in \mathcal{L}(E)$ diagonalisable.

Montrez l'équivalence

$$f^* = f^2 \iff f \text{ est un projecteur orthogonal.}$$

Correction 9.22

\Leftarrow Si f est un projecteur orthogonal, alors $f = f^*$ (cf. Exercice 9.12) et $f = f^2$ donc $f^* = f^2$.

\Rightarrow

En base orthonormée, la matrice A de f vérifie $A^\top = A^2$.

Donc $A = (A^2)^\top = (A^\top)^2 = (A^2)^2 = A^4$.

Donc $X^4 - X$ est un polynôme annulateur de A .

Donc μ_A , polynôme minimal de A , divise $X^4 - X$ et est scindé à racines simples dans $\mathbb{R}[X]$.

Or $X^4 - X = X(X-1)(X^2 + X + 1)$.

Donc $\mu_A = X$ ou $\mu_A = X - 1$ ou $\mu_A = X(X-1)$ i.e. $f = 0$ ou $f = \text{id}_E$ ou $f^2 = f$.

Dans tous les cas, f est projecteur.

9.1.4 Stabilité de sous-espaces vectoriels

Une propriété remarquable et utile pour la suite du cours.

Proposition 9.23

Soient $f \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E .

Si F est stable par f , alors F^\perp est stable par f^* .

Démonstration 9.24

On veut montrer que pour tout $x \in F^\perp$, $f^*(x) \in F^\perp$.

Soient $x \in F^\perp$ et $y \in F$.

On a $\langle f^*(x) | y \rangle = \langle x | f(y) \rangle$.

Or $y \in F$ et F est stable par f donc $f(y) \in F$.

Or $x \in F^\perp$ donc $\langle x | f(y) \rangle = 0$.

Donc $f^*(x) \perp y$ donc $f^*(x) \in F^\perp$. ■

9.2 Orientation d'un \mathbb{R} -espace vectoriel de dimension finie

Soit E un \mathbb{R} -espace vectoriel de dimension finie $n \geq 1$.

Définition 9.25

On dit que deux bases $\mathcal{B}, \mathcal{B}'$ de E ont la même orientation quand $\det_{\mathcal{B}} \mathcal{B}' > 0$, sinon on dit qu'elles sont d'orientations contraires.

Orienter E , c'est choisir une base de référence et déclarer directes toutes les bases qui ont la même orientation que cette base de référence. Les bases de l'autre classe d'équivalence sont dites indirectes (ou rétrogrades).

En géométrie classique, dans le plan ou l'espace, on convient systématiquement d'une orientation.

Dans toute la suite, E désigne un espace euclidien de dimension n . On suppose aussi que E est orienté.

9.3 Isométries vectorielles

Définition 9.26

On appelle isométrie vectorielle (ou automorphisme orthogonal) tout endomorphisme de E qui conserve la norme : pour tout $x \in E$, $\|f(x)\| = \|x\|$.

Remarque 9.27

L'appellation « automorphisme » n'est pas usurpée.

Démonstration 9.28

Si f est une isométrie vectorielle, alors $\ker f = \{0\}$ car si $f(x) = 0$, alors $\|f(x)\| = 0$ donc $\|x\| = 0$ donc $x = 0$.

Comme E est de dimension finie, f est un automorphisme. ■

L'ensemble des isométries vectorielles de E est noté $\mathcal{O}(E)$.

Proposition 9.29

$\mathcal{O}(E)$ est un sous-groupe de $(\text{GL}(E), \circ)$.

Démonstration 9.30

On vient de montrer que $\mathcal{O}(E) \subseteq \text{GL}(E)$.

De plus, $\text{id}_E \in \mathcal{O}(E)$.

Enfin, pour $(f, g) \in \mathcal{O}(E)$ et $x \in E$, on a

$$\begin{aligned} \|g \circ f(x)\| &= \|g(f(x))\| \\ &= \|f(x)\| \\ &= \|x\|. \end{aligned} \quad \begin{array}{l} \left. \begin{array}{l} \\ \end{array} \right\} g \text{ conserve la norme} \\ \left. \right\} f \text{ conserve la norme} \end{array}$$

Donc $g \circ f \in \mathcal{O}(E)$.

De plus, pour $f \in \mathcal{O}(E)$, on a

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

donc $f^{-1} \in \mathcal{O}(E)$.

Donc $\mathcal{O}(E)$ est un sous-groupe de $(\text{GL}(E), \circ)$. ■

Les symétries orthogonales sont des isométries vectorielles. Parmi celles-ci, on distingue les symétries orthogonales par rapport à un hyperplan : on les appelle les réflexions.

On peut caractériser les isométries vectorielles de diverses façons.

Proposition 9.31

Soit $f \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :

- (1) f est une isométrie vectorielle
- (2) f conserve le produit scalaire : pour tout $(x, y) \in E^2$, $\langle f(x) | f(y) \rangle = \langle x | y \rangle$
- (3) f transforme toute base orthonormée en base orthonormée
- (4) f est un automorphisme et $f^* = f^{-1}$, ou, ce qui revient au même : $f^* \circ f = \text{id}_E$.

Démonstration 9.32 ((1) \implies (2))

Soient $f \in \mathcal{O}(E)$ et $(x, y) \in E^2$.

On a, d'après Al-Kashi

$$\begin{aligned} \langle f(x) | f(y) \rangle &= \frac{\|f(x) + f(y)\|^2 - \|f(x)\|^2 - \|f(y)\|^2}{2} \\ &= \frac{\|x + y\|^2 - \|x\|^2 - \|y\|^2}{2} \\ &= \langle x | y \rangle. \end{aligned} \quad \left. \begin{array}{l} \text{)} \\ \text{)} \end{array} \right\} \begin{array}{l} f \text{ conserve la norme} \\ \text{Al-Kashi} \end{array} \quad \text{■}$$

Démonstration 9.33 ((2) \implies (1))

On a $\forall (x, y) \in E^2$, $\langle f(x) | f(y) \rangle = \langle x | y \rangle$.

Donc $\forall x \in E$, $\langle f(x) | f(x) \rangle = \langle x | x \rangle$.

Donc $\forall x \in E$, $\|f(x)\|^2 = \|x\|^2$.

Donc $\forall x \in E$, $\|f(x)\| = \|x\|$.

Donc f est une isométrie vectorielle. ■

Démonstration 9.34 ((1) \implies (3))

Soient $f \in \mathcal{O}(E)$ et (e_1, \dots, e_n) une base orthonormée de E .

f conserve la norme donc pour tout $i \in \llbracket 1 ; n \rrbracket$, $\|f(e_i)\| = \|e_i\| = 1$.

f conserve le produit scalaire donc pour tout $(i, j) \in \llbracket 1 ; n \rrbracket^2$, si $i \neq j$, alors $\langle f(e_i) | f(e_j) \rangle = \langle e_i | e_j \rangle = 0$.

Donc $(f(e_1), \dots, f(e_n))$ est une famille orthonormée de E , espace de dimension n , donc une base orthonormée de E . ■

Démonstration 9.35 ((3) \implies (1))

Soient $f \in \mathcal{L}(E)$ et $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E telle que $\mathcal{B}' = (f(e_1), \dots, f(e_n))$ soit aussi une base orthonormée de E .

Soit $x \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$. On a $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ car \mathcal{B} est orthonormée.

Or $x = \sum_{i=1}^n x_i e_i$ donc $f(x) = \sum_{i=1}^n x_i f(e_i)$.

Donc $f(x) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}'}$.

Or \mathcal{B}' est orthonormée donc

$$\|f(x)\| = \sqrt{\sum_{i=1}^n x_i^2} = \|x\|.$$

Donc $f \in \mathcal{O}(E)$. ■

Démonstration 9.36 ((1) \iff (4))

On a

$$\begin{aligned} f \in \mathcal{O}(E) &\iff \forall (x, y) \in E^2, \langle f(x) | f(y) \rangle = \langle x | y \rangle \\ &\iff \forall (x, y) \in E^2, \langle x | f^* \circ f(y) \rangle = \langle x | y \rangle \\ &\iff \forall y \in E, f^* \circ f(y) = y \\ &\iff f^* \circ f = \text{id}_E \\ &\iff f^* = f^{-1}. \end{aligned} \quad \left. \vphantom{\begin{aligned} f \in \mathcal{O}(E) &\iff \forall (x, y) \in E^2, \langle f(x) | f(y) \rangle = \langle x | y \rangle \\ &\iff \forall (x, y) \in E^2, \langle x | f^* \circ f(y) \rangle = \langle x | y \rangle \\ &\iff \forall y \in E, f^* \circ f(y) = y \end{aligned}} \right\} \text{Lemme 9.3}$$

■

Exercice 9.37

Soient E un espace euclidien, $a \in E \setminus \{0\}$ et $k \in \mathbb{R}$. On pose $f : x \mapsto x + k \langle x | a \rangle a$.

Montrez que f est linéaire, puis déterminez les conditions sur a et k pour que f soit une isométrie vectorielle.

Dans ce cas, reconnaissez-la.

Correction 9.38

Soient $(x, y) \in E^2$ et $\lambda \in \mathbb{R}$.

On a

$$\begin{aligned} f(\lambda x + y) &= \lambda x + y + k \langle \lambda x + y | a \rangle a \\ &= \lambda (x + k \langle x | a \rangle a) + y + k \langle y | a \rangle a \\ &= \lambda f(x) + f(y). \end{aligned}$$

Si $k = 0$, alors $f = \text{id}_E \in \mathcal{O}(E)$.

Supposons $k \neq 0$.

Pour tout $x \in a^\perp$, on a $f(x) = x$.

On a $f(a) = a + k \|a\|^2 a = (1 + k \|a\|^2) a$.

Si f conserve la norme, alors $\|f(a)\| = \|a\|$ donc $1 + k \|a\|^2 = -1$ donc $k = \frac{-2}{\|a\|^2}$.

Réciproquement, si $k = \frac{-2}{\|a\|^2}$ alors $\begin{cases} \forall x \in a^\perp, f(x) = x \\ f(a) = -a \end{cases}$ donc f est la réflexion par rapport à a^\perp .

9.4 Matrices orthogonales

Proposition 9.39

Soient \mathcal{B} une base orthonormée de E et $f \in \mathcal{L}(E)$. On pose $A = \text{Mat}_{\mathcal{B}}(f)$.

On a $f \in \mathcal{O}(E) \iff A^\top A = I_n$.

Attention ! Ceci n'est valable que si la base \mathcal{B} est orthonormée.

Définition 9.40

Une matrice carrée A est dite orthogonale quand $A^\top A = I_n$, ce qui est équivalent à $AA^\top = I_n$ ou A est inversible et $A^{-1} = A^\top$.

Proposition 9.41

Une matrice de $\mathcal{M}_n(\mathbb{R})$ est orthogonale quand ses colonnes sont de norme 1 et deux à deux orthogonales pour le produit scalaire canonique de $\mathcal{M}_n(\mathbb{R})$.

Cela est également valable pour les lignes de la matrice.

Démonstration 9.42

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$.

On pose $A^T A = (b_{i,j})$, où $\forall (i, j) \in \llbracket 1 ; n \rrbracket^2$, $b_{i,j} = \sum_{k=1}^n a_{k,i} a_{k,j}$.

Alors A est orthogonale ssi $A^T A = I_n$, i.e.

$$\forall (i, j) \in \llbracket 1 ; n \rrbracket^2, \quad b_{i,j} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{sinon} \end{cases}$$

Or

$$\begin{aligned} b_{i,i} = 1 &\iff \sum_{k=1}^n a_{k,i}^2 = 1 \\ &\iff \text{la norme de la } i\text{-ème colonne est } 1. \end{aligned}$$

Pour $i \neq j$, $b_{i,j} = 0$ ssi le produit scalaire canonique des colonnes i et j est nul. ■

Exercice 9.43

Vérifiez que la matrice $M = \frac{1}{7} \begin{pmatrix} -2 & 6 & -3 \\ 6 & 3 & 2 \\ -3 & 2 & 6 \end{pmatrix}$ est une matrice orthogonale, puis montrez qu'elle est la matrice d'une symétrie orthogonale donc vous préciserez les éléments caractéristiques.

Correction 9.44

On a

$$\frac{1}{7} \left\| \begin{pmatrix} -2 \\ 6 \\ -3 \end{pmatrix} \right\| = 1 \quad \frac{1}{7} \left\| \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix} \right\| = 1 \quad \frac{1}{7} \left\| \begin{pmatrix} -3 \\ 2 \\ 6 \end{pmatrix} \right\| = 1$$

et

$$\left\langle \begin{pmatrix} -2 \\ 6 \\ -3 \end{pmatrix} \middle| \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix} \right\rangle = -12 + 18 - 6 = 0 \quad \left\langle \begin{pmatrix} -2 \\ 6 \\ -3 \end{pmatrix} \middle| \begin{pmatrix} -3 \\ 2 \\ 6 \end{pmatrix} \right\rangle = 6 + 12 - 18 = 0 \quad \left\langle \begin{pmatrix} 6 \\ 3 \\ 2 \end{pmatrix} \middle| \begin{pmatrix} -3 \\ 2 \\ 6 \end{pmatrix} \right\rangle = -18 + 6 + 12 = 0.$$

Donc M est orthogonale : on a $M^T = M^{-1}$.

On remarque que M est symétrique : $M^T = M$.

Donc $M = M^{-1}$ i.e. $M^2 = I_3$.

Si on note f l'endomorphisme de matrice M dans une base orthonormée, f est une isométrie vectorielle et une symétrie, donc f est une symétrie orthogonale.

Les valeurs propres de f sont 1 et -1 .

f est la symétrie orthogonale par rapport à $\text{sep}(f, 1)$.

Or $n = 3$ et $\text{tr } f = 1$ donc $\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

f est donc une réflexion par rapport au plan d'équation $3x - 2y + z = 0$.

Exercice 9.45

Déterminez les réels a et b tels que la matrice $A = \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}$ soit orthogonale. Reconnaissez la nature de l'isométrie vectorielle de matrice A dans une base orthonormée \mathcal{B} .

Correction 9.46

La matrice A est orthogonale ssi

$$\begin{cases} a^2 + 2b^2 = 1 \\ 2ab + b^2 = 0 \end{cases} \iff \begin{cases} a^2 + 2b^2 = 1 \\ b(b + 2a) = 0 \end{cases}$$

Les solutions sont les couples $(1, 0)$, $(-1, 0)$, $(1/3, -2/3)$ et $(-1/3, 2/3)$.

On a donc $A = I_3$ ou $A = -I_3$ ou (1) $A = \frac{1}{3} \begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}$ ou (2) $A = \frac{-1}{3} \begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}$.

Soit $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Dans le cas (1), on a $AX = X \iff x + y + z = 0$ donc A est la matrice de la réflexion par rapport au plan d'équation $x + y + z = 0$.

Dans le cas (2), on a $AX = X \iff \begin{cases} y = x \\ z = x \end{cases}$ donc A est la matrice de la symétrie orthogonale par rapport à la droite $\text{Vect}((1, 1, 1))$.

L'ensemble des matrices orthogonales est noté $\mathcal{O}_n(\mathbb{R})$.

Proposition 9.47

$\mathcal{O}_n(\mathbb{R})$ est un sous-groupe compact de $(\text{GL}_n(\mathbb{R}), \times)$.

Démonstration 9.48

Soient E un espace euclidien de dimension n , \mathcal{B} une base orthonormée de E et $f \in \mathcal{L}(E)$ de matrice $A \in \mathcal{M}_n(\mathbb{R})$ dans la base \mathcal{B} .

On a $A \in \mathcal{O}_n(\mathbb{R}) \iff f \in \mathcal{O}(E)$.

Comme $\mathcal{O}(E)$ est un groupe pour \circ , $\mathcal{O}_n(\mathbb{R})$ est un groupe pour \times via l'isomorphisme de groupes

$$\begin{array}{ccc} (\mathcal{O}(E), \circ) & \xrightarrow{\sim} & (\mathcal{O}_n(\mathbb{R}), \times) \\ f & \mapsto & A \end{array}$$

Si $A = (a_{i,j}) \in \mathcal{O}_n(\mathbb{R})$, on a

$$\forall j \in \llbracket 1 ; n \rrbracket, \sum_{i=1}^n a_{i,j}^2 = 1$$

donc

$$\forall (i, j) \in \llbracket 1 ; n \rrbracket^2, 0 \leq a_{i,j}^2 \leq 1 \text{ i.e. } |a_{i,j}| \leq 1$$

donc $\|A\|_\infty \leq 1$, donc $\mathcal{O}_n(\mathbb{R})$ est borné.

De plus, on a $\mathcal{O}_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) \mid M^\top M = I_n\}$.

On pose $\varphi : M \mapsto M^\top M$ continue sur $\mathcal{M}_n(\mathbb{R})$.

Or $\mathcal{O}_n(\mathbb{R}) = \varphi^{-1}(\{I_n\})$ et $\{I_n\}$ fermé donc $\mathcal{O}_n(\mathbb{R})$ fermé.

Enfin, comme $\mathcal{M}_n(\mathbb{R})$ est un \mathbb{R} -espace vectoriel de dimension finie, $\mathcal{O}_n(\mathbb{R})$ est un compact de $\mathcal{M}_n(\mathbb{R})$. ■

9.4.1 Déterminant d'une isométrie vectorielle

Proposition 9.49

Si $f \in \mathcal{O}(E)$, alors $\det f \in \{-1, 1\}$.

Démonstration 9.50

Soit $f \in \mathcal{O}(E)$.

Dans une base orthonormée \mathcal{B} de E , $A = \underset{\mathcal{B}}{\text{Mat}}(f)$ est orthogonale.

Donc $A^\top A = I_n$.

Or $\det A^\top = \det A$ donc

$$\det(A^\top A) = \det A^\top \times \det A = (\det A)^2 = 1.$$

Donc $\det A \in \{-1, 1\}$. ■

La réciproque est bien sûr fausse.

Les isométries vectorielles de déterminant 1 sont celles qui conservent l'orientation : les transforment les bases orthonormées directes en bases orthonormées directes. On les appelle les isométries vectorielles directes ou positives.

On note $\mathcal{SO}_n(\mathbb{R})$ l'ensemble des matrices orthogonales de déterminant 1 et $\mathcal{SO}(E)$ l'ensemble des isométries vectorielles positives.

Proposition 9.51

$\mathcal{SO}(E)$ est un sous-groupe de $\mathcal{O}(E)$, appelé groupe spécial orthogonal de E .

$\mathcal{SO}_n(\mathbb{R})$ est un sous-groupe de $\mathcal{O}_n(\mathbb{R})$, appelé groupe spécial orthogonal d'ordre n .

Les réflexions sont des isométries négatives.

9.4.2 Changements de bases orthonormées

Proposition 9.52

Soient \mathcal{B} et \mathcal{B}' deux bases orthonormées de E .

La matrice de passage de \mathcal{B} à \mathcal{B}' est une matrice orthogonale.

Démonstration 9.53

On note $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$.

On note $\mathcal{P}_{\mathcal{B} \rightarrow \mathcal{B}'} = (p_{i,j})$.

Pour tout $j \in \llbracket 1 ; n \rrbracket$, on a $e'_j = \sum_{i=1}^n p_{i,j} e_i$.

Les bases \mathcal{B} et \mathcal{B}' étant orthonormées, on a

$$\forall j \in \llbracket 1 ; n \rrbracket, \quad \|e'_j\| = \sqrt{\sum_{i=1}^n p_{i,j}^2} = 1$$

et

$$\forall (j, k) \in \llbracket 1 ; n \rrbracket^2, \quad j \neq k \implies \langle e'_j | e'_k \rangle = \sum_{i=1}^n p_{i,j} p_{i,k} = 0.$$

Donc $\mathcal{P}_{\mathcal{B} \rightarrow \mathcal{B}'} \in \mathcal{O}_n(\mathbb{R})$. ■

Remarque 9.54

De plus, si \mathcal{B} et \mathcal{B}' ont la même orientation, alors $\det \mathcal{P}_{\mathcal{B} \rightarrow \mathcal{B}'} > 0$ donc $\det \mathcal{P}_{\mathcal{B} \rightarrow \mathcal{B}'} = 1$.

Si elles sont de sens contraires, alors $\det \mathcal{P}_{\mathcal{B} \rightarrow \mathcal{B}'} < 0$ donc $\det \mathcal{P}_{\mathcal{B} \rightarrow \mathcal{B}'} = -1$.

L'intérêt de ce genre de changement de bases est que la difficulté liée au calcul de l'inverse de la matrice de passage disparaît :

$X = PX'$ est équivalent à $X' = P^\top X$ donc $A' = P^{-1}AP$ devient $A' = P^\top AP$.

9.4.3 Produit mixte

Proposition 9.55

Soit (v_1, \dots, v_n) une famille de n vecteurs de E .

Le déterminant de (v_1, \dots, v_n) dans une base orthonormée directe ne dépend pas du choix de cette base.

Démonstration 9.56

Soient \mathcal{B} et \mathcal{B}' deux bases orthonormées directes de E .

D'après la Remarque 9.54, on a $\det_{\mathcal{B}} \mathcal{B}' = 1$.

Donc

$$\begin{aligned} \det_{\mathcal{B}} (v_1, \dots, v_n) &= \det_{\mathcal{B}} \mathcal{B}' \times \det_{\mathcal{B}'} (v_1, \dots, v_n) \\ &= \det_{\mathcal{B}'} (v_1, \dots, v_n). \end{aligned}$$

■

Dans ce cas, on appelle produit mixte de (v_1, \dots, v_n) le déterminant de cette famille dans n'importe quelle base orthonormée directe : il est noté habituellement $\text{Det}(v_1, \dots, v_n)$ ou $[v_1, \dots, v_n]$.

Une conséquence directe de la définition du produit mixte est la caractérisation des bases directes.

Proposition 9.57

Soit (v_1, \dots, v_n) une famille de n vecteurs de E .

La famille (v_1, \dots, v_n) est une base directe de E ssi $[v_1, \dots, v_n] > 0$.

9.4.4 Produit vectoriel en dimension 3

Dans ce paragraphe, $n = 3$.

Proposition 9.58

Soit $(u, v) \in E^2$.

Il existe un unique vecteur $w \in E$ tel que pour tout $x \in E$, $[u, v, x] = \langle w | x \rangle$.

Ce vecteur est appelé le produit vectoriel de u et v et est noté $u \wedge v$ ou $u \times v$.

Démonstration 9.59

Pour $(u, v) \in E^2$, l'application $x \mapsto [u, v, x]$ est une forme linéaire sur E .

Ainsi, d'après le théorème de représentation de Riesz, il existe un unique vecteur w tel que

$$\forall x \in E, [u, v, x] = \langle w | x \rangle.$$

■

En base orthonormée directe, les coordonnées du produit vectoriel se calculent facilement. En base quelconque, c'est beaucoup plus pénible.

Notons quelques propriétés algébriques et géométriques du produit vectoriel.

Proposition 9.60

- ▷ L'application \wedge est bilinéaire et antisymétrique.
- ▷ $u \wedge v = 0$ ssi u et v sont colinéaires.
- ▷ Si u et v ne sont pas colinéaires, alors $u \wedge v$ est un vecteur normal au plan $\text{Vect}(u, v)$ et la famille $(u, v, u \wedge v)$ est une base directe de E .
- ▷ Si u et v sont unitaires et orthogonaux, alors la famille $(u, v, u \wedge v)$ est une base orthonormée directe de E .

Démonstration 9.61

On reprend les mêmes notations.

- ▷ L'application \wedge est bilinéaire antisymétrique car le produit mixte est trilinéaire alterné.
- ▷ Si u et v sont colinéaires, on a $\forall x \in E, [u, v, x] = 0$ donc $u \wedge v = 0$.

Réciproquement, si u et v ne sont pas colinéaires alors $\text{Vect}(u, v)$ est un plan.

On choisit $x \in E$ normal à ce plan et on obtient une base de $E : (u, v, x)$.

Alors $[u, v, x] \neq 0$, donc $\langle u \wedge v | x \rangle \neq 0$ donc $u \wedge v \neq 0$.

► Si u et v ne sont pas colinéaires, on a

$$\langle u \wedge v \mid u \rangle = [u, v, u] = 0 \text{ donc } u \wedge v \perp u$$

et

$$\langle u \wedge v \mid v \rangle = [u, v, v] = 0 \text{ donc } u \wedge v \perp v$$

donc $u \wedge v \perp \text{Vect}(u, v)$.

De plus, on a

$$[u, v, u \wedge v] = \langle u \wedge v \mid u \wedge v \rangle = \|u \wedge v\|^2 > 0$$

donc $(u, v, u \wedge v)$ est une base directe de E .

► Pour tout $(u, v) \in E^2$, $\|u \wedge v\|^2 + \langle u \mid v \rangle^2 = \|u\|^2 \|v\|^2$.

On choisit une base orthonormée directe \mathcal{B} de E .

On note $u \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{B}}$ et $v \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\mathcal{B}}$.

En posant $X = \begin{vmatrix} y & y' \\ z & z' \end{vmatrix}$, $Y = -\begin{vmatrix} x & x' \\ z & z' \end{vmatrix}$ et $Z = \begin{vmatrix} x & x' \\ y & y' \end{vmatrix}$, on a alors $u \wedge v \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_{\mathcal{B}}$ et

$$\begin{aligned} \|u \wedge v\|^2 + \langle u \mid v \rangle^2 &= (yz' - y'z)^2 + (xz' - x'z)^2 + (xy' - x'y)^2 + (xx' + yy' + zz')^2 \\ &= (x^2 + y^2 + z^2)(x'^2 + y'^2 + z'^2). \end{aligned}$$

■

9.5 Étude en dimension 2

Proposition 9.62

$\mathcal{O}_2(\mathbb{R})$ contient exclusivement les matrices suivantes :

► les matrices de rotation $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$

► les matrices de réflexions $\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$

où θ est un réel quelconque.

Démonstration 9.63

Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathcal{O}_2(\mathbb{R})$ de déterminant 1. On a

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \\ ad - bc = 1 \end{cases}$$

donc il existe $\theta, \alpha \in \mathbb{R}$ tels que

$$\begin{cases} a = \cos \theta \\ b = \sin \theta \end{cases} \quad \text{et} \quad \begin{cases} c = \cos \alpha \\ d = \sin \alpha \end{cases}$$

On a alors

$$\begin{cases} ac + bd = 0 = \cos(\alpha - \theta) \\ ad - bc = 1 = \sin(\alpha - \theta) \end{cases}$$

Donc $\alpha - \theta \equiv \frac{-\pi}{2} [2\pi]$ i.e. $\alpha \equiv \frac{-\pi}{2} + \theta [2\pi]$, d'où

$$\begin{cases} c = -\sin \theta \\ d = \cos \theta \end{cases}$$

$$\text{i.e. } A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

De même, si $\det A = -1$, on obtient $A = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$.

Et réciproquement. ■

L'ensemble des matrices de rotation forme le sous-groupe $\mathcal{SO}_2(\mathbb{R})$: c'est l'ensemble des matrices orthogonales de déterminant 1.

Il est remarquable que ce groupe est commutatif, car en dimension $n \geq 3$, ce n'est plus le cas. En effet, il est facile de constater que l'application $\theta \mapsto R(\theta)$ est un morphisme surjectif de groupes de $(\mathbb{R}, +)$ dans $(\mathcal{SO}_2(\mathbb{R}), \times)$ (dont le noyau est le sous-groupe $2\pi\mathbb{Z}$ de $(\mathbb{R}, +)$).

Autrement dit, l'application $\begin{array}{ccc} \mathbb{U} & \longrightarrow & \mathcal{SO}_2(\mathbb{R}) \\ e^{i\theta} & \longmapsto & R(\theta) \end{array}$ est un isomorphisme de groupes.

Démonstration 9.64

On a donc bien

$$\mathcal{SO}_2(\mathbb{R}) = \left\{ R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$$

et on remarque

$$\begin{cases} R(\theta + \theta') = R(\theta) R(\theta') = R(\theta') R(\theta) \\ R(-\theta) = R(\theta)^{-1} \end{cases}$$

■

Proposition 9.65

En dimension 2, les isométries vectorielles sont :

- ▷ les rotations vectorielles
- ▷ les réflexions vectorielles.

Remarque 9.66

On pose, pour $\theta \in \mathbb{R}$, $S(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$.

$S(\theta)$ est orthogonale et symétrique donc $S(\theta)^2 = I_2$, donc $S(\theta)$ est la matrice d'une symétrie orthogonale.

Donc comme $S(\theta)$ n'est ni I_2 ni $-I_2$, il s'agit de la matrice d'une réflexion.

Or on a

$$\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta/2 \\ \sin \theta/2 \end{pmatrix} = \begin{pmatrix} \cos(\theta - \theta/2) \\ \sin(\theta - \theta/2) \end{pmatrix} = \begin{pmatrix} (\cos \theta)/2 \\ (\sin \theta)/2 \end{pmatrix}.$$

9.6 Réduction des isométries vectorielles ou des matrices orthogonales

9.6.1 Réduction des isométries vectorielles

D'abord, deux résultats généraux sur les isométries vectorielles.

Proposition 9.67

Soit $f \in \mathcal{O}(E)$. On a :

- ▷ $\text{Sp}(f) \subseteq \{-1, 1\}$
- ▷ Si F est un sous-espace vectoriel de E stable par f , alors F^\perp est aussi un sous-espace vectoriel de E stable par f .

Démonstration 9.68

► Soient λ une valeur propre de f (s'il en existe) et x un vecteur propre associé.

On a $f(x) = \lambda x$ donc

$$\|f(x)\| = |\lambda| \|x\| = \|x\| \neq 0.$$

Donc $|\lambda| = 1$ i.e. $\lambda \in \{-1, 1\}$.

► Soit F un sous-espace vectoriel de E stable par f .

D'après la Proposition 9.23, F^\perp est stable par f^* .

Or $f \in \mathcal{O}(E)$ donc $f^* = f^{-1}$, donc $f^{-1}(F^\perp) \subseteq F^\perp$.

Or f^{-1} est un automorphisme de E donc f^{-1} conserve la dimension, donc $\dim f^{-1}(F^\perp) = \dim F^\perp$.

D'où $f^{-1}(F^\perp) = F^\perp$ et donc $F^\perp = f(F^\perp)$. ■

Lemme 9.69

Si E est un \mathbb{R} -espace vectoriel de dimension finie n et $f \in \mathcal{L}(E)$, alors il existe une droite ou un plan de E stable par f .

Démonstration 9.70

Si f possède une valeur propre réelle, la droite vectorielle dirigée par n'importe quel vecteur propre associé est stable par f .

Sinon, on choisit une base \mathcal{B} de E et on pose $A = \text{Mat}_{\mathcal{B}}(f) \in \mathcal{M}_n(\mathbb{R})$.

A ne possède aucune valeur propre réelle mais a au moins une valeur propre complexe $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

On choisit $Z \in \mathcal{M}_{n,1}(\mathbb{C})$ un vecteur propre de A associé à λ .

On a $Z = X + iY$ où $(X, Y) \in \mathcal{M}_{n,1}(\mathbb{R})^2$.

On pose $x, y \in E$ tels que x et y aient pour coordonnées X et Y dans la base \mathcal{B} .

On veut montrer que $\text{Vect}(x, y)$ est stable par f et est un plan de E .

On note $\lambda = a + ib$ où $(a, b) \in \mathbb{R} \times \mathbb{R}^*$.

On a $AZ = \lambda Z$ donc

$$A(X + iY) = (a + ib)(X + iY)$$

$$AX + iAY = aX - bY + i(aY + bX)$$

donc

$$\begin{cases} AX = aX - bY \\ AY = bX + aY \end{cases}$$

i.e.

$$\begin{cases} f(x) = ax - by \in \text{Vect}(x, y) \\ f(y) = bx + ay \in \text{Vect}(x, y) \end{cases}$$

Donc $\text{Vect}(x, y)$ est stable par f .

Supposons maintenant par l'absurde que x et y sont colinéaires.

Comme $Z \neq 0$, x ou y est non-nul.

Supposons, par exemple, que $y = \alpha x$ où $\alpha \in \mathbb{R}$ et $x \neq 0$.

On a $AZ = \lambda Z$ donc

$$\begin{aligned} A(X + iY) &= A(X + i\alpha X) \\ &= (a + ib)(X + i\alpha X) \end{aligned}$$

donc

$$\begin{aligned} (1 + i\alpha)AX &= (1 + i\alpha)(a + ib)X \\ AX &= (a + ib)X \end{aligned}$$

$$\text{donc } AX + 0i = AX + ibX \text{ donc } \begin{cases} AX = aX \\ 0 = bX \end{cases}$$

Or $b \neq 0$ et $x \neq 0$: contradiction.

De même si $x = \alpha y$ avec $y \neq 0$.

Donc x et y ne sont pas colinéaires *i.e.* $\text{Vect}(x, y)$ est un plan. ■

De ces propriétés découlent le théorème suivant.

Théorème 9.71

Soit $f \in \mathcal{O}(E)$.

Il existe une base orthonormée de E dans laquelle la matrice de f est diagonale par blocs, les blocs étant des scalaires 1 ou -1 ou des matrices $(2, 2)$ de rotation.

Les matrices diagonales par blocs sont donc du type suivant :

$$D = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & \cos \theta_1 & -\sin \theta_1 \\ & & & & & & \sin \theta_1 & \cos \theta_1 \\ & & & & & & & \ddots & & \\ & & & & & & & & \cos \theta_k & -\sin \theta_k \\ & & & & & & & & \sin \theta_k & \cos \theta_k \end{pmatrix}.$$

Démonstration 9.72

On raisonne par récurrence sur $n = \dim E$.

On pose $\mathcal{P}(n)$: « si E est de dimension n et si $f \in \mathcal{O}(E)$, alors il existe une base orthonormée ... »

► $n = 1$

Si $f \in \mathcal{O}(E)$ alors $f = \text{id}_E$ ou $f = -\text{id}_E$ donc dans n'importe quelle base \mathcal{B} , $\text{Mat}_{\mathcal{B}}(f) = (1)$ ou (-1) .

Donc $\mathcal{P}(1)$ est vraie.

► $n = 2$

Si $f \in \mathcal{O}(E)$, d'après la Proposition 9.62, f est une rotation et il existe $\theta \in \mathbb{R}$ tel que dans n'importe quelle base orthonormée directe, $\text{Mat}_{\mathcal{B}}(f) = R(\theta)$; ou f est une réflexion, *i.e.* une symétrie orthogonale par rapport à une droite donc il existe une base orthonormée dans laquelle $\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Donc $\mathcal{P}(2)$ est vraie.

► Pour $n \geq 3$, supposons que $\mathcal{P}(n-2)$ et $\mathcal{P}(n-1)$ sont vraies, que $\dim E = n$ et que $f \in \mathcal{O}(E)$.

D'après le Lemme 9.69, f possède une droite stable ou un plan stable et d'après la Proposition 9.67, l'orthogonal de ce sous-espace vectoriel stable est aussi stable par f .

— Si f possède une droite stable D :

Alors D est une droite propre associée à une valeur propre 1 ou -1 de f .

g , l'endomorphisme induit par f dans l'hyperplan D^\perp , est une isométrie vectorielle de D^\perp .

Or $\dim D^\perp = n - 1$ donc d'après $\mathcal{P}(n-1)$, il existe une base orthonormée \mathcal{B}^\perp de D^\perp dans laquelle

$$\text{Mat}_{\mathcal{B}^\perp}(g) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \\ & & & & & & R(\theta_1) & & \\ & & & & & & & \ddots & \\ & & & & & & & & R(\theta_k) \end{pmatrix}.$$

En choisissant un vecteur directeur unitaire u de D , on obtient $\mathcal{B} = (u, \mathcal{B}^\perp)$ une base

orthonormée de E telle que

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \pm 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & -1 & & \\ & & & & & \ddots & \\ & & & & & & -1 \\ & & & & & & & R(\theta_1) & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & R(\theta_k) \end{pmatrix}.$$

— Si f possède un plan stable P :

Alors P^\perp est stable par f .

L'endomorphisme induit par f dans P est une isométrie vectorielle de P .

Or $\dim P = 2$ donc il existe une base orthonormée de P dans laquelle l'endomorphisme induit par f dans P a pour matrice $R(\theta)$ ou $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

L'endomorphisme induit par f dans P^\perp est une isométrie vectorielle de P^\perp , or $\dim P^\perp = n-2$.

D'après $\mathcal{P}(n-2)$, ...

En concaténant deux bases orthonormées de P et P^\perp , on obtient une base \mathcal{B} dans laquelle

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} R(\theta) & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & -1 & & \\ & & & & & \ddots & \\ & & & & & & -1 \\ & & & & & & & R(\theta_1) & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & R(\theta_k) \end{pmatrix}$$

ou

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & & & & & \\ & -1 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & -1 & \\ & & & & & & \ddots & \\ & & & & & & & -1 & \\ & & & & & & & & R(\theta_1) & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & R(\theta_k) \end{pmatrix}$$

Dans les deux cas, $\mathcal{P}(n)$ est vraie. ■

9.6.2 Réduction des matrices orthogonales

Définition 9.73

Soient A, B deux matrices de $\mathcal{M}_n(\mathbb{R})$.

On dit que A et B sont orthogonalement semblables (ou orthosemblables) quand il existe $P \in \mathcal{O}_n(\mathbb{R})$ telle que $B = P^{-1}AP = P^{\top}AP$.

Deux matrices sont orthosemblables quand elles représentent le même endomorphisme dans des bases orthonormées différentes.

Le théorème de réduction précédent a une traduction matricielle.

Théorème 9.74

Toute matrice orthogonale est orthosemblable à une matrice diagonale par blocs du type ci-dessus.

Pour tout $A \in \mathcal{O}_n(\mathbb{R})$, il existe $P \in \mathcal{O}_n(\mathbb{R})$ et D diagonale par blocs comme ci-dessus telles que $A = P^{\top}DP$.

9.6.3 Étude en dimension 3

À l'aide de ce résultat, on peut classer les isométries vectorielles de E en dimension 3. Seule la réduction des rotations est au programme.

Dans la suite de cette section, E est un espace euclidien de dimension 3 et orienté.

Proposition 9.75

Soit $f \in \mathcal{O}(E)$. On pose $F = \ker(f - \text{id}_E)$. Alors

- ▷ si $\dim F = 3$, alors $f = \text{id}_E$
- ▷ si $\dim F = 2$, alors f est la réflexion par rapport à F
- ▷ si $\dim F = 1$, alors f est une rotation d'axe F
- ▷ si $\dim F = 0$, alors f est une antirotation, c'est-à-dire la composée d'une rotation et d'une réflexion dont l'axe et le plan de base respectifs sont orthogonaux.

Démonstration 9.76

Si $\dim F = 1$, d'après le Théorème 9.71, il existe une base orthonormée \mathcal{B} telle que

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & ? \end{pmatrix}$$

où le bloc $?$ est $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = R(\pi)$ ou $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = R(\theta)$.

On pose $\mathcal{B} = (e_1, e_2, e_3)$.

On a $f(e_1) = e_1$.

Dans le plan F^\perp , stable par f , l'endomorphisme induit par f est une rotation « d'angle θ » (ou $-\theta$ selon l'orientation choisie de F^\perp).

On choisit l'orientation de F^\perp de sorte que \mathcal{B} soit directe.

Ce qui caractérise f , c'est le couple (e_1, θ) : e_1 dirige F et θ est l'angle associé.

Pour déterminer e_1 , on résout $f(e_1) = e_1$ et on choisit une solution de norme 1.

Pour déterminer l'angle θ associé, on choisit un vecteur $v \neq 0$ orthogonal à e_1 puis on calcule $f(v)$ et $e_1 \wedge v$.

Comme $\|e_1\| = 1$, on a $\|e_1 \wedge v\| = \|v\| = \|f(v)\|$.

Alors $f(v) = \cos(\theta)v + \sin(\theta)(e_1 \wedge v)$.

On en déduit $\cos \theta$ et $\sin \theta$ et enfin θ (modulo 2π). ■

En étudiant les différents cas, on constate un lien entre le type de f et son déterminant.

Proposition 9.77

Soit $f \in \mathcal{O}(E)$ telle que $f \neq \text{id}_E$.

f est une rotation ssi $\det f = 1$.

Dans le cas où $\det f = -1$, cette information ne suffit pas à connaître le type de f . Cependant, si on connaît la matrice A de f dans une base orthonormée, alors on peut distinguer les cas 1 et 3.

Proposition 9.78

Soient $f \in \mathcal{L}(E)$ et A la matrice de f dans une base orthonormée.

Si A est une matrice orthogonale et symétrique, alors f est une symétrie orthogonale.

► Si $\det f = 1$, alors A est un demi-tour (une rotation d'angle π).

► Si $\det f = -1$, alors A est une réflexion.

Donc, si A est orthogonale de déterminant -1 et non-symétrique, alors f est une antirotation.

Exercice 9.79

Reconnaissez la nature de l'endomorphisme dont la matrice dans une base orthonormée \mathcal{B} est

$\frac{1}{15} \begin{pmatrix} -11 & 10 & 2 \\ -2 & -5 & 14 \\ 10 & 10 & 5 \end{pmatrix}$ et précisez ses éléments caractéristiques.

Correction 9.80

On note A la matrice donnée et C_1, C_2, C_3 ses colonnes.

On a

$$\|C_1\|^2 = \frac{1}{15^2} \left((-11)^2 + (-2)^2 + 10^2 \right) = 1 = \|C_2\|^2 = \|C_3\|^2$$

et

$$\langle C_1 | C_2 \rangle = \langle C_1 | C_3 \rangle = \langle C_2 | C_3 \rangle = 0$$

donc $A \in \mathcal{O}_3(\mathbb{R})$.

Or \mathcal{B} est orthonormée donc $f \in \mathcal{O}(E)$.

On note $u \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{B}} \in E$ et on a

$$\begin{aligned} f(u) = u &\iff \begin{cases} -11x + 10y + 2z = 15x \\ -2x - 5y + 14z = 15y \\ 10x + 10y + 5z = 15z \end{cases} \\ &\iff \begin{cases} -13x + 5y + z = 0 \\ -x - 10y + 7z = 0 \\ x + y - z = 0 \end{cases} \\ &\iff \begin{cases} x + y - z = 0 \\ -9y + 6z = 0 \\ 18y - 12z = 0 \end{cases} \\ &\iff \begin{cases} x = \frac{1}{3}z \\ y = \frac{2}{3}z \end{cases} \end{aligned}$$

Donc $\ker(f - \text{id}_E)$ est la droite vectorielle dirigée par $e_1 \begin{pmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{pmatrix}_{\mathcal{B}}$ de norme 1, donc f est une rotation.

On pose $v \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}_{\mathcal{B}}$ de sorte que $v \perp e_1$.

On a $f(v) \begin{pmatrix} 32/15 \\ -1/15 \\ -10/15 \end{pmatrix}$ et $e_1 \wedge v \begin{pmatrix} -3/\sqrt{14} \\ -6/\sqrt{14} \\ 5/\sqrt{14} \end{pmatrix}_{\mathcal{B}}$.

Alors

$$\begin{aligned} f(v) = \cos(\theta) v + \sin(\theta) (e_1 \wedge v) &\iff \begin{cases} -2 \cos \theta - \frac{3}{\sqrt{14}} \sin \theta = \frac{32}{15} \\ \cos \theta - \frac{6}{\sqrt{14}} \sin \theta = \frac{-1}{15} \\ \frac{5}{\sqrt{14}} \sin \theta = \frac{-10}{15} \end{cases} \\ &\iff \begin{cases} \sin \theta = \frac{-2\sqrt{14}}{15} \\ \cos \theta = \frac{-1}{15} + \frac{6}{\sqrt{14}} \sin \theta = \frac{-13}{15} \end{cases} \end{aligned}$$

Enfin, on obtient

$$\theta \equiv \pi + \text{Arctan} \frac{2\sqrt{14}}{15} [2\pi] .$$

Exercice 9.81

Même exercice avec la matrice $\frac{1}{7} \begin{pmatrix} 6 & -2 & -3 \\ 3 & 6 & 2 \\ 2 & -3 & 6 \end{pmatrix}$.

Correction 9.82

On note A la matrice donnée et on vérifie aisément que A est orthogonale.

On a $\det A = 1$ donc f est une rotation.

On pose $u \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{B}}$ et on a

$$f(u) = u \iff \begin{cases} 6x - 2y - 3z = 7x \\ 3x + 6y + 2z = 7y \\ 2x - 3y + 6z = 7z \end{cases}$$

$$\iff \begin{cases} -x - 2y - 3z = 0 \\ 3x - y + 2z = 0 \\ 2x - 3y - z = 0 \end{cases}$$

$$\iff \begin{cases} x + 2y + 3z = 0 \\ -7y - 7z = 0 \\ -7y - 7z = 0 \end{cases}$$

$$\iff \begin{cases} y = -z \\ x = y \end{cases}$$

$$\text{Donc } \ker(f - \text{id}_E) = \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right) = \text{Vect} \left(\begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}_{\mathcal{B}} \right) = \text{Vect}(e_1).$$

On pose $v \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}_{\mathcal{B}} \perp e_1$ et on a

$$f(v) \begin{pmatrix} -5/7 \\ 8/7 \\ 3/7 \end{pmatrix}_{\mathcal{B}} \quad \text{et} \quad e_1 \wedge v \begin{pmatrix} 2/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}_{\mathcal{B}}.$$

Alors

$$f(v) = \cos(\theta)v + \sin(\theta)(e_1 \wedge v) \iff \begin{cases} \frac{2}{\sqrt{3}} \sin \theta = \frac{-5}{7} \\ \cos \theta - \frac{1}{\sqrt{3}} \sin \theta = \frac{8}{7} \\ \cos \theta + \frac{1}{\sqrt{3}} \sin \theta = \frac{3}{7} \end{cases}$$

$$\iff \begin{cases} \sin \theta = \frac{-5\sqrt{3}}{14} \\ \cos \theta = \frac{8}{7} + \frac{1}{\sqrt{3}} \left(\frac{-5\sqrt{3}}{14} \right) = \frac{8}{7} - \frac{5}{14} = \frac{11}{14} \end{cases}$$

Donc f est la rotation d'axe orienté par e_1 et d'angle associé $\theta = \text{Arctan} \frac{-5\sqrt{3}}{11}$.

Exercice 9.83

Même exercice avec la matrice $\frac{1}{7} \begin{pmatrix} 6 & -3 & -2 \\ 3 & 2 & 6 \\ 2 & 6 & -3 \end{pmatrix}$.

Correction 9.84

On note A la matrice donnée. Comme ses colonnes sont de norme 1 et leurs produits scalaires deux à deux tous nuls, A est orthogonale et donc f est une isométrie vectorielle.

On a $\det A = -1$ donc f est une antirotation.

Exercice 9.85

Soit $\mathcal{B} = (i, j, k)$ une base orthonormée directe de E de dimension 3.

Déterminez la matrice dans la base \mathcal{B} de la rotation d'axe orienté par $i + j + k$ et d'angle $\frac{\pi}{3}$.

Correction 9.86

f est la rotation d'axe orienté par $i + j + k$ et d'angle $\frac{\pi}{3}$ donc il existe une base orthonormée $\mathcal{B}' = (e_1, e_2, e_3)$ telle que

$$A' = \text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \pi/3 & -\sin \pi/3 \\ 0 & \sin \pi/3 & \cos \pi/3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & 1/2 \end{pmatrix}.$$

En notant $P = \mathcal{P}_{\mathcal{B} \rightarrow \mathcal{B}'}$, on a $A' = P^{-1}AP$ donc

$$A = PA'P^{-1} = PA'P^{\top}.$$

On a $e_1 = \frac{i+j+k}{\sqrt{3}} \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$ donc $P = \begin{pmatrix} 1/\sqrt{3} & ? & ? \\ 1/\sqrt{3} & ? & ? \\ 1/\sqrt{3} & ? & ? \end{pmatrix}$.

On a $e_2 \perp e_1$ et $\|e_2\| = 1$, par exemple $e_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}$, donc $P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & ? \\ 1/\sqrt{3} & -1/\sqrt{2} & ? \\ 1/\sqrt{3} & 0 & ? \end{pmatrix}$.

Puis $e_3 = e_1 \wedge e_2 = \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}$ donc $P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}$.

On en déduit A .

9.7 Endomorphismes auto-adjoints

9.7.1 Définition et propriétés

Définition 9.87

On dit qu'un endomorphisme f de E est auto-adjoint quand $f^* = f$, autrement dit quand

$$\forall (x, y) \in E^2, \langle f(x) | y \rangle = \langle x | f(y) \rangle.$$

On rencontre encore très souvent le mot « symétrique » pour « auto-adjoint ».

Exemple 9.88

- Les projecteurs orthogonaux sont des endomorphismes auto-adjoints (mais pas des endomorphismes orthogonaux!).
- Les symétries orthogonales sont aussi des endomorphismes auto-adjoints.

Démonstration 9.89

Soit s une symétrie orthogonale. On lui associe un projecteur p tel que $s = 2p - \text{id}_E$.

p est un projecteur orthogonal donc auto-adjoint.

Pour tout $(x, y) \in E^2$, on a

$$\begin{aligned} \langle s(x) | y \rangle &= \langle 2p(x) - x | y \rangle \\ &= 2 \langle p(x) | y \rangle - \langle x | y \rangle \\ &= 2 \langle x | p(y) \rangle - \langle x | y \rangle \\ &= \langle x | 2p(y) - y \rangle \\ &= \langle x | s(y) \rangle. \end{aligned}$$

Donc s est auto-adjoint. ■

Proposition 9.90

Soient \mathcal{B} une base orthonormée de E et $f \in \mathcal{L}(E)$.

f est un endomorphisme auto-adjoint ssi sa matrice dans la base \mathcal{B} est symétrique.

Corollaire 9.91

L'ensemble des endomorphismes auto-adjoints est un sous-espace vectoriel de $\mathcal{L}(E)$, de dimension $\frac{n(n+1)}{2}$.

Il est noté $\mathcal{S}(E)$.

9.7.2 Théorème spectral

Il y a essentiellement un seul résultat à connaître sur les endomorphismes auto-adjoints ! On commence par deux lemmes.

Lemme 9.92

Le polynôme caractéristique d'un endomorphisme auto-adjoint est scindé sur \mathbb{R} .

Démonstration 9.93

Soient $f \in \mathcal{S}(E)$ et \mathcal{B} une base orthonormée de E .

Comme f est auto-adjoint, on a $A = \text{Mat}_{\mathcal{B}}(f) \in \mathcal{S}_n(\mathbb{R})$.

χ_A est scindé sur \mathbb{C} donc pour montrer que χ_A est scindé sur \mathbb{R} , on montre que toutes les valeurs propres complexes de A sont en fait réelles.

Soient $\lambda \in \text{Sp}_{\mathbb{C}}(A)$ et Z un vecteur propre associé.

Comme $A^T = A$, on a $(AZ)^T = Z^T A^T = Z^T A = (\lambda Z)^T = \lambda Z^T$.

Alors

$$\begin{aligned} \overline{Z}^T AZ &= \left(\overline{Z}^T A \right) Z \\ &= \overline{Z^T A Z} \\ &= \overline{\lambda Z^T Z} \\ &= \overline{\lambda} \times \overline{Z^T Z} \end{aligned}$$

et $\overline{Z}^T AZ = \overline{Z}^T (AZ) = \lambda \overline{Z}^T Z$ donc

$$\lambda \overline{Z}^T Z = \overline{\lambda} \times \overline{Z}^T Z.$$

Or, avec $Z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$, on a $\overline{Z}^T Z = \sum_{i=1}^n z_i \overline{z_i} = \sum_{i=1}^n |z_i|^2 > 0$.

Donc $\overline{Z}^T Z \neq 0$ donc $\lambda = \overline{\lambda}$ donc $\lambda \in \mathbb{R}$. ■

Lemme 9.94

Si un sous-espace vectoriel F est stable par un endomorphisme auto-adjoint, alors F^\perp l'est aussi.

Théorème 9.95

Les sous-espaces propres d'un endomorphisme auto-adjoint sont deux à deux orthogonaux et leur somme directe est E .

Autrement dit, tout endomorphisme auto-adjoint est diagonalisable en base orthonormée, c'est-à-dire qu'il existe une base orthonormée de vecteurs propres.

Démonstration 9.96

On pose $\mathcal{P}(n)$: l'énoncé du Théorème 9.95 avec $\dim E = n$.

▷ $n = 1$

On a $\mathcal{P}(1)$ car tout endomorphisme est auto-adjoint et tout vecteur non-nul est propre.

▷ Supposons $\mathcal{P}(n-1)$ et soit $f \in \mathcal{S}(E)$ où $\dim E = n$.

D'après le Lemme 9.92, f possède au moins une valeur propre λ et un vecteur propre associé u .

$D = \text{Vect}(u)$ est stable par f (c'est une droite propre) donc d'après le Lemme 9.94, D^\perp est stable par f .

Or $\dim D^\perp = n - 1$ donc par hypothèse de récurrence, l'endomorphisme induit par f dans D^\perp est auto-adjoint et on peut donc trouver dans D^\perp une base orthonormée de vecteurs propres pour cet endomorphisme : (e_2, \dots, e_n) .

$\mathcal{B} = \left(\frac{u}{\|u\|}, e_2, \dots, e_n \right)$ est alors une base orthonormée de E dont tous les vecteurs sont propres pour f .

Donc $\mathcal{P}(n)$ est vraie. ■

On dit que les endomorphismes auto-adjoints sont orthodiagonalisables.

Remarque 9.97

La réciproque est vraie et presque évidente : si un endomorphisme est orthodiagonalisable, alors il est auto-adjoint.

Exercice 9.98 (Un grand classique à savoir refaire)

Soit u un endomorphisme auto-adjoint de E , B la boule-unité fermée de E et S la sphère-unité de E .

On pose α la plus petite des valeurs propres de u et β la plus grande.

Montrez que $\inf_{x \in S} \langle x | u(x) \rangle = \alpha$ et $\sup_{x \in B} \langle x | u(x) \rangle = \sup_{x \in S} \langle x | u(x) \rangle = \beta$.

Correction 9.99

u est auto-adjoint donc orthodagonalisable : il existe une base orthonormée $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que $\text{Mat}_{\mathcal{B}}(u)$ soit diagonale.

On pose $D = \text{Mat}_{\mathcal{B}}(u) = \text{diag}(\lambda_1, \dots, \lambda_n)$.

On peut choisir d'ordonner les valeurs propres par ordre croissant :

$$\alpha = \lambda_1 \leq \dots \leq \lambda_n = \beta.$$

Soit $x \in E$ de coordonnées $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$.

On a $\langle x | u(x) \rangle = X^T D X = \sum_{i=1}^n \lambda_i x_i^2$.

Si $x \in B$, alors $\|x\| \leq 1$.

Pour tout $i \in \llbracket 1 ; n \rrbracket$, on a $\lambda_i \leq \beta$ donc $\lambda_i x_i^2 \leq \beta x_i^2$, et donc

$$\begin{aligned} \langle x | u(x) \rangle &= \sum_{i=1}^n \lambda_i x_i^2 \\ &\leq \sum_{i=1}^n \beta x_i^2 \\ &= \beta \|x\|^2 \\ &\leq \beta. \end{aligned}$$

On a montré $\forall x \in B, \langle x | u(x) \rangle \leq \beta$.

Pour $x = e_n$, on a $u(e_n) = \lambda_n e_n = \beta e_n$ donc

$$\langle e_n | u(e_n) \rangle = \langle e_n | \beta e_n \rangle = \beta \|e_n\|^2 = \beta.$$

Ceci prouve $\max_{x \in B} \langle x | u(x) \rangle = \beta$.

Si on remplace B par S , idem.

De même, soit $x \in S$.

Pour tout $i \in \llbracket 1 ; n \rrbracket$, on a $\lambda_i \geq \alpha$ donc $\lambda_i x_i^2 \geq \alpha x_i^2$.

Donc

$$\langle x | u(x) \rangle = \sum_{i=1}^n \lambda_i x_i^2 \geq \alpha \sum_{i=1}^n x_i^2 = \alpha \|x\|^2 = \alpha.$$

On a montré $\forall x \in S, \langle x | u(x) \rangle \geq \alpha$, avec égalité en $x = e_1$.

Donc $\alpha = \min_{x \in S} \langle x | u(x) \rangle$.

Exercice 9.100 (Un prolongement de l'exercice précédent)

Montrez que l'application $N : \mathcal{S}(E) \longrightarrow \mathbb{R}_+$ définie par $N(u) = \sup_{x \in B} |\langle x | u(x) \rangle|$ est une norme sur $\mathcal{S}(E)$.

Le théorème précédent a une version matricielle.

Théorème 9.101

Une matrice réelle est orthosemblable à une matrice diagonale ssi elle est symétrique.

On dit que les matrices symétriques réelles sont orthodiagonalisables.

Exercice 9.102

Orthodiagonalisez la matrice suivante :

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

Remarque 9.103

La condition « réelle » est indispensable dans le théorème spectral !

9.8 Endomorphismes auto-adjoints positifs, définis-positifs

9.8.1 Endomorphismes auto-adjoints positifs

Définition 9.104

Soit $f \in \mathcal{S}(E)$.

On dit que f est un endomorphisme auto-adjoint positif quand pour tout $x \in E$, $\langle f(x) | x \rangle \geq 0$.

On dit que f est un endomorphisme auto-adjoint défini-positif quand pour tout $x \in E \setminus \{0\}$, $\langle f(x) | x \rangle > 0$.

On note $\mathcal{S}^+(E)$ l'ensemble des endomorphismes auto-adjoints positifs et $\mathcal{S}^{++}(E)$ celui des endomorphismes auto-adjoints définis positifs. Attention, ces deux ensembles ne sont pas des espaces vectoriels et ne sont pas stables par composition.

Ces endomorphismes sont couramment présents dans les théories physiques et sont l'objet de propriétés spécifiques.

Remarque 9.105

Si $f \in \mathcal{S}^{++}(E)$, $(x, y) \mapsto \langle x | f(y) \rangle$ est aussi un produit scalaire.

On donne par exemple une caractérisation simple à l'aide de valeurs propres.

Proposition 9.106

Soit $f \in \mathcal{S}(E)$.

On a $f \in \mathcal{S}^+(E)$ ssi les valeurs propres de f sont positives.

De même, $f \in \mathcal{S}^{++}(E)$ ssi les valeurs propres de f sont strictement positives.

Démonstration 9.107

Supposons $f \in \mathcal{S}^+(E)$.

Soient $\lambda \in \text{Sp}(f)$ et x un vecteur propre associé.

Alors $\langle x | f(x) \rangle = \langle x | \lambda x \rangle = \lambda \|x\|^2 \geq 0$.

Or $\|x\|^2 > 0$ donc $\lambda \geq 0$.

Réciproquement, si $f \in \mathcal{S}(E)$ et $\text{Sp}(f) \subseteq [0 ; +\infty[$, on choisit une base \mathcal{B} orthonormée de diagonalisation de f telle que

$$\text{Mat}_{\mathcal{B}}(f) = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Avec $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$, on a $\langle x | f(x) \rangle = \sum_{i=1}^n \lambda_i x_i^2 \geq 0$.

Idem pour $\mathcal{S}^{++}(E)$ en mettant des inégalités strictes. ■

En particulier, $\mathcal{S}^{++}(E) = \mathcal{S}^+(E) \cap \text{GL}(E)$.

9.8.2 Matrices symétriques positives

Définition 9.108

Soit $A \in \mathcal{S}_n(\mathbb{R})$.

On dit que A est une matrice symétrique positive quand pour tout $X \in \mathbb{R}^n$, $X^\top A X \geq 0$.

On dit que A est une matrice symétrique définie-positive quand pour tout $X \in \mathbb{R}^n \setminus \{0\}$, $X^\top A X > 0$.

Les matrices symétriques positives (respectivement définies-positives) sont donc les matrices dans des bases orthonormées des endomorphismes auto-adjoints positifs (respectivement définis-positifs).

On note $\mathcal{S}_n^+(\mathbb{R})$ l'ensemble des matrices symétriques positives et $\mathcal{S}_n^{++}(\mathbb{R})$ celui des matrices symétriques définies-positives. Attention, ces deux ensembles ne sont pas des espaces vectoriels et ne sont pas stables par produit.

Proposition 9.109

Soit $A \in \mathcal{S}_n(\mathbb{R})$.

On a $A \in \mathcal{S}_n^+(\mathbb{R})$ ssi les valeurs propres de A sont positives.

De même, $A \in \mathcal{S}_n^{++}(\mathbb{R})$ ssi les valeurs propres de A sont strictement positives.

Chapitre 10

Fonctions vectorielles

Sommaire

10.1	Dérivée en un point	. 307
10.1.1	Dérivabilité en un point	307
10.1.2	Interprétation géométrique, développement limité d'ordre 1, continuité	308
10.1.3	Dérivées à gauche, dérivées à droite	309
10.1.4	Lien avec les coordonnées	309
10.1.5	Théorèmes opératoires	310
10.2	Fonction dérivée	. 312
10.3	Dérivées successives	. 315
10.3.1	Définitions et exemples	315
10.3.2	Théorèmes opératoires pour les dérivées successives	316
10.4	Intégrales	. 317
10.4.1	Définition	317
10.4.2	Propriétés	318
10.4.3	Primitives d'une fonction continue	319
10.4.4	Formules de Taylor	320

Dans tout le chapitre, I et J désignent des intervalles de \mathbb{R} contenant au moins deux points et n désigne un entier naturel non-nul.

E et F désignent des espaces vectoriels normés de dimension finie. Par choix d'une base et sachant que les normes sont équivalentes, on peut se ramener à une étude sur \mathbb{R}^n .

10.1 Dérivée en un point

10.1.1 Dérivabilité en un point

Définition 10.1

Soient $f : I \longrightarrow E$ et $a \in I$.

On appelle (fonction) taux d'accroissement de f en a l'application

$$\begin{array}{ccc} \tau_a : I \setminus \{a\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{f(x) - f(a)}{x - a} \end{array} .$$

On dit que f est dérivable en a quand $\frac{f(x) - f(a)}{x - a}$ tend vers une limite $\ell \in E$ quand $x \rightarrow a$.

Si f est dérivable en a , on appelle dérivée de f en a le vecteur

$$f'(a) = \ell = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}.$$

Remarque 10.2

Lorsqu'on étudie « à la main » la limite du taux d'accroissement, on effectue très souvent le changement d'origine $h = x - a$ et on étudie $\frac{f(a+h) - f(a)}{h}$ quand $h \rightarrow 0$ (ce qui permet d'utiliser les équivalents ou les développements limités usuels).

10.1.2 Interprétation géométrique, développement limité d'ordre 1, continuité

Proposition 10.3 (Développement limité d'ordre 1 d'une fonction dérivable)

Soient $f : I \rightarrow E$ et $a \in I$.

f est dérivable en a ssi il existe $m \in E$ tel que $f(x) \underset{x \rightarrow a}{=} f(a) + (x - a)m + o(x - a)$.

Lorsque ces énoncés sont vrais, on a $f'(a) = m$.

On peut utiliser un développement limité à un ordre au moins 1 en a pour montrer que la fonction est dérivable en a .

Proposition 10.4

Si une fonction est dérivable en a , alors elle est continue en a .

Remarque 10.5

Évidemment, la réciproque est fausse !

Dans le cas où $f'(a) \neq 0$: un vecteur directeur de la droite passant par $f(a)$ et $f(x)$ est par exemple $f(x) - f(a)$, mais aussi $\frac{f(x) - f(a)}{x - a}$ quand $x \neq a$.

Autrement dit, f est dérivable en a quand ce vecteur directeur de la droite passant par $f(a)$ et $f(x)$ a une limite dans E , autrement dit, quand la courbe décrite par f a une tangente en a : c'est la droite passant par $f(a)$ et dirigée par $f'(a)$.

Si x représente une variable de temps, le vecteur dérivée $f'(a)$ est le vecteur vitesse instantanée au point a . Son sens donne le sens de parcours de la courbe.

10.1.3 Dérivées à gauche, dérivées à droite

Définition 10.6

Soient $f : I \longrightarrow E$ et $a \in I$.

On dit que f est dérivable à droite (respectivement à gauche) en a quand $\frac{f(x) - f(a)}{x - a}$ admet une limite à droite (respectivement à gauche) dans \mathbb{R}^n quand x tend vers a .

Lorsque f est dérivable à droite (respectivement à gauche) en a , on appelle cette limite la dérivée à droite (respectivement à gauche) de f en a et on la note $f'_d(a)$ (respectivement $f'_g(a)$).

Proposition 10.7

Si f est dérivable à droite (respectivement à gauche) en a , alors f est continue à droite (respectivement à gauche) en a .

En outre, f est dérivable en a ssi f est dérivable à droite et à gauche en a et $f'_d(a) = f'_g(a)$.

Lorsque c'est le cas, $f'(a)$ est égale à la valeur commune de $f'_d(a)$ et $f'_g(a)$.

On peut parler, dans le cas de vecteurs non-nuls, de demi-tangentes à gauche ou à droite.

10.1.4 Lien avec les coordonnées

E étant de dimension finie, on choisit une base $\mathcal{B} = (e_1, \dots, e_n)$ de E . Soit $f : I \longrightarrow E$.

On a pour tout $t \in I$, $f(t) = \sum_{i=1}^n f_i(t) e_i$.

On associe ainsi à f ses n fonctions-coordonnées dans la base \mathcal{B} (qui sont des fonctions à valeurs dans \mathbb{R}).

Proposition 10.8

Avec les mêmes notations, il y a équivalence entre « f est dérivable en a » et « les fonctions f_i sont dérivables en a ».

Dans ce cas, on a $f'(a) = \sum_{i=1}^n f'_i(a) e_i$.

Autrement dit, travailler avec une fonction à valeurs dans E revient à travailler avec n fonctions numériques simultanément.

10.1.5 Théorèmes opératoires

Proposition 10.9

Soient $f : I \longrightarrow E$, $g : I \longrightarrow E$, $a \in I$ et $\lambda \in \mathbb{K}$.

Si f et g sont dérivables en a , alors $f + g$ et λf sont dérivables en a et

$$(f + g)'(a) = f'(a) + g'(a) \quad (\lambda f)'(a) = \lambda f'(a).$$

Proposition 10.10

Soient $\varphi : I \longrightarrow \mathbb{R}$ et $f : J \longrightarrow E$ telles que $\varphi(I) \subseteq J$ et $a \in I$.

Si φ est dérivable en a et f est dérivable en $\varphi(a)$, alors $f \circ \varphi$ est dérivable en a et

$$(f \circ \varphi)'(a) = \varphi'(a) f'(\varphi(a)).$$

Proposition 10.11

Soient $f : I \longrightarrow E$, $L \in \mathcal{L}(E, F)$ et $a \in I$.

Si f est dérivable en a , alors $L \circ f$ l'est aussi et

$$(L \circ f)'(a) = L \circ f'(a).$$

Démonstration 10.12

Par linéarité de L , pour $t \in I$, on a

$$\frac{L \circ f(t) - L \circ f(a)}{t - a} = \frac{L(f(t) - f(a))}{t - a} = L \left(\frac{f(t) - f(a)}{t - a} \right).$$

Comme E est de dimension finie, L est continue donc

$$\lim_{t \rightarrow a} \frac{L \circ f(t) - L \circ f(a)}{t - a} = L \left(\lim_{t \rightarrow a} \frac{f(t) - f(a)}{t - a} \right) = L(f'(a))$$

i.e. $L \circ f$ est dérivable en a et $(L \circ f)'(a) = L \circ f'(a)$. ■

Exemple 10.13

Si $A \in \mathcal{M}_n(\mathbb{R})$ et X est une fonction de I dans $\mathcal{M}_{n1}(\mathbb{R})$ dérivable en a , alors $Y : t \mapsto AX(t)$ est dérivable en a et $Y'(a) = AX'(a)$.

Proposition 10.14

Soient $f, g : I \longrightarrow E$ et $B : E^2 \longrightarrow F$ bilinéaire.

Si f et g sont dérivables en a , alors (f, g) l'est aussi et $(B(f, g))'(a) = B(f', g)(a) + B(f, g')(a)$.

Démonstration 10.15

Pour $t \in I$, on a

$$\begin{aligned}
B(f(t), g(t)) - B(f(a), g(a)) &= B(f(t) - f(a), g(t)) + B(f(a), g(t)) - B(f(a), g(a)) \\
&= B(f(t) - f(a), g(t)) + B(f(a), g(t) - g(a)).
\end{aligned}$$

Comme E est de dimension finie, B est continue donc

$$\begin{aligned}
\lim_{t \rightarrow a} \frac{B(f(t), g(t)) - B(f(a), g(a))}{t - a} &= \lim_{t \rightarrow a} \left[B\left(\frac{f(t) - f(a)}{t - a}, g(t)\right) + B\left(f(a), \frac{g(t) - g(a)}{t - a}\right) \right] \\
&= B(f'(a), g(a)) + B(f(a), g'(a)).
\end{aligned}$$

■

Exemple 10.16

- Si A et B sont deux fonctions de I dans $\mathcal{M}_n(\mathbb{R})$ dérivables en a , alors $M : t \mapsto A(t)B(t)$ est aussi dérivable en a et $M'(a) = A'(a)B(a) + A(a)B'(a)$.
- Si u et v sont deux fonctions à valeurs dans E , espace euclidien, et dérivables en a , alors $p : t \mapsto \langle u(t) | v(t) \rangle$ est dérivable en a et $p'(a) = \langle u'(a) | v(a) \rangle + \langle u(a) | v'(a) \rangle$.

Proposition 10.17

Soient $f_1, \dots, f_p : I \rightarrow E$, $a \in I$ et M une application p -linéaire de E^p dans F .

Si f_1, \dots, f_p sont dérivables en a , alors $M(f_1, \dots, f_p)$ l'est aussi et

$$(M(f_1, \dots, f_p))'(a) = \sum_{i=1}^p M(f_1, \dots, f_i', \dots, f_p)(a).$$

Démonstration 10.18

Si on considère $f, g, h : I \rightarrow E$ dérivables en a et $T : E^3 \rightarrow F$ trilinéaire, pour $t \in I$, on a

$$\begin{aligned}
&T(f(t), g(t), h(t)) - T(f(a), g(a), h(a)) \\
&= T(f(t) - f(a), g(t), h(t)) + T(f(a), g(t), h(t)) - T(f(a), g(a), h(a)) \\
&= T(f(t) - f(a), g(t), h(t)) + T(f(a), g(t) - g(a), h(t)) + T(f(a), g(a), h(t)) \\
&\quad - T(f(a), g(a), h(a)) \\
&= T(f(t) - f(a), g(t), h(t)) + T(f(a), g(t) - g(a), h(t)) + T(f(a), g(a), h(t) - h(a)).
\end{aligned}$$

On en déduit le cas général : pour $t \in I$, on a

$$\begin{aligned}
&M(f_1(t), \dots, f_p(t)) - M(f_1(a), \dots, f_p(a)) \\
&= \sum_{j=1}^p M(f_1(a), \dots, f_{j-1}(a), f_j(t) - f_j(a), f_{j+1}(a), \dots, f_p(a)).
\end{aligned}$$

Et il s'ensuit la conclusion, de la même manière que dans les cas linéaire et bilinéaire.

■

Exemple 10.19

Si M est une fonction de I dans $\mathcal{M}_n(\mathbb{R})$ dérivable en $a \in I$, alors $d : t \mapsto \det M(t)$ est aussi dérivable en a et $d'(a) = \sum_{k=1}^n \det M_k^c(a)$ où $M_k^c(a)$ est la matrice obtenue à partir de $M(a)$ en remplaçant sa k -ème colonne par sa dérivée en a .

On a aussi $d'(a) = \sum_{k=1}^n \det M_k^\ell(a)$ où $M_k^\ell(a)$ est la matrice obtenue à partir de $M(a)$ en remplaçant sa k -ème ligne par sa dérivée en a .

10.2 Fonction dérivée

Définition 10.20

Soit $f : I \longrightarrow E$.

- On dit que f est dérivable sur I si f est dérivable en tout point $a \in I$, ce qui revient à dire, après choix d'une base, que les fonctions coordonnées de f sont dérivables sur I .
- Si f est dérivable sur I , on définit sa fonction dérivée par

$$\begin{aligned} f' : I &\longrightarrow \mathbb{R} \\ x &\longmapsto f'(x) \end{aligned}$$

Proposition 10.21 (Théorèmes d'opérations sur les fonctions dérivables)

Soient f, g deux fonctions définies sur I et $\lambda \in \mathbb{K}$.

Si f et g sont dérivables sur I , alors $f + g$ et λf sont dérivables sur I et

$$(f + g)' = f' + g' \quad (\lambda f)' = \lambda f'.$$

Proposition 10.22

Soient $\varphi : I \longrightarrow \mathbb{R}$ et $f : J \longrightarrow E$ telles que $\varphi(I) \subseteq J$.

Si φ est dérivable sur I et f est dérivable sur J , alors $f \circ \varphi$ est dérivable sur I et

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

Proposition 10.23

Soient $f : I \longrightarrow E$ et $L \in \mathcal{L}(E, F)$.

Si f est dérivable sur I , alors $L \circ f$ l'est aussi et $(L \circ f)' = L \circ f'$.

Proposition 10.24

Soient $f, g : I \longrightarrow E$ et $B : E^2 \longrightarrow F$ bilinéaire.

Si f et g sont dérivables sur I , alors $B(f, g)$ l'est aussi et $(B(f, g))' = B(f', g) + B(f, g')$.

Proposition 10.25

Soient $f_1, \dots, f_p : I \longrightarrow E$ et $M : E^p \longrightarrow F$ p -linéaire.

Si f_1, \dots, f_p sont dérivables sur I , alors $M(f_1, \dots, f_p)$ l'est aussi et

$$(M(f_1, \dots, f_p))' = \sum_{k=1}^p M(f_1, \dots, f_k', \dots, f_p).$$

Exercice 10.26

Montrez que si un point M se déplace sur une sphère de centre A , sa vitesse est toujours orthogonale au vecteur \overrightarrow{AM} .

Montrez que la réciproque est vraie : si A est un point fixe et si la vitesse de M est toujours orthogonale au vecteur \overrightarrow{AM} , alors M se déplace sur une sphère.

Correction 10.27

$M : I \longrightarrow E$ est dérivable sur I .

On suppose $t \longmapsto AM(t)$ constante i.e. $t \longmapsto AM(t)^2$ constante.

Pour $t \in I$, on remarque

$$AM(t)^2 = \|\overrightarrow{AM(t)}\|^2 = \left\langle \overrightarrow{AM(t)} \mid \overrightarrow{AM(t)} \right\rangle$$

donc

$$\begin{aligned} \frac{d}{dt} (AM(t)^2) &= \left\langle \frac{d}{dt} (\overrightarrow{AM(t)}) \mid \overrightarrow{AM(t)} \right\rangle + \left\langle \overrightarrow{AM(t)} \mid \frac{d}{dt} (\overrightarrow{AM(t)}) \right\rangle \\ &= 2 \left\langle \overrightarrow{AM(t)} \mid \frac{d}{dt} (\overrightarrow{AM(t)}) \right\rangle. \end{aligned}$$

Or $\frac{d}{dt} (AM(t)^2) = 0$, d'où l'orthogonalité.

Et réciproquement en remontant les calculs.

Exercice 10.28

Soit $S : I \longrightarrow \mathcal{M}_n(\mathbb{R})$ dérivable sur I telle que pour tout $t \in I$, $S(t)$ est une matrice de symétrie.

Montrez que pour tout $t \in I$, $\text{tr}(S(t)S'(t)) = 0$.

Correction 10.29

Soit $t \in I$.

$S(t)$ est une matrice de symétrie donc $S(t)^2 = I_n$.

Donc $S(t) S'(t) + S'(t) S(t) = 0$.

Donc $\text{tr}(S(t) S'(t)) + \text{tr}(S'(t) S(t)) = 0$.

Donc $2 \text{tr}(S(t) S'(t)) = 0$.

Donc $\text{tr}(S(t) S'(t)) = 0$.

Exercice 10.30

Soit $A : I \longrightarrow \text{GL}_n(\mathbb{R})$ dérivable sur I .

Montrez que $B : t \longmapsto A^{-1}(t)$ est dérivable sur I et calculez sa dérivée en fonction de celle de A .

Correction 10.31

Pour $t \in I$, on a

$$A^{-1}(t) = \frac{1}{\det A(t)} (\text{Com } A(t))^{\top}.$$

A étant dérivable sur I , ses coefficients le sont aussi, donc $\text{Com } A$ et $\det A$ sont dérivables sur I , et donc A^{-1} l'est aussi.

Sachant que A et A^{-1} sont dérivables sur I et pour tout $t \in I$, $A(t) B(t) = I_n$, par dérivation, pour $t \in I$, on obtient

$$A'(t) B(t) + A(t) B'(t) = 0$$

donc

$$A^{-1}(t) A'(t) B(t) + B'(t) = 0.$$

D'où

$$(A^{-1})' = -A^{-1} A' A^{-1}.$$

Exercice 10.32

Soit $M : I \longrightarrow \mathcal{M}_n(\mathbb{R})$ dérivable sur I , intervalle contenant 0. On pose $f : t \longmapsto \det(I_n + tM(t))$.

Justifiez que f est dérivable sur I et calculez $f'(0)$.

Correction 10.33

La fonction $t \mapsto M(t)$ est dérivable sur I donc tous ses coefficients le sont aussi donc f est dérivable par sommes et produits de fonctions dérivables.

Si M est constante : $M = (m_{i,j})$, on a

$$f(t) = \begin{vmatrix} 1 + tm_{1,1} & tm_{1,2} & \dots & tm_{1,n} \\ tm_{2,1} & 1 + tm_{2,2} & \dots & tm_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ tm_{n,1} & \dots & tm_{n,n-1} & 1 + tm_{n,n} \end{vmatrix}.$$

On obtient donc $f'(t)$ en sommant tous les déterminants obtenus en dérivant colonne après colonne.

On a donc $f'(0) = m_{1,1} + \dots + m_{n,n} = \text{tr } M$.

Plus généralement, on a donc $f'(0) = \text{tr } M(0)$.

10.3 Dérivées successives

10.3.1 Définitions et exemples

Si f est dérivable sur I , f' est une fonction définie sur I . On peut donc essayer de la dériver : quand c'est possible, on obtient la dérivée seconde f'' , et ainsi de suite...

Définition 10.34

Soit $f : I \longrightarrow E$. On définit, par récurrence, les notions suivantes :

- Par convention, on dit que f est toujours dérivable 0 fois sur I et on définit la dérivée d'ordre 0 de f par $f^{(0)} = f$.
- Soit $k \in \mathbb{N}^*$.

On dit que f est dérivable k fois sur I si elle est dérivable $(k-1)$ fois sur I et que sa dérivée d'ordre $(k-1)$, la fonction $f^{(k-1)}$, est dérivable sur I .

On définit alors la dérivée d'ordre k par

$$f^{(k)} = \left(f^{(k-1)} \right)'.$$

Remarque 10.35

$f^{(0)}$ désigne f , $f^{(1)}$ désigne f' et $f^{(2)}$ est aussi notée f'' . À partir de trois dérivations, on n'utilise plus de primes.

La dérivée d'ordre k de f est également notée $\frac{d^k f}{dt^k}$.

Enfin, il est facile de montrer que f est $(p+q)$ fois dérivable sur I ssi f est p fois dérivable et $f^{(p)}$ est q fois dérivable sur I . Dans ce cas, on a l'égalité

$$\left(f^{(p)}\right)^{(q)} = f^{(p+q)}.$$

Définition 10.36

Soit $k \in \mathbb{N}$.

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

On dit que f est de classe \mathcal{C}^∞ sur I si f est dérivable k fois sur I quel que soit $k \in \mathbb{N}$.

Pour tout $k \in \mathbb{N} \cup \{\infty\}$, on notera $\mathcal{C}^k(I, E)$ l'ensemble des fonctions de classe \mathcal{C}^k sur I .

Remarque 10.37

« f est de classe \mathcal{C}^0 sur I » signifie « f est continue sur I ».

« f est de classe \mathcal{C}^1 sur I » signifie « f est dérivable sur I et f' est continue sur I ».

Proposition 10.38

Soient f une fonction définie sur I et $k \in \mathbb{N}^*$.

- ▷ Si f est de classe \mathcal{C}^k sur I , alors pour tout $p \in \llbracket 0 ; k \rrbracket$, f est de classe \mathcal{C}^p sur I .
- ▷ f est de classe \mathcal{C}^k sur I ssi f est dérivable sur I et f' est de classe \mathcal{C}^{k-1} sur I , ou, ce qui revient au même, f est de classe \mathcal{C}^{k-1} sur I et $f^{(k-1)}$ est de classe \mathcal{C}^1 sur I .

Les ensembles $\mathcal{C}^k(I, E)$ forment donc une chaîne d'inclusions :

$$\mathcal{C}^\infty(I, E) \subseteq \dots \subseteq \mathcal{C}^{k+1}(I, E) \subseteq \mathcal{C}^k(I, E) \subseteq \dots \subseteq \mathcal{C}^1(I, E) \subseteq \mathcal{C}^0(I, E).$$

10.3.2 Théorèmes opératoires pour les dérivées successives

Proposition 10.39

Soient $k \in \mathbb{N} \cup \{\infty\}$, $f, g : I \longrightarrow E$ deux fonctions de classe \mathcal{C}^k sur I et $\lambda \in \mathbb{K}$.

On a :

- ▷ $f + g$ est de classe \mathcal{C}^k sur I et $(f + g)^{(k)} = f^{(k)} + g^{(k)}$
- ▷ λf est de classe \mathcal{C}^k sur I et $(\lambda f)^{(k)} = \lambda f^{(k)}$.

Proposition 10.40

Soient $k \in \mathbb{N} \cup \{\infty\}$, $\varphi : I \longrightarrow \mathbb{R}$ et $f : J \longrightarrow E$.

Si $\varphi(I) \subseteq J$ et φ et f sont de classe \mathcal{C}^k sur I et J respectivement, alors $f \circ \varphi$ est de classe \mathcal{C}^k sur I .

Proposition 10.41

Soient $k \in \mathbb{N} \cup \{\infty\}$, $f : I \longrightarrow E$ et $L \in \mathcal{L}(E, F)$.

Si f est une fonction de classe \mathcal{C}^k sur I , alors $L \circ f$ l'est aussi et $(L \circ f)^{(k)} = L \circ f^{(k)}$.

Proposition 10.42

Soient $k \in \mathbb{N} \cup \{\infty\}$, $f, g : I \longrightarrow E$ et $B : E^2 \longrightarrow F$ bilinéaire.

Si f et g sont de classe \mathcal{C}^k sur I , alors $B(f, g)$ l'est aussi et, d'après la formule de Leibniz :

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

10.4 Intégrales

10.4.1 Définition

Définition 10.43

Soit $f : [a ; b] \longrightarrow E$ une fonction définie sur le segment $[a ; b]$.

On dit que f est continue par morceaux sur $[a ; b]$ quand il existe une subdivision (c_0, \dots, c_n) de $[a ; b]$ telle que :

- ▷ pour tout $i \in \llbracket 0 ; n-1 \rrbracket$, f est continue sur l'intervalle ouvert $]c_i ; c_{i+1}[$
- ▷ f a une limite réelle en a à droite, en b à gauche et des limites réelles à gauche et à droite en chaque point c_i tel que $1 \leq i \leq n-1$.

Toute subdivision qui convient dans cette définition est dite adaptée à f .

On choisit une base $\mathcal{B} = (e_1, \dots, e_n)$ de E . Il est alors évident qu'une fonction à valeurs dans E est continue par morceaux ssi ses fonctions coordonnées dans la base \mathcal{B} le sont aussi.

Définition 10.44

Soit $f : [a ; b] \longrightarrow E$ continue par morceaux sur $[a ; b]$.

On note f_1, \dots, f_n ses fonctions coordonnées dans la base \mathcal{B} , i.e. $f = \sum_{i=1}^n f_i e_i$.

On pose alors

$$\int_{[a;b]} f = \sum_{i=1}^n \left(\int_{[a;b]} f_i \right) e_i.$$

Bien sûr, les notations classiques pour les intégrales sont conservées : $\int_a^b f$, $\int_a^b f(t) dt$, etc.

On pose encore $\int_b^a f = - \int_a^b f$.

Cette définition est *a priori* ambiguë car elle dépend de la base \mathcal{B} choisie. On montre aisément qu'en fait ce n'est pas le cas : on obtient toujours le même vecteur intégrale, indépendamment de la base choisie.

10.4.2 Propriétés

En se ramenant aux coordonnées dans une base, on retrouve les propriétés essentielles de l'intégrale.

Proposition 10.45

L'application $\mathcal{C}_m^0([a ; b]) \longrightarrow E$ est linéaire.

$$f \longmapsto \int_a^b f$$

Proposition 10.46

La relation de Chasles reste valable : pour tout $(a, b, c) \in I^3$, si f est continue par morceaux sur I , alors $\int_a^b f = \int_a^c f + \int_c^b f$.

On retrouve une inégalité triangulaire avec la norme (n'importe laquelle!).

Proposition 10.47

Soit $f : [a ; b] \longrightarrow E$ continue par morceaux sur $[a ; b]$.

La fonction $t \longmapsto \|f(t)\|$ est continue par morceaux sur $[a ; b]$ et à valeurs réelles.

De plus, on a

$$\left\| \int_a^b f \right\| \leq \int_a^b \|f\|.$$

On retrouve la notion de sommes de Riemann.

Proposition 10.48

Soit $f : [a ; b] \longrightarrow E$ continue par morceaux.

Pour tout $n \in \mathbb{N}^*$, on introduit les sommes de Riemann associées à la fonction f sur $[a ; b]$: pour $k \in \llbracket 0 ; n \rrbracket$, on pose $c_k = a + k \frac{b-a}{n}$ et

$$S_n = \frac{b-a}{n} \sum_{k=0}^{n-1} f(c_k) \quad S'_n = \frac{b-a}{n} \sum_{k=1}^n f(c_k) \quad S''_n = \frac{b-a}{n} \sum_{k=0}^n f(c_k).$$

Les suites (S_n) , (S'_n) et (S''_n) tendent toutes les trois vers $\int_{[a;b]} f(x) dx$.

Enfin, une petite nouveauté.

Proposition 10.49

Soient $f : [a ; b] \longrightarrow E$ continue par morceaux et $L \in \mathcal{L}(E, F)$.

La fonction $L \circ f = L(f)$ est continue par morceaux et $\int_a^b L(f) = L\left(\int_a^b f\right)$.

10.4.3 Primitives d'une fonction continue

Proposition 10.50

Soient $f : I \longrightarrow E$ continue et $a \in I$. On pose $\Phi : x \longmapsto \int_a^x f(t) dt$.

Φ est l'unique primitive de f sur I qui s'annule en a .

Remarque 10.51

On a donc montré que si f est continue sur I , alors la fonction $\Phi : x \longmapsto \int_a^x f$ est de classe \mathcal{C}^1 sur I , car pour tout $x \in I$, $\Phi'(x) = f(x)$, et non pas $\Phi'(x) = f(x) - f(a)$!

Corollaire 10.52

Toute fonction continue sur un intervalle y admet des primitives.

On en déduit l'inégalité des accroissements finis pour une fonction de classe \mathcal{C}^1 (dérivable ne suffit pas).

Proposition 10.53

Soit $f : I \longrightarrow E$ une fonction de classe \mathcal{C}^1 sur I .

Pour tout $(a, b) \in I^2$, $\|f(b) - f(a)\| \leq |b - a| \sup_{[a;b]} \|f'\|$.

10.4.4 Formules de Taylor

On retrouve encore par utilisation des fonctions coordonnées les formules de Taylor usuelles.

La formule de Taylor avec reste intégral.

Proposition 10.54

Soient f une fonction de classe \mathcal{C}^{n+1} sur un intervalle I et $a, b \in I$.

On a

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(a)}{2}(b-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t) dt.$$

L'inégalité de Taylor-Lagrange qui s'en déduit.

Proposition 10.55

Soient f une fonction de classe \mathcal{C}^{n+1} sur un intervalle I et $a, b \in I$.

On a

$$\left\| f(b) - f(a) - f'(a)(b-a) - \frac{f''(a)}{2}(b-a)^2 - \cdots - \frac{f^{(n)}(a)}{n!}(b-a)^n \right\| \leq \frac{|b-a|^{n+1}}{(n+1)!} \sup_{[a,b]} \|f^{(n+1)}\|.$$

Et la formule de Taylor-Young.

Proposition 10.56

Soit $f : I \longrightarrow E$ de classe \mathcal{C}^n sur I .

Pour tout $t_0 \in I$, f possède un développement limité en t_0 à l'ordre n : il existe une fonction ε définie sur I et à valeurs dans E telle que

$$\begin{cases} \text{pour tout } t \in I, & f(t) = f(t_0) + f'(t_0)(t-t_0) + \frac{f''(t_0)}{2}(t-t_0)^2 + \cdots + \frac{f^{(n)}(t_0)}{n!}(t-t_0)^n + (t-t_0)^n \varepsilon(t) \\ \varepsilon(t) \xrightarrow[t \rightarrow t_0]{} 0 \end{cases}$$

Chapitre 11

Suites et séries de fonctions

Sommaire

11.1	Convergence d'une suite de fonctions	. 321
11.1.1	Convergence simple	322
11.1.2	Convergence uniforme	322
11.2	Convergence d'une série de fonctions	. 329
11.2.1	Convergence simple	329
11.2.2	Convergence uniforme	331
11.2.3	Convergence normale	333
11.3	Propriétés de la fonction limite	. 337
11.3.1	Monotonie	337
11.3.2	Continuité	338
11.3.3	Interversion de limite et d'intégrale	346
11.3.4	Interversion de limites	351
11.3.5	Dérivabilité	353
11.3.6	Dérivation à un ordre plus élevé	356
11.4	Généralisation	. 357
11.4.1	Convergence simple	357
11.4.2	Convergence uniforme	357
11.4.3	Convergence normale des séries	358
11.4.4	Résultats préservés	358
11.5	Approximation uniforme	. 358
11.5.1	Densité des fonctions en escaliers dans les fonctions continues par morceaux	358
11.5.2	Densité des polynômes sur un segment dans les fonctions continues	360

Dans ce chapitre, \mathbb{K} désigne \mathbb{R} ou \mathbb{C} . Toutes les fonctions considérées vont de \mathbb{R} dans \mathbb{K} dans un premier temps. Dans un second temps, on généralisera les définitions et résultats à des fonctions de E dans F , deux espaces vectoriels normés de dimension finies.

11.1 Convergence d'une suite de fonctions

Comme son nom l'indique, une suite de fonctions est une suite (f_n) où f_n est une fonction de \mathbb{R} dans \mathbb{K} .

11.1.1 Convergence simple

Définition 11.1

Soient A une partie de \mathbb{R} et (f_n) une suite de fonctions définies sur A .

On dit que la suite (f_n) converge simplement sur A quand pour tout $x \in A$, la suite numérique $(f_n(x))$ converge.

Dans ce cas, on peut définir une fonction f sur A en posant pour tout $x \in A$, $f(x) = \lim_{n \rightarrow +\infty} f_n(x)$.

La fonction f est alors appelée limite simple sur A de la suite (f_n) et on dit que la suite (f_n) converge simplement vers f sur A .

On observe la définition formelle de la convergence simple sur A :

$$\forall x \in A, \forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, |f_n(x) - f(x)| \leq \varepsilon.$$

Le rang n_0 dépend à la fois de x et de ε .

Exercice 11.2

Étudiez la convergence simple de la suite de fonctions $f_n : x \mapsto \frac{ne^{-x} + x^2}{n + x}$, où $n > 0$, sur $[0 ; +\infty[$.

Exercice 11.3

Même question avec la suite de fonctions $f_n : x \mapsto \frac{x^n}{1 + x^n}$, où $n > 0$, sur $[0 ; +\infty[$.

Exercice 11.4

Même question avec la suite de fonctions $f_n : x \mapsto n^\alpha x^n (1 - x)$, où $n > 0$ et $\alpha \in \mathbb{R}_+^*$, sur $[0 ; 1]$.

11.1.2 Convergence uniforme

Rappel 11.5

Si g est une fonction bornée sur une partie A de \mathbb{R} , on pose $\|g\|_\infty^A = \sup_{x \in A} |g(x)|$.

$\|\cdot\|_\infty^A$ est une norme sur le \mathbb{K} -espace vectoriel des fonctions bornées sur A , appelée norme infinie ou norme uniforme sur A .

Un résultat essentiel : la majoration de la norme uniforme.

Proposition 11.6

Soient f une fonction bornée sur une partie A de \mathbb{R} et K un réel positif.

On a

$$\|f\|_{\infty}^A \leq K \iff \forall x \in A, |f(x)| \leq K.$$

Définition 11.7

Soient A une partie de \mathbb{R} , (f_n) une suite de fonctions définies sur A et f une fonction définie sur A .

On dit que la suite (f_n) converge uniformément vers f sur A quand pour tout $n \in \mathbb{N}$, la fonction $f - f_n$ est bornée sur A et la suite réelle $(\|f - f_n\|_{\infty}^A)$ converge vers 0.

La fonction f est alors appelée limite uniforme sur A de la suite (f_n) .

On observe la définition formelle de la convergence uniforme sur A :

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, \forall x \in A, |f_n(x) - f(x)| \leq \varepsilon.$$

Le rang n_0 dépend seulement de ε mais pas de x : c'est le même n_0 pour toutes les valeurs de x , en ce sens, il est uniforme.

Souvent, on ne sait pas calculer les normes uniformes des fonctions $f_n - f$. Une simple majoration suffit, qu'on appelle majoration uniforme.

Définition 11.8

Soient A une partie de \mathbb{R} et (f_n) une suite de fonctions définies sur A .

On appelle majoration uniforme sur A toute proposition du type

$$\forall n \in \mathbb{N}, \forall x \in A, |f_n(x)| \leq K_n$$

où K_n est une constante indépendante de x .

Exemple 11.9

► La proposition

$$\forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}, \left| \frac{\sin(nx)}{n} \right| \leq \frac{1}{n}$$

est une majoration uniforme des fonctions $x \mapsto \frac{\sin(nx)}{n}$ sur \mathbb{R} .

► La proposition

$$\forall n \in \mathbb{N}^*, \forall x \in \mathbb{R}, \left| \frac{\sin(nx)}{n+x^2} \right| \leq \frac{1}{n+x^2}$$

n'est pas une majoration uniforme des fonctions $x \mapsto \frac{\sin(nx)}{n+x^2}$ sur \mathbb{R} car le majorant dépend de x .

Exercice 11.10

Donnez une majoration uniforme des fonctions $x \mapsto \frac{\ln(1+nx)}{x}$ sur $]0; +\infty[$.

Correction 11.11

On a

$$\forall n \in \mathbb{N}, \quad \forall x \in]0; +\infty[, \quad 0 \leq \frac{\ln(1+nx)}{x} \leq \frac{nx}{x} = n.$$

Exercice 11.12

Même exercice avec les fonctions $x \mapsto \frac{\sin(nx)}{\sin x}$ sur $]-\pi; \pi[\setminus \{0\}$.

Correction 11.13

Pour $n \in \mathbb{N}$, on pose $f_n :]-\pi; \pi[\setminus \{0\} \longrightarrow \mathbb{R}$

$$x \longmapsto \frac{\sin(nx)}{\sin x}$$

Soit $n \in \mathbb{N}$.

Pour $x \in \left[\frac{\pi}{6}; \frac{5\pi}{6}\right]$, on a $\frac{1}{2} \leq \sin x$ donc

$$|f_n| = \frac{|\sin(nx)|}{\sin x} \leq 2 |\sin x| \leq 2.$$

Pour $x \in \left]0; \frac{\pi}{6}\right]$, comme \sin est concave sur $\left]0; \frac{\pi}{6}\right]$, on a $\frac{3}{\pi}x \leq \sin x$ et donc

$$|f_n(x)| = \frac{|\sin(nx)|}{\sin x} \leq \frac{|\sin(nx)|}{x} \times \frac{\pi}{3} \leq \frac{|nx|}{x} \times \frac{\pi}{3} = \frac{n\pi}{3}.$$

Pour $x \in \left[\frac{5\pi}{6}; \pi\right]$, on pose $t = \pi - x \in \left]0; \frac{\pi}{6}\right]$ et on a

$$\frac{\sin(nx)}{\sin x} = \frac{\sin(n(\pi - t))}{\sin(\pi - t)} = \frac{(-1)^n \sin(-nt)}{\sin t} = \frac{(-1)^{n+1} \sin(nt)}{\sin t}$$

donc

$$|f_n(x)| = \left| \frac{\sin(nt)}{\sin t} \right| \leq \frac{n\pi}{3}.$$

Finalement, pour tout $x \in]0; \pi[$, $|f_n(x)| \leq \max\left(2, \frac{n\pi}{3}\right)$.

De plus, $|f_n|$ est paire donc ceci reste vrai sur $]-\pi; 0[$.

Proposition 11.14

Avec les mêmes hypothèses, il y a équivalence entre

- (1) la suite (f_n) converge uniformément vers f sur A
- (2) il existe une suite positive (α_n) telle que pour tout $n \in \mathbb{N}$, $\|f_n - f\|_\infty^A \leq \alpha_n$ et $\alpha_n \xrightarrow{n \rightarrow +\infty} 0$
- (3) il existe une majoration uniforme des fonctions $f_n - f$ sur A par les termes d'une suite numérique positive qui converge vers 0, i.e. on a une proposition du type

$$\forall n \in \mathbb{N}, \forall x \in A, |f_n(x) - f(x)| \leq \alpha_n \quad \text{et} \quad \alpha_n \xrightarrow{n \rightarrow +\infty} 0$$

où α_n ne dépend pas de x .

Démonstration 11.15

(1) \implies (2) : on prend $\alpha_n = \|f_n - f\|_\infty^A$.

(2) \implies (1) : théorème d'encadrement.

(2) \iff (3) : $\|f_n - f\|_\infty^A \leq \alpha_n \iff \forall x \in A, |f_n(x) - f(x)| \leq \alpha_n$. ■

Il y a un lien entre convergence simple et uniforme, dans un seul sens!

Théorème 11.16

Si une suite (f_n) converge uniformément vers f sur A , alors (f_n) converge simplement vers f sur A .

Démonstration 11.17

On a $\forall x \in A, |f_n(x) - f(x)| \leq \|f_n - f\|_\infty^A$.

Donc si (f_n) converge uniformément vers f sur A , $\|f_n - f\|_\infty^A \xrightarrow{n \rightarrow +\infty} 0$.

Donc $|f_n(x) - f(x)| \xrightarrow{n \rightarrow +\infty} 0$ par encadrement, i.e. $f_n(x) \xrightarrow{n \rightarrow +\infty} f(x)$ i.e. (f_n) converge simplement vers f sur A . ■

La réciproque est fausse! Contre-exemple : la suite de fonctions $(x \mapsto x^n)$ sur $[0 ; 1]$.

Démonstration 11.18

On pose $f_n : x \mapsto x^n$ sur $[0 ; 1]$.

(f_n) converge simplement vers $f : x \mapsto \begin{cases} 0 & \text{si } x \in [0 ; 1[\\ 1 & \text{sinon} \end{cases}$

On a

$$\|f_n - f\|_\infty^{[0;1]} = \sup_{x \in [0;1]} |x^n - f(x)| = 1 \not\xrightarrow{n \rightarrow +\infty} 0.$$

Donc (f_n) ne converge pas uniformément sur $[0 ; 1]$. ■

Conséquence : pour étudier la convergence uniforme d'une suite de fonctions, on commence par étudier sa convergence simple, car d'abord on détermine sa limite simple f , puis on cherche à savoir si elle est limite uniforme.

Exercice 11.19

Étudiez la convergence uniforme de la suite de fonctions $f_n : x \mapsto \frac{ne^{-x} + x^2}{n+x}$, où $n > 0$, sur $[0 ; 1]$.

Correction 11.20

Pour tout $x \in [0 ; 1]$, $f_n(x) = \frac{ne^{-x} + x^2}{n+x} \xrightarrow{n \rightarrow +\infty} e^{-x}$ donc (f_n) converge simplement vers $f : x \mapsto e^{-x}$ sur $[0 ; 1]$.

Pour $n \in \mathbb{N}^*$ et $x \in [0 ; 1]$, on a

$$\begin{aligned} |f_n(x) - f(x)| &= \left| \frac{ne^{-x} + x^2}{n+x} - e^{-x} \right| \\ &= \left| \frac{x^2 - xe^{-x}}{n+x} \right| \\ &\leq \frac{|x^2| + |xe^{-x}|}{n+x} \\ &= \frac{2}{n}. \end{aligned}$$

Donc $\|f_n - f\|_{\infty}^{[0;1]} \leq \frac{2}{n}$ et, par encadrement, on a

$$\|f_n - f\|_{\infty}^{[0;1]} \xrightarrow{n \rightarrow +\infty} 0$$

i.e. (f_n) converge uniformément vers f sur $[0 ; 1]$.

De même, on peut montrer la convergence uniforme sur $[0 ; b]$ pour tout $b > 0$.

En revanche, il n'y a pas convergence uniforme sur $[0 ; +\infty[$ car $\lim_{x \rightarrow +\infty} |f_n(x) - f(x)| = +\infty$ donc $f_n - f$ n'est pas bornée sur $[0 ; +\infty[$ donc $\|f_n - f\|_{\infty}^{[0;+\infty[}$ n'est pas défini.

Remarque 11.21

Cet exemple prouve que la convergence uniforme n'est pas préservée par réunion d'intervalles.

Exercice 11.22

Même question avec la suite de fonctions $f_n : x \mapsto \frac{x^n}{1+x^n}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.23

(f_n) converge simplement vers $f : x \mapsto \begin{cases} 0 & \text{si } x < 1 \\ 1/2 & \text{si } x = 1 \\ 1 & \text{sinon} \end{cases}$ sur $[0 ; +\infty[$.

On a $f_n - f : x \mapsto \begin{cases} \frac{x^n}{1+x^n} & \text{si } x < 1 \\ 0 & \text{si } x = 1 \\ \frac{-1}{1+x^n} & \text{sinon} \end{cases}$ donc $|f_n - f| : x \mapsto \begin{cases} \frac{x^n}{1+x^n} & \text{si } x < 1 \\ 0 & \text{si } x = 1 \\ \frac{1}{1+x^n} & \text{sinon} \end{cases}$

On remarque $\lim_{x \rightarrow 1^+} |f_n(x) - f(x)| = \frac{1}{2}$ donc $\|f_n - f\|_{\infty}^{[0;+\infty[} \geq \frac{1}{2}$.

Donc (f_n) ne converge pas uniformément sur $[0 ; +\infty[$.

Exercice 11.24

Même question avec la suite de fonctions $f_n : x \mapsto n^\alpha x^n (1-x)$, où $n > 0$ et $\alpha \in \mathbb{R}_+^*$, sur $[0 ; 1]$.

Correction 11.25

Si $x < 1$, alors $\frac{1}{x} > 1$ et donc $n^\alpha = o\left(\frac{1}{x^n}\right)$, donc $\lim_{n \rightarrow +\infty} n^\alpha x^n = 0$ et enfin $\lim_{n \rightarrow +\infty} f_n(x) = 0$.

Si $x = 1$, on a $f_n(1) = 0 \xrightarrow{n \rightarrow +\infty} 0$.

Donc (f_n) converge simplement vers la fonction nulle sur $[0 ; 1]$.

Comme f_n est continue sur le compact $[0 ; 1]$, on a

$$\|f_n\|_{\infty}^{[0;1]} = \sup_{[0;1]} |f_n| = \sup_{[0;1]} f_n = \max_{[0;1]} f_n.$$

De plus, pour $x \in [0 ; 1]$, on a

$$\begin{aligned} f'_n(x) &= n^\alpha (nx^{n-1} - (n+1)x^n) \\ &= n^\alpha x^{n-1} (n - (n+1)x). \end{aligned}$$

On a donc le tableau de variations suivant :

x	0	$\frac{n}{n+1}$	1
$f'_n(x)$	0	+	0
f_n	0		0

Donc

$$\|f_n\|_{\infty}^{[0;1]} = f_n\left(\frac{n}{n+1}\right) = n^{\alpha} \left(\frac{n}{n+1}\right)^n \left(\frac{1}{n+1}\right).$$

De plus, on a $\left(\frac{n}{n+1}\right)^n = \exp\left(n \ln \frac{n}{n+1}\right)$.

Or

$$n \ln \frac{n}{n+1} = n \ln \frac{1}{1+1/n} = -n \ln \left(1 + \frac{1}{n}\right) \underset{n \rightarrow +\infty}{\sim} \frac{-n}{n} \xrightarrow{n \rightarrow +\infty} -1.$$

Donc $\exp\left(n \ln \frac{n}{n+1}\right) \xrightarrow{n \rightarrow +\infty} e^{-1}$.

Donc $\|f_n\|_{\infty}^{[0;1]} \underset{n \rightarrow +\infty}{\sim} \frac{n^{\alpha-1}}{e}$.

Donc $\|f_n\|_{\infty}^{[0;1]} \xrightarrow{n \rightarrow +\infty} 0$ ssi $\alpha < 1$.

Enfin, (f_n) converge uniformément vers la fonction nulle sur $[0 ; 1]$ ssi $\alpha < 1$.

Pour montrer la non-convergence uniforme, on dispose d'un critère simple suffisant dans bien des cas.

Proposition 11.26

Soit (f_n) une suite de fonctions qui converge simplement vers f sur une partie A de \mathbb{R} .

S'il existe une suite (x_n) à termes dans A telle que la suite $(f(x_n) - f_n(x_n))$ ne converge pas vers 0, alors la suite (f_n) ne converge pas uniformément sur A .

Démonstration 11.27

On a $\|f_n - f\|_{\infty}^A \geq |f(x_n) - f_n(x_n)|$, d'où la proposition. ■

Exercice 11.28

Montrez que la suite de fonctions $f_n : x \mapsto x^2 \cos \frac{x}{n}$ converge simplement sur \mathbb{R} vers une fonction f à préciser.

Montrez que la convergence n'est pas uniforme sur \mathbb{R} , mais qu'elle l'est sur tout segment $[0 ; a]$.

Correction 11.29

(f_n) converge simplement vers $f : x \mapsto x^2$ sur \mathbb{R} .

On pose $x_n = n\pi$ et on a

$$f_n(x_n) - f(x_n) = (n\pi)^2 \cos(\pi) - (n\pi)^2 = -2n^2\pi^2 \xrightarrow{n \rightarrow +\infty} -\infty.$$

Donc $(f_n(x_n) - f(x_n))$ ne converge pas vers 0 et donc (f_n) ne converge pas uniformément vers f sur \mathbb{R} .

Soit $a > 0$. Pour $x \in [0 ; a]$, on a

$$\begin{aligned} \left| x^2 \cos \frac{x}{n} - x^2 \right| &= \left| x^2 \left(\cos \frac{x}{n} - 1 \right) \right| \\ &= |x^2| \left| \cos \frac{x}{n} - 1 \right| \\ &\leq a^2 \left| \cos \frac{x}{n} - 1 \right|. \end{aligned}$$

Or $\cos(0) = 1$ et $|\cos'| = |-\sin| \leq 1$ donc d'après l'inégalité des accroissements finis, on a

$$\left| \cos \frac{x}{n} - \cos(0) \right| \leq 1 \left| \frac{x}{n} - 0 \right|.$$

Donc $\left| \cos \frac{x}{n} - 1 \right| \leq \frac{x}{n} \leq \frac{a}{n}$ et donc

$$\left| x^2 \cos \frac{x}{n} - x^2 \right| \leq \frac{a^3}{n}.$$

Or $\frac{a^3}{n} \xrightarrow{n \rightarrow +\infty} 0$ donc $\|f - f_n\|_{\infty}^{[0;a]} \xrightarrow{n \rightarrow +\infty} 0$.

Donc (f_n) converge uniformément vers f sur $[0 ; a]$.

11.2 Convergence d'une série de fonctions

11.2.1 Convergence simple

Définition 11.30

Soient A une partie de \mathbb{R} et (f_n) une suite de fonctions définies sur A .

On dit que la série $\sum_{n \geq 0} f_n$ converge simplement sur A quand pour tout $x \in A$, la série numérique $\sum_{n \geq 0} f_n(x)$ converge.

Autrement dit, la série de fonctions $\sum_{n \geq 0} f_n$ converge simplement sur A quand la suite des sommes partielles $\left(\sum_{k=0}^n f_k \right)$ converge simplement sur A .

Dans ce cas, on peut définir une fonction f sur A en posant pour tout $x \in A$, $f(x) = \sum_{n=0}^{+\infty} f_n(x)$.

La fonction f est alors appelée (fonction) somme sur A de la série $\sum_{n \geq 0} f_n$.

Exercice 11.31

Étudiez la convergence simple de la série de fonctions $f_n : x \mapsto \frac{nx^2}{n^3 + x^2}$, où $n \geq 0$, sur $[0 ; +\infty[$.

Correction 11.32

On a $f_n(x) \underset{n \rightarrow +\infty}{\sim} \frac{x^2}{n^2}$ donc par théorème de comparaison des séries à termes positifs, $\sum_n f_n(x)$ converge.

Donc $\sum_n f_n$ converge simplement sur $[0 ; +\infty[$.

Exercice 11.33

Même question avec la série de fonctions $f_n : x \mapsto \frac{x^n}{1 + x^n}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.34

Si $x \in [0 ; 1[$, on a $f_n(x) \underset{n \rightarrow +\infty}{\sim} x^n$ donc par théorème de comparaison des séries à termes positifs, $\sum_n f_n(x)$ converge et donc $\sum_n f_n$ converge simplement.

Si $x = 1$, $\sum_n f_n(x) = \sum_n \frac{1}{2}$ diverge grossièrement.

Si $x \in [1 ; +\infty[$, $f_n(x) \xrightarrow{n \rightarrow +\infty} 1$ donc $\sum_n f_n(x)$ diverge grossièrement.

Exercice 11.35

Même question avec la série de fonctions $f_n : x \mapsto \frac{\sin(nx)}{n^3 + x^3}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.36

On a

$$|f_n(x)| = \frac{|\sin(nx)|}{n^3 + x^3} \leq \frac{1}{n^3}$$

donc par théorème de comparaison des séries à termes positifs, $\sum_n f_n(x)$ converge absolument et donc converge, donc $\sum_n f_n$ converge simplement.

11.2.2 Convergence uniforme

On a la même définition en remplaçant « simple » par « uniforme ».

Définition 11.37

Soient A une partie de \mathbb{R} et (f_n) une suite de fonctions définies sur A .

On dit que la série $\sum_{n \geq 0} f_n$ converge uniformément sur A quand la suite des sommes partielles $\left(\sum_{k=0}^n f_k\right)$ converge uniformément sur A .

La fonction somme f est alors appelée limite uniforme sur A de la série $\sum_{n \geq 0} f_n$.

Encore une fois, il y a un lien entre la convergence simple et uniforme, dans un seul sens !

Théorème 11.38

Si une série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur A , alors $\sum_{n \geq 0} f_n$ converge simplement sur A .

La réciproque est fausse ! Contre-exemple : la série de fonctions $\sum_{n \geq 0} (x \mapsto x^n)$.

Démonstration 11.39

On a $\sum_{k=0}^n f_k(x) = \frac{1-x^{n+1}}{1-x}$ et $\sum_{n=0}^{+\infty} f_n(x) = \frac{1}{1-x} = f(x)$ donc

$$\sum_{k=0}^n f_k(x) - f(x) = \frac{-x^{n+1}}{1-x}$$

$$\left| \sum_{k=0}^n f_k(x) - f(x) \right| = \frac{x^{n+1}}{1-x}$$

$$\xrightarrow{x \rightarrow 1^-} +\infty.$$

Donc $\sum_{k=0}^n f_k - f$ n'est pas bornée sur $[0 ; 1[$. ■

On peut réécrire la définition à l'aide des restes partiels.

Proposition 11.40

Soit A une partie de \mathbb{R} .

La série $\sum_{n \geq 0} f_n$ converge uniformément sur A quand la série $\sum_{n \geq 0} f_n$ converge simplement sur A et la suite des restes partiels $\left(\sum_{k=n+1}^{+\infty} f_k \right)$ converge uniformément vers 0 sur A , i.e. $\left\| \sum_{k=n+1}^{+\infty} f_k \right\|_{\infty}^A \xrightarrow{n \rightarrow +\infty} 0$.

Il suffit donc qu'il existe une suite positive (α_n) telle que $\left\| \sum_{k=n+1}^{+\infty} f_k \right\|_{\infty}^A \leq \alpha_n$ et $\alpha_n \xrightarrow{n \rightarrow +\infty} 0$ pour que la série $\sum_{n \geq 0} f_n$ converge uniformément vers f sur A .

Autrement dit, il suffit de trouver une majoration uniforme des restes partiels par les termes d'une suite positive qui converge vers 0.

Exercice 11.41

Étudiez la convergence uniforme de la série de fonctions $f_n : x \mapsto \frac{(-1)^n}{x+n}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.42

Pour $x \in [0 ; +\infty[$, la suite $\left(\frac{1}{x+n} \right)_{n \geq 1}$ est positive, décroissante et converge vers 0 donc d'après le critère spécial des séries alternées, $\sum_n f_n(x)$ converge et donc $\sum_n f_n$ converge simplement.

Pour $n \geq 1$, on a

$$\left| \sum_{k=0}^{+\infty} f_k(x) - \sum_{k=0}^n f_k(x) \right| = \left| \sum_{k=n+1}^{+\infty} f_k(x) \right| \leq \frac{1}{x+n+1} \leq \frac{1}{n+1}.$$

(Majoration uniforme par une suite convergent vers 0).

Exercice 11.43

Même question avec la série de fonctions $f_n : x \mapsto \frac{(-1)^n e^{-nx}}{n}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.44

Idem, pour $x \in [0 ; +\infty[$, la suite $\left(\frac{e^{-nx}}{n} \right)$ est positive, décroissante et converge vers 0 donc d'après le critère spéciale des séries alternées, $\sum_n f_n(x)$ converge.

De plus, on a

$$\left| \sum_{k=n+1}^{+\infty} \frac{(-1)^k e^{-kx}}{k} \right| \leq \frac{e^{-(n+1)x}}{n+1} \leq \frac{1}{n+1}.$$

De manière analogue aux séries numériques, on note le lien entre la convergence uniforme d'une série de fonctions $\sum_{n \geq 0} u_n$ et la convergence uniforme de la suite (u_n) .

Proposition 11.45

Soient A une partie de \mathbb{R} et (u_n) une suite de fonctions définies sur A .

Si la série $\sum_{n \geq 0} u_n$ converge uniformément sur A , alors la suite (u_n) converge uniformément vers 0.

Démonstration 11.46

On a $u_n = \sum_{k=n}^{+\infty} f_k - \sum_{k=n+1}^{+\infty} f_k$ donc

$$\|u_n\|_{\infty}^A \leq \left\| \sum_{k=n}^{+\infty} f_k \right\|_{\infty}^A + \left\| \sum_{k=n+1}^{+\infty} f_k \right\|_{\infty}^A.$$

Donc si $\sum_n f_n$ converge uniformément sur A , alors

$$\left\| \sum_{k=n}^{+\infty} f_k \right\|_{\infty}^A \xrightarrow{n \rightarrow +\infty} 0.$$

Donc, par encadrement, $\|u_n\|_{\infty}^A \xrightarrow{n \rightarrow +\infty} 0$ i.e. (u_n) converge uniformément vers 0. ■

Évidemment, la réciproque est fausse.

11.2.3 Convergence normale

Ce dernier type de convergence est spécifique aux séries de fonctions.

Définition 11.47

Soient A une partie de \mathbb{R} et (f_n) une suite de fonctions définies sur A .

On dit que la série $\sum_{n \geq 0} f_n$ converge normalement sur A quand la série numérique $\sum_{n \geq 0} \|f_n\|_{\infty}^A$ converge.

Souvent, on ne sait pas calculer les normes uniformes des fonctions f_n . Une simple majoration suffit.

Proposition 11.48

Avec les mêmes hypothèses, il y a équivalence entre

- ▷ la série $\sum_{n \geq 0} f_n$ converge normalement sur A
- ▷ il existe une suite positive (α_n) telle que pour tout $n \in \mathbb{N}$, $\|f_n\|_\infty^A \leq \alpha_n$ et la série $\sum_{n \geq 0} \alpha_n$ converge
- ▷ il existe une majoration uniforme des fonctions f_n sur A par les termes d'une série positive convergente.

Il y a un lien entre convergence uniforme et normale, dans un seul sens !

Théorème 11.49

Si une série de fonctions converge normalement sur A , alors elle converge uniformément sur A . Elle converge aussi absolument sur A .

Démonstration 11.50

Si $\sum_n f_n$ converge normalement sur A , alors

$$\forall n \in \mathbb{N}, \forall x \in A, |f_n(x)| \leq \|f_n\|_\infty^A.$$

Or $\|f_n\|_\infty^A$ est le terme général d'une série convergente donc $\sum_n f_n(x)$ converge absolument.

Donc $\sum_n f_n$ converge simplement sur A .

De plus,

$$\left\| \sum_{k=n+1}^{+\infty} f_k \right\|_\infty^A \leq \sum_{k=n+1}^{+\infty} \|f_k\|_\infty^A \xrightarrow{n \rightarrow +\infty} 0.$$

) *reste partiel d'une série convergente*

Donc, par encadrement, $\left\| \sum_{k=n+1}^{+\infty} f_k \right\|_\infty^A \xrightarrow{n \rightarrow +\infty} 0$ i.e. $\sum_n f_n$ converge uniformément sur A . ■

La réciproque est fausse ! Contre-exemple : la série de fonctions $\sum_{n \geq 2} \left(x \mapsto \frac{x^n}{\ln n} - \frac{x^{n+1}}{\ln(n+1)} \right)$ converge uniformément sur $[0 ; 1]$ mais pas normalement.

En pratique, il est souvent plus simple de montrer la convergence normale que la convergence uniforme d'une série. On commence donc par étudier la convergence simple, puis la convergence normale et, en cas d'échec, la convergence uniforme.

Exercice 11.51

Étudiez la convergence uniforme de la série de fonctions $f_n : x \mapsto \frac{nx^2}{n^3 + x^2}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.52

D'après la Correction 11.32, $\sum_n f_n$ converge simplement sur $[0 ; +\infty[$.

$\sum_n \|f_n\|_\infty^{[0;+\infty[}$ converge-t-elle ?

Pour $n \geq 1$ et $x \in [0 ; +\infty[$, on a

$$f'_n(x) = \frac{2nx(n^3 + x^2) - nx^2(2x)}{(n^3 + x^2)^2} = \frac{2n^4x}{(n^3 + x^2)^2}.$$

Donc on a le tableau de variations suivant :

x	0	$+\infty$
$f'_n(x)$	0	+
f_n	0	n

Donc $\|f_n\|_\infty^{[0;+\infty[} \geq n$ et donc on n'a pas convergence normale sur $[0 ; +\infty[$.

Soit $a > 0$. $\sum_n \|f_n\|_\infty^{[0;a]}$ converge-t-elle ?

Pour $n \in \mathbb{N}^*$, f_n est croissante sur $[0 ; a]$ donc

$$\|f_n\|_\infty^{[0;a]} = f_n(a) = \frac{na^2}{n^3 + a^2}$$

donc $\sum_n f_n$ converge normalement sur $[0 ; a]$ et donc converge uniformément sur $[0 ; a]$.

En posant $R_n(x) = \sum_{k=n+1}^{+\infty} \frac{kx^2}{k^3 + x^2}$ et $S_n(x) = \sum_{k=n+1}^{2n} \frac{kx^2}{k^3 + x^2}$, on a

$$R_n(x) \geq S_n(x).$$

De plus, on a

$$\begin{aligned} \lim_{x \rightarrow +\infty} \sum_{k=n+1}^{2n} \frac{kx^2}{k^3 + x^2} &= \sum_{k=n+1}^{2n} \lim_{x \rightarrow +\infty} \frac{kx^2}{k^3 + x^2} \\ &= \sum_{k=n+1}^{2n} k \\ &\geq n^2. \end{aligned}$$

On veut montrer $\|R_n\|_\infty^{[0;+\infty[} \not\underset{n \rightarrow +\infty}{\rightarrow} 0$.

On a $R_n(x) \geq S_n(x)$ et $\lim_{x \rightarrow +\infty} S_n(x) \geq n^2$ donc il existe $a_n > 0$ tel que

$$\forall x \in [a_n ; +\infty[, \quad S_n(x) \geq \frac{n^2}{2}$$

donc

$$\forall x \in [a_n ; +\infty[, \quad R_n(x) \geq \frac{n^2}{2}$$

donc

$$\|R_n\|_\infty^{[0;+\infty[} \geq \frac{n^2}{2}.$$

Donc la convergence n'est pas uniforme sur $[0 ; +\infty[$.

Exercice 11.53

Même question avec la série de fonctions $f_n : x \mapsto \frac{x^n}{1+x^n}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.54

D'après la Correction 11.34, $\sum_n f_n$ converge simplement sur $[0 ; 1[$.

On étudie d'abord la convergence normale de la série $\sum_n f_n$.

Pour $n \in \mathbb{N}^*$, on a

$$\begin{aligned} \frac{d}{dx} \frac{x^n}{x^n + 1} &= \frac{nx^{n-1}(1+x^n) - x^n(nx^{n-1})}{(1+x^n)^2} \\ &= \frac{nx^{n-1} + nx^{2n-1} - nx^{2n-1}}{(1+x^n)^2} \\ &= \frac{nx^{n-1}}{(1+x^n)^2}. \end{aligned}$$

On a donc le tableau de variations suivant :

x	0	1
f_n	0	$\frac{1}{2}$

Donc $\|f_n\|_\infty^{[0;1]} = \frac{1}{2}$.

Donc $\sum_n \|f_n\|_\infty^{[0;1]}$ diverge grossièrement.

Donc $\sum_n f_n$ ne converge pas normalement.

Pour $a \in [0 ; 1[$, $\|f_n\|_\infty^{[0;a]} = f_n(a)$ et $\sum_n f_n(a)$ converge donc $\sum_n f_n$ converge normalement sur $[0 ; a]$ et donc uniformément sur $[0 ; a]$.

Exercice 11.55

Même question avec la série de fonctions $f_n : x \mapsto \frac{x^2}{1+n^3x^3}$, où $n > 0$, sur $[0 ; +\infty[$.

Correction 11.56

Si $x = 0$, $f_n(0) = 0$ donc $\sum_n f_n(0)$ converge.

Si $x > 0$, $f_n(x) \underset{n \rightarrow +\infty}{\sim} \frac{1}{n^3x}$ donc par théorème de comparaison des séries à termes positifs, $\sum_n f_n(x)$ converge.

Donc $\sum_n f_n$ converge simplement sur $[0 ; +\infty[$.

Pour $x > 0$, $|f_n(x)| = f_n(x) \leq \frac{1}{n^3x}$.

Si $a > 0$, pour tout $x \in [a ; +\infty[$, on a $|f_n(x)| \leq \frac{1}{an^3}$ donc

$$\|f_n\|_\infty^{[a;+\infty[} \leq \frac{1}{an^3}$$

donc $\sum_n f_n$ converge normalement et donc uniformément sur $[a ; +\infty[$.

Pour tout $x \geq 0$, on a $1 \leq 1 + n^3x^3$ donc $\frac{1}{1+n^3x^3} \leq 1$ donc $f_n(x) \leq x^2$.

Donc pour tout $x \in \left[0 ; \frac{1}{n}\right]$, $0 \leq f_n(x) \leq x^2 \leq \frac{1}{n^2}$, et pour tout $x \geq \frac{1}{n}$, idem.

Donc $\|f_n\|_\infty^{[0;+\infty[} \leq \frac{1}{n^2}$.

Donc $\sum_n f_n$ converge normalement et donc uniformément sur $[0 ; +\infty[$.

11.3 Propriétés de la fonction limite

11.3.1 Monotonie

Proposition 11.57

Soient I un intervalle de \mathbb{R} et (f_n) une suite de fonctions définies sur I .

Si la suite de fonctions (f_n) converge simplement vers une fonction f sur I et si pour tout $n \in \mathbb{N}$, f_n est croissante sur I , alors f est croissante sur I .

Démonstration 11.58

Si les fonctions f_n sont croissantes sur I alors pour tout $(x, y) \in I^2$ tel que $x \leq y$,

$$\forall n \in \mathbb{N}, f_n(x) \leq f_n(y).$$

Or $f_n(x) \xrightarrow{n \rightarrow +\infty} f(x)$ et $f_n(y) \xrightarrow{n \rightarrow +\infty} f(y)$ et, par passage à la limite, on a

$$f(x) \leq f(y).$$

Donc f est croissante sur I . ■

On a évidemment le même résultat avec l'hypothèse et la conclusion de décroissance.

11.3.2 Continuité

Soit (f_n) une suite de fonctions qui converge simplement vers f sur A .

En général, même si les fonctions f_n sont des fonctions continues, on ne peut rien affirmer à propos de la continuité de f . Contre-exemple : la suite de fonctions $(x \mapsto x^n)$ sur $[0 ; 1]$.

Le théorème suivant donne une condition suffisante pour que f soit continue sur A .

Théorème 11.59

Soient I un intervalle de \mathbb{R} et (f_n) une suite de fonctions définies sur I .

Si

- la suite de fonctions (f_n) converge uniformément vers une fonction f sur I
- pour tout $n \in \mathbb{N}$, f_n est continue sur I

alors f est continue sur I .

Démonstration 11.60

On suppose que (f_n) converge uniformément vers f sur I et que pour tout $n \in \mathbb{N}$, f_n est continue sur I .

Soit $x_0 \in I$. On veut montrer $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

Soit $\varepsilon > 0$.

On a $\|f_n - f\|_\infty^I \xrightarrow{n \rightarrow +\infty} 0$ donc il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \|f_n - f\|_\infty^I \leq \varepsilon$$

i.e.

$$\forall n \geq n_0, \forall x \in I, |f_n(x) - f(x)| \leq \varepsilon.$$

On choisit un entier $n \geq n_0$.

f_n est continue en x_0 donc il existe $\alpha > 0$ tel que

$$|x - x_0| \leq \alpha \implies |f_n(x) - f_n(x_0)| \leq \varepsilon.$$

Donc

$$\begin{aligned} |f(x) - f(x_0)| &= |f(x) - f_n(x) + f_n(x) - f(x_0)| \\ &= |f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)| \\ &\leq \underbrace{|f(x) - f_n(x)|}_{\leq \|f_n - f\|_\infty^I \leq \varepsilon} + \underbrace{|f_n(x) - f_n(x_0)|}_{\leq \varepsilon \text{ si } |x - x_0| \leq \alpha} + \underbrace{|f_n(x_0) - f(x_0)|}_{\leq \|f - f_n\|_\infty^I \leq \varepsilon}. \end{aligned}$$

Donc si $|x - x_0| \leq \alpha$, alors $|f(x) - f(x_0)| \leq 3\varepsilon$.

Donc f est continue. ■

On en déduit la version correspondante de ce théorème sur les séries de fonctions.

Corollaire 11.61

Soient I un intervalle de \mathbb{R} et (f_n) une suite de fonctions définies sur I .

Si

- ▷ la série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur I
- ▷ pour tout $n \in \mathbb{N}$, f_n est continue sur I

alors $\sum_{n=0}^{+\infty} f_n$ est continue sur I .

Exercice 11.62

Montrez que la fonction $f : x \mapsto \sum_{n=1}^{+\infty} \frac{\text{Arctan}(nx)}{n^2}$ est définie sur \mathbb{R} et qu'elle y est continue.

Correction 11.63

On pose $f_n : x \mapsto \frac{\text{Arctan}(nx)}{n^2}$ pour $n \in \mathbb{N}^*$.

On a $|\text{Arctan}| \leq \frac{\pi}{2}$ sur \mathbb{R} donc $\|f_n\|_{\infty}^{\mathbb{R}} \leq \frac{\pi}{2n^2}$.

Donc $\sum_n f_n$ converge normalement sur \mathbb{R}

▷ donc $\sum_n f_n$ converge simplement et absolument sur \mathbb{R} : f est définie sur \mathbb{R}

▷ donc $\sum_n f_n$ converge uniformément sur \mathbb{R} et les fonctions f_n sont continues donc f est continue sur \mathbb{R} .

Exercice 11.64

Même exercice avec $f : x \mapsto \sum_{n=1}^{+\infty} (-1)^n \frac{e^{-nx}}{n}$ sur \mathbb{R}_+ .

Correction 11.65

Pour $n \in \mathbb{N}^*$, on pose $f_n : x \mapsto (-1)^n \frac{e^{-nx}}{n}$.

Pour $x \geq 0$, $\left(\frac{e^{-nx}}{n}\right)_n$ est positive, décroissante et converge vers 0 donc d'après le critère spécial des séries alternées, $\sum_n f_n(x)$ converge, donc $\sum_n f_n$ converge simplement et donc f est définie sur \mathbb{R}_+ .

On a $\|f_n\|_{\infty}^{[0;+\infty[} = \frac{1}{n}$ donc $\sum_n f_n$ ne converge pas normalement sur $[0 ; +\infty[$.

D'après le critère spécial des séries alternées, pour $x \geq 0$ et $n \in \mathbb{N}^*$, on a

$$\begin{aligned} \left| \sum_{k=n+1}^{+\infty} f_k(x) \right| &\leq |f_{n+1}(x)| \\ &= \frac{e^{-(n+1)x}}{n+1} \\ &\leq \frac{1}{n+1}. \end{aligned}$$

Donc $\left\| \sum_{k=n+1}^{+\infty} f_k \right\|_{\infty}^{[0;+\infty[} \leq \frac{1}{n+1} \xrightarrow{n \rightarrow +\infty} 0$.

Donc $\sum_n f_n$ converge uniformément sur $[0 ; +\infty[$.

Or les fonctions f_n sont continues sur \mathbb{R}_+ donc f est continue sur \mathbb{R}_+ .

Exercice 11.66

Montrez que la fonction $f : x \mapsto e^{-x} \sum_{n=1}^{+\infty} \operatorname{Arctan} \left(\frac{e^x}{n^2} \right)$ est bien définie et continue sur $[0 ; +\infty[$, que f est intégrable sur $[0 ; +\infty[$, et que $\int_0^{+\infty} f = \sum_{n=1}^{+\infty} \left(\operatorname{Arctan} \frac{1}{n^2} + \frac{1}{2n^2} \ln(1+n^4) \right)$.

Correction 11.67

► Pour $x \geq 0$, $\frac{e^x}{n^2} \xrightarrow{n \rightarrow +\infty} 0$ et $\operatorname{Arctan} u \underset{u \rightarrow 0}{\sim} u$ donc

$$\operatorname{Arctan} \frac{e^x}{n^2} \underset{n \rightarrow +\infty}{\sim} \frac{e^x}{n^2}$$

donc, par théorème de comparaison des séries à termes positifs, $\sum_n \operatorname{Arctan} \frac{e^x}{n^2}$ converge et donc f est définie sur \mathbb{R}_+ .

De plus, Arctan est concave sur \mathbb{R}_+ donc pour tout $x \in \mathbb{R}_+$, $0 \leq \operatorname{Arctan} x \leq x$ donc

$$\text{pour tout } x \in \mathbb{R}_+, \quad 0 \leq \operatorname{Arctan} \frac{e^x}{n^2} \leq \frac{e^x}{n^2}$$

$$0 \leq e^{-x} \operatorname{Arctan} \frac{e^x}{n^2} \leq \frac{1}{n^2}.$$

On pose $f_n : x \mapsto e^{-x} \operatorname{Arctan} \frac{e^x}{n^2}$.

On a montré $\|f_n\|_{\infty}^{[0;+\infty[} \leq \frac{1}{n^2}$.

Donc $\sum_n f_n$ converge normalement et donc uniformément.

Or les fonctions f_n sont continues donc f est continue.

► Pour $n \in \mathbb{N}^*$, on a

$$\forall x \in [0 ; +\infty[, \quad |f_n(x)| \leq \frac{\pi}{2} e^{-x}$$

or $x \mapsto e^{-x}$ est intégrable sur $[0 ; +\infty[$ donc par comparaison de fonctions positives, f_n est intégrable sur $[0 ; +\infty[$.

Montrons que $\sum_n \int_0^{+\infty} |f_n|$ converge.

Les fonctions f_n étant positives, on a $\int_0^{+\infty} |f_n| = \int_0^{+\infty} f_n$.

En faisant le changement de variable $u = e^{-x}$ et en intégrant par parties (sous réserve de convergence), on a

$$\begin{aligned} \int_0^{+\infty} e^{-x} \operatorname{Arctan} \left(\frac{e^x}{n^2} \right) dx &= \int_1^{+\infty} \frac{1}{u} \operatorname{Arctan} \left(\frac{u}{n^2} \right) \frac{1}{u} du \\ &= \int_1^{+\infty} \frac{1}{u^2} \operatorname{Arctan} \left(\frac{u}{n^2} \right) du \\ &= \left[\frac{-1}{t} \operatorname{Arctan} \frac{t}{n^2} \right]_{t=1}^{+\infty} - \int_1^{+\infty} \frac{-1}{t} \times \frac{1}{n^2} \times \frac{1}{1 + (t/n^2)^2} dt. \end{aligned}$$

Or on a $\left| \frac{-1}{t} \times \frac{1}{1 + (t/n^2)^2} \right| \underset{t \rightarrow +\infty}{\sim} \left| \frac{-n^4}{t^3} \right|$ et $t \mapsto \frac{1}{t^3}$ est intégrable sur $[1 ; +\infty[$ donc par comparaison de fonctions positives, $\int_1^{+\infty} \frac{-1}{t} \times \frac{1}{n^2} \times \frac{1}{1 + (t/n^2)^2} dt$ converge et donc l'intégration par parties est licite.

On a donc

$$\int_0^{+\infty} f_n = \operatorname{Arctan} \frac{1}{n^2} + \frac{1}{n^2} \int_1^{+\infty} \frac{1}{t \left(1 + (t/n^2)^2 \right)} dt.$$

En décomposant $\frac{1}{t \left(1 + (t/n^2)^2 \right)}$ en éléments simples, on obtient

$$\frac{1}{t \left(1 + (t/n^2)^2 \right)} = \frac{1}{t} + \frac{(-1/n^4) t}{1 + (t/n^2)^2}.$$

Donc

$$\begin{aligned} \int_1^{+\infty} \frac{1}{t \left(1 + (t/n^2)^2 \right)} dt &= \int_1^{+\infty} \left(\frac{1}{t} - \frac{1}{n^4} \times \frac{t}{1 + (t/n^2)^2} \right) dt \\ &= \left[\ln t - \frac{1}{2} \ln (t^2 + n^4) \right]_{t=1}^{+\infty} \\ &= \left[\frac{1}{2} \ln \frac{t^2}{n^4 + t^2} \right]_{t=1}^{+\infty} \\ &= \frac{1}{2} \ln (n^4 + 1). \end{aligned}$$

Finalement, on a

$$\int_0^{+\infty} f_n = \operatorname{Arctan} \frac{1}{n^2} + \frac{1}{2n^2} \ln (1 + n^4).$$

Or on a $\operatorname{Arctan} \frac{1}{n^2} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n^2}$ donc $\sum_n \operatorname{Arctan} \frac{1}{n^2}$ converge par théorème de comparaison des séries à termes positifs.

De plus, on a

$$\frac{1}{2n^2} \ln (1 + n^4) \underset{n \rightarrow +\infty}{\sim} \frac{\ln n^4}{2n^2} = \frac{2 \ln n}{n^2} \underset{n \rightarrow +\infty}{=} o \left(\frac{1}{n^{3/2}} \right)$$

donc $\sum_n \frac{1}{2n^2} \ln (1 + n^4)$ converge.

On a donc $\sum_n \int_0^{+\infty} |f_n|$ convergente.

Donc, d'après le théorème d'intégration terme à terme, f est intégrable sur $[0 ; +\infty[$ et on a

$$\int_0^{+\infty} f = \sum_{n=1}^{+\infty} \int_0^{+\infty} f_n$$

ce qui conclut.

La continuité étant une propriété locale, il est souvent inutile d'avoir une convergence uniforme globale pour conclure. En général, on peut se contenter de convergence uniforme sur des parties plus petites que I , en général les segments inclus dans I ou toute famille recouvrante de parties de I .

Définition 11.68

Soit I un intervalle.

Une famille \mathcal{F} de parties de I est dite recouvrante quand $I = \bigcup_{X \in \mathcal{F}} X$.

Proposition 11.69

Soit I un intervalle.

La famille des segments inclus dans I est recouvrante : $I = \bigcup_{(a,b) \in I^2} [a ; b]$.

On en déduit alors le théorème suivant, dont il vaut mieux, à mon avis, présenter le détail des idées sur chaque exercice.

Théorème 11.70

Soient I un intervalle de \mathbb{R} , (f_n) une suite de fonctions définies sur I et \mathcal{F} une famille d'intervalles recouvrant I .

Si

- ▶ pour toute partie $X \in \mathcal{F}$, la suite de fonctions (f_n) converge uniformément vers une fonction f sur X
- ▶ pour tout $n \in \mathbb{N}$, f_n est continue sur I

alors f est continue sur I .

En pratique, on choisit pour \mathcal{F} une famille d'intervalles adaptés au problème (voir les nombreux exemples en exercice).

Exercice 11.71

Pour $x > 0$, on pose $u_0(x) = x$ et pour tout $n \in \mathbb{N}$, $u_{n+1}(x) = \text{Arctan}(x + u_n(x))$.

Montrez que la suite $(u_n(x))$ converge vers un réel $\varphi(x)$.

Montrez que pour tout $n \in \mathbb{N}$, $|u_n(x) - \varphi(x)| \leq \frac{x + \pi/2}{(1+x^2)^n}$.

Déduisez-en que φ est continue sur $]0; +\infty[$.

Correction 11.72

► Pour $x > 0$, on pose $f_x : t \mapsto \text{Arctan}(x + t)$ et on a

$$\forall n \in \mathbb{N}, u_{n+1}(x) = f_x(u_n(x)) \leq \frac{\pi}{2}.$$

La suite $(u_n(x))_n$ est majorée par $\frac{\pi}{2}$ à partir du rang 1 donc par récurrence immédiate, $(u_n(x))_n$ est à termes positifs.

On veut montrer que l'équation $f_x(t) = t$ d'inconnue t a une unique solution.

On pose $g_x : t \mapsto f_x(t) - t$ et on a

$$\forall t \in \mathbb{R}_+, g'_x(t) = \frac{1}{1+(x+t)^2} - 1 = \frac{-(x+t)^2}{1+(x+t)^2} < 0.$$

On en déduit le tableau de variations suivant :

t	0	$\varphi(x)$	$+\infty$
$g_x(t)$	<div style="display: flex; align-items: center; justify-content: space-between;"> <div style="text-align: center;">Arctan x</div> <div style="text-align: center;"> \searrow 0 \searrow </div> <div style="text-align: center;">$-\infty$</div> </div>		

On l'interprète ainsi : g_x est continue et strictement décroissante sur $[0; +\infty[$ et change de signe donc d'après le théorème des valeurs intermédiaires, il existe un unique $t \in [0; +\infty[$ tel que $g_x(t) = 0$. On pose $\varphi(x)$ cette unique solution.

De plus, pour $t \in \mathbb{R}_+$, on a

$$0 \leq f'_x(t) \leq \frac{1}{1+x^2} < 1$$

donc d'après l'inégalité des accroissements finis, pour $(t, t') \in [0; +\infty[$, on a

$$|f_x(t) - f_x(t')| \leq \frac{1}{1+x^2} |t - t'|$$

et en particulier

$$\begin{aligned} \forall n \in \mathbb{N}, |f_x(u_n(x)) - f_x(\varphi(x))| &\leq \frac{1}{1+x^2} |u_n(x) - \varphi(x)| \\ |u_{n+1}(x) - \varphi(x)| &\leq \frac{1}{1+x^2} |u_n(x) - \varphi(x)|. \end{aligned} \quad \left. \vphantom{\begin{aligned} \forall n \in \mathbb{N}, |f_x(u_n(x)) - f_x(\varphi(x))| &\leq \frac{1}{1+x^2} |u_n(x) - \varphi(x)| \\ |u_{n+1}(x) - \varphi(x)| &\leq \frac{1}{1+x^2} |u_n(x) - \varphi(x)|. \end{aligned}} \right\} \varphi(x) \text{ point fixe}$$

Par récurrence immédiate, on a donc

$$\forall n \in \mathbb{N}, \quad |u_n(x) - \varphi(x)| \leq \left(\frac{1}{1+x^2} \right)^n |u_0(x) - \varphi(x)|.$$

Puis, par théorème d'encadrement, on a

$$u_n(x) \xrightarrow{n \rightarrow +\infty} \varphi(x).$$

► De plus, on a

$$\begin{aligned} |u_0(x) - \varphi(x)| &\leq |x - \varphi(x)| \\ &\leq |x| + |\varphi(x)| \\ &= x + \varphi(x) \\ &\leq x + \operatorname{Arctan}(x + \varphi(x)) \\ &\leq x + \frac{\pi}{2} \end{aligned}$$

donc

$$\forall n \in \mathbb{N}, \quad |u_n(x) - \varphi(x)| \leq \frac{x + \pi/2}{(1+x^2)^n}.$$

► Pour $n \in \mathbb{N}$, on a

$$u_{n+1} = \operatorname{Arctan} \circ (u_n + \operatorname{id})$$

donc, par récurrence immédiate, les fonctions u_n sont continues sur $]0 ; +\infty[$.

Soit $(a, b) \in]0 ; +\infty[^2$ tel que $0 < a < b$.

Pour $x \in [a ; b]$, on a la majoration uniforme

$$|u_n(x) - \varphi(x)| \leq \frac{b + \pi/2}{(1+a^2)^n}$$

donc

$$\|u_n - \varphi\|_{\infty}^{[a;b]} \leq \frac{b + \pi/2}{(1+a^2)^n}.$$

Or $\frac{b + \pi/2}{(1+a^2)^n} \xrightarrow{n \rightarrow +\infty} 0$ donc (u_n) converge uniformément vers φ sur $[a ; b]$.

D'après le Théorème 11.59, φ est donc continue sur $[a ; b]$.

Enfin, φ est continue sur $]0 ; +\infty[= \bigcup_{0 < a < b} [a ; b]$.

Exercice 11.73

Montrez que la fonction $f : x \mapsto \sum_{n=1}^{+\infty} \frac{\ln(1+nx)}{n^2}$ est définie et continue sur $[0 ; +\infty[$.

Correction 11.74

► Pour $x \in [0 ; +\infty[$, on a

$$0 \leq \frac{\ln(1+nx)}{n^2} \underset{n \rightarrow +\infty}{\sim} \frac{\ln(nx)}{n^2} \underset{n \rightarrow +\infty}{=} o\left(\frac{1}{n^{3/2}}\right)$$

donc $\sum_n \frac{\ln(1+nx)}{n^2}$ converge.

► Soit $a > 0$. Pour $x \in [0 ; a]$, on a

$$\left| \frac{\ln(1+nx)}{n^2} \right| \leq \frac{\ln(1+na)}{n^2}$$

donc en posant $f_n : x \mapsto \frac{\ln(1+nx)}{n^2}$, on a

$$\|f_n\|_{\infty}^{[0;a]} \leq f_n(a).$$

Or $\sum_n f_n(a)$ converge donc $\sum_n f_n$ converge normalement et donc uniformément sur $[0 ; a]$.

De plus, les fonctions f_n sont continues donc f est continue sur $[0 ; a]$.

Enfin, f est continue sur $[0 ; +\infty[= \bigcup_{a>0} [0 ; a]$.

11.3.3 Interversion de limite et d'intégrale

Soit (f_n) une suite de fonctions qui converge simplement vers f sur un segment $[a ; b]$.

En général, même si les fonctions f_n sont des fonctions continues par morceaux, on ne peut rien affirmer à propos de la continuité par morceaux de f . Et même si on le peut (*cf.* théorème précédent), il n'y a *a priori* aucun rapport entre $\lim_{n \rightarrow +\infty} \int_a^b f_n$ et $\int_a^b f = \int_a^b \left(\lim_{n \rightarrow +\infty} f_n \right)$: on ne peut en général pas intervertir une limite et une intégrale.

Exercice 11.75

Pour $n \in \mathbb{N}$ et $x \in [0 ; 1]$, on pose $f_n(x) = n^2 x^n (1-x)$.

Montrez que la suite de fonctions (f_n) converge simplement vers une fonction f et vérifiez que $\int_0^1 f = \int_0^1 \left(\lim_{n \rightarrow +\infty} f_n \right) \neq \lim_{n \rightarrow +\infty} \int_0^1 f_n$.

Le théorème suivant donne une condition suffisante pour que cela soit faisable.

Théorème 11.76

Soient I un intervalle de \mathbb{R} , (f_n) une suite de fonctions définies qui converge simplement vers une fonction f sur I et $a \in I$.

Si

- ▷ la suite de fonctions (f_n) converge uniformément vers f sur tout segment inclus dans I
- ▷ pour tout $n \in \mathbb{N}$, f_n est continue sur I

alors f est continue sur I et la suite des primitives qui s'annulent en a , i.e. $\left(x \mapsto \int_a^x f_n\right)$, converge uniformément sur tout segment inclus dans I vers la primitive de f qui s'annule en a , i.e. $x \mapsto \int_a^x f$.

Démonstration 11.77

Pour $n \in \mathbb{N}$, on pose $F_n : x \mapsto \int_a^x f_n$ et $F : x \mapsto \int_a^x f$.

Soit $(c, d) \in I^2$ tel que $c < d$. On veut montrer

$$\|F_n - F\|_{\infty}^{[c;d]} \xrightarrow{n \rightarrow +\infty} 0.$$

Pour $x \in [c ; d]$, on a

$$|F_n(x) - F(x)| = \left| \int_a^x (f_n - f) \right| \leq s_x \int_a^x |f_n(t) - f(t)| dt$$

où $s_x = 1$ si $x \geq a$ et -1 sinon.

Si on prend $x \in [c ; d]$ et t tel que $\begin{cases} a \leq t \leq x & \text{si } x \geq a \\ x \leq t \leq a & \text{sinon} \end{cases}$ on a $\begin{cases} a \leq t \leq d & \text{si } x \geq a \\ c \leq t \leq a & \text{sinon} \end{cases}$ i.e.

$$\underbrace{\min(c, a)}_{\alpha} \leq t \leq \underbrace{\max(d, a)}_{\beta}.$$

Donc $|f_n(t) - f(t)| \leq \|f_n - f\|_{\infty}^{[\alpha;\beta]}$.

Donc

$$\begin{aligned} s_x \int_a^x |f_n(t) - f(t)| dt &\leq s_x \int_a^x \|f_n - f\|_{\infty}^{[\alpha;\beta]} dt \\ &= s_x (x - a) \|f_n - f\|_{\infty}^{[\alpha;\beta]}. \end{aligned}$$

De plus, comme $x \in [c ; d]$, on a $|x - a| \leq K$ où K est une constante et donc

$$\begin{aligned} |F_n(x) - F(x)| &\leq |x - a| \|f_n - f\|_{\infty}^{[\alpha;\beta]} \\ &\leq \underbrace{K \|f_n - f\|_{\infty}^{[\alpha;\beta]}}_{\xrightarrow{n \rightarrow +\infty} 0}. \end{aligned}$$

Donc (F_n) converge uniformément vers F sur $[c ; d]$. ■

On en déduit les corollaires sur l'interversion de limite et d'intégrale.

Corollaire 11.78

Soient a, b deux réels et (f_n) une suite de fonctions définies sur $[a ; b]$.

Si

- la suite de fonctions (f_n) converge uniformément vers une fonction f sur $[a ; b]$
- pour tout $n \in \mathbb{N}$, f_n est continue sur $[a ; b]$

alors f est continue sur $[a ; b]$ et

$$\int_a^b f = \int_a^b \left(\lim_{n \rightarrow +\infty} f_n \right) = \lim_{n \rightarrow +\infty} \int_a^b f_n.$$

On rappelle qu'un symbole $\sum_{+\infty}$ est une limite : $\sum_{+\infty} = \lim_{N \rightarrow +\infty} \sum^N$. On ne peut donc pas intervertir sans précaution un symbole $\sum_{+\infty}$ et une intégrale.

Corollaire 11.79

Soient a, b deux réels et (f_n) une suite de fonctions définies sur $[a ; b]$.

Si

- la série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur $[a ; b]$
- pour tout $n \in \mathbb{N}$, f_n est continue sur $[a ; b]$

alors $\sum_{n=0}^{+\infty} f_n$ est continue sur $[a ; b]$, la série des intégrales $\sum_{n \geq 0} \int_a^b f_n$ est convergente et

$$\int_a^b f = \int_a^b \sum_{n=0}^{+\infty} f_n = \sum_{n=0}^{+\infty} \int_a^b f_n.$$

On appelle aussi ce résultat le théorème d'intégration terme à terme sur un segment.

Remarque 11.80

Ces deux théorèmes ne sont valables que sur des segments, donc pour de vraies intégrales. Ils sont faux si on se place sur des intervalles autres et donc si on parle d'intégrales généralisées.

Exercice 11.81

Quelle est la limite de $\int_0^1 \frac{n \cos x/n}{n + x^n} dx$ quand $n \rightarrow +\infty$?

Correction 11.82

Pour $n \in \mathbb{N}^*$, on pose $f_n : x \mapsto \frac{n \cos x/n}{n + x^n}$ continue sur $[0 ; 1]$.

Pour $x \in [0 ; 1]$, on a $\cos \frac{x}{n} \xrightarrow[n \rightarrow +\infty]{} 1$ et $n + x^n \underset{n \rightarrow +\infty}{\sim} n$ donc

$$f_n(x) \xrightarrow[n \rightarrow +\infty]{} 1$$

i.e. la suite (f_n) converge simplement vers la fonction constante égale à 1 sur $[0 ; 1]$.

De plus, pour $x \in [0 ; 1]$ et $n \in \mathbb{N}^*$, on a

$$|f_n(x)| = \frac{n |\cos x/n|}{n + x^n} \leq \frac{n}{n + x^n} \leq 1.$$

Or la fonction $x \mapsto 1$ est intégrable sur $[0 ; 1]$ donc d'après le théorème de convergence dominée, on a

$$\int_0^1 f_n \xrightarrow[n \rightarrow +\infty]{} \int_0^1 dt = 1.$$

Exercice 11.83

Montrez que la fonction $f : x \mapsto \sum_{n=1}^{+\infty} \frac{(-1)^n x^{n+1}}{n}$ est définie et continue sur $[0 ; 1]$, puis justifiez

$$\int_0^1 f = \frac{-1}{4}.$$

Correction 11.84

Pour $x \in [0 ; 1]$, la suite $\left(\frac{x^{n+1}}{n}\right)_{n \geq 1}$ est positive, décroissante et converge vers 0 donc d'après le critère spécial des séries alternées, la série $\sum_n \frac{(-1)^n x^{n+1}}{n}$ converge et donc la fonction f est définie sur $[0 ; 1]$.

Pour $x \in [0 ; 1]$ et $n \in \mathbb{N}^*$, on a

$$\left| \sum_{k=n+1}^{+\infty} \frac{(-1)^k x^{k+1}}{k} \right| \leq \left| \frac{(-1)^{n+1} x^{n+2}}{n+1} \right| \leq \frac{1}{n+1}$$

$$\text{donc } \left\| \sum_{k=n+1}^{+\infty} f_k \right\|_{\infty}^{[0;1]} \leq \frac{1}{n+1} \xrightarrow[n \rightarrow +\infty]{} 0.$$

Donc $\sum_n f_n$ converge uniformément sur $[0 ; 1]$.

Or les fonctions f_n sont continues sur $[0 ; 1]$ donc f est continue sur $[0 ; 1]$.

De plus, d'après le théorème d'intégration terme à terme, on a

$$\begin{aligned} \int_0^1 f &= \sum_{n=1}^{+\infty} \int_0^1 \frac{(-1)^n x^{n+1}}{n} dx \\ &= \sum_{n=1}^{+\infty} \frac{(-1)^n}{n(n+2)} \\ &= s. \end{aligned}$$

La série $\sum_n \frac{(-1)^n}{n(n+2)}$ est absolument convergente donc la famille $\left(\frac{(-1)^n}{n(n+2)}\right)_{n \geq 1}$ est sommable.

Donc, d'après le théorème de sommation par paquets, on a

$$\begin{aligned} s &= \sum_{\substack{n=1 \\ n \text{ impair}}}^{+\infty} \frac{(-1)^n}{n(n+2)} + \sum_{\substack{n=1 \\ n \text{ pair}}}^{+\infty} \frac{(-1)^n}{n(n+2)} \\ &= - \sum_{k=0}^{+\infty} \frac{1}{(2k+1)(2k+3)} + \sum_{k=1}^{+\infty} \frac{1}{2k(2k+2)}. \end{aligned}$$

Or on a

$$\begin{aligned} \sum_{k=1}^{+\infty} \frac{1}{2k(2k+2)} &= \frac{1}{4} \sum_{k=1}^{+\infty} \left(\frac{1}{k} - \frac{1}{k+1} \right) \\ &= \frac{1}{4} \left(1 - \lim_{k \rightarrow +\infty} \frac{1}{k+1} \right) \quad \left. \vphantom{\sum_{k=1}^{+\infty}} \right\} \text{ série télescopique} \\ &= \frac{1}{4} \end{aligned}$$

et

$$\begin{aligned} \sum_{k=0}^{+\infty} \frac{1}{(2k+1)(2k+3)} &= \sum_{k=0}^{+\infty} \left(\frac{1}{2(2k+1)} - \frac{1}{2(2k+3)} \right) \\ &= \frac{1}{2} - \lim_{k \rightarrow +\infty} \frac{1}{2(2k+3)} \quad \left. \vphantom{\sum_{k=0}^{+\infty}} \right\} \text{ série télescopique} \\ &= \frac{1}{2}. \end{aligned}$$

Enfin, on a bien $\int_0^1 f = \frac{-1}{4}$.

Exercice 11.85

Montrez que $\sum_{n=1}^{+\infty} \frac{1}{n2^n} = \int_0^1 \frac{1}{2-x} dx$ et donnez-en une valeur concrète.

Exercice 11.86

On pose $f_n : x \mapsto \frac{x^n e^{-x}}{n!}$.

Montrez que la suite de fonctions (f_n) converge uniformément sur $[0 ; +\infty[$.

Justifiez la convergence des intégrales $\int_0^{+\infty} f_n$ et déterminez leur limite quand $n \rightarrow +\infty$.

11.3.4 Intersion de limites

De même, il est faux en général que $\lim_{x \rightarrow a} \left(\lim_{n \rightarrow +\infty} f_n(x) \right) = \lim_{n \rightarrow +\infty} \left(\lim_{x \rightarrow a} f_n(x) \right)$: on ne peut en général pas intervertir deux limites.

Exemple 11.87

On pose, pour tout $(x, n) \in [0 ; 1[\times \mathbb{N}$, $u_n(x) = \frac{x + x^n}{1 + nx^n}$.

D'une part, pour tout $x \in [0 ; 1[$, $\lim_{n \rightarrow +\infty} u_n(x) = x$ donc $\lim_{x \rightarrow 1} \left(\lim_{n \rightarrow +\infty} u_n(x) \right) = 1$.

D'autre part, pour tout $n \in \mathbb{N}$, $\lim_{x \rightarrow 1} u_n(x) = \frac{2}{n+1}$ donc $\lim_{n \rightarrow +\infty} \left(\lim_{x \rightarrow 1} u_n(x) \right) = 0$.

Encore une fois, grâce à la convergence uniforme, c'est possible. Cependant, les résultats sont admis car plus difficiles à montrer : ce sont les théorèmes de double limite.

Théorème 11.88

Soient I un intervalle de \mathbb{R} , a une extrémité de I et (f_n) une suite de fonctions définies sur I .

Si

- la suite de fonctions (f_n) converge uniformément vers une fonction f sur I
- pour tout $n \in \mathbb{N}$, f_n a pour limite finie ℓ_n en a

alors la suite (ℓ_n) a une limite finie ℓ et f a pour limite ℓ en a :

$$\lim_{x \rightarrow a} \left(\lim_{n \rightarrow +\infty} f_n(x) \right) = \lim_{n \rightarrow +\infty} \left(\lim_{x \rightarrow a} f_n(x) \right) = \ell.$$

On en déduit la version correspondante sur les séries de fonctions.

Corollaire 11.89

Soient I un intervalle de \mathbb{R} , a une extrémité de I et (f_n) une suite de fonctions définies sur I .

Si

- la série de fonctions $\sum_{n \geq 0} f_n$ converge uniformément sur I
- pour tout $n \in \mathbb{N}$, f_n a pour limite finie ℓ_n en a

alors la série $\sum_{n \geq 0} \ell_n$ converge et, en notant ℓ la somme de cette série, la fonction $\sum_{n=0}^{+\infty} f_n$ a pour limite ℓ en a :

$$\lim_{x \rightarrow a} \sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty} \left(\lim_{x \rightarrow a} f_n(x) \right) = \sum_{n=0}^{+\infty} \ell_n.$$

Exercice 11.90

Montrez que la fonction $f : x \mapsto \sum_{n=1}^{+\infty} \frac{(-1)^n \operatorname{Arctan}(nx)}{n^4}$ est définie et continue sur \mathbb{R} , puis déterminez la valeur de sa limite en $+\infty$ (on admettra que $\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$).

Correction 11.91

Pour $n \in \mathbb{N}^*$, on pose $f_n : x \mapsto \frac{(-1)^n \operatorname{Arctan}(nx)}{n^4}$.

Pour $x \in \mathbb{R}$ et $n \in \mathbb{N}^*$, on a $|f_n(x)| \leq \frac{\pi}{2n^4}$.

Donc $\|f_n\|_{\infty}^{\mathbb{R}} \leq \frac{\pi}{2n^4}$ donc $\sum_n f_n$ converge normalement sur \mathbb{R} .

En particulier, $\sum_n f_n$ converge simplement sur \mathbb{R} donc f est définie sur \mathbb{R} .

De plus, $\sum_n f_n$ converge uniformément sur \mathbb{R} et les fonctions f_n sont continues donc f est continue sur \mathbb{R} .

Pour $n \in \mathbb{N}^*$, on a $f_n(x) \xrightarrow{x \rightarrow +\infty} \frac{(-1)^n \pi/2}{n^4}$ donc d'après le théorème de la double limite, on a

$$f(x) \xrightarrow{x \rightarrow +\infty} \sum_{n=1}^{+\infty} \frac{(-1)^n \pi/2}{n^4} = \frac{\pi}{2} \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^4}.$$

Or la série $\sum_n \frac{(-1)^n}{n^4}$ converge absolument donc

$$\begin{aligned} \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^4} &= \sum_{k=1}^{+\infty} \frac{1}{(2k)^4} - \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^4} \\ &= \frac{1}{16} \sum_{k=1}^{+\infty} \frac{1}{k^4} - \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^4} \\ &= \frac{\pi^4}{16 \times 90} - \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^4}. \end{aligned}$$

De même, on a $\sum_{k=1}^{+\infty} \frac{1}{k^4} = \sum_{k=1}^{+\infty} \frac{1}{(2k)^4} + \sum_{k=0}^{+\infty} \frac{1}{(2k+1)^4}$.

Or $\sum_{k=1}^{+\infty} \frac{1}{k^4} = \frac{\pi^4}{90}$ et $\sum_{k=1}^{+\infty} \frac{1}{(2k)^4} = \frac{\pi^4}{16 \times 90}$ donc $\sum_{k=0}^{+\infty} \frac{1}{(2k+1)^4} = \frac{15\pi^4}{16 \times 90}$.

Donc $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n^4} = \frac{-7\pi^4}{720}$.

Finalement, on a $\lim_{+\infty} f = \frac{-7\pi^5}{1440}$.

11.3.5 Dérivabilité

La situation est encore plus délicate : une dérivée étant une limite, intervertir une dérivation et une limite est en général impossible.

D'abord, la limite simple d'une suite de fonctions dérivables n'est pas forcément dérivable. Il est raisonnable de se dire que si on remplace la limite simple par une limite uniforme, comme dans les théorèmes précédents, on obtient le résultat.

Et bien non !

On peut montrer qu'il existe des suites de fonctions dérivables qui converge uniformément vers des fonctions non-dérivables (difficilement, car il faut construire des fonctions qui sont continues mais dérivables nulle part).

De même, il existe une suite de fonctions dérivables (f_n) qui converge uniformément vers une fonction dérivable f , mais la suite des dérivées (f'_n) ne converge pas vers la dérivée f' : par exemple la suite des fonctions $x \mapsto \frac{\sin(nx)}{n}$ sur $\left[\frac{\pi}{4}; \frac{3\pi}{4}\right]$.

Pour obtenir un résultat vraiment utile, il faut supposer plus que dérivable, *i.e.* de classe \mathcal{C}^1 , et imposer la convergence uniforme sur les dérivées.

Théorème 11.92

Soient I un intervalle de \mathbb{R} et (f_n) une suite de fonctions définies sur I .

Si

- ▷ les fonctions f_n sont de classe \mathcal{C}^1 sur I
- ▷ la suite de fonctions (f_n) converge simplement sur I
- ▷ la suite des dérivées (f'_n) converge uniformément sur I

alors $f = \lim_{n \rightarrow +\infty} f_n$ est de classe \mathcal{C}^1 sur I et

$$f' = \lim_{n \rightarrow +\infty} f'_n.$$

Démonstration 11.93

La suite (f'_n) converge uniformément sur I vers une fonction g .

Les fonctions f'_n sont continues car les fonctions f_n sont de classe \mathcal{C}^1 .

D'après le Théorème 11.76, la suite $\left(x \mapsto \int_a^x f'_n\right)$ converge uniformément vers $x \mapsto \int_a^x g$ sur tout segment *i.e.* la suite $(f_n - f_n(a))$ converge uniformément vers $x \mapsto \int_a^x g$ sur tout segment.

En notant $f = \lim_{n \rightarrow +\infty} f_n$, on obtient

$$f - f(a) = x \mapsto \int_a^x g$$

i.e.

$$\forall x \in I, \quad f(x) = f(a) + \int_a^x g.$$

Donc f est \mathcal{C}^1 sur I et $f' = g$ i.e. $\left(\lim_{n \rightarrow +\infty} f_n\right)' = \lim_{n \rightarrow +\infty} f_n'$. ■

On en déduit la version correspondante sur les séries de fonctions.

Corollaire 11.94

Soient I un intervalle de \mathbb{R} et (f_n) une suite de fonctions définies sur I .

Si

- ▷ les fonctions f_n sont de classe \mathcal{C}^1 sur I
- ▷ la série de fonctions $\sum_{n \geq 0} f_n$ converge simplement sur I
- ▷ la série des dérivées $\sum_{n \geq 0} f_n'$ converge uniformément sur I

alors $\sum_{n=0}^{+\infty} f_n$ est de classe \mathcal{C}^1 sur I et

$$\left(\sum_{n=0}^{+\infty} f_n\right)' = \sum_{n=0}^{+\infty} f_n'.$$

On appelle aussi ce résultat le théorème de dérivation terme à terme.

Exercice 11.95

Montrez que la fonction $f : x \mapsto \sum_{n=1}^{+\infty} \frac{\ln(1+nx)}{n^2}$ est définie et de classe \mathcal{C}^1 sur $[1; +\infty[$. Que vaut $f'(1)$?

Correction 11.96

- ▷ Pour $x > 0$, on a $\ln(1+nx) \underset{n \rightarrow +\infty}{\sim} \ln(nx) = \ln n + \ln x \underset{n \rightarrow +\infty}{\sim} \ln n$.

$$\text{Donc } f_n(x) \underset{n \rightarrow +\infty}{\sim} \frac{\ln n}{n^2} \underset{n \rightarrow +\infty}{=} o\left(\frac{1}{n^{3/2}}\right).$$

Donc $\sum_n f_n(x)$ converge absolument.

► Les fonctions f_n sont \mathcal{C}^1 sur $[1 ; +\infty[$.

► Pour $x \in [1 ; +\infty[$ et $n \in \mathbb{N}^*$, on a

$$f'_n(x) = \frac{n}{(1+nx)n^2} = \frac{1}{n(1+nx)}.$$

On a donc $|f'_n(x)| \leq \frac{1}{n(1+n)}$ et donc $\sum_n f'_n$ converge normalement et donc uniformément sur $[1 ; +\infty[$.

D'après le théorème de dérivation terme à terme, f est donc \mathcal{C}^1 sur $[1 ; +\infty[$ et, pour $x \in [1 ; +\infty[$, on a

$$f'(x) = \sum_{n=1}^{+\infty} \frac{1}{n(1+nx)}.$$

En particulier, on a

$$f'(1) = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) = 1.$$

Remarque 11.97

Là encore, le caractère \mathcal{C}^1 d'une fonction est une propriété locale. Donc tout ce qui a été signalé à propos de la continuité sur des parties d'une famille recouvrante reste valable. En pratique, on choisit des intervalles adaptés au problème.

Exercice 11.98

Montrez que la fonction de l'exercice précédent est de classe \mathcal{C}^1 sur $]0 ; +\infty[$.

Correction 11.99

Soit $a > 0$.

Pour $x \in [a ; +\infty[$, on a

$$|f'_n(x)| \leq \frac{1}{n(1+na)}.$$

Or $\frac{1}{n(1+na)} \underset{n \rightarrow +\infty}{\sim} \frac{1}{an^2}$ donc $\sum_n \frac{1}{n(1+na)}$ converge.

Donc $\sum_n f'_n$ converge normalement et donc uniformément sur $[a ; +\infty[$.

Donc f est \mathcal{C}^1 sur $[a ; +\infty[$.

Finalement, f est \mathcal{C}^1 sur $]0 ; +\infty[= \bigcup_{a>0} [a ; +\infty[$.

11.3.6 Dérivation à un ordre plus élevé

Théorème 11.100

Soient I un intervalle de \mathbb{R} , $k \in \mathbb{N}^*$ et (f_n) une suite de fonctions définies sur I .

Si

- ▷ les fonctions f_n sont de classe \mathcal{C}^k sur I
- ▷ pour tout $j \in \llbracket 0 ; k-1 \rrbracket$, la suite de fonctions $(f_n^{(j)})$ converge simplement sur I
- ▷ la suite des dérivées k -èmes $(f_n^{(k)})$ converge uniformément sur I

alors $f = \lim_{n \rightarrow +\infty} f_n$ est de classe \mathcal{C}^k sur I et pour tout $j \in \llbracket 1 ; k \rrbracket$,

$$f^{(j)} = \lim_{n \rightarrow +\infty} f_n^{(j)}.$$

Démonstration 11.101

Par récurrence, en utilisant le Théorème 11.92. ■

On en déduit la version correspondante sur les séries de fonctions.

Corollaire 11.102

Soient I un intervalle de \mathbb{R} , $k \in \mathbb{N}^*$ et (f_n) une suite de fonctions définies sur I .

Si

- ▷ les fonctions f_n sont de classe \mathcal{C}^k sur I
- ▷ pour tout $j \in \llbracket 0 ; k-1 \rrbracket$, la série de fonctions $\sum_{n \geq 0} f_n^{(j)}$ converge simplement sur I
- ▷ la série des dérivées k -èmes $\sum_{n \geq 0} f_n^{(k)}$ converge uniformément sur I

alors $\sum_{n=0}^{+\infty} f_n$ est de classe \mathcal{C}^k sur I et pour tout $j \in \llbracket 1 ; k \rrbracket$,

$$\left(\sum_{n=0}^{+\infty} f_n \right)^{(j)} = \sum_{n=0}^{+\infty} f_n^{(j)}.$$

11.4 Généralisation

Dans cette section, E et F sont deux espaces vectoriels normés de dimension finie et les fonctions considérées vont de E dans F .

On étend sans peine les définitions de convergence simple et uniforme. Il suffit de remplacer les valeurs absolues ou modules par les normes.

11.4.1 Convergence simple

Définition 11.103

Soient A une partie de E et (f_n) une suite de fonctions définies sur A .

On dit que la suite (f_n) converge simplement sur A quand pour tout $x \in A$, la suite vectorielle $(f_n(x))$ converge dans F .

Dans ce cas, on peut définir une fonction f sur A en posant pour tout $x \in A$, $f(x) = \lim_{n \rightarrow +\infty} f_n(x)$.

La fonction f est alors appelée limite simple sur A de la suite (f_n) et on dit que la suite (f_n) converge simplement vers f sur A .

On adapte de même la définition de convergence simple d'une série de fonctions à valeurs dans F .

11.4.2 Convergence uniforme

Rappel 11.104

Si g est une fonction bornée sur une partie A de E , on pose $\|g\|_\infty^A = \sup_{x \in A} \|g(x)\|$.

$\|\cdot\|_\infty^A$ est une norme sur le \mathbb{K} -espace vectoriel des fonctions bornées sur A , appelée norme infinie ou norme uniforme sur A .

Définition 11.105

Soient A une partie de E , (f_n) une suite de fonctions définies sur A et f une fonction définie sur A .

On dit que la suite (f_n) converge uniformément vers f sur A quand pour tout $n \in \mathbb{N}$, $f - f_n$ est bornée sur A et la suite réelle $(\|f - f_n\|_\infty^A)$ converge vers 0.

La fonction f est alors appelée limite uniforme sur A de la suite (f_n) .

On adapte de même la définition de convergence uniforme d'une série de fonctions à valeurs dans F .

11.4.3 Convergence normale des séries

Définition 11.106

Soient A une partie de E et (f_n) une suite de fonctions définies sur A .

On dit que la série $\sum_{n \geq 0} f_n$ converge normalement sur A quand la série numérique $\sum_{n \geq 0} \|f_n\|_\infty^A$ converge.

Théorème 11.107

Si une série de fonctions converge normalement sur A , alors elle converge uniformément (et donc simplement) sur A .

11.4.4 Résultats préservés

Tous les résultats sur la continuité et la double limite restent valables pour des fonctions de E quand F .

Tous les résultats sur la primitivation, l'intégration sur des segments, la dérivation \mathcal{C}^1 ou, plus généralement \mathcal{C}^k , restent valables pour des fonctions de \mathbb{R} dans F .

11.5 Approximation uniforme

11.5.1 Densité des fonctions en escaliers dans les fonctions continues par morceaux

On a déjà rencontré des exemples d'approximation uniforme en première année pour des fonctions à valeurs réelles. En travaillant sur les coordonnées, on peut généraliser ce résultat.

Théorème 11.108

Toute fonction continue par morceaux sur un segment est approchable uniformément par des fonctions en escaliers.

Soit f une fonction continue par morceaux sur un segment $[a ; b]$ et à valeurs dans F . Il existe une suite de fonctions en escaliers sur $[a ; b]$ qui converge uniformément vers f sur $[a ; b]$.

Avec le vocabulaire des espaces vectoriels normés, on peut retraduire ce résultat.

Proposition 11.109

Dans l'espace des fonctions bornées $\mathcal{B}([a ; b], F)$, muni de la norme infinie, le sous-espace des fonctions en escaliers $\text{Esc}([a ; b], F)$ est dense dans le sous-espace des fonctions continues par morceaux $\mathcal{C}_m^0([a ; b], F)$.

Autrement dit

$$\forall f \in \mathcal{C}_m^0([a ; b], F), \quad \forall \varepsilon > 0, \quad \exists \varphi \in \text{Esc}([a ; b], F), \quad \|f - \varphi\|_\infty^{[a;b]} \leq \varepsilon.$$

En application, un exercice classique : le lemme de Lebesgue (Centrale - Mines).

Exercice 11.110

Soit $f : [a ; b] \longrightarrow \mathbb{K}$ continue par morceaux. Montrez que $\lim_{n \rightarrow +\infty} \int_a^b f(t) \sin(nt) dt = 0$.

Correction 11.111

- Si f est en escaliers sur $[a ; b]$, il existe une subdivision $a = c_0 < \dots < c_p = b$ du segment $[a ; b]$ telle que

$$\text{pour tout } i \in \llbracket 0 ; p-1 \rrbracket, \quad f|_{]c_i; c_{i+1}[} = \lambda_i \text{ constante.}$$

Donc on a

$$\begin{aligned} \int_a^b f(t) \sin(nt) dt &= \sum_{i=0}^{p-1} \int_{c_i}^{c_{i+1}} \lambda_i \sin(nt) dt \\ &= \sum_{i=0}^{p-1} \lambda_i \left[\frac{-\cos(nt)}{n} \right]_{c_i}^{c_{i+1}} \\ &\xrightarrow{n \rightarrow +\infty} 0. \end{aligned} \quad \left(\left| \left[\frac{-\cos(nt)}{n} \right]_{c_i}^{c_{i+1}} \right| \leq \frac{2}{n} \right)$$

- Si f est continue par morceaux sur $[a ; b]$:

Soit $\varepsilon > 0$. Il existe g une fonction en escaliers sur $[a ; b]$ telle que

$$\|f - g\|_{\infty}^{[a;b]} < \varepsilon.$$

On a $\int_a^b f(t) \sin(nt) dt = \int_a^b (f(t) - g(t)) \sin(nt) dt + \int_a^b g(t) \sin(nt) dt$ donc

$$\begin{aligned} \left| \int_a^b f(t) \sin(nt) dt \right| &\leq \int_a^b |f(t) - g(t)| dt + \left| \int_a^b g(t) \sin(nt) dt \right| \\ &\leq \varepsilon(b-a) + \left| \int_a^b g(t) \sin(nt) dt \right|. \end{aligned}$$

Or, d'après le cas précédent, on a $\lim_{n \rightarrow +\infty} \int_a^b g(t) \sin(nt) dt = 0$ donc il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \quad \left| \int_a^b g(t) \sin(nt) dt \right| \leq \varepsilon(b-a).$$

Donc

$$\forall n \geq n_0, \quad \left| \int_a^b f(t) \sin(nt) dt \right| \leq 2\varepsilon(b-a).$$

Donc $\lim_{n \rightarrow +\infty} \int_a^b f(t) \sin(nt) dt = 0$.

11.5.2 Densité des polynômes sur un segment dans les fonctions continues

Il s'agit du théorème de Stone-Weierstrass.

Théorème 11.112

Toute fonction continue sur un segment est approchable uniformément par des fonctions polynômes.

Soit f une fonction continue sur un segment $[a ; b]$ et à valeurs dans \mathbb{K} . Il existe une suite de fonctions polynômes sur $[a ; b]$ qui converge uniformément vers f sur $[a ; b]$.

Avec le vocabulaire des espaces vectoriels normés, on peut retraduire ce résultat (en identifiant les polynômes et les fonctions polynômes sur un segment, ce qui est loisible car un segment est un ensemble infini).

Proposition 11.113

Dans l'espace des fonctions bornées $\mathcal{B}([a ; b], \mathbb{K})$, muni de la norme infinie, le sous-espace des fonctions polynômes $\mathcal{P}([a ; b], \mathbb{K}) = \mathbb{K}[X]$ est dense dans le sous-espace des fonctions continues $\mathcal{C}^0([a ; b], \mathbb{K})$.

Autrement dit

$$\forall f \in \mathcal{C}^0([a ; b], \mathbb{K}), \forall \varepsilon > 0, \exists P \in \mathbb{K}[X], \|f - P\|_{\infty}^{[a;b]} \leq \varepsilon.$$

En application, un exercice classique (CCP).

Exercice 11.114

Soit $f : [a ; b] \rightarrow \mathbb{K}$ continue telle que pour tout $n \in \mathbb{N}$, $\int_a^b t^n f(t) dt = 0$.

Montrez que $f = 0$.

Correction 11.115

On a $\forall n \in \mathbb{N}$, $\int_a^b t^n f(t) dt = 0$ donc par linéarité de l'intégrale, on a

$$\forall P \in \mathbb{K}[X], \int_a^b P(t) f(t) dt = 0.$$

f étant continue sur $[a ; b]$, d'après le théorème de Stone-Weierstrass, il existe une suite de polynômes $(P_k)_k$ qui converge uniformément vers f sur $[a ; b]$.

Donc pour tout $k \in \mathbb{N}$, $\int_a^b P_k(t) f(t) dt = 0$.

Pour $t \in [a ; b]$, on a

$$\begin{aligned} |P_k(t) f(t) - f^2(t)| &= |f(t)| \times |P_k(t) - f(t)| \\ &\leq \sup_{[a;b]} |f| \times \|P_k - f\|_{\infty}^{[a;b]}. \end{aligned} \quad \left. \vphantom{\int_a^b} \right\} \text{bornes atteintes}$$

Or $\|P_k - f\|_\infty^{[a;b]} \xrightarrow[k \rightarrow +\infty]{} 0$ donc, par encadrement, on a

$$\|P_k f - f^2\|_\infty^{[a;b]} \xrightarrow[k \rightarrow +\infty]{} 0$$

i.e. la suite $(P_k f)_k$ converge uniformément vers f^2 sur $[a ; b]$.

Donc, d'après le Corollaire 11.78, on a

$$0 = \int_a^b P_k f \xrightarrow[k \rightarrow +\infty]{} \int_a^b f^2.$$

Donc $\int_a^b f^2 = 0$, or $f^2 \geq 0$ et continue donc, d'après le théorème de stricte positivité de l'intégrale, $f^2 = 0$ sur $[a ; b]$.

D'où $f = 0$.

Chapitre 12

Séries entières

Sommaire

12.1	Convergence simple d'une série entière.	362
12.1.1	Rayon de convergence	362
12.1.2	Détermination du rayon de convergence	364
12.1.3	Comparaison de séries entières	368
12.1.4	Opérations sur les séries entières	371
12.2	Propriétés de la fonction somme d'une série entière.	373
12.2.1	Convergence uniforme et continuité	373
12.2.2	Primitivation et dérivation	374
12.2.3	Convergence radiale	378
12.3	Fonction développable en série entière.	379
12.3.1	Généralités	379
12.3.2	Unicité du développement en série entière	382
12.3.3	Série de Taylor d'une fonction	383
12.3.4	Développements en série entière usuels	386

Dans ce chapitre, on s'intéresse aux séries de fonctions de la forme $x \mapsto a_n x^n$, qu'on appelle des séries entières.

Par abus de notations, on les note sous la forme $\sum_{n \geq 0} a_n x^n$.

Lorsque la variable est complexe, on la note systématiquement z et on parle des séries $\sum_{n \geq 0} a_n z^n$.

12.1 Convergence simple d'une série entière

12.1.1 Rayon de convergence

On commence par le lemme d'Abel.

Proposition 12.1

Soient (a_n) une suite de complexes et z_0 un complexe.

Si la suite $(a_n z_0^n)$ est bornée, alors pour $z \in \mathbb{C}$ tel que $|z| < |z_0|$, la série numérique $\sum_{n \geq 0} a_n z^n$ converge absolument.

Démonstration 12.2

On suppose que $(a_n z_0^n)_n$ est bornée et on se donne $z \in \mathbb{C}$ tel que $|z| < |z_0|$.

On a $a_n z^n = a_n z_0^n \left(\frac{z}{z_0}\right)^n$ et il existe $M > 0$ tel que

$$\forall n \in \mathbb{N}, \quad |a_n z_0^n| \leq M$$

donc

$$\forall n \in \mathbb{N}, \quad |a_n z^n| \leq M \left| \frac{z}{z_0} \right|^n.$$

Or $\left| \frac{z}{z_0} \right| < 1$ donc $\sum \left| \frac{z}{z_0} \right|^n$ converge.

Donc, par théorème de comparaison des séries à termes positifs, $\sum a_n z^n$ converge absolument. ■

Définition 12.3

Soit (a_n) une suite de complexes. On pose $M = \{r \in \mathbb{R}_+ \mid \text{la suite } (a_n r^n) \text{ est bornée}\}$.

M est une partie non-vide de \mathbb{R} car elle contient 0.

Si elle est majorée, on appelle rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$ le réel $R = \sup M$.

Sinon, on pose $R = +\infty$.

On peut préciser la convergence simple d'une série entière.

Proposition 12.4

Soit (a_n) une suite de complexes. On pose R le rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$.

Pour tout $z \in \mathbb{C}$,

- ▷ si $|z| < R$, alors la série numérique $\sum_{n \geq 0} a_n z^n$ converge absolument
- ▷ si $|z| > R$, alors la série numérique $\sum_{n \geq 0} a_n z^n$ diverge grossièrement
- ▷ si $|z| = R$, alors on ne peut rien dire a priori.

Démonstration 12.5

Supposons $R \in \mathbb{R}_+$.

- Si $|z| < R$, par définition de R , $|z|$ n'est pas un majorant de $\{r \geq 0 \mid (a_n r^n) \text{ bornée}\}$.

Donc il existe $r \geq 0$ tel que $(a_n r^n)$ est bornée et $|z| < r$.

Donc, d'après la Proposition 12.1, $\sum a_n z^n$ converge (absolument).

Idem si $R = +\infty$.

- Si $|z| > R$ alors $(a_n z^n)$ n'est pas bornée et ne converge donc pas vers 0, donc $\sum a_n z^n$ diverge grossièrement. ■

Définition 12.6

Soit (a_n) une suite de complexes. On pose R le rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$.

Si $R > 0$, on appelle disque ouvert de convergence l'ensemble $D(0, R) \subseteq \mathbb{C}$ et intervalle ouvert de convergence l'ensemble $] -R ; R[\subseteq \mathbb{R}$.

Exemple 12.7

- Le rayon de convergence de la série entière $\sum_{n \geq 0} z^n$ est égal à 1 et pour tout $z \in D(0, 1)$, $\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$.
- Le rayon de convergence de la série entière $\sum_{n \geq 0} \frac{z^n}{n!}$ est égal à $+\infty$ et pour tout $z \in \mathbb{C}$, $\sum_{n=0}^{+\infty} \frac{z^n}{n!} = e^z$.

12.1.2 Détermination du rayon de convergence

La réciproque de la Proposition 12.4 est vraie.

Proposition 12.8

Soient (a_n) une suite de complexes et R un réel strictement positif.

Si

- pour tout $z \in \mathbb{C}$ tel que $|z| < R$, la série numérique $\sum_{n \geq 0} a_n z^n$ converge
- pour tout $z \in \mathbb{C}$ tel que $|z| > R$, la série numérique $\sum_{n \geq 0} a_n z^n$ diverge

alors R est le rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$.

Démonstration 12.9

Soit $R > 0$ tel que
$$\begin{cases} \sum a_n z^n \text{ converge} & \text{si } |z| < R \\ \sum a_n z^n \text{ diverge} & \text{si } |z| > R \end{cases}$$

On note R_0 le rayon de convergence de la série entière $\sum a_n z^n$.

Si $|z| < R$, $\sum a_n z^n$ converge donc $a_n z^n \xrightarrow{n \rightarrow +\infty} 0$, donc $(a_n z^n)$ est bornée, donc $|z| \in \{r \geq 0 \mid (a_n z^n) \text{ bornée}\}$, donc $|z| \leq R_0$, *i.e.*

$$D(0, R) \subseteq \overline{D}(0, R_0)$$

donc $R \leq R_0$.

Réciproquement, si $R < R_0$, alors on choisit $z \in \mathbb{C}$ tel que $R < |z| < R_0$.

On a $|z| < R_0$ donc d'après la Proposition 12.4, $\sum a_n z^n$ converge.

Or $R < |z|$ donc $\sum a_n z^n$ diverge : contradiction.

Donc $R = R_0$. ■

De même, on peut donner quelques résultats pratiques découlant de la définition.

Proposition 12.10

Soient (a_n) une suite de complexes et R un réel strictement positif.

Si

- *pour tout $z \in \mathbb{C}$ tel que $|z| < R$, la série numérique $\sum_{n \geq 0} a_n z^n$ converge*
- *il existe $z_0 \in \mathbb{C}$ tel que $|z_0| = R$ et la série $\sum_{n \geq 0} a_n z_0^n$ diverge*

alors R est le rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$.

Démonstration 12.11

Soit $R > 0$ vérifiant les conditions énoncées.

On note R_0 le rayon de convergence de la série entière $\sum a_n z^n$.

Si $|z| < R$, $\sum a_n z^n$ converge donc $a_n z^n \xrightarrow{n \rightarrow +\infty} 0$, donc $(a_n z^n)$ est bornée, donc $|z| \in \{r \geq 0 \mid (a_n z^n) \text{ bornée}\}$, donc $|z| \leq R_0$, *i.e.*

$$D(0, R) \subseteq \overline{D}(0, R_0)$$

donc $R \leq R_0$.

Réciproquement, si $R < R_0$, alors $R = |z_0| < R_0$.

On a $|z_0| < R_0$ donc d'après la Proposition 12.4, $\sum a_n z_0^n$ converge.

Or $R = |z_0|$ donc $\sum a_n z_0^n$ diverge : contradiction.

Donc $R = R_0$. ■

Proposition 12.12

Soient (a_n) une suite de complexes et R un réel strictement positif.

Si

► il existe $z_0 \in \mathbb{C}$ tel que $|z_0| = R$ et la série $\sum_{n \geq 0} a_n z_0^n$ converge

► pour tout $z \in \mathbb{C}$ tel que $|z| > R$, la série numérique $\sum_{n \geq 0} a_n z^n$ diverge

alors R est le rayon de convergence de la série entière $\sum_{n \geq 0} a_n z^n$.

Démonstration 12.13

Soit $R > 0$ vérifiant les conditions énoncées.

On note R_0 le rayon de convergence de la série entière $\sum a_n z^n$.

$\sum a_n z_0^n$ converge donc $a_n z_0^n \xrightarrow{n \rightarrow +\infty} 0$, donc $(a_n z_0^n)$ est bornée, donc

$$\forall z \in \mathbb{C}, |z| < |z_0| = R \implies \sum a_n z^n \text{ converge.}$$

Donc d'après la Proposition 12.8, $R = R_0$. ■

Exercice 12.14

Donnez le rayon de convergence des séries entières suivantes :

$$\sum_{n \geq 0} n z^n \quad \sum_{n \geq 1} \frac{z^n}{n} \quad \sum_{n \geq 0} n! z^n \quad \sum_{n \geq 0} 2^n z^{2n} \quad \sum_{n \geq 0} \cos^2(n) z^n.$$

Correction 12.15

▷ $(a_n) = (n)$

Si $|z| > 1$, on a $\forall n \in \mathbb{N}$, $|a_n z^n| \geq n$ donc $\sum n z^n$ diverge grossièrement.

Si $|z| < 1$, comme $n^3 \underset{n \rightarrow +\infty}{=} o\left(\left(\frac{1}{|z|}\right)^n\right)$, on a $n|z|^n \underset{n \rightarrow +\infty}{=} o\left(\frac{1}{n^2}\right)$ donc $\sum n z^n$ converge absolument par théorème de comparaison des séries à termes positifs.

Ainsi, d'après la Proposition 12.8, $R = 1$.

Remarque : si $|z| = 1$, de même, $\sum n z^n$ diverge grossièrement donc

$$\left\{z \in \mathbb{C} \mid \sum n z^n \text{ converge} \right\} = D(0, 1).$$

▷ $(a_n) = \left(\frac{1}{n}\right)$

$\sum \frac{(-1)^n}{n}$ converge par critère spécial des séries alternées et

$$\forall z \in \mathbb{C}, |z| > 1 \implies \left| \frac{z^n}{n} \right| = \frac{|z|^n}{n} \underset{n \rightarrow +\infty}{=} o(|z|^n)$$

donc $\sum \frac{z^n}{n}$ diverge.

Ainsi, d'après la Proposition 12.8, $R = 1$.

Remarque : on peut montrer que si $z = e^{i\theta}$ et $\theta \not\equiv 0 [2\pi]$, alors $\sum \frac{e^{in\theta}}{n}$ converge, donc

$$\left\{z \in \mathbb{C} \mid \sum \frac{z^n}{n} \text{ converge} \right\} = \overline{D}(0, 1) \setminus \{1\}.$$

▷ $(a_n) = (n!)$

$\sum n! z^n$ ne converge que si $z = 0$ car pour tout $a \in \mathbb{R}_+$, $a^n \underset{n \rightarrow +\infty}{=} o(n!)$ donc pour tout $z \in \mathbb{C}^*$, $|n! z^n| \xrightarrow{n \rightarrow +\infty} +\infty$.

Donc $R = 0$.

▷ On a $\sum 2^n z^{2n} = \sum a_n z^n$ où $\begin{cases} a_n = 0 & \text{si } n \text{ impair} \\ a_n = 2^{n/2} & \text{sinon} \end{cases}$

On a $2^n z^{2n} = (2z^2)^n$ or

$$\sum (2z^2)^n \text{ converge} \iff |2z^2| < 1$$

$$\iff |z|^2 < \frac{1}{2}$$

$$\iff |z| < \frac{1}{\sqrt{2}}.$$

Donc $R = \frac{1}{\sqrt{2}}$.

Remarque : une telle série entière (où des termes sont nuls) est dite lacunaire.

▷ $(a_n) = (\cos^2 n)$

Pour tout $z \in D(0, 1)$, $|\cos^2(n) z^n| \leq |z|^n$ donc par théorème de comparaison des séries à termes positifs, $\sum \cos^2(n) z^n$ converge (absolument).

La suite (a_n) diverge donc $\sum a_n$ diverge grossièrement et donc $\sum a_n z^n$ diverge quand $z = 1$.

Donc $R = 1$.

12.1.3 Comparaison de séries entières

Proposition 12.16

Soient $(a_n), (b_n)$ deux suites de complexes.

On pose R_a, R_b les rayons de convergence des séries entières $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} b_n z^n$.

Si $a_n \underset{n \rightarrow +\infty}{=} \mathcal{O}(b_n)$, alors $R_a \geq R_b$.

En particulier, si $|a_n| \leq |b_n|$, alors $R_a \geq R_b$.

Si $a_n \underset{n \rightarrow +\infty}{\sim} b_n$, alors $R_a = R_b$.

Démonstration 12.17

▷ Si $a_n \underset{n \rightarrow +\infty}{=} \mathcal{O}(b_n)$, il existe $K > 0$ et $n_0 \in \mathbb{N}$ tels que

$$\forall n \geq n_0, |a_n| \leq K |b_n|.$$

Donc $\forall n \geq n_0, \forall z \in \mathbb{C}, |a_n z^n| \leq K |b_n z^n|$.

Si $|z| < R_b$ alors $\sum |b_n z^n|$ converge, donc par théorème de comparaison des séries à termes positifs, $\sum |a_n z^n|$ converge, et donc $|z| < R_a$.

Donc $D(0, R_b) \subseteq \overline{D}(0, R_a)$ i.e. $R_b \leq R_a$.

▷ Si $a_n \underset{n \rightarrow +\infty}{\sim} b_n$, alors $a_n \underset{n \rightarrow +\infty}{=} \mathcal{O}(b_n)$ et $b_n \underset{n \rightarrow +\infty}{=} \mathcal{O}(a_n)$.

Donc $R_a \geq R_b$ et $R_b \geq R_a$, i.e. $R_a = R_b$. ■

Remarque 12.18

Comme le rayon de convergence se calcule en référence à la convergence absolue de la série, il suffit que $|a_n| \underset{n \rightarrow +\infty}{\sim} |b_n|$ pour avoir $R_a = R_b$.

Exercice 12.19

Donnez le rayon de convergence de la série entière $\sum_{n \geq 0} \frac{n}{n^2 + 1} z^n$.

Correction 12.20

On a $\frac{n}{n^2 + 1} \underset{n \rightarrow +\infty}{\sim} \frac{1}{n}$, or $R\left(\sum \frac{z^n}{n}\right) = 1$.

Donc $R\left(\sum \frac{n}{n^2 + 1} z^n\right) = 1$.

Exercice 12.21

Même question avec la série entière $\sum_{n \geq 0} \left(1 + \frac{1}{n}\right)^{n^2} z^n$.

Correction 12.22

On a $a_n = \left(1 + \frac{1}{n}\right)^{n^2} = e^{n^2 \ln(1+1/n)}$.

Or $\frac{1}{n} \xrightarrow{n \rightarrow +\infty} 0$ donc

$$\ln\left(1 + \frac{1}{n}\right) \underset{n \rightarrow +\infty}{=} \frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right)$$

$$n^2 \ln\left(1 + \frac{1}{n}\right) \underset{n \rightarrow +\infty}{=} n - \frac{1}{2} + o(1)$$

$$\exp\left(n^2 \ln\left(1 + \frac{1}{n}\right)\right) \underset{n \rightarrow +\infty}{\sim} e^n e^{-1/2}$$

$$\left(1 + \frac{1}{n}\right)^{n^2} \underset{n \rightarrow +\infty}{\sim} \frac{e^n}{\sqrt{e}}.$$

Donc $R\left(\sum a_n z^n\right) = R\left(\sum \frac{1}{\sqrt{e}} e^n z^n\right) = \frac{1}{e}$.

Un cas favorable très courant qui permet de calculer le rayon de convergence par utilisation de la règle de D'Alembert, particulièrement utile quand le coefficient a_n s'écrit à l'aide de produits ou de quotients.

Rappel 12.23 (Règle de D'Alembert)

Soit $(u_n) \in (\mathbb{R}^*)^{\mathbb{N}}$. On suppose $\frac{u_{n+1}}{u_n} \xrightarrow{n \rightarrow +\infty} \ell$.

Si $\ell < 1$, $\sum u_n$ converge.

Si $\ell > 1$, $\sum u_n$ diverge grossièrement.

Si $\ell = 1$, on ne peut rien dire.

Proposition 12.24

Soit (a_n) une suite de complexes qui ne s'annule pas.

Si la suite $\left(\frac{|a_{n+1}|}{|a_n|}\right)$ converge

▷ vers un réel $\ell > 0$, alors $R_a = \frac{1}{\ell}$

▷ vers 0, alors $R_a = +\infty$.

Démonstration 12.25

Soit $z \in \mathbb{C}^*$.

La suite $(|a_n z^n|)$ est à termes strictement positifs.

Donc si $\frac{|a_{n+1} z^{n+1}|}{|a_n z^n|} = \frac{|a_{n+1}|}{|a_n|} |z| \xrightarrow{n \rightarrow +\infty} L < 1$, alors $\sum |a_n z^n|$ converge.

De plus, si $\frac{|a_{n+1}|}{|a_n|} |z| \xrightarrow{n \rightarrow +\infty} L > 1$, alors $\sum |a_n z^n|$ diverge grossièrement.

Donc si $\left|\frac{a_{n+1}}{a_n}\right| \xrightarrow{n \rightarrow +\infty} \ell$, $\left|\frac{a_{n+1}}{a_n}\right| |z| \xrightarrow{n \rightarrow +\infty} \ell |z|$.

Donc si $\ell |z| < 1$, $\sum a_n z^n$ converge absolument et $R = \frac{1}{\ell}$ (Proposition 12.8) et si $\ell |z| > 1$, $\sum a_n z^n$ diverge. ■

Exercice 12.26

Donnez le rayon de convergence de la série entière $\sum_{n \geq 0} \frac{2^n + 1}{n^2 + 1} z^n$.

On pose $(a_n) = \left(\frac{2^n + 1}{n^2 + 1} \right)$.

On a

$$\begin{aligned} \left| \frac{a_{n+1}}{a_n} \right| &= \frac{(2^{n+1} + 1)(n^2 + 1)}{((n+1)^2 + 1)(2^n + 1)} \\ &\underset{n \rightarrow +\infty}{\sim} \frac{2^{n+1}n^2}{(n+1)^2 2^n} \\ &\underset{n \rightarrow +\infty}{\sim} \frac{2n^2}{(n+1)^2} \\ &\xrightarrow[n \rightarrow +\infty]{} 2. \end{aligned}$$

Donc, d'après la Proposition 12.24, $R = \frac{1}{2}$.

Exercice 12.28

Même question avec la série entière $\sum_{n \geq 0} \binom{2n}{n} z^{2n}$.

12.1.4 Opérations sur les séries entières

Proposition 12.29

Soient $(a_n), (b_n)$ deux suites de complexes et $\lambda \in \mathbb{C}$.

On pose R_a, R_b les rayons de convergence des séries entières $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} b_n z^n$.

On a

$$R_{a+b} \geq \min(R_a, R_b)$$

et

$$R_{\lambda a} \geq R_a$$

avec égalité si $\lambda \neq 0$

Démonstration 12.30

Pour tout $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$, les séries $\sum a_n z^n$ et $\sum b_n z^n$ convergent donc la série $\sum (a_n + b_n) z^n$ converge.

Donc $D(0, \min(R_a, R_b)) \subseteq \overline{D}(0, R_{a+b})$ i.e. $\min(R_a, R_b) \leq R_{a+b}$.

De même pour la multiplication par un scalaire. ■

Par conséquent, l'ensemble des séries entières qui convergent sur un disque $D(0, R)$ est un \mathbb{C} -espace vectoriel.

Proposition 12.31

Soient $(a_n), (b_n)$ deux suites de complexes.

On pose R_a, R_b les rayons de convergence des séries entières $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} b_n z^n$.

On pose (c_n) le produit de Cauchy des suites (a_n) et (b_n) .

Alors la série entière $\sum_{n \geq 0} c_n z^n$ est le produit de Cauchy des séries $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} b_n z^n$ et $R_c \geq \min(R_a, R_b)$.

Démonstration 12.32

$$\text{On a } \sum_{k=0}^n (a_k z^k) (b_{n-k} z^{n-k}) = \left(\sum_{k=0}^n a_k b_{n-k} \right) z^n.$$

Pour tout $z \in \mathbb{C}$ tel que $|z| \leq \min(R_a, R_b)$, les séries $\sum a_n z^n$ et $\sum b_n z^n$ convergent absolument donc la série $\sum c_n z^n$ où $c_n = \sum_{k=0}^n a_k b_{n-k}$ converge absolument, i.e. $R_c \geq \min(R_a, R_b)$. ■

Exercice 12.33

Calculez le produit de Cauchy de la série $\sum_{n \geq 0} z^n$ avec elle-même.

Donnez la valeur de $\sum_{n=0}^{+\infty} n z^n$ quand elle a un sens.

Correction 12.34

La série entière $\sum z^n$ a pour rayon de convergence 1.

Le produit de Cauchy de cette série avec elle-même est la série $\sum (n+1) z^n$ (car $\sum_{k=0}^n 1 \times 1 = n+1$), qui a aussi pour rayon de convergence 1.

Pour $z \in D(0, 1)$, on a $\sum_{n=0}^{+\infty} (n+1) z^n = \left(\sum_{n=0}^{+\infty} z^n \right)^2$.

Donc

$$\sum_{n=0}^{+\infty} n z^n = \left(\frac{1}{1-z} \right)^2 - \frac{1}{1-z} = \frac{z}{(1-z)^2}.$$

Remarque 12.35

Le résultat énoncé ici est utilisable avec des séries $\sum_{n \geq 0} a_n z^n$: on commence à $n = 0$!

Dans le cas général, on s'y ramène en ajoutant des termes nuls au début des séries : si on veut faire le produit de Cauchy de deux séries $\sum_{n \geq n_1} a_n z^n$ et $\sum_{n \geq n_2} b_n z^n$, on pose $a_0 = \dots = a_{n_1-1} = 0$ et $b_0 = \dots = b_{n_2-1} = 0$, puis on peut appliquer le résultat précédent : pour tout $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$,

$$\sum_{n=n_1}^{+\infty} a_n z^n \times \sum_{n=n_2}^{+\infty} b_n z^n = \sum_{n=0}^{+\infty} a_n z^n \times \sum_{n=0}^{+\infty} b_n z^n = \sum_{n=0}^{+\infty} c_n z^n$$

où pour tout $n \in \mathbb{N}$, $c_n = \sum_{k=0}^n a_k b_{n-k} = a_0 b_n + a_1 b_{n-1} + \dots + a_{n-1} b_1 + a_n b_0$.

Maintenant on peut prendre en compte les termes nuls.

Si $n < n_1 + n_2$, alors dans cette somme qu'on coupe en deux, tous les termes sont nuls :

$$c_n = \underbrace{a_0 b_n + \dots + a_{n_1-1} b_{n+1-n_1}}_{\forall i \in \llbracket 0; n_1-1 \rrbracket, a_i=0} + \underbrace{a_{n_1} b_{n-n_1} + \dots + a_n b_0}_{\forall i \in \llbracket n_1; n \rrbracket, n-i \leq n-n_1 \text{ donc } b_{n-i}=0} = 0$$

donc $\sum_{n=n_1}^{+\infty} a_n z^n \times \sum_{n=n_2}^{+\infty} b_n z^n = \sum_{n=n_1+n_2}^{+\infty} c_n z^n$ où pour tout $n \geq n_1 + n_2$, $c_n = \sum_{k=n_1}^{n-n_2} a_k b_{n-k}$.

Ne retenons pas ça ! Retenons l'idée et ré-appliquons-la à chaque fois car, en pratique, on a très souvent $(n_1, n_2) = (1, 0)$ ou $(1, 1)$.

12.2 Propriétés de la fonction somme d'une série entière

12.2.1 Convergence uniforme et continuité

Théorème 12.36

Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence $R > 0$.

La série de fonctions $\sum_{n \geq 0} a_n z^n$ converge normalement (donc uniformément) sur tout compact inclus dans $D(0, R)$.

Démonstration 12.37

Soit K un compact non-vidé inclus dans $D(0, R)$.

On veut montrer qu'il existe $r \in [0 ; R[$ tel que $K \subseteq \overline{D}(0, r) \subseteq D(0, R)$.

La fonction $z \mapsto |z|$ étant continue sur le compact K , d'après le théorème des bornes atteintes, elle possède un maximum r sur K .

Il existe donc $z_0 \in K$ tel que $r = |z_0|$.

Or $K \subseteq D(0, R)$ donc $|z_0| = r < R$.

Pour tout $n \in \mathbb{N}$ et $z \in K$, on a $|a_n z^n| \leq |a_n| r^n = |a_n z_0^n|$.

Donc $\|z \mapsto a_n z^n\|_\infty^K \leq |a_n z_0^n|$.

Or $\sum a_n z_0^n$ converge absolument car $|z_0| < R$.

Donc, d'après le théorème de comparaison des séries à termes positifs, $\sum \|z \mapsto a_n z^n\|_\infty^K$ converge, i.e. $\sum a_n z^n$ converge normalement. ■

D'après le théorème de continuité des séries de fonctions, on en déduit la continuité de la fonction somme.

Théorème 12.38

Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence $R > 0$.

La fonction somme $z \mapsto \sum_{n=0}^{+\infty} a_n z^n$ est continue sur $D(0, R)$.

12.2.2 Primitivation et dérivation

Lemme 12.39

Soit (a_n) une suite de complexes.

Les séries entières $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} n a_n z^n$ ont le même rayon de convergence.

Plus généralement, pour tout polynôme P non-nul, la série entière $\sum_{n \geq 0} P(z) a_n z^n$ a le même rayon de convergence.

Démonstration 12.40

Soit $\sum a_n z^n$ une série entière de rayon de convergence R .

On note R' le rayon de convergence de $\sum n a_n z^n$.

- Pour tout $n \geq 1$, $|a_n z^n| \leq |n a_n z^n|$ donc si $|z| < R'$, alors $\sum n a_n z^n$ converge absolument et donc, par théorème de comparaison des séries à termes positifs, $\sum a_n z^n$ converge. Donc $|z| \leq R$.

Donc $D(0, R') \subseteq \overline{D}(0, R)$, et donc $R' \leq R$.

► Si $|z| < R$, on choisit z' tel que $|z| < |z'| < R$.

On a

$$\begin{aligned} |na_n z^n| &= |na_n z'^n| \times \left| \frac{z}{z'} \right|^n \\ &= |a_n z'^n| \times n \left| \frac{z}{z'} \right|^n. \end{aligned}$$

Or $\left| \frac{z}{z'} \right| < 1$ donc $n \left| \frac{z}{z'} \right|^n \xrightarrow{n \rightarrow +\infty} 0$ et donc $\left(n \left| \frac{z}{z'} \right|^n \right)$ est bornée : on choisit K un majorant.

On a alors $\forall n \in \mathbb{N}, |na_n z^n| \leq K |a_n z'^n|$.

Or $|z'| < R$ donc $\sum a_n z'^n$ est absolument convergente et donc, par théorème de comparaison des séries à termes positifs, $\sum na_n z^n$ converge (absolument).

Donc $|z| \leq R'$.

Donc $D(0, R) \subseteq \overline{D}(0, R')$ et donc $R \leq R'$.

► D'où $R = R'$.

Par récurrence immédiate, pour tout $k \in \mathbb{N}$, la série $\sum n^k a_n z^n$ a pour rayon de convergence R .

Donc pour tout $P \in \mathbb{C}[X]$, comme $P(n) \underset{n \rightarrow +\infty}{\sim} \lambda n^k$, d'après la Proposition 12.16, le rayon de convergence de $\sum P(n) a_n z^n$ est celui de $\sum \lambda n^k a_n z^n$ i.e. R . ■

D'après le théorème de dérivabilité des séries de fonctions, on en déduit la dérivabilité de la fonction somme.

Théorème 12.41

Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence $R > 0$.

La fonction somme $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est de classe \mathcal{C}^∞ sur $] -R ; R[$.

De plus, on a

$$\forall k \in \mathbb{N}, \forall x \in] -R ; R[, f^{(k)}(x) = \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} a_n x^{n-k}.$$

On en déduit que pour tout $k \in \mathbb{N}$, $a_k = \frac{f^{(k)}(0)}{k!}$.

Démonstration 12.42

Pour $n \in \mathbb{N}$, on pose $u_n : x \mapsto a_n x^n$ et u_n est de classe \mathcal{C}^1 .

Pour tout $n \in \mathbb{N}$, on a $u'_n(x) = n a_n x^{n-1}$.

D'après le Lemme 12.39, la série entière $\sum n a_n x^{n-1}$ a pour rayon de convergence R .

Donc, pour tout $r \in [0 ; R[$, la série $\sum u'_n$ converge normalement (et donc uniformément) sur $[-r ; r]$ d'après le Théorème 12.36.

D'après le théorème de dérivation terme à terme, la fonction $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est de classe \mathcal{C}^1 sur $[-r ; r]$.

Par réunion d'intervalles, f est \mathcal{C}^1 sur $] -R ; R[$ et pour tout $x \in] -R ; R[$, on a $f'(x) = \sum_{n=0}^{+\infty} n a_n x^{n-1}$.

Par récurrence, f est de classe \mathcal{C}^∞ et on a

$$\begin{aligned} f^{(k)}(x) &= \sum_{n=0}^{+\infty} n(n-1)\dots(n-k+1) a_n x^{n-k} \\ &= \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} a_n x^{n-k}. \end{aligned}$$

En évaluant en 0, on a $f^{(k)}(0) = \frac{k!}{0!} a_k$ et donc $a_k = \frac{f^{(k)}(0)}{k!}$. ■

On en déduit le théorème de primitivation.

Théorème 12.43

Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence $R > 0$. On note f sa somme.

La fonction somme $x \mapsto \sum_{n=0}^{+\infty} a_n \frac{x^{n+1}}{n+1}$ est de classe \mathcal{C}^∞ sur $] -R ; R[$ et est une primitive de f sur $] -R ; R[$ (c'est d'ailleurs la primitive de f qui s'annule en 0).

Exercice 12.44

Donnez le rayon de convergence de la série entière $\sum_{n \geq 0} n x^n$ et retrouvez la valeur de sa somme.

Correction 12.45

La série entière $\sum nx^n$ a pour rayon de convergence 1.

On pose $g : x \mapsto \sum_{n=0}^{+\infty} nx^n$.

Pour $x \in]-1 ; 1[$, on a

$$\begin{aligned} g(x) &= x \sum_{n=0}^{+\infty} nx^{n-1} \\ &= x \frac{d}{dx} \left(\sum_{n=0}^{+\infty} x^n \right) \\ &= x \frac{d}{dx} \left(\frac{1}{1-x} \right) \\ &= \frac{x}{(1-x)^2}. \end{aligned}$$

Exercice 12.46

Même exercice avec $\sum_{n \geq 0} n^2 x^n$.

Correction 12.47

On reprend la fonction g de la Correction 12.45.

La série entière $\sum n^2 x^n$ a pour rayon de convergence 1 et pour $x \in]-1 ; 1[$, on a

$$\begin{aligned} \sum_{n=0}^{+\infty} n^2 x^n &= x \sum_{n=0}^{+\infty} n^2 x^{n-1} \\ &= x g'(x) \\ &= x \frac{(1-x)^2 - x(2x-2)}{(1-x)^4} \\ &= x \frac{1-2x+x^2-2x^2+2x}{(1-x)^4} \\ &= \frac{-x(x^2-1)}{(1-x)^4} \\ &= \frac{x(x+1)}{(1-x)^3}. \end{aligned}$$

Exercice 12.48

Même exercice avec $\sum_{n \geq 1} \frac{1}{n} x^n$.

Correction 12.49

La série entière $\sum \frac{1}{n} x^n$ a pour rayon de convergence 1.

La fonction $f : x \mapsto \sum_{n=1}^{+\infty} \frac{x^n}{n}$ est de classe \mathcal{C}^1 sur $] -1 ; 1[$ et pour $x \in] -1 ; 1[$, on a

$$f'(x) = \sum_{n=1}^{+\infty} x^{n-1} = \frac{1}{1-x}.$$

On remarque $f(0) = 0$.

On a donc $\forall x \in] -1 ; 1[$, $f(x) = -\ln(1-x)$.

On pourra retenir :

$$\forall x \in] -1 ; 1[, \sum_{n=1}^{+\infty} \frac{x^n}{n} = -\ln(1-x).$$

12.2.3 Convergence radiale

Il existe plusieurs types de résultats concernant ce qui se passe au bord du disque ouvert de convergence. Le programme ne cite que le théorème de convergence radiale d'Abel.

Théorème 12.50

Soit $\sum_{n \geq 0} a_n z^n$ une série entière de rayon de convergence $R > 0$. On note f sa somme, définie sur $[0 ; R[$.

Si la série $\sum a_n R^n$ converge, alors f est définie sur $[0 ; R]$ et est continue à gauche en R , donc sur $[0 ; R]$ tout entier :

$$\lim_{x \rightarrow R^-} \sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} a_n R^n.$$

La démonstration de ce théorème est hors-programme. Il y a deux cas particuliers faciles à démontrer : si les coefficients a_n sont tous positifs ou quand la série $\sum a_n R^n$ converge absolument. Dans le cas général, c'est plus difficile à justifier.

Démonstration 12.51 (Dans le cas où $\sum a_n R^n$ converge absolument)
 Pour $x \in [0 ; R]$, on a $|a_n x^n| \leq |a_n| R^n = |a_n R^n|$.

Donc $\|x \mapsto a_n x^n\|_{\infty}^{[0;R]} \leq |a_n R^n|$.

Par théorème de comparaison des séries à termes positifs, $\sum (x \mapsto a_n x^n)$ converge normalement et donc uniformément sur $[0 ; R]$.

NB : le cas $(a_n) \geq 0$ est inclus dans celui-ci. ■

Remarque 12.52

Le théorème reste valable si $\sum a_n (-R)^n$ converge : dans ce cas, f est continue sur $[-R ; 0]$.

12.3 Fonction développable en série entière

12.3.1 Généralités

Définition 12.53

Soient $f : \mathbb{R} \longrightarrow \mathbb{C}$ et $r > 0$.

On dit que f est développable en série entière sur $] -r ; r[$ quand il existe une suite complexe (a_n) telle que la série entière $\sum_{n \geq 0} a_n z^n$ soit de rayon de convergence $R \geq r$ et pour tout $x \in] -r ; r[$, $f(x) =$

$$\sum_{n=0}^{+\infty} a_n x^n.$$

La somme $\sum_{n=0}^{+\infty} a_n x^n$ est appelée le développement en série entière de f .

On dispose d'un théorème d'opérations sur les fonctions développables en série entière.

Théorème 12.54

Soient f, g deux fonctions développables en série entière sur $] -r ; r[$, où $r > 0$.

- ▷ Les fonctions $f + g$ et fg sont développables en série entière sur $] -r ; r[$.
- ▷ Si $f(0) = 0$, alors $g \circ f$ est développable en série entière sur un intervalle $] -r' ; r'[$, où $r' > 0$.
- ▷ Si $f(0) \neq 0$, alors $\frac{1}{f}$ est développable en série entière sur un intervalle $] -r' ; r'[$ où $r' > 0$.

Exemple 12.55

Les fonctions rationnelles qui n'ont pas 0 comme pôle sont développables en série entière.

Correction 12.56 (Prolongement de l'Exercice 12.48)

On a

$$\forall x \in]-1 ; 1[, \sum_{n=1}^{+\infty} \frac{x^n}{n} = -\ln(1-x)$$

donc

$$\forall x \in]-1 ; 1[, \sum_{n=1}^{+\infty} \frac{(-x)^n}{n} = -\ln(1+x).$$

Donc

$$\forall x \in]-1 ; 1[, \ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1} x^n}{n}.$$

Or la série harmonique alternée $\sum \frac{(-1)^n}{n}$ converge (CSSA) donc d'après le théorème de convergence radiale, la fonction $x \mapsto \sum_{n=1}^{+\infty} \frac{(-1)^{n-1} x^n}{n}$ est continue sur $]-1 ; 1]$.

Or $x \mapsto \ln(1+x)$ est aussi continue sur $]-1 ; 1]$.

On a donc ce résultat à connaître :

$$\forall x \in]-1 ; 1] , \ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1} x^n}{n}.$$

En particulier, on a $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = -\ln 2$.

Exercice 12.57

Développez en série entière la fonction $x \mapsto \frac{1}{x^2 - 3x + 2}$ et précisez le rayon de convergence.

Correction 12.58

En effectuant la décomposition en éléments simples de la fraction $\frac{1}{X^2 - 3X + 2}$, on obtient

$$\frac{1}{X^2 - 3X + 2} = \frac{1}{1-X} - \frac{1}{2} \frac{1}{1-X/2}.$$

Sur $]-1 ; 1[$, les fonctions $x \mapsto \frac{1}{1-x}$ et $x \mapsto \frac{1}{1-x/2}$ sont développables en séries entières :

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \quad \text{et} \quad \frac{1}{1-x/2} = \sum_{n=0}^{+\infty} \frac{1}{2^n} x^n.$$

Ainsi, pour $x \in]-1 ; 1[$, on a

$$\frac{1}{x^2 - 3x + 2} = \sum_{n=0}^{+\infty} \left(1 - \frac{1}{2^{n+1}}\right) x^n.$$

Exercice 12.59

Même exercice avec la fonction $x \mapsto \frac{1}{1+x+x^2}$.

Correction 12.60

En effectuant une décomposition en éléments simples, on obtient

$$\frac{1}{1+X+X^2} = \frac{1}{j-j^2} \left(\frac{1}{j^2-X} - \frac{1}{j-X} \right).$$

De plus, on a

$$j - j^2 = j(1-j) = e^{2i\pi/3} (1 - e^{2i\pi/3}) = e^{2i\pi/3} \times e^{i\pi/3} (e^{-i\pi/3} - e^{i\pi/3}) = - (e^{-i\pi/3} - e^{i\pi/3}) = 2i \sin \frac{\pi}{3}.$$

On a donc, pour $x \in]-1 ; 1[$:

$$\begin{aligned} \frac{1}{1+x+x^2} &= \frac{1}{j-j^2} \sum_{n=0}^{+\infty} \left(\frac{1}{j^{2n+2}} - \frac{1}{j^{n+1}} \right) x^n \\ &= \sum_{n=0}^{+\infty} \frac{(-1)^n \sin((n+1)\pi/3)}{\sin \pi/3} x^n. \end{aligned}$$

Exercice 12.61

Même exercice avec la fonction $f : x \mapsto \ln(1+x+x^2)$.

Correction 12.62

En effectuant une décomposition en éléments simples, on obtient

$$\frac{2X+1}{1+X+X^2} = \frac{(2j+1)/(j^2-j)}{j-X} + \frac{(2j+1)/(j-j^2)}{j^2-X}.$$

Pour $x \in]-1 ; 1[$, on a donc

$$\begin{aligned} f'(x) &= \frac{1}{j-j^2} \left(\frac{2j^2+1}{j^2} \sum_{n=0}^{+\infty} \left(\frac{x}{j^2} \right)^n - \frac{2j+1}{j} \sum_{n=0}^{+\infty} \left(\frac{x}{j} \right)^n \right) \\ &= \sum_{n=0}^{+\infty} \underbrace{\left(\frac{2j^2+1}{1-j} \frac{1}{j^{2n}} + \frac{2j+1}{1-j^2} \frac{1}{j^n} \right)}_{a_n} x^n. \end{aligned}$$

Donc

$$\begin{aligned} f(x) &= \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1} + f(0) \\ &= \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}. \end{aligned}$$

Si une fonction f est développable en série entière sur un intervalle $]-r ; r[$ (où $r > 0$), alors, d'après ce qui précède, f est de classe \mathcal{C}^∞ sur $]-r ; r[$ et pour tout $x \in]-r ; r[$, $f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$.

La réciproque est fautive : une fonction de classe \mathcal{C}^∞ sur un intervalle $]-r ; r[$ n'est pas forcément développable en série entière.

De même, si une fonction est développable en série entière sur un intervalle $]-r ; r[$ (où $r > 0$), alors elle possède un développement limité à tout ordre en 0, obtenu en tronquant le développement en série entière à l'ordre voulu.

Là encore, la réciproque est fautive.

Démonstration 12.63

Soient f une fonction développable en série entière sur $]-r ; r[$ et $N \in \mathbb{N}$.

Montrons que f admet un développement limité en 0 à l'ordre N .

Pour $x \in]-r ; r[$, on a

$$\begin{aligned} f(x) &= \sum_{n=0}^{+\infty} a_n x^n \\ &= \sum_{n=0}^N a_n x^n + \sum_{n=N+1}^{+\infty} a_n x^n \\ &= \sum_{n=0}^N a_n x^n + x^N \underbrace{\sum_{n=N+1}^{+\infty} a_n x^{n-N}}_{\xrightarrow{x \rightarrow 0} 0}. \end{aligned}$$

D'où le développement limité. ■

12.3.2 Unicité du développement en série entière

Proposition 12.64

Soient (a_n) une suite de complexes et $r > 0$.

Si pour tout $x \in]-r ; r[$, $\sum_{n=0}^{+\infty} a_n x^n = 0$, alors $(a_n) = 0$.

La fonction $x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est nulle donc ses dérivées sont nulles et donc

$$\forall n \in \mathbb{N}, \quad a_n = \frac{f^{(n)}(0)}{n!} = 0.$$

■

Le développement en série entière d'une fonction, s'il existe, est unique. Autrement dit, on peut identifier les coefficients de deux développements en série entière égaux.

Corollaire 12.66

Soient $(a_n), (b_n)$ deux suites de complexes.

Si les deux séries entières $\sum_{n \geq 0} a_n z^n$ et $\sum_{n \geq 0} b_n z^n$ ont un rayon de convergence au moins égal à $r > 0$ et si pour tout $x \in]-r ; r[$, $\sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} b_n x^n$, alors $(a_n) = (b_n)$.

Corollaire 12.67

Soit f une fonction développable en série entière sur $]-r ; r[$, où $r > 0$.

Si f est paire, alors les coefficients d'indices impairs du développement en série entière de f sont nuls.

Si f est impaire, alors les coefficients d'indices pairs du développement en série entière de f sont nuls.

Exercice 12.68

On admet momentanément que $f : x \mapsto \sqrt{1+x}$ est développable en série entière sur $]-1 ; 1[$ et on note alors $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ pour $x \in]-1 ; 1[$.

Déterminez une équation différentielle linéaire homogène du premier ordre dont f est solution.

Déduisez-en que pour tout $n \in \mathbb{N}$, $a_{n+1} = \frac{2n+1}{2n+2} a_n$, puis une expression de a_n en fonction de n .

12.3.3 Série de Taylor d'une fonction

Définition 12.69

Soient $r > 0$ et f une fonction de classe \mathcal{C}^∞ sur $] -r ; r[$.

On appelle série de Taylor de f la série $\sum_{n \geq 0} \frac{f^{(n)}(0)}{n!} x^n$.

Autrement dit, f est développable en série entière ssi f est égale à la somme de sa série de Taylor.

Proposition 12.70

Soient $r > 0$ et f une fonction de classe \mathcal{C}^∞ au voisinage de 0.

f est égale à sa série de Taylor sur $] -r ; r[$ ssi $\forall x \in] -r ; r[$, $\lim_{n \rightarrow +\infty} \int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt = 0$.

La démonstration est élémentaire et repose sur l'égalité de Taylor avec reste intégral.

Comme l'inégalité de Taylor-Lagrange découle de cette égalité, on a une condition suffisante pour qu'une fonction soit développable en série entière.

Proposition 12.71

Soient $r > 0$ et f une fonction de classe \mathcal{C}^∞ sur $] -r ; r[$.

Si $\lim_{n \rightarrow +\infty} \frac{r^n}{n!} \sup_{[-r;r]} |f^{(n)}| = 0$, alors f est développable en série entière sur $] -r ; r[$.

Démonstration 12.72

D'après l'inégalité de Taylor-Lagrange, on a

$$\begin{aligned} \left| f(x) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k \right| &\leq \frac{|x|^{n+1}}{(n+1)!} \sup_{[-x;x]} |f^{(n+1)}| \\ &\leq \frac{r^{n+1}}{(n+1)!} \sup_{[-r;r]} |f^{(n+1)}|. \end{aligned}$$

Il s'ensuit le résultat énoncé. ■

Exercice 12.73

Soit $a \in]0 ; 1[$. On pose $f : x \mapsto (1+x)^a$.

Montrez que f est développable en série entière sur $\left] -\frac{1}{2} ; \frac{1}{2} \right[$.

Correction 12.74

On a $f(x) = (1+x)^a = e^{a \ln(1+x)}$ et f est de classe \mathcal{C}^∞ sur $] -1 ; 1[$.

Pour $x \in] -1 ; 1[$, on a

$$f'(x) = a(1+x)^{a-1} \quad \text{et} \quad f''(x) = a(a-1)(1+x)^{a-2}.$$

Par récurrence immédiate, on a donc

$$\forall k \in \mathbb{N}, \quad f^{(k)}(x) = a(a-1)\dots(a-k+1)(1+x)^{a-k}.$$

Soient $r \in]0 ; 1[$ et $x \in [-r ; r]$.

$$\text{On a } \left| f^{(n)}(x) \right| \leq |a(a-1)\dots(a-n+1)(1+x)^{a-n}|.$$

Pour $n > a$, on a $a-n < 0$ donc

$$\begin{aligned} 1-r &\leq 1+x \leq 1+r \\ \ln(1-r) &\leq \ln(1+x) \leq \ln(1+r) \\ (a-n)\ln(1+r) &\leq (a-n)\ln(1+x) \leq (a-n)\ln(1-r). \end{aligned}$$

$$\text{Donc } (1+x)^{a-n} \leq (1-r)^{a-n}.$$

$$\text{Donc } \sup_{[-r;r]} \left| f^{(n)} \right| \leq K_n (1-r)^{a-n} \text{ où } K_n = |a(a-1)\dots(a-n+1)|.$$

$$\text{Donc } \frac{r^n}{n!} \sup_{[-r;r]} \left| f^{(n)} \right| \leq \frac{K_n}{n!} r^n (1-r)^{a-n}.$$

$$\text{Or } \frac{K_n}{n!} = \frac{|a||1-a|\dots|n-a+1|}{1 \times 2 \times \dots \times n} \text{ donc}$$

$$\begin{aligned} \ln \frac{K_n}{n!} &= \sum_{k=0}^{n-1} \ln |a-k| - \sum_{k=0}^{n-1} \ln (k+1) \\ &= \sum_{k=0}^{n-1} \ln |1+a-(k+1)| - \sum_{k=0}^{n-1} \ln (k+1) \\ &= \sum_{k=0}^{n-1} \ln \left| 1 - \frac{1-a}{k+1} \right| \\ &\leq \sum_{k=0}^{n-1} \frac{|1+a|}{k+1} \\ &= |1+a| \underbrace{\sum_{k=1}^n \frac{1}{k}}_{\sim \ln n \text{ quand } n \rightarrow +\infty}. \end{aligned}$$

Donc, à partir d'un certain rang, $\ln \frac{K_n}{n!} \leq 2|1+a| \ln n$.

Donc

$$\begin{aligned} \frac{r^n}{n!} \sup_{[-r;r]} |f^{(n)}| &\leq \frac{K_n}{n!} r^n (1-r)^{a-n} \\ &\leq Ln (1-r)^a \left(\frac{r}{1-r} \right)^n \end{aligned}$$

où L est une constante.

Or $\frac{r}{1-r} < 1$ quand $r < \frac{1}{2}$.

Donc si $r < \frac{1}{2}$, on a $n \left(\frac{r}{1-r} \right)^n \xrightarrow{n \rightarrow +\infty} 0$.

D'où

$$r < \frac{1}{2} \implies \left[\forall x \in [-r; r], (1+x)^a = \sum_{n=0}^{+\infty} \frac{a(a-1)\dots(a-n+1)}{n!} x^n \right].$$

12.3.4 Développements en série entière usuels

Soient a un réel et n un entier naturel.

On pose $\binom{a}{n} = \frac{a(a-1)(a-2)\dots(a-n+1)}{n!}$.

Les fonctions usuelles suivantes sont développables en série entière avec un rayon de convergence $R > 0$:

- ▷ $e^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$ avec $R = +\infty$
- ▷ $\cos x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n}$ avec $R = +\infty$
- ▷ $\sin x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$ avec $R = +\infty$
- ▷ $\operatorname{ch} x = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^{2n}$ avec $R = +\infty$
- ▷ $\operatorname{sh} x = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)!} x^{2n+1}$ avec $R = +\infty$
- ▷ $\operatorname{Arctan} x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$ avec $R = 1$
- ▷ $\ln(1+x) = \sum_{n=0}^{+\infty} \frac{(-1)^{n-1}}{n} x^n$ avec $R = 1$
- ▷ $(1+x)^a = \sum_{n=0}^{+\infty} \binom{a}{n} x^n$ avec $R = 1$.

Exercice 12.75

Développez en série entière la fonction $x \mapsto \ln \frac{1-x}{1+x}$.

Exercice 12.76

Même exercice avec $x \mapsto (x+2) \operatorname{ch} x$.

Exercice 12.77

Même exercice avec $x \mapsto \sin^2 x$.

Exercice 12.78

Même exercice avec $x \mapsto e^x \ln(1+x)$.

Chapitre 13

Probabilités

Sommaire

13.1	Dénombrabilité	. 388
13.1.1	Vocabulaire	388
13.1.2	Exemples	389
13.1.3	Quelques propriétés	392
13.2	Espace probabilisé	. 393
13.2.1	Univers d'une expérience aléatoire	393
13.2.2	Tribu d'événements	393
13.2.3	Probabilité	395
13.2.4	Propriétés	396
13.2.4.1	Continuité	396
13.2.4.2	Sous-additivité	400
13.2.4.3	Événements négligeables ou presque sûrs	401
13.2.5	Probabilité discrète	402
13.3	Probabilités conditionnelles	. 404
13.3.1	Généralités	404
13.3.2	Systèmes complets d'événements	407
13.3.3	Formule des probabilités totales	408
13.3.4	Formule de Bayes	409
13.4	Indépendance	. 412
13.4.1	Indépendance de deux événements	412
13.4.2	Indépendance mutuelle	414

13.1 Dénombrabilité

13.1.1 Vocabulaire

Définition 13.1

On dit qu'un ensemble E est dénombrable quand il existe une bijection de \mathbb{N} dans E .

Dans ce cas, cela signifie qu'on peut numéroter les éléments de E par les entiers naturels et donc qu'on peut écrire E en extension sous une forme $E = \{x_n \mid n \in \mathbb{N}\}$ sans jamais répéter deux fois le même élément : on dit qu'on a énuméré les éléments de E .

Un ensemble fini est de la forme $\{x_0, \dots, x_n\}$: on peut aussi l'écrire sous la forme $\{x_n \mid n \in \mathbb{N}\}$ en répétant une infinité de fois au moins un élément. C'est pourquoi on voit souvent dans les raisonnements apparaître la locution « ensemble fini ou dénombrable » ou « ensemble au plus dénombrable ».

Évidemment, tout ensemble en bijection avec un ensemble dénombrable est aussi dénombrable, car une composée de bijections est une bijection.

13.1.2 Exemples

Proposition 13.2

- ▷ \mathbb{N} , \mathbb{N}^* , plus généralement $[n_0 ; +\infty[\cap \mathbb{N}$ (pour tout $n_0 \in \mathbb{N}$), et encore plus généralement toute partie infinie de \mathbb{N} sont dénombrables.
- ▷ \mathbb{Z} est dénombrable.
- ▷ Pour tout $p \in \mathbb{N}^*$, \mathbb{N}^p est dénombrable.
- ▷ \mathbb{Q} est dénombrable.
- ▷ Si $(u_i)_{i \in I}$ est une famille sommable de complexes, alors le support de la famille $\{i \in I \mid u_i \neq 0\}$ est dénombrable.

Démonstration 13.3

- ▷ Si A est une partie infinie de \mathbb{N} , on construit une énumération de A par « l'algorithme » suivant :

$$u_0 = \min A \quad \text{et} \quad \forall n \in \mathbb{N}, \quad u_{n+1} = \min (A \setminus \{u_0, \dots, u_n\}).$$

- ▷ On peut représenter \mathbb{Z} de la manière suivante, avec en bas les éléments de \mathbb{Z} et en haut les étiquettes :

$$\begin{array}{cccccccc} 6 & 4 & 2 & 0 & 1 & 3 & 5 & 7 \\ | & | & | & | & | & | & | & | \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 \end{array}$$

On en déduit que $\varphi : \mathbb{N} \longrightarrow \mathbb{Z}$ est une bijection de \mathbb{N} dans \mathbb{Z} dont la bijection

$$n \longmapsto (-1)^{n+1} \left\lfloor \frac{n+1}{2} \right\rfloor$$

réciproque est $\mathbb{Z} \longrightarrow \mathbb{N}$

$$p \longmapsto \begin{cases} -2p & \text{si } p \leq 0 \\ 2p - 1 & \text{sinon} \end{cases}$$

- ▷ On peut représenter \mathbb{N}^2 de la manière suivante :

4	(0, 4) 14	(1, 4) 19	(2, 4) 25	(3, 4) 32
3	(0, 3) 9	(1, 3) 13	(2, 3) 18	(3, 3) 24
2	(0, 2) 5	(1, 2) 8	(2, 2) 12	(3, 2) 17
1	(0, 1) 2	(1, 1) 4	(2, 1) 7	(3, 1) 11
0	(0, 0) 0	(1, 0) 1	(2, 0) 3	(3, 0) 6
$\begin{array}{c} p \\ \hline n \end{array}$	0	1	2	3

On en déduit que $\varphi : \mathbb{N}^2 \longrightarrow \mathbb{N}$ est une bijection de \mathbb{N}^2 dans \mathbb{N}
 $(n, p) \longmapsto \frac{(n+p)(n+p+1)}{2} + p$
dont la réciproque est définie comme suit.

Pour $N \in \mathbb{N}$, on pose $k = \max \left\{ i \in \mathbb{N} \left| \frac{i(i+1)}{2} \leq N \right. \right\}$ puis $p = N - k$ et $n = k - p$, et enfin $N \longmapsto (n, p)$.

Autre bijection possible : on a unicité de la décomposition $\forall n \geq 1, n = 2^a (2b+1)$ où $(a, b) \in \mathbb{N}^2$
i.e. $\psi : \mathbb{N}^* \longrightarrow \mathbb{N}^2$ est une bijection et donc $\mathbb{N} \xrightarrow{\sim} \mathbb{N}^2$.
 $n \longmapsto (a, b)$ $n \longmapsto \psi(n+1)$

► Si on considère φ_2 une bijection de \mathbb{N}^2 dans \mathbb{N} ,

$$\begin{array}{ccccccc} \mathbb{N}^3 & \xrightarrow{\sim} & \mathbb{N} \times \mathbb{N}^2 & \xrightarrow{\sim} & \mathbb{N} \times \mathbb{N} & \xrightarrow{\sim} & \mathbb{N} \\ (a, b, c) & \longmapsto & (a, (b, c)) & \longmapsto & (a, \varphi_2(b, c)) & \longmapsto & \varphi_2(a, \varphi_2(b, c)) \end{array}$$

est une bijection de \mathbb{N}^3 dans \mathbb{N} .

Par récurrence, si φ_p est une bijection de \mathbb{N}^p dans \mathbb{N} , alors

$$\begin{array}{ccc} \varphi_{p+1} : & \mathbb{N}^{p+1} & \xrightarrow{\sim} \mathbb{N} \\ & (a_1, \dots, a_{p+1}) & \longmapsto \varphi_2(a_1, \varphi_p(a_2, \dots, a_{p+1})) \end{array}$$

est une bijection de \mathbb{N}^{p+1} dans \mathbb{N} .

► Tout $x \in \mathbb{Q}^*$ s'écrit de façon unique sous forme irréductible $x = \frac{p}{q}$ où $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ tel que $p \wedge q = 1$, donc

$$\begin{array}{ccc} \mathbb{Z} \times \mathbb{N}^* & \longrightarrow & \mathbb{Q}^* \\ (p, q) & \longmapsto & \frac{p}{q} \end{array}$$

est bijective.

Donc $\mathbb{Q} \simeq \mathbb{Q}^* \simeq \mathbb{Z} \times \mathbb{N}^* \simeq \mathbb{N}^2 \simeq \mathbb{N}$ (où \simeq est la relation d'équipotence).

► Soit $(u_i)_{i \in I}$ une famille sommable de complexes.

On pose $S = \text{Supp}(u_i) = \{i \in I \mid u_i \neq 0\}$.

— Si S est fini, alors S est au plus dénombrable.

— Si S est infini, pour $k \in \mathbb{N}^*$, on pose $S_k = \left\{i \in I \mid |u_i| > \frac{1}{k}\right\}$ et on a $S = \bigcup_{k \in \mathbb{N}^*} S_k$.

Montrons que pour tout $k \in \mathbb{N}^*$, S_k est fini.

Soient $k \in \mathbb{N}^*$ et $(i_1, \dots, i_p) \in S_k^p$ distincts.

Pour $j \in \llbracket 1 ; p \rrbracket$, on a $|u_{i_j}| \geq \frac{1}{k}$ donc

$$\sum_{j=1}^p |u_{i_j}| \geq \frac{p}{k}.$$

Or $(u_i)_{i \in I}$ est sommable donc il existe $M > 0$ tel que pour tout $J \in \mathcal{P}_f(I)$, $\sum_{j \in J} |u_j| \leq M$.

On a donc $\frac{p}{k} \leq \sum_{j=1}^p |u_{i_j}| \leq M$, et donc $p \leq Mk$.

Si S_k est infini, on peut choisir $\lfloor Mk \rfloor + 1$ éléments distincts de S_k donc $\lfloor Mk \rfloor + 1 \leq Mk$ d'après ce qui précède : contradiction.

Donc S_k est fini.

On peut donc numéroter les éléments de S_k via $\varphi_k : S_k \longrightarrow \llbracket 0 ; N_k \rrbracket$.

Pour $i \in S$, on pose $m_i = \min \{k \in \mathbb{N}^* \mid i \in S_k\}$ et à i on peut associer le couple $(m_i, \varphi_{m_i}(i))$.

On construit ainsi une injection de S dans \mathbb{N}^2 , donc par composition avec une bijection de \mathbb{N}^2 dans \mathbb{N} , une injection de S dans \mathbb{N} , *i.e.* S est en bijection avec une partie infinie de \mathbb{N} .

Donc S est en bijection avec \mathbb{N} . ■

En revanche, il existe des ensembles infinis non-dénombrables, comme \mathbb{R} ou tout intervalle de longueur non-nulle. Un ensemble non-dénombrable est trop gros pour qu'on puisse ordonner ses éléments et les numéroter.

Démonstration 13.4 (\mathbb{R} n'est pas dénombrable)

Tout réel x peut s'écrire de manière unique sous la forme

$$x = p + \sum_{k=0}^{+\infty} \frac{c_k}{10^k}$$

où $p = \lfloor x \rfloor \in \mathbb{Z}$ et $(c_k) \in \llbracket 0 ; 9 \rrbracket^{\mathbb{N}}$ n'est pas stationnaire en 9. On peut écrire $x = p, c_1 c_2 \dots c_n \dots$

Par l'absurde, si $[0 ; 1[$ est dénombrable, on peut numéroter les réels de $[0 ; 1[: x_0, x_1, \dots, x_n, \dots$

On a

$$x_0 = 0, c_{0,1}c_{0,2}c_{0,3} \dots$$

$$x_1 = 0, c_{1,1}c_{1,2}c_{1,3} \dots$$

$$\vdots$$

$$x_n = 0, c_{n,1}c_{n,2}c_{n,3} \dots c_{n,n+1}$$

$$\vdots$$

On choisit $d_1 \in \llbracket 0 ; 8 \rrbracket$ tel que $d_1 \neq c_{0,1}$, $d_2 \in \llbracket 0 ; 8 \rrbracket$ tel que $d_2 \neq c_{1,2}$, etc., $d_n \in \llbracket 0 ; 8 \rrbracket$ tel que $d_n \neq c_{n,n+1}$ puis on pose

$$y = 0, d_1 d_2 \dots d_n \dots = \sum_{k=1}^{+\infty} \frac{d_k}{10^k}.$$

Donc $y \in [0 ; 1[$.

Or, par construction, $y \notin \{x_0, \dots, x_n, \dots\} = [0 ; 1[$: contradiction. ■

13.1.3 Quelques propriétés

Proposition 13.5

Toute partie infinie d'un ensemble dénombrable est dénombrable.

Si E est dénombrable, alors pour toute injection de F dans E , F est au plus dénombrable.

Pourvu qu'on ne considère pas trop d'ensembles, les réunions d'ensembles dénombrables le sont aussi.

Proposition 13.6

Si E_1, \dots, E_n sont (au plus) dénombrables, alors $E_1 \cup \dots \cup E_n$ l'est aussi.

Si $(E_i)_{i \in I}$ est une famille (au plus) dénombrable d'ensembles (au plus) dénombrables, alors $\bigcup_{i \in I} E_i$ l'est aussi.

En revanche, pour les produits cartésiens, il faut se contenter d'un nombre fini d'ensembles.

Proposition 13.7

Si E_1, \dots, E_n sont dénombrables, alors $E_1 \times \dots \times E_n$ est dénombrable.

En revanche, un produit cartésien quelconque d'ensembles dénombrables ne l'est pas en général : par exemple, $\{0, 1\}^{\mathbb{N}}$ n'est pas dénombrable.

13.2 Espace probabilisé

13.2.1 Univers d'une expérience aléatoire

Une expérience aléatoire est une expérience dont on connaît les résultats possibles (les issues possibles) mais dont on ne peut pas connaître à l'avance le résultat. On modélise l'expérience par la donnée de l'ensemble Ω des issues possibles.

Définition 13.8

L'ensemble des issues possibles est appelé univers des possibles (ou univers). Il est souvent noté Ω .

Exemple 13.9

- ▶ On jette un dé non-truqué : les issues possibles sont les six entiers $1, \dots, 6$; l'univers est donc $\Omega = \{1, \dots, 6\}$.
- ▶ On jette n fois un dé : les issues possibles sont les suites de n entiers de l'ensemble $\{1, \dots, 6\}$; donc $\Omega = \llbracket 1 ; 6 \rrbracket^n$.
- ▶ On lance une pièce une infinité de fois : les issues possibles sont les suites infinies de 0 ou 1 ; donc $\Omega = \{0, 1\}^{\mathbb{N}}$.

13.2.2 Tribu d'événements

De manière informelle, un événement est une partie de l'univers Ω . Mais cette définition est trop vague. Si on autorise toutes les parties de l'univers à être des événements, alors quand l'ensemble Ω est infini non-dénombrable, les seules probabilités sur Ω sont des probabilités discrètes (théorème d'Ulam), ce qui exclut tout un tas de probabilités intéressantes. Donc on doit en général restreindre la notion d'événement à certaines parties de Ω : certaines parties n'ont donc pas le droit d'être nommées « événement ».

Définition 13.10

Soit Ω un univers.

On appelle tribu sur Ω une partie \mathcal{T} de $\mathcal{P}(\Omega)$ telle que :

- ▶ $\Omega \in \mathcal{T}$
- ▶ pour tout $A \in \mathcal{T}$, $\bar{A} \in \mathcal{T}$
- ▶ pour toute suite $(A_n) \in \mathcal{T}^{\mathbb{N}}$, $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{T}$.

Modéliser une expérience aléatoire, c'est choisir l'univers et une tribu : les éléments de la tribu sont appelés les événements. On dit qu'un événement est réalisé quand l'issue de l'expérience aléatoire appartient à cet événement. Le couple (Ω, \mathcal{T}) est appelé espace probabilisable.

Exemple 13.11

- ▶ L'ensemble $\{\emptyset, \Omega\}$ est une tribu, c'est la plus petite tribu envisageable. Elle est inutilisable en pratique car elle ne comporte pas assez d'événements pour décrire des situations issues de la vie réelle.
- ▶ L'ensemble $\mathcal{P}(\Omega)$ est une tribu : quand l'univers Ω est fini ou dénombrable, c'est la tribu utilisée systématiquement, mais quand Ω est infini non-dénombrable, c'est une tribu trop grosse pour pouvoir y définir une probabilité vraiment utile.
- ▶ En école d'ingénieur, vous entendrez parler de la tribu des boréliens : c'est celle qui est couramment utilisée quand $\Omega = \mathbb{R}$.

Proposition 13.12

Soient Ω un ensemble et \mathcal{T} une tribu sur Ω .

Alors on a :

- ▶ $\emptyset \in \mathcal{T}$
- ▶ pour toute suite $(A_n) \in \mathcal{T}^{\mathbb{N}}$, $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{T}$
- ▶ pour toute famille finie (A_1, \dots, A_n) d'événements, $\bigcup_{k=0}^n A_k$ et $\bigcap_{k=0}^n A_k$ sont des événements.

Démonstration 13.13

Soit $(A_n) \in \mathcal{T}^{\mathbb{N}}$.

On a $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{T}$ car

$$\bigcap_{n \in \mathbb{N}} A_n = \overline{\bigcup_{n \in \mathbb{N}} \overline{A_n}}.$$

■

Définition 13.14

Ω est appelé l'événement certain, \emptyset est appelé l'événement impossible.

Deux événements A et B sont dits incompatibles quand ils sont disjoints, *i.e.* $A \cap B = \emptyset$.

13.2.3 Probabilité

Définition 13.15

Soient Ω un univers et \mathcal{T} une tribu sur Ω .

On appelle probabilité sur (Ω, \mathcal{T}) toute application $\mathbb{P} : \mathcal{T} \longrightarrow [0 ; 1]$ telle que :

- $\mathbb{P}(\Omega) = 1$
- pour toute suite $(A_n) \in \mathcal{T}^{\mathbb{N}}$ d'événements deux à deux incompatibles, la série de terme général $\mathbb{P}(A_n)$ est convergente et $\mathbb{P}\left(\bigsqcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$.

Le triplet $(\Omega, \mathcal{T}, \mathbb{P})$ est appelé espace probabilisé.

En pratique, sauf dans certains cas où Ω est fini ou dénombrable, nous postulons toujours l'existence d'un espace probabilisé qui modélise la situation, car c'est hors de notre portée de construire concrètement un tel espace. Dès que les expériences aléatoires peuvent avoir une infinité non-dénombrable de résultats possibles, il est souvent difficile de reprendre les idées développées en première année et définir, par exemple, des probabilités à partir d'événements élémentaires $\{\omega\}$ (définir Ω est en général à notre portée, mais construire la tribu et la probabilité est inaccessible à nos moyens). Nous admettons donc toujours l'existence d'un espace probabilisé représentant notre expérience.

Dans moult cas, nous définirons les événements à partir d'événements « primitifs » : par exemple, dans le cas d'une suite de lancer de pièce, l'univers est simple : $\{0, 1\}^{\mathbb{N}}$, mais il n'est pas dénombrable ; les événements primitifs sont les événements P_i , où $i \in \mathbb{N}$ et

$$P_i = \{\omega = (\omega_n) \in \{0, 1\}^{\mathbb{N}} \mid \omega_i = 1\}$$

modélise l'événement au sens concret « obtenir pile au i -ème lancer ».

Exercice 13.16

On lance une pièce une infinité de fois.

Exprimez par une phrase ce que représentent les événements suivants :

$$A_n = \bigcap_{i=0}^n P_i \quad B_n = \bigcup_{i=0}^n \left(P_i \cap \bigcap_{\substack{0 \leq j \leq n \\ j \neq i}} \overline{P_j} \right) \quad C = \bigcup_{i \in \mathbb{N}} \bigcap_{j \geq i} P_j.$$

Réciproquement, définissez à l'aide des P_i les événements suivants :

- on obtient pile un nombre fini de fois
- on obtient pile une infinité de fois.

Correction 13.17

▷ $\omega \in A_n \iff$ « on obtient pile lors des $n + 1$ premiers lancers ».

▷

$$\begin{aligned}\omega \in B_n &\iff \exists i \in \llbracket 0 ; n \rrbracket, \omega \in P_i \cap \bigcap_{\substack{0 \leq j \leq n \\ i \neq j}} \overline{P_j} \\ &\iff \exists i \in \llbracket 0 ; n \rrbracket, [\omega \in P_i \text{ et } \forall j \in \llbracket 0 ; n \rrbracket \setminus \{i\}, \omega \in \overline{P_j}] \\ &\iff \text{« parmi les } n + 1 \text{ premiers lancers, un seul donne pile ».}\end{aligned}$$

▷ $\omega \in C \iff$ « à partir d'un certain rang, on obtient uniquement pile ».

▷ « On obtient pile un nombre fini de fois » $\iff \omega \in \bigcup_{i \in \mathbb{N}} \bigcap_{j \geq i} \overline{P_j} = D$.

▷ « On obtient pile une infinité de fois » $\iff \omega \in \overline{D} = \bigcap_{i \in \mathbb{N}} \bigcup_{j \geq i} P_j$.

13.2.4 Propriétés

On retrouve les propriétés vues en première année.

Proposition 13.18

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

Alors \mathbb{P} vérifie les propriétés suivantes :

▷ pour tout $A \in \mathcal{F}$, $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$; en particulier $\mathbb{P}(\emptyset) = 0$

▷ pour tout $(A, B) \in \mathcal{F}^2$, $A \subseteq B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$

▷ pour tout $(A, B) \in \mathcal{F}^2$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Mais on en a d'autres, liées à la notion de suite dénombrable.

13.2.4.1 Continuité

Proposition 13.19 (Continuité croissante)

Si (A_n) est une suite croissante d'événements, c'est-à-dire pour tout $n \in \mathbb{N}$, $A_n \subseteq A_{n+1}$, alors

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n).$$

Démonstration 13.20

Soit (A_n) une suite croissante d'événements.

On pose $B_0 = A_0$ et pour tout $n \in \mathbb{N}$, $B_n = A_n \setminus A_{n-1}$.

On vérifie d'abord que $\bigcup_{i \in \mathbb{N}} B_i = \bigcup_{i \in \mathbb{N}} A_i$:

► Pour tout $n \in \mathbb{N}$, $B_n \subseteq A_n$ donc $\bigcup_{n \in \mathbb{N}} B_n \subseteq \bigcup_{n \in \mathbb{N}} A_n$.

► Soit $\omega \in \bigcup_{i \in \mathbb{N}} A_i$.

Alors $\{k \in \mathbb{N} \mid \omega \in A_k\}$ est une partie non-vide de \mathbb{N} et possède donc un minimum $i \in \mathbb{N}$.

Donc, si $i \geq 1$, alors $i - 1$ n'appartient pas à cet ensemble.

Ainsi, ou bien $i = 0$ et $\omega \in B_0$, ou bien $i \neq 0$ et $\omega \in A_i$ et $\omega \notin A_{i-1}$ donc $\omega \in B_i$.

Donc $\omega \in \bigcup_{n \in \mathbb{N}} B_n$ et donc $\bigcup_{n \in \mathbb{N}} A_n \subseteq \bigcup_{n \in \mathbb{N}} B_n$.

De plus, les événements B_i sont deux à deux incompatibles *i.e.*

$$\forall (i, j) \in \mathbb{N}^2, \quad i \neq j \implies B_i \cap B_j = \emptyset.$$

En effet, soit $(i, j) \in \mathbb{N}^2$ tel que $i > j$.

Si $\omega \in B_i \cap B_j$, alors $\omega \in B_j \subseteq A_j \subseteq \dots \subseteq A_{i-1}$.

Or $\omega \in B_i = A_i \setminus A_{i-1}$ donc $\omega \notin A_{i-1}$: contradiction.

Donc $B_i \cap B_j = \emptyset$.

D'après la deuxième propriété de \mathbb{P} , $\mathbb{P}\left(\bigsqcup_{i \in \mathbb{N}} B_i\right) = \sum_{i \in \mathbb{N}} \mathbb{P}(B_i)$ donc

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \mathbb{P}(B_i) = \sum_{i=0}^{+\infty} \mathbb{P}(B_i).$$

Or $B_i = A_i \setminus A_{i-1}$ et $A_{i-1} \subseteq A_i$ donc

$$A_i = A_{i-1} \sqcup B_i.$$

Donc $\mathbb{P}(B_0) = \mathbb{P}(A_0)$ et pour $i \geq 1$ on a

$$\mathbb{P}(A_i) = \mathbb{P}(A_{i-1}) + \mathbb{P}(B_i)$$

$$\mathbb{P}(A_i) - \mathbb{P}(A_{i-1}) = \mathbb{P}(B_i).$$

Donc

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) &= \lim_{N \rightarrow +\infty} \sum_{i=0}^N \mathbb{P}(B_i) \\
&= \lim_{N \rightarrow +\infty} \left[\sum_{i=1}^N (\mathbb{P}(A_i) - \mathbb{P}(A_{i-1})) + \mathbb{P}(A_0) \right] \\
&= \lim_{N \rightarrow +\infty} (\mathbb{P}(A_0) + \mathbb{P}(A_N) - \mathbb{P}(A_0)) \\
&= \lim_{N \rightarrow +\infty} \mathbb{P}(A_N).
\end{aligned}$$

■

Proposition 13.21 (Continuité décroissante)

Si (A_n) est une suite décroissante d'événements, c'est-à-dire pour tout $n \in \mathbb{N}$, $A_{n+1} \subseteq A_n$, alors

$$\mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n).$$

Démonstration 13.22

Soit (A_n) une suite décroissante d'événements.

Alors $(\overline{A_n})$ est croissante.

$$\text{Donc } \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} \overline{A_n}\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(\overline{A_n}).$$

$$\text{Donc } 1 - \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} \overline{A_n}\right) = \lim_{n \rightarrow +\infty} (1 - \mathbb{P}(\overline{A_n})).$$

$$\text{Donc } \mathbb{P}\left(\overline{\bigcup_{n \in \mathbb{N}} \overline{A_n}}\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n).$$

$$\text{Donc } \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n).$$

■

Proposition 13.23

Soit (A_n) une suite quelconque d'événements. On a

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{N \rightarrow +\infty} \mathbb{P}\left(\bigcup_{n=0}^N A_n\right) \quad \text{et} \quad \mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{N \rightarrow +\infty} \mathbb{P}\left(\bigcap_{n=0}^N A_n\right).$$

Démonstration 13.24

Pour $n \in \mathbb{N}$, on pose $A'_n = \bigcup_{k=0}^n A_k$.

La suite (A'_n) est croissante et $\bigcup_{n \in \mathbb{N}} A'_n = \bigcup_{n \in \mathbb{N}} A_n$.

D'après la proposition de continuité croissante, on a donc

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A'_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A'_n).$$

Idem pour l'intersection avec la continuité décroissante. ■

Exercice 13.25

On considère l'univers des suites infinies de lancers indépendants d'une pièce équilibrée : $\Omega = \{0, 1\}^{\mathbb{N}}$.

Montrez que les événements élémentaires $\{\omega\}$ sont de probabilité nulle.

Déduisez-en que l'univers n'est pas dénombrable.

Quelle est la probabilité d'obtenir un nombre fini de face ?

Correction 13.26

► On considère par exemple l'événement « on n'obtient que des piles » $= A$.

$A = \bigcap_{i \in \mathbb{N}} P_i = \{(1, 1, 1, \dots, 1, \dots)\}$ est un événement élémentaire.

D'après la Proposition 13.23, on a $\mathbb{P}(A) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcap_{i=0}^n P_i\right)$.

Or les événements P_i sont indépendants et $\mathbb{P}(P_i) = \frac{1}{2}$ donc

$$\mathbb{P}\left(\bigcap_{i=0}^n P_i\right) = \prod_{i=0}^n \mathbb{P}(P_i) = \left(\frac{1}{2}\right)^{n+1}.$$

Donc $\mathbb{P}(A) = \lim_{n \rightarrow +\infty} \frac{1}{2^{n+1}} = 0$.

► Soient $\omega \in \Omega$ et $A = \{\omega\} = \{(\omega_0, \omega_1, \dots, \omega_n, \dots)\} = \bigcap_{i \in \mathbb{N}} B_i$ où $B_i = \begin{cases} P_i & \text{si } \omega_i = 1 \\ \overline{P_i} & \text{si } \omega_i = 0 \end{cases}$

D'après la Proposition 13.23, on a

$$\mathbb{P}(A) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcap_{i=0}^n B_i\right) = \lim_{n \rightarrow +\infty} \frac{1}{2^{n+1}} = 0.$$

► Supposons $\Omega = \{0, 1\}^{\mathbb{N}}$ dénombrable.

On a $\Omega = \bigsqcup_{\omega \in \Omega} \{\omega\}$.

L'union étant dénombrable, on a

$$1 = \mathbb{P}(\Omega) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 0$$

contradiction.

Donc $\{0, 1\}^{\mathbb{N}}$ n'est pas dénombrable.

► Soit B l'événement « on obtient un nombre fini de face ». On a

$$B = \bigcup_{i \in \mathbb{N}} \underbrace{\bigcap_{j \geq i} P_j}_{Q_i}.$$

Q_i est l'événement « on obtient que des pile à partir du rang i ».

Si $\omega \in Q_i$, $\forall j \geq i$, $\omega_j = 1$ donc $\forall j \geq i+1$, $\omega_j = 1$, donc $\omega \in Q_{i+1}$ i.e. $Q_i \subseteq Q_{i+1}$.

Donc (Q_i) est croissante et on a $\mathbb{P}(B) = \lim_{i \rightarrow +\infty} \mathbb{P}(Q_i)$.

Or $Q_i = \bigcap_{j \geq i} P_j$ donc

$$\begin{aligned} \mathbb{P}(Q_i) &= \lim_{N \rightarrow +\infty} \mathbb{P}\left(\bigcap_{j=i}^N P_j\right) \\ &= \lim_{N \rightarrow +\infty} \left(\frac{1}{2}\right)^{N-i+1} \\ &= 0. \end{aligned}$$

Donc $\mathbb{P}(B) = 0$.

13.2.4.2 Sous-additivité

Proposition 13.27 (Sous-additivité)

Si (A_n) est une suite d'événements, alors

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

Démonstration 13.28

D'après la Proposition 13.18, on a

$$\forall (A, B) \in \mathcal{T}^2, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \leq \mathbb{P}(A) + \mathbb{P}(B).$$

Puis, par récurrence, on a

$$\forall n \in \mathbb{N}^*, \forall (A_1, \dots, A_n) \in \mathcal{T}^n, \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i).$$

Donc

$$\forall (A_n) \in \mathcal{T}^{\mathbb{N}}, \forall k \in \mathbb{N}, \mathbb{P}\left(\bigcup_{i=0}^k A_i\right) \leq \sum_{i=0}^k \mathbb{P}(A_i) \leq \sum_{i=0}^{+\infty} \mathbb{P}(A_i).$$

Or, d'après la Proposition 13.23, on a $\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \lim_{k \rightarrow +\infty} \mathbb{P}\left(\bigcup_{i=0}^k A_i\right)$, donc par passage à la limite, on a

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i=0}^{+\infty} \mathbb{P}(A_i).$$

■

Dans ce résultat, le symbole $\sum_{n=0}^{+\infty} \mathbb{P}(A_n)$ signifie $\lim_{N \rightarrow +\infty} \sum_{n=0}^N \mathbb{P}(A_n)$: c'est un réel si la série $\sum \mathbb{P}(A_n)$ est convergente, et $+\infty$ sinon (puisque la série est à termes positifs).

13.2.4.3 Événements négligeables ou presque sûrs

Définition 13.29

Un événement est dit négligeable quand sa probabilité est nulle.

Un événement est dit presque sûr quand sa probabilité est 1.

Proposition 13.30

Toute réunion ou intersection au plus dénombrable d'événements négligeables est négligeable.

Toute réunion ou intersection au plus dénombrable d'événements presque sûrs est presque sûre.

Démonstration 13.31

Application directe de la Proposition 13.27.

■

13.2.5 Probabilité discrète

Définition 13.32

Soit Ω un univers.

On appelle distribution de probabilité discrète sur Ω toute famille de réels positifs indexée par Ω , sommable et de somme totale 1.

Si Ω est un ensemble fini, on retrouve la définition de l'an dernier.

Proposition 13.33

Si $(p_\omega)_{\omega \in \Omega}$ est une distribution de probabilité discrète sur Ω , alors son support $\{\omega \in \Omega \mid p_\omega > 0\}$ est au plus dénombrable.

À toute distribution de probabilité discrète sur Ω , on peut associer une probabilité sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$.

Proposition 13.34

Si $(p_\omega)_{\omega \in \Omega}$ est une distribution de probabilité discrète sur Ω , alors il existe une unique probabilité \mathbb{P} sur la tribu $\mathcal{P}(\Omega)$ telle que pour tout $\omega \in \Omega$, $\mathbb{P}(\{\omega\}) = p_\omega$.

Démonstration 13.35

Soit $(p_\omega)_{\omega \in \Omega}$ une distribution de probabilité discrète.

Pour $A \in \mathcal{P}(\Omega)$, on pose $\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$ (au sens des familles sommables).

On a $A \subseteq \Omega$ donc $0 \leq \mathbb{P}(A) \leq \mathbb{P}(\Omega) = \sum_{\omega \in \Omega} p_\omega = 1$.

Donc \mathbb{P} est une application de $\mathcal{P}(\Omega)$ dans $[0 ; 1]$ qui vérifie $\mathbb{P}(\Omega) = 1$.

Soit $(A_n)_{n \in \mathbb{N}}$ une famille dénombrable de $\mathcal{P}(\Omega)$ telle que pour tout $(m, n) \in \mathbb{N}^2$, $m \neq n \implies A_m \cap A_n = \emptyset$.

On veut montrer $\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$.

D'après le théorème de sommation par paquets, on a

$$\begin{aligned} \sum_{n \in \mathbb{N}} \mathbb{P}(A_n) &= \sum_{n \in \mathbb{N}} \sum_{\omega \in A_n} p_\omega \\ &= \sum_{\omega \in \bigcup_{n \in \mathbb{N}} A_n} p_\omega \\ &= \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right). \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} (A_n) \text{ partition de } \bigcup_{n \in \mathbb{N}} A_n$$

\mathbb{P} est donc une probabilité sur $\mathcal{P}(\Omega)$ telle que pour tout $\omega \in \Omega$, $\mathbb{P}(\{\omega\}) = p_\omega$.

Ceci prouve donc l'existence de \mathbb{P} .

Pour montrer l'unicité, soit Q une probabilité qui vérifie les conditions de l'énoncé.

On pose S le support $Q : S = \{\omega \in \Omega \mid p_\omega > 0\}$.

Soit $A \in \mathcal{P}(\Omega)$.

On a $A = (A \cap S) \sqcup (A \cap \bar{S})$ donc

$$\begin{aligned} Q(A) &= Q(A \cap S) + Q(A \cap \bar{S}) \\ &= Q\left(\bigsqcup_{\omega \in A \cap S} \{\omega\}\right) + Q(A \cap \bar{S}) \\ &= \sum_{\omega \in A \cap S} Q(\{\omega\}) + Q(A \cap \bar{S}) \\ &= \sum_{\omega \in A \cap S} p_\omega + Q(A \cap \bar{S}). \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} A \cap S \text{ dénombrable}$$

En particulier, si $A = \Omega$, on a

$$Q(\Omega) = 1 = \sum_{\omega \in S} p_\omega + Q(\bar{S}).$$

Or, d'après le théorème de sommation par paquets, on a

$$\sum_{\omega \in \Omega} p_\omega = 1 = \sum_{\omega \in S} p_\omega + \underbrace{\sum_{\omega \in \bar{S}} p_\omega}_{=0}.$$

Donc $Q(\bar{S}) = 0$.

Donc $0 \leq Q(A \cap \bar{S}) \leq Q(\bar{S}) = 0$ et donc $Q(A \cap \bar{S}) = 0$.

Donc $Q(A) = \sum_{\omega \in A \cap S} p_\omega$.

Ceci est aussi valable pour la probabilité \mathbb{P} donc $\mathbb{P}(A) = Q(A)$.

Donc $\mathbb{P} = Q$. ■

Quand l'univers est infini non-dénombrable, le résultat précédent donne toutes les probabilités discrètes sur $(\Omega, \mathcal{P}(\Omega))$. Mais si on choisit des tribus plus petites (ce qui est très relatif : en général, ce sont des ensembles énormes, dont la puissance dépasse celle du continu !), alors on peut créer d'autres types de probabilités (comme les probabilités dites continues).

Quand l'univers Ω est fini ou dénombrable, alors toutes les probabilités sont discrètes : on choisit toujours la tribu $\mathcal{P}(\Omega)$, ce qui est toujours sous-entendu.

Proposition 13.36

Soit $\Omega = \{\omega_1, \dots, \omega_n\}$ un ensemble fini.

Alors pour tout n -uplet $(p_1, \dots, p_n) \in [0; 1]^n$ tel que $\sum_{i=1}^n p_i = 1$, il existe une unique probabilité \mathbb{P} sur Ω telle que pour tout $i \in \llbracket 1; n \rrbracket$, $\mathbb{P}(\{\omega_i\}) = p_i$.

Proposition 13.37

Soit $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$ un ensemble dénombrable.

Alors pour toute suite $(p_n) \in [0; 1]^{\mathbb{N}}$ telle que la série $\sum_{n \geq 1} p_n$ converge et $\sum_{n=1}^{+\infty} p_n = 1$, il existe une unique probabilité \mathbb{P} sur Ω telle que pour tout $i \in \mathbb{N}^*$, $\mathbb{P}(\{\omega_i\}) = p_i$.

Exercice 13.38

Déterminez l'unique constante λ telle qu'on puisse définir une probabilité sur \mathbb{N} en posant

$$\text{pour tout } n \in \mathbb{N}, \quad \mathbb{P}(\{n\}) = \frac{\lambda}{n!}.$$

Correction 13.39

La famille $\left(\frac{\lambda}{n!}\right)_{n \in \mathbb{N}}$ est une distribution de probabilité discrète sur \mathbb{N} si et seulement si

$$\begin{cases} \sum_{n \in \mathbb{N}} \frac{\lambda}{n!} = 1 \\ \forall n \in \mathbb{N}, \quad \frac{\lambda}{n!} \geq 0 \end{cases}$$

Or on a $\sum_{n \in \mathbb{N}} \frac{\lambda}{n!} = \lambda \sum_{n=0}^{+\infty} \frac{1}{n!} = \lambda e$ donc la seule solution est $\lambda = \frac{1}{e}$.

Dans toute la suite, on suppose donné un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

13.3 Probabilités conditionnelles

13.3.1 Généralités

Définition 13.40

Soit $A \in \mathcal{F}$ un événement non-négligeable.

Pour $B \in \mathcal{F}$, on pose $\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$, appelé probabilité sachant A de B .

On voit aussi la notation $\mathbb{P}(B | A)$, mais attention, cette notation est trompeuse, elle peut laisser penser qu'il existe un événement qui s'appellerait « B sachant A », ce qui n'a aucun sens.

L'idée derrière la notion de probabilité conditionnelle est que lorsqu'on dispose d'une information partielle sur le résultat de l'expérience, notre perception des probabilités s'en trouve modifiée.

Théorème 13.41

Sous les mêmes hypothèses, \mathbb{P}_A est une probabilité sur (Ω, \mathcal{F}) , appelée probabilité conditionnelle relative à A .

Démonstration 13.42

On a $\mathbb{P}_A(\Omega) = \frac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$.

Pour toute famille $(B_n)_{n \in \mathbb{N}}$ d'événements deux à deux incompatibles, on a

$$\begin{aligned}\mathbb{P}_A\left(\bigsqcup_{n \in \mathbb{N}} B_n\right) &= \frac{1}{\mathbb{P}(A)} \mathbb{P}\left(A \cap \bigsqcup_{n \in \mathbb{N}} B_n\right) \\ &= \frac{1}{\mathbb{P}(A)} \mathbb{P}\left(\bigsqcup_{n \in \mathbb{N}} (A \cap B_n)\right) \\ &= \frac{1}{\mathbb{P}(A)} \sum_{n \in \mathbb{N}} \mathbb{P}(A \cap B_n) \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}_A(B_n).\end{aligned}$$

Donc \mathbb{P}_A est une probabilité sur (Ω, \mathcal{F}) . ■

En général, on connaît plutôt $\mathbb{P}(A)$ et $\mathbb{P}_A(B)$, ce qui permet de calculer $\mathbb{P}(A \cap B)$:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}_A(B).$$

On peut généraliser.

Théorème 13.43 (Formule des probabilités composées)

Soit (A_1, \dots, A_n) une famille d'événements tels que $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$.

Alors $\mathbb{P}(A_1 \cap \dots \cap A_{n-1} \cap A_n) = \mathbb{P}(A_1) \mathbb{P}_{A_1}(A_2) \mathbb{P}_{A_1 \cap A_2}(A_3) \dots \mathbb{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n)$.

En général, on utilise ce résultat lorsque des événements (au sens naturel) se succèdent et que la connaissance de chaque événement permet de déterminer l'état du système.

Exercice 13.44

On dispose d'une urne contenant une boule blanche et n boules noires (avec $n \geq 1$).

On effectue une suite de tirages jusqu'à obtenir la boule blanche en respectant le protocole suivant : si on tire une boule noire, on la remplace par deux boules noires.

Calculez la probabilité d'obtenir la boule blanche à l'issue du k -ème tirage et la probabilité de ne jamais tirer la boule blanche.

Correction 13.45

Pour $k \in \mathbb{N}^*$, on note A_k l'événement « on tire une boule noire au k -ème tirage » et B_k l'événement « on tire la boule blanche au k -ème tirage ».

On a

$$B_k = \bigcap_{i=1}^{k-1} A_i \cap \overline{A_k}.$$

D'après la formule des probabilités composées, on a donc

$$\mathbb{P}(B_k) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \dots \times \mathbb{P}_{A_1 \cap \dots \cap A_{k-2}}(A_{k-1}) \times \mathbb{P}_{A_1 \cap \dots \cap A_{k-1}}(\overline{A_k}).$$

En faisant l'hypothèse que les tirages dans l'urne sont, à chaque étape, équiprobables, on a

$$\mathbb{P}(A_1) = \frac{n}{n+1} \quad \mathbb{P}_{A_1}(A_2) = \frac{n+1}{n+2} \quad \dots \quad \mathbb{P}_{A_1 \cap \dots \cap A_{k-2}}(A_{k-1}) = \frac{n+k-2}{n+k-1}$$

et

$$\mathbb{P}_{A_1 \cap \dots \cap A_{k-1}}(\overline{A_k}) = \frac{1}{n+k}.$$

On en déduit

$$\mathbb{P}(B_k) = \frac{n}{(n+k-1)(n+k)}.$$

De plus, en notant C l'événement « ne jamais tirer la boule blanche », on a $C = \bigcap_{n \in \mathbb{N}^*} A_n$ et donc

$$\mathbb{P}(C) = \lim_{N \rightarrow +\infty} \mathbb{P}\left(\bigcap_{k=1}^N A_k\right) = \lim_{N \rightarrow +\infty} \frac{n}{n+1} \times \dots \times \frac{n+N-1}{n+N} = \lim_{N \rightarrow +\infty} \frac{n}{n+N} = 0.$$

13.3.2 Systèmes complets d'événements

Définition 13.46

Soit $(A_i)_{i \in I}$ une famille d'événements.

On dit que la famille $(A_i)_{i \in I}$ est un système complet d'événements si, et seulement si :

- I est fini ou dénombrable (en pratique, on a souvent $I = \llbracket 1 ; n \rrbracket$ ou $I = \mathbb{N}$)
- les événements sont deux à deux incompatibles :

$$\text{pour tout } (i, j) \in I^2, \ i \neq j \implies A_i \cap A_j = \emptyset$$

- $\bigcup_{i \in I} A_i = \Omega.$

Exemple 13.47

- Si A est un événement, le couple (A, \overline{A}) est un système complet d'événements.
- Si Ω est fini ou dénombrable, la famille de tous les événements élémentaires est un système complet d'événements.

Les systèmes complets d'événements interviennent lorsqu'on est tenté de faire une disjonction de cas : on est dans un cas, ou alors dans un autre, etc., mais sans que jamais deux cas soient simultanément possibles.

Si $(A_i)_{i \in I}$ est un système complet d'événements, alors $\sum_{i \in I} \mathbb{P}(A_i) = 1.$

Remarque 13.48

D'une manière générale, dans toute la suite du cours, si on voit apparaître une somme $\sum_{i \in I} \dots$, alors elle signifiera :

- une vraie somme quand I est fini
- une somme d'une famille sommable quand I est infini
- cas particulier : dans le cas de réels positifs, on peut oublier la condition de sommabilité et prendre les sommes dans $[0 ; +\infty]$ (c'est presque toujours le cas tant qu'on ne manipule que des probabilités, qui sont des réels positifs).

Définition 13.49

Pour définir un système quasi-complet d'événements, on remplace la condition $\bigcup_{i \in I} A_i = \Omega$ par la condition $\sum_{i \in I} \mathbb{P}(A_i) = 1$.

Un système complet d'événements est donc un système quasi-complet d'événements.

Réciproquement, si $(A_i)_{i \in I}$ est un système quasi-complet d'événements, on pose $B = \overline{\bigcup_{i \in I} A_i}$: B est alors un événement négligeable et le système d'événements constitué des événements A_i auxquels on ajoute l'événement B est alors un système complet.

Conclusion : à un événement négligeable près, les deux notions sont identiques. La suite du cours montre que la différence entre les deux notions n'est pas fondamentale en pratique.

Démonstration 13.50

Si $(A_i)_{i \in I}$ est un système quasi-complet d'événements, alors $\Omega = \bigsqcup_{i \in I} A_i \sqcup \overline{\bigsqcup_{i \in I} A_i}$ donc

$$\begin{aligned} \mathbb{P}(\Omega) = 1 &= \mathbb{P}\left(\bigsqcup_{i \in I} A_i\right) + \mathbb{P}\left(\overline{\bigsqcup_{i \in I} A_i}\right) \\ &= \underbrace{\sum_{i \in I} \mathbb{P}(A_i)}_{=1} + \mathbb{P}\left(\overline{\bigsqcup_{i \in I} A_i}\right). \end{aligned} \quad \begin{array}{l} \curvearrowright \\ I \text{ dénombrable} \end{array}$$

Donc $\mathbb{P}\left(\overline{\bigsqcup_{i \in I} A_i}\right) = 0$. ■

13.3.3 Formule des probabilités totales

Théorème 13.51

Soit $(A_i)_{i \in I}$ un système complet d'événements.

Alors pour tout événement B , $\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B \cap A_i)$.

Si, de plus, tous les événements A_i sont de probabilité non-nulle, alors $\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}_{A_i}(B) \mathbb{P}(A_i)$.

Si A est un événement négligeable, alors $A \cap B$ en est un aussi : formellement, la probabilité conditionnelle $\mathbb{P}_A(B)$ n'est pas définie ; on lui donne alors une valeur arbitraire (souvent 0, en fait peu importe) et on accepte quand même l'égalité $\mathbb{P}(A \cap B) = \mathbb{P}_A(B) \mathbb{P}(A) = 0$, car, dans ce cas, cette égalité est vraie puisque $\mathbb{P}(A \cap B) = \mathbb{P}(A) = 0$.

Avec cette convention, on peut étendre la formule des probabilités totales à tout système quasi-complet d'événements.

Théorème 13.52

Soit $(A_i)_{i \in I}$ un système complet d'événements.

Alors pour tout événement B , $\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}_{A_i}(B) \mathbb{P}(A_i)$.

Exercice 13.53

Soit $p \in]0 ; 1[$.

Dans \mathbb{N} , on définit la probabilité \mathbb{P} par $\mathbb{P}(\{n\}) = (1-p)p^n$.

Justifiez que cette égalité définit effectivement une probabilité sur \mathbb{N} .

On tire un entier N au hasard selon cette probabilité, puis on remplit une urne avec une boule noire et N boules blanches. On prélève enfin une boule dans l'urne.

Quelle est la probabilité d'obtenir une boule noire dans cette expérience ?

Correction 13.54

On a $\sum_{n \in \mathbb{N}} (1-p)p^n = (1-p) \sum_{n \in \mathbb{N}} p^n = \frac{1-p}{1-p} = 1$ donc d'après la Proposition 13.37, \mathbb{P} est bien une probabilité sur \mathbb{N} .

Pour $n \in \mathbb{N}$, on note A_n l'événement « on a choisi $n \in \mathbb{N}$ à la première étape » et on note B l'événement « on obtient la boule noire à la deuxième étape ».

$(A_n)_{n \in \mathbb{N}}$ est un système complet d'événements donc d'après la formule des probabilités totales, on a

$$\begin{aligned} \mathbb{P}(B) &= \sum_{n \in \mathbb{N}} \mathbb{P}_{A_n}(B) \mathbb{P}(A_n) \\ &= \sum_{n=0}^{+\infty} \frac{1}{n+1} p^n (1-p) \\ &= (1-p) \sum_{n=0}^{+\infty} \frac{p^n}{n+1} \\ &= \frac{1-p}{p} \sum_{n=1}^{+\infty} \frac{p^n}{n} \\ &= \frac{p-1}{p} \ln(1-p). \end{aligned}$$

13.3.4 Formule de Bayes

Proposition 13.55

Soient A et B deux événements de probabilité non-nulle.

$$\text{Alors } \mathbb{P}_B(A) = \frac{\mathbb{P}(A) \mathbb{P}_A(B)}{\mathbb{P}(B)}.$$

Cette formule est appelée la formule de probabilité des causes.

On déduit de cela et de la formule des probabilités totales la formule de Bayes.

Théorème 13.56

Soit $(A_i)_{i \in I}$ un système quasi-complet d'événements.

$$\text{Alors pour tout événement } B \text{ tel que } \mathbb{P}(B) \neq 0, \text{ on a } \mathbb{P}_B(A_i) = \frac{\mathbb{P}(A_i) \mathbb{P}_{A_i}(B)}{\sum_{j \in I} \mathbb{P}(A_j) \mathbb{P}_{A_j}(B)}.$$

Remarque 13.57

L'énoncé est donné ici dans sa pleine version mais, en pratique, on retrouve la formule à chaque fois en refaisant la démonstration dans le cas qui nous préoccupe.

Exercice 13.58

Une proportion p d'une population est atteinte d'une maladie donnée (*prévalence* de la maladie) pour laquelle un test de dépistage existe.

Appliqué à un individu atteint, le test donne un résultat positif avec une probabilité s (*sensibilité* du test).

Appliqué à un individu indemne, il donne un résultat négatif avec une probabilité s' (*spécificité* du test).

Calculez la probabilité qu'un patient soit effectivement atteint lorsque son test est positif (*valeur prédictive positive*) ; qu'il soit effectivement indemne lorsque son test est négatif (*valeur prédictive négative*).

Correction 13.59

On note M l'événement « l'individu est malade » et T l'événement « le test est positif ».

On a $\mathbb{P}(M) = p$, $\mathbb{P}_M(T) = s$ et $\mathbb{P}_{\overline{M}}(\overline{T}) = s'$.

On a $\mathbb{P}(M \cap T) = \mathbb{P}_T(M) \mathbb{P}(T) = \mathbb{P}_M(T) \mathbb{P}(M)$ donc

$$\mathbb{P}_T(M) = \frac{\mathbb{P}(T \cap M)}{\mathbb{P}(T)}.$$

Or (M, \overline{M}) est un système complet d'événements donc

$$\begin{aligned}\mathbb{P}(T) &= \mathbb{P}_M(T) \mathbb{P}(M) + \mathbb{P}(\overline{M}) \mathbb{P}_{\overline{M}}(T) \\ &= sp + \left(1 - \mathbb{P}_{\overline{M}}(\overline{T})\right) (1 - \mathbb{P}(M)) \\ &= sp + (1 - s') (1 - p) .\end{aligned}$$

Donc

$$\mathbb{P}_T(M) = \frac{sp}{sp + (1 - s') (1 - p)} .$$

Idem pour l'autre probabilité.

Exercice 13.60

Jean-Eudes lance une pièce équilibrée jusqu'à obtenir pile, il compte le nombre de lancers nécessaires, noté k , puis il remplit une urne avec k boules numérotées de 1 à k . Enfin, il tire une boule dans l'urne.

Il nous annonce qu'il a obtenu 1, mais ne nous donne pas la valeur de k .

Déterminez la probabilité qu'il n'ait fait qu'un seul lancer de pièce.

Correction 13.61

Pour $k \in \mathbb{N}^*$, on note A_k l'événement « nécessiter k lancers pour obtenir pile » et B l'événement « tirer la boule 1 ».

On cherche $\mathbb{P}_B(A_1)$.

D'après la formule de probabilité des causes, on a

$$\mathbb{P}_B(A_1) = \frac{\mathbb{P}_{A_1}(B) \mathbb{P}(A_1)}{\mathbb{P}(B)} .$$

On a $\mathbb{P}(A_1) = \frac{1}{2}$ (pièce équilibrée) et $\mathbb{P}_{A_1}(B) = 1$.

La famille $(A_k)_{k \in \mathbb{N}^*}$ est un système quasi-complet d'événements car $((A_k)_{k \in \mathbb{N}^*}, \{\omega_F\})$, où $\omega_F = (F, F, \dots, F, \dots)$, est une partition de Ω telle que $\mathbb{P}(\{\omega_F\}) = 0$, donc $\sum_{k \in \mathbb{N}^*} \mathbb{P}(A_k) = 1$.

On note F_k l'événement « obtenir face au k -ème lancer » et on a

$$A_k = \bigcap_{i=1}^{k-1} F_i \cap \overline{F_k} .$$

Les F_i étant supposés indépendants, on a

$$\mathbb{P}(A_k) = \prod_{i=1}^{k-1} \mathbb{P}(F_i) \times \mathbb{P}(\overline{F_k}) = \frac{1}{2^k}$$

et on remarque $\sum_{k=1}^{+\infty} \frac{1}{2^k} = 1$.

D'après la formule des probabilités totales, on a donc

$$\mathbb{P}(B) = \sum_{k=1}^{+\infty} \mathbb{P}(A_k) \mathbb{P}_{A_k}(B) = \sum_{k=1}^{+\infty} \frac{1}{2^k k} = \ln 2.$$

Finalement, $\mathbb{P}_B(A_1) = \frac{1}{2 \ln 2}$.

13.4 Indépendance

13.4.1 Indépendance de deux événements

Définition 13.62

Soient A et B deux événements.

On dit que A et B sont indépendants (pour la probabilité \mathbb{P}) ssi $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$.

Proposition 13.63

Si A et B sont deux événements indépendants, alors les événements A et \overline{B} sont indépendants, les événements \overline{A} et B sont indépendants et les événements \overline{A} et \overline{B} sont indépendants.

Remarque 13.64

- Attention ! Ne pas confondre « événements indépendants » et « événements incompatibles » !
- Le fait que deux événements soient indépendants ou pas n'est pas seulement lié aux événements eux-mêmes, mais dépend aussi de la probabilité. Quand plusieurs probabilités sont utilisées (par exemple des probabilités conditionnelles), il est essentiel de préciser pour quelle probabilité les événements sont indépendants.

Proposition 13.65

Soient A et B deux événements tels que A n'est pas négligeable.

Alors A et B sont indépendants ssi $\mathbb{P}(B) = \mathbb{P}_A(B)$.

Intuitivement, deux événements sont indépendants si le fait de savoir que l'un des deux est réalisé n'apporte aucune information sur le fait de savoir que l'autre le soit ou non : dans le cas où A et B sont deux événements tels que $\mathbb{P}(A) \neq 0$ et $\mathbb{P}(B) \neq 0$, alors $\mathbb{P}_B(A) = \mathbb{P}(A)$ et $\mathbb{P}_A(B) = \mathbb{P}(B)$. Et non pas, comme c'est souvent indiqué dans la littérature, le fait qu'un événement n'influe pas sur un autre ! Mais tout ça n'est que baratin. **Il faut se méfier de l'intuition quand on fait des calculs de probabilités.**

Exercice 13.66

Une famille a n enfants (avec $n \geq 2$).

Quelle est la probabilité qu'il n'y ait que des enfants du même sexe ?

Quelle est la probabilité qu'il y ait au moins deux filles ?

Montrez que ces deux événements sont indépendants ssi $n = 3$.

Correction 13.67

Pour $i \in \llbracket 1 ; n \rrbracket$, on note F_i l'événement « le i -ème enfant est une fille ».

► On note A l'événement « il n'y a que des enfants du même sexe » *i.e.*

$$A = \bigcap_{i=1}^n F_i \sqcup \bigcap_{i=1}^n \overline{F_i}.$$

On a

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}\left(\bigcap_{i=1}^n F_i\right) + \mathbb{P}\left(\bigcap_{i=1}^n \overline{F_i}\right) \\ &= \prod_{i=1}^n \mathbb{P}(F_i) + \prod_{i=1}^n \mathbb{P}(\overline{F_i}) \\ &= \frac{1}{2^n} + \frac{1}{2^n} \\ &= \frac{1}{2^{n-1}}. \end{aligned}$$

► On note B l'événement « il y a au plus une fille » *i.e.*

$$B = \bigcap_{i=1}^n \overline{F_i} \sqcup \bigsqcup_{i=1}^n \left(F_i \cap \bigcap_{\substack{1 \leq j \leq n \\ j \neq i}} \overline{F_j} \right).$$

Donc

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}\left(\bigcap_{i=1}^n \overline{F_i}\right) + \sum_{i=1}^n \mathbb{P}\left(F_i \cap \bigcap_{\substack{1 \leq j \leq n \\ i \neq j}} \overline{F_j}\right) \\ &= \frac{1}{2^n} + \frac{n}{2^n} \\ &= \frac{n+1}{2^n}. \end{aligned}$$

► On a $\mathbb{P}(A \cap B) = \mathbb{P}\left(\bigcap_{i=1}^n \overline{F_i}\right) = \frac{1}{2^n}.$

On a alors

$$\begin{aligned} A \text{ et } \overline{B} \text{ indépendants} &\iff A \text{ et } B \text{ indépendants} \\ &\iff \mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B) \\ &\iff \frac{1}{2^n} = \frac{1}{2^{n-1}} \times \frac{n+1}{2^n} \\ &\iff 2^{n-1} = n+1. \end{aligned}$$

$n = 3$ est une solution évidente et on montre que $x \mapsto 2^{x-1} - x - 1$ est strictement croissante sur $[3 ; +\infty[$ donc l'équation $2^{n-1} = n+1$ n'a pas de solution dans $\llbracket 4 ; +\infty \rrbracket$.

De plus, $n = 2$ n'est pas solution.

Donc A et \overline{B} sont indépendants ssi $n = 3$.

13.4.2 Indépendance mutuelle

Définition 13.68

Soit $(A_i)_{i \in I}$ une famille finie ou dénombrable d'événements.

On dit que les événements $(A_i)_{i \in I}$ sont (mutuellement) indépendants ssi pour tout sous-ensemble fini

$$J \subseteq I, \quad \mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

Le résultat précédent peut être généralisé.

Proposition 13.69

Si $(A_i)_{i \in I}$ est une famille d'événements (mutuellement) indépendants, alors toute famille d'événements $(C_i)_{i \in I}$ où pour tout $i \in I$, on choisit $C_i = A_i$ ou $C_i = \overline{A_i}$, est aussi une famille d'événements mutuellement indépendants.

Remarque 13.70

On s'intéresse rarement (pour ne pas dire jamais!) à des familles d'événements deux à deux indépendants seulement : cette notion n'implique pas l'indépendance mutuelle, qui est la seule notion vraiment utile.

Remarque 13.71

L'indépendance mutuelle d'un grand nombre d'événements est presque toujours une propriété postulée lors de la modélisation et rarement une propriété démontrée car c'est une preuve difficile en général.

On finit par le lemme des coalitions sur les événements.

Théorème 13.72

Soit $(A_i)_{i \in I}$ une famille d'événements mutuellement indépendants.

Si l'événement B est le résultat d'opérations ensemblistes sur une sous-famille $(A_j)_{j \in J}$ (où $J \subseteq I$) et C est le résultat d'opérations ensemblistes sur la sous-famille complémentaire $(A_i)_{i \in I \setminus J}$, alors B et C sont indépendants.

Chapitre 14

Variables aléatoires discrètes

Sommaire

14.1	Variables aléatoires discrètes	417
14.1.1	Définition	417
14.1.2	Probabilité-image d'une variable aléatoire discrète	417
14.1.3	Loi d'une variable aléatoire discrète	418
14.1.4	Cas des variables aléatoires discrètes réelles ou complexes	421
14.2	Espérance	423
14.2.1	Définitions	423
14.2.2	Propriétés	425
14.2.3	Théorème de transfert	428
14.2.4	Inégalité de Markov	430
14.3	Variance d'une variable réelle	431
14.3.1	Moments d'ordre 2	431
14.3.2	Variance et écart-type	433
14.3.3	Inégalité de Bienaymé-Tchebychev	434
14.3.4	Généralisation	436
14.4	Lois classiques	437
14.4.1	Loi uniforme	437
14.4.2	Loi de Bernoulli	437
14.4.3	Loi binomiale	438
14.4.4	Loi géométrique	438
14.4.5	Loi de Poisson	439
14.5	Couples de variables aléatoires	440
14.5.1	Généralités	440
14.5.2	Lois marginales	440
14.5.3	Lois conditionnelles	443
14.5.4	Covariance	443
14.6	Indépendance de variables aléatoires	445
14.6.1	Généralités	445
14.6.2	Espérance et indépendance	446
14.6.3	Généralisation	446
14.6.4	Théorème de réalisation	448
14.6.5	Somme de variables indépendantes identiquement distribuées	449
14.7	Loi faible des grands nombres	450
14.8	Fonction génératrice d'une variable aléatoire à valeurs entières	451
14.8.1	Généralités	451

14.8.2	Lien entre espérance et fonction génératrice	454
14.8.3	Fonction génératrice d'une somme de variables indépendantes	455

Dans tout ce chapitre, même si ce n'est pas rappelé à chaque fois, on suppose que $(\Omega, \mathcal{T}, \mathbb{P})$ est un espace probabilisé.

Dans ce chapitre, si E est un ensemble fini ou dénombrable et $(u_x)_{x \in E}$ une famille de réels indexée par E , la notation $\sum_{x \in E} u_x$ désigne la somme au sens des familles sommables ou des familles positives selon les cas.

14.1 Variables aléatoires discrètes

14.1.1 Définition

Définition 14.1

On appelle variable aléatoire discrète toute application X de Ω dans un ensemble quelconque telle que son ensemble-image $X(\Omega)$ soit fini ou dénombrable et telle que pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathcal{T}$.

On appelle variable aléatoire discrète réelle (respectivement complexe) toute variable aléatoire discrète à valeurs dans \mathbb{R} (respectivement \mathbb{C}).

En général, on note avec des lettres majuscules droites les variables aléatoires : X, Y, S , etc.

Si $U \subseteq X(\Omega)$, on note plutôt $X^{-1}(U)$ sous la forme préférée des probabilistes $\{X \in U\}$ ou même $(X \in U)$.

Remarque 14.2

- Comme son nom ne l'indique pas, une variable aléatoire n'est pas une variable, mais une fonction ! La terminologie a été fixée à une époque ancienne où la notion n'était pas encore parfaitement claire.
- Toute fonction constante est une variable aléatoire, appelée variable aléatoire certaine.
- Si A est un événement, la variable aléatoire 1_A est appelée fonction indicatrice de A .

14.1.2 Probabilité-image d'une variable aléatoire discrète

Proposition 14.3

Soit X une variable aléatoire discrète à valeurs dans E .

À toute partie U de $X(\Omega)$, on associe $\mathbb{P}_X(U) = \mathbb{P}(\{X \in U\}) = \mathbb{P}(X \in U)$.

Alors \mathbb{P}_X est une probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$, appelée probabilité-image de X ou loi de X .

Démonstration 14.4

On a $(X \in X(\Omega)) = \{\omega \in \Omega \mid X(\omega) \in X(\Omega)\} = \Omega$ donc $\mathbb{P}_X(X(\Omega)) = 1$.

Si $(A_i)_{i \in \mathbb{N}}$ est une famille de parties de $X(\Omega)$ deux à deux disjointes, on a

$$\begin{aligned}\mathbb{P}_X\left(\bigsqcup_{i \in \mathbb{N}} A_i\right) &= \mathbb{P}\left(X^{-1}\left(\bigsqcup_{i \in \mathbb{N}} A_i\right)\right) \\ &= \mathbb{P}\left(\bigsqcup_{i \in \mathbb{N}} X^{-1}(A_i)\right) \\ &= \sum_{i \in \mathbb{N}} \mathbb{P}(X^{-1}(A_i)) \\ &= \sum_{i \in \mathbb{N}} \mathbb{P}_X(A_i).\end{aligned}$$

Donc \mathbb{P}_X est une probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$. ■

Remarque 14.5

L'intérêt de la notion de variable aléatoire est de déplacer les calculs de probabilité dans l'univers Ω souvent inconnu, muni d'une tribu inconnue, dans un ensemble fini ou dénombrable $X(\Omega)$ bien plus agréable, muni de la tribu simple $\mathcal{P}(X(\Omega))$.

Dans notre pratique des probabilités, nous supposons toujours l'existence des variables aléatoires qu'on considère ! Voyons néanmoins un résultat de réalisation effective.

14.1.3 Loi d'une variable aléatoire discrète

On rappelle qu'une distribution discrète (de probabilités) est une famille de réels positifs $(p_x)_{x \in L}$ où L est un ensemble fini ou dénombrable (non vide) telle que $\sum_{x \in L} p_x = 1$.

Les exemples suivants sont fondamentaux et sont à connaître.

Exemple 14.6

Avec L un ensemble fini :

- Distribution uniforme sur L : $\mathcal{U}(L) = (p_x)_{x \in L}$ où $p_x = \frac{1}{\text{Card } L}$.
- Distribution de Bernoulli de paramètre $p \in [0 ; 1]$: $\mathcal{B}(p) = (p_0, p_1)$ où $p_1 = p$ et $p_0 = 1 - p$.
- Distribution binomiale de paramètre $(n, p) \in \mathbb{N} \times [0 ; 1]$: $\mathcal{B}(n, p) = (p_k)_{k \in \llbracket 0; n \rrbracket}$ où $p_k = \binom{n}{k} p^k (1 - p)^{n-k}$.

Avec L dénombrable :

- ▷ Distribution géométrique de paramètre $p \in]0 ; 1[$: $\mathcal{G}(p) = (p_x)_{x \in \mathbb{N}^*}$ où $p_x = (1-p)^{x-1} p$.
- ▷ Distribution de Poisson de paramètre $\lambda > 0$: $\mathcal{P}(\lambda) = (p_k)_{k \in \mathbb{N}}$ où $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$.

Définition 14.7

On dit qu'une variable aléatoire discrète X sur Ω suit la loi discrète \mathcal{L} quand

$$X(\Omega) \subseteq L \quad \text{et} \quad \forall x \in L, \mathbb{P}(X = x) = p_x$$

où $(p_x)_{x \in L}$ est la distribution de probabilités nommée par le même nom que \mathcal{L} .

Quand on demande la loi d'une variable aléatoire X , on demande donc de déterminer un ensemble L tel que $X(\Omega) \subseteq L$ et pour chaque $a \in L$, la valeur de $\mathbb{P}(X = a)$. En général, on a souvent $X(\Omega) = L$, mais parfois ce n'est pas ce qui est le plus simple.

Par abus de langage, on confond la loi et la distribution de probabilité associée.

Proposition 14.8

Soit X une variable aléatoire qui suit la loi $\mathcal{L} = (p_x)_{x \in L}$.

La famille d'événements $(\{X = x\})_{x \in L}$ est un système quasi-complet d'événements, appelé système quasi-complet d'événements associé à X .

En particulier, on a $\sum_{x \in L} \mathbb{P}(X = x) = 1$.

Exercice 14.9

Une urne contient $n + 1$ boules numérotées de 0 à n .

On procède à deux tirages avec remise et on note S la somme des numéros des deux boules.

Déterminez la loi de S .

Correction 14.10

On a $S(\Omega) = \llbracket 0 ; 2n \rrbracket$.

Pour $k \in \llbracket 0 ; 2n \rrbracket$, on cherche $\mathbb{P}(S = k)$.

On pose R la variable aléatoire égale au couple (a_1, a_2) des résultats des tirages.

On a $R(\Omega) = \llbracket 0 ; n \rrbracket^2$ et on remarque $R \sim \mathcal{U}(\llbracket 0 ; n \rrbracket^2)$.

De plus, on remarque

$$(S = 0) = (R = (0, 0)) \quad \text{et} \quad (S = 1) = (R = (0, 1)) \sqcup (R = (1, 0)).$$

On en déduit

$$(S = k) = \begin{cases} \bigcup_{\ell=0}^k (R = (\ell, k - \ell)) & \text{pour } k \in \llbracket 0 ; n \rrbracket \\ \bigcup_{\ell=k-n}^n (R = (\ell, k - \ell)) & \text{pour } k \in \llbracket n + 1 ; 2n \rrbracket \end{cases}$$

et donc

$$\mathbb{P}(S = k) = \begin{cases} \sum_{\ell=0}^k \mathbb{P}(R = (\ell, k - \ell)) = \sum_{\ell=0}^k \frac{1}{(n+1)^2} = \frac{k+1}{(n+1)^2} & \text{si } k \in \llbracket 0 ; n \rrbracket \\ \sum_{\ell=k-n}^n \mathbb{P}(R = (\ell, k - \ell)) = \frac{2n - k + 1}{(n+1)^2} & \text{si } k \in \llbracket n + 1 ; 2n \rrbracket \end{cases}$$

Exercice 14.11

On lance une infinité de fois une pièce dont la probabilité de tomber sur pile vaut $p \in]0 ; 1[$.

On note X la variable aléatoire qui compte le nombre de lancers nécessaires pour obtenir une première fois pile.

Déterminez la loi de X en faisant des hypothèses raisonnables sur les lancers.

Correction 14.12

On a $X(\Omega) = \mathbb{N}^*$.

On note F_i l'événement « obtenir face au i -ème lancer ».

Pour $k \in \mathbb{N}^*$, on a $(X = k) = \bigcap_{i=1}^{k-1} F_i \cap \overline{F_k}$.

Par indépendance des F_i , on a donc

$$\mathbb{P}(X = k) = \prod_{i=1}^{k-1} \mathbb{P}(F_i) \times \mathbb{P}(\overline{F_k}) = (1-p)^{k-1} p.$$

On remarque $X \sim \mathcal{G}(p)$.

On pourra retenir : **loi géométrique = première réussite d'une expérience de Bernoulli indépendante.**

Théorème de réalisation.

Théorème 14.13

Soit \mathcal{L} une distribution discrète de probabilités.

Il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et une variable aléatoire X sur Ω tels que X suive la loi \mathcal{L} .

Démonstration 14.14

On note $\mathcal{L} = (p_x)_{x \in L}$ et on pose $\Omega = L$, $\mathcal{T} = \mathcal{P}(L)$ et \mathbb{P} la probabilité sur L qui suit la distribution \mathcal{L} (cf. Proposition 13.34).

En posant $X = \text{id}_L$, on a bien

$$\begin{aligned}\mathbb{P}(X = x) &= \mathbb{P}(\{\omega \in L \mid X(\omega) = x\}) \\ &= \mathbb{P}(\{\omega \in L \mid \omega = x\}) \\ &= \mathbb{P}(\{x\}) \\ &= p_x.\end{aligned}$$

■

On note $X \sim \mathcal{L}$ pour signifier que X suit la loi \mathcal{L} , ou $X \sim Y$ pour signifier que X et Y suivent la même loi (ce qui ne présuppose rien sur les univers de départ de X et Y : on a seulement $X(\Omega) = Y(\Omega')$ et les égalités numériques $\forall x \in X(\Omega), \mathbb{P}(X = x) = \mathbb{P}(Y = x)$).

14.1.4 Cas des variables aléatoires discrètes réelles ou complexes

Proposition 14.15

L'ensemble des variables aléatoires discrètes réelles sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ est une \mathbb{R} -algèbre : on peut donc additionner ou multiplier par un scalaire des variables aléatoires discrètes réelles, ou les additionner entre elles.

Si X est une variable aléatoire discrète réelle sur Ω et si f est une fonction de \mathbb{R} dans \mathbb{R} définie sur $X(\Omega)$, alors la composée $f \circ X$ est une variable aléatoire discrète réelle définie sur Ω , qu'on note plutôt $f(X)$.

On a bien sûr le même résultat avec les variables complexes.

La loi de $f(X)$ est donc théoriquement donnée par :

$$\forall y \in f(X)(\Omega), \mathbb{P}(f(X) = y) = \sum_{x \in f^{-1}(\{y\}) \cap X(\Omega)} \mathbb{P}(X = x).$$

En pratique, cette loi est souvent difficilement calculable.

Exercice 14.16

Soient X une variable aléatoire discrète suivant une loi de Poisson $\mathcal{P}(\lambda)$ et $p \in \mathbb{N}^*$.

On pose $Y = X^2 - (2p + 1)X$.

Déterminez la loi de Y .

Correction 14.17

On a $X \sim \mathcal{P}(\lambda)$ donc $X(\Omega) = \mathbb{N}$ et $\forall k \in \mathbb{N}$, $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$.

On a donc $Y(\Omega) \subseteq \mathbb{Z}$ et pour $\ell \in \mathbb{Z}$, on a

$$(Y = \ell) = \{\omega \in \Omega \mid X^2(\omega) - (2p+1)X(\omega) - \ell = 0\}.$$

On considère l'équation $x^2 - (2p+1)x - \ell = 0$ d'inconnue $x \in \mathbb{N}$ et de discriminant $\Delta = (2p+1)^2 + 4\ell$.

Si $\ell < \frac{-(2p+1)^2}{4}$ alors $\Delta < 0$ donc $(Y = \ell) = \emptyset$ et donc $\mathbb{P}(Y = \ell) = 0$.

Si $\ell = \frac{-(2p+1)^2}{4}$ alors ℓ n'est pas entier donc $\mathbb{P}(Y = \ell) = 0$.

Si $\ell > \frac{-(2p+1)^2}{4}$, les racines de l'équation sont

$$x_1 = \frac{2p+1 + \sqrt{\Delta}}{2} \quad \text{et} \quad x_2 = \frac{2p+1 - \sqrt{\Delta}}{2}.$$

On a $x_1 + x_2 = 2p+1$ donc si l'une des racines est entière, l'autre l'est aussi.

Δ est un entier impair donc si $\sqrt{\Delta}$ est un entier, c'est un entier impair et donc x_1 et x_2 sont deux entiers.

Réciproquement, si x_1 est un entier alors $\sqrt{\Delta} = 2x_1 - (2p+1)$ est un entier.

Donc $x_1 \in \mathbb{N} \iff \sqrt{\Delta} \in \mathbb{N}$.

Donc

$$\begin{aligned} \sqrt{\Delta} \in \mathbb{N} &\iff \exists a \in \mathbb{N}, (2p+1)^2 + 4\ell = a^2 \\ &\iff \ell = \frac{a^2 - (2p+1)^2}{4} \text{ où } a \in \mathbb{N} \text{ impair} \\ &\iff \ell = \frac{2b+1 - (2p+1)^2}{4} \text{ où } b \in \mathbb{N} \\ &\iff \ell = \frac{2(b-p) \times 2(b+p+1)}{4} \\ &\iff \ell = (b-p)(b+p+1) \text{ où } b \in \mathbb{N}. \end{aligned}$$

Si $b \in \llbracket 0 ; p \rrbracket$ alors $\ell \leq 0$ donc $x_1 \in \mathbb{N}$ et $x_2 \in \mathbb{N}$ et

$$\begin{aligned} (Y = \ell) &= \{\omega \in \Omega \mid X(\omega) = x_1 \text{ ou } X(\omega) = x_2\} \\ &= (X = x_1) \sqcup (X = x_2). \end{aligned}$$

Donc

$$\mathbb{P}(Y = \ell) = e^{-\lambda} \left(\frac{\lambda^{x_1}}{x_1!} + \frac{\lambda^{x_2}}{x_2!} \right)$$

où $x_1 = b+p+1$ et $x_2 = p-b$.

Si $b \in \llbracket p+1 ; +\infty \rrbracket$, alors $\ell > 0$ et donc $x_1 > 0$ et $x_2 < 0$.

Donc

$$\mathbb{P}(Y = \ell) = \mathbb{P}(X = x_1) = e^{-\lambda} \frac{\lambda^{x_1}}{x_1!}.$$

14.2 Espérance

14.2.1 Définitions

Définition 14.18

Soit X une variable aléatoire discrète à valeurs dans $[0 ; +\infty]$.

On appelle alors espérance de X le nombre $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \mathbb{P}(X = x)$ (au sens des familles positives) : c'est un réel si la famille $(x \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable et $+\infty$ sinon.

Une variable aléatoire réelle positive possède donc toujours une espérance, éventuellement infinie. Dans le cas général, on ne peut pas toujours définir l'espérance.

Définition 14.19

Soit X une variable aléatoire discrète complexe.

On dit que X possède une espérance finie quand la famille $(x \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas, on appelle espérance de X le nombre $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \mathbb{P}(X = x)$ (au sens des familles sommables).

On note L^1 l'ensemble des variables à espérance finie. On voit souvent écrit $X \in L^1$ pour signifier en abrégé que la variable X possède une espérance finie.

L'espérance de X est donc la moyenne de ses valeurs possibles, pondérées par leurs probabilités respectives, quand cela a un sens.

Exemple 14.20

- L'espérance d'une variable aléatoire certaine égale à a est a .
- Si A est un événement, l'espérance de son indicatrice est sa probabilité : $\mathbb{E}(1_A) = \mathbb{P}(A)$.

Exercice 14.21

Soient $\lambda \in \mathbb{R}$ et $\alpha > 0$.

Déterminez une CNS sur λ et α pour qu'existe une variable aléatoire X qui suit une loi $\left(\frac{\lambda}{n^\alpha}\right)_{n \in \mathbb{N}^*}$.

À quelle condition X possède-t-elle une espérance finie ?

► $\mathcal{L} = \left(\frac{\lambda}{n^\alpha} \right)_{n \in \mathbb{N}^*}$ est une distribution de probabilité discrète ssi $\begin{cases} \lambda \geq 0 \\ \sum_n \frac{\lambda}{n^\alpha} \text{ converge et } \sum_{n=1}^{+\infty} \frac{\lambda}{n^\alpha} = 1 \end{cases}$
i.e.
 $\alpha > 1 \quad \text{et} \quad \lambda = \frac{1}{\sum_{n=1}^{+\infty} 1/n^\alpha} = \frac{1}{\zeta(\alpha)}.$

Dans ce cas, il existe une variable aléatoire X telle que $X \sim \mathcal{L}$.

► On a

$$\begin{aligned} X \in L^1 &\iff (x\mathbb{P}(X=x))_{x \in X(\Omega)} \text{ est sommable} \\ &\iff \sum_{x \in X(\Omega)} |x| \mathbb{P}(X=x) < +\infty \\ &\iff \sum_{n \in \mathbb{N}^*} n \mathbb{P}(X=n) < +\infty \\ &\iff \sum_{n \in \mathbb{N}^*} \frac{\lambda}{n^{\alpha-1}} < +\infty \\ &\iff \sum_n \frac{\lambda}{n^{\alpha-1}} \text{ converge} \\ &\iff \alpha - 1 > 1 \\ &\iff \alpha > 2. \end{aligned}$$

Exercice 14.23

Calculez l'espérance d'une variable X qui suit une loi géométrique de paramètre $p \in]0 ; 1[$.

Correction 14.24

Soit $X \sim \mathcal{G}(p)$ où $p \in]0 ; 1[$. On a $X \geq 0$.

On a

$$\begin{aligned} \mathbb{E}(X) &= \sum_{n=1}^{+\infty} n (1-p)^{n-1} p \\ &= p \sum_{n=1}^{+\infty} n (1-p)^{n-1}. \end{aligned}$$

Or pour $x \in]-1 ; 1[$, on a $\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$ donc

$$\sum_{n=1}^{+\infty} n x^{n-1} = \frac{d}{dx} \left(\frac{1}{1-x} \right) = \frac{1}{(1-x)^2}.$$

Donc $\mathbb{E}(X) = \frac{1}{p}$.

Proposition 14.25

Toute variable aléatoire finie possède une espérance finie.

Toute variable aléatoire discrète complexe bornée possède une espérance finie.

Démonstration 14.26

Si X est une variable aléatoire bornée, alors on choisit $M > 0$ tel que $|X| \leq M$.

On a $\forall x \in X(\Omega), |x| \mathbb{P}(X = x) \leq M \mathbb{P}(X = x)$.

Or la famille $(\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable de somme 1.

Donc, par comparaison, la famille $(|x| \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable *i.e.* X possède une espérance finie. ■

14.2.2 Propriétés

Une propriété qui découle des théorèmes de comparaison classiques sur les familles sommables.

Proposition 14.27

Soient X et Y deux variables aléatoires discrètes complexes.

Si $|X| \leq |Y|$ et $Y \in L^1$, alors $X \in L^1$.

Si $X \in L^1$, alors $|X| \in L^1$ et $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$ (inégalité triangulaire).

La linéarité découle encore de la théorie des familles sommables.

Proposition 14.28

Soient X, Y deux variables aléatoires discrètes complexes admettant chacune une espérance finie et λ un complexe.

Alors $X + Y$ et λX admettent une espérance finie et

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y) \quad \text{et} \quad \mathbb{E}(\lambda X) = \lambda \mathbb{E}(X).$$

En particulier : $\forall (a, b) \in \mathbb{C}^2, \mathbb{E}(aX + b) = a \mathbb{E}(X) + b$.

Si X est une variable aléatoire réelle positive, alors $\mathbb{E}(X) \geq 0$.

Si X et Y sont deux variables aléatoires réelles telles que $X \leq Y$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

Démonstration 14.29 (Linéarité de l'espérance (somme))

On note $L = X(\Omega)$ et $M = Y(\Omega)$, qui sont par hypothèse des ensembles au plus dénombrables.

La variable aléatoire $X + Y$ est à valeurs dans $S = L + M = \{\ell + m \mid (\ell, m) \in L \times M\}$, qui est lui aussi un ensemble au plus dénombrable.

On pose $A = \sum_{s \in S} |s| \mathbb{P}(X + Y = s)$.

La famille $((X = \ell))_{\ell \in L}$ étant un système complet d'événements, d'après la formule des probabilités totales, on a

$$\forall s \in S, \quad \mathbb{P}(X + Y = s) = \sum_{\ell \in L} \mathbb{P}((X = \ell) \cap (X + Y = s)).$$

Or $(X = \ell) \cap (X + Y = s) = (X = \ell) \cap (Y = s - \ell)$ donc

$$A = \sum_{s \in S} |s| \sum_{\ell \in L} \mathbb{P}((X = \ell) \cap (X + Y = s)) = \sum_{\substack{\ell \in L \\ s \in S}} |s| \mathbb{P}((X = \ell) \cap (Y = s - \ell)).$$

Or, pour $s \in S$ et $\ell \in L$, soit il existe $b \in M$ tel que $s = \ell + b$ et dans ce cas

$$\mathbb{P}((X = \ell) \cap (Y = s - \ell)) = \mathbb{P}((X = \ell) \cap (Y = b))$$

soit $s - \ell \notin M$ et donc

$$\mathbb{P}((X = \ell) \cap (Y = s - \ell)) = 0.$$

Par changement de variables $m = s - \ell$, on a donc

$$A = \sum_{\substack{\ell \in L \\ m \in M}} |\ell + m| \mathbb{P}((X = \ell) \cap (Y = m))$$

car la première expression a les mêmes termes que la seconde, avec éventuellement des termes nuls supplémentaires.

Les familles $((X = \ell))_{\ell \in L}$ et $((Y = m))_{m \in M}$ étant des système complet d'événements, d'après la formule des probabilités totales, on a donc

$$\begin{aligned} A &\leq \sum_{\substack{\ell \in L \\ m \in M}} (|\ell| + |m|) \mathbb{P}((X = \ell) \cap (Y = m)) \\ &= \sum_{\ell \in L} \sum_{m \in M} |\ell| \mathbb{P}((X = \ell) \cap (Y = m)) + \sum_{m \in M} \sum_{\ell \in L} |m| \mathbb{P}((X = \ell) \cap (Y = m)) \\ &= \sum_{\ell \in L} |\ell| \sum_{m \in M} \mathbb{P}((X = \ell) \cap (Y = m)) + \sum_{m \in M} |m| \sum_{\ell \in L} \mathbb{P}((X = \ell) \cap (Y = m)) \\ &= \sum_{\ell \in L} |\ell| \mathbb{P}(X = \ell) + \sum_{m \in M} |m| \mathbb{P}(Y = m). \end{aligned}$$

Par sommabilité des familles $(\ell \mathbb{P}(X = \ell))_{\ell \in L}$ et $(m \mathbb{P}(Y = m))_{m \in M}$, on obtient

$$A < +\infty.$$

Ceci prouve la sommabilité de la famille $(s\mathbb{P}(X+Y=s))_{s \in S}$ i.e. la variable aléatoire $X+Y$ possède une espérance finie.

Maintenant que l'on a prouvé la sommabilité, on peut tout refaire sans les valeurs absolues et on a alors

$$\begin{aligned}\mathbb{E}(X+Y) &= \sum_{s \in S} s\mathbb{P}(X+Y=s) \\ &= \dots \\ &= \sum_{\substack{\ell \in L \\ m \in M}} (\ell+m) \mathbb{P}((X=\ell) \cap (Y=m)) \\ &= \dots \\ &= \sum_{\ell \in L} \ell \mathbb{P}(X=\ell) + \sum_{m \in M} m \mathbb{P}(Y=m) \\ &= \mathbb{E}(X) + \mathbb{E}(Y). \quad \blacksquare\end{aligned}$$

Démonstration 14.30 (Linéarité de l'espérance (multiplication par un scalaire))

En adaptant la preuve précédente, on montre que λX a une espérance finie et $\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$.

Si $\lambda = 0$, c'est évident. On suppose donc $\lambda \neq 0$.

Avec les mêmes notations, λX prend ses valeurs dans $S = \lambda L = \{\lambda \ell \mid \ell \in L\}$, qui est aussi au plus dénombrable.

On remarque $\forall \ell \in L, (\lambda X = \lambda \ell) = (X = \ell)$.

Donc $A = \sum_{s \in S} |s| \mathbb{P}(\lambda X = s) = \sum_{\ell \in L} |\lambda \ell| \mathbb{P}(\lambda X = \lambda \ell)$ car $t \mapsto \lambda t$ est une bijection de L dans S .

Donc $A = |\lambda| \sum_{\ell \in L} |\ell| \mathbb{P}(X = \ell) < +\infty$.

La famille $(s\mathbb{P}(\lambda X = s))_{s \in S}$ est donc sommable i.e. λX possède une espérance finie.

On peut donc reprendre les calculs précédents sans valeurs absolues et obtenir $\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$. ■

Proposition 14.31

Soit X une variable aléatoire discrète réelle positive.

Alors X a une espérance nulle ssi X est presque sûrement nulle, i.e. $\mathbb{P}(X=0)=1$.

Démonstration 14.32

\Rightarrow

Supposons $\mathbb{E}(X) = 0$.

Alors $\sum_{x \in X(\Omega)} \underbrace{x \mathbb{P}(X=x)}_{\geq 0} = 0$.

Donc $\forall x \in X(\Omega), \mathbb{P}(X = x) = 0$.

Donc $\forall x \in X(\Omega) \setminus \{0\}, \mathbb{P}(X = x) = 0$.

Or $\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = 1$ donc $\mathbb{P}(X = 0) = 1$.



Trivial. ■

Définition 14.33

On dit qu'une variable aléatoire discrète complexe X est centrée ssi $\mathbb{E}(X) = 0$.

À toute variable aléatoire discrète complexe X , on associe une variable aléatoire discrète complexe centrée : $X - \mathbb{E}(X)$.

14.2.3 Théorème de transfert

Théorème 14.34

Soient X une variable aléatoire discrète complexe et f une fonction de \mathbb{C} dans \mathbb{C} définie sur $X(\Omega)$.

La variable aléatoire discrète complexe $f(X)$ possède une espérance finie ssi la famille $(f(x) \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable et, dans ce cas, on a

$$\mathbb{E}(f(X)) = \sum_{x \in X(\Omega)} f(x) \mathbb{P}(X = x).$$

Démonstration 14.35

On note $L = X(\Omega)$ qui est au plus dénombrable et donc $S = f \circ X(\Omega) = \{f(\ell) \mid \ell \in L\}$ est au plus dénombrable.

Pour $s \in S$, on note A_s l'ensemble des antécédents dans L de s par $f : A_s = f^{-1}(\{s\})$.

On suppose que $f(X)$ a une espérance finie i.e. $B = \sum_{s \in S} |s| \mathbb{P}(f(X) = s) = \sum_{s \in S} |s| \mathbb{P}(X \in A_s) < +\infty$.

Or $(X \in A_s) = \bigsqcup_{\ell \in A_s} (X = \ell)$ (réunion au plus dénombrable) donc $\mathbb{P}(X \in A_s) = \sum_{\ell \in A_s} \mathbb{P}(X = \ell)$.

Donc

$$B = \sum_{s \in S} \sum_{\ell \in A_s} |s| \mathbb{P}(X = \ell) = \sum_{s \in S} \sum_{\ell \in A_s} |f(\ell)| \mathbb{P}(X = \ell) = \sum_{\ell \in L} |f(\ell)| \mathbb{P}(X = \ell)$$

car la famille de parties $(A_s)_{s \in S}$ est une partition de L donc le théorème de sommation par paquets s'applique.

Ceci prouve que la famille $(f(\ell) \mathbb{P}(X = \ell))_{\ell \in L}$ est sommable, donc on peut reprendre les calculs précédents sans valeurs absolues :

$$\sum_{\ell \in L} f(\ell) \mathbb{P}(X = \ell) = \sum_{s \in S} \sum_{\ell \in A_s} s \mathbb{P}(X = \ell) = \sum_{s \in S} s \mathbb{P}(X \in A_s) = \sum_{s \in S} s \mathbb{P}(f(X) = s) = \mathbb{E}(f(X)).$$

■

Ce théorème permet de calculer directement l'espérance de $f(X)$ sans devoir calculer la loi de $f(X)$: il suffit de connaître celle de X .

Exercice 14.36

Soient $p \in]0 ; 1[$ et $X \sim \mathcal{G}(p)$.

Calculez $\mathbb{E}(e^{-X})$.

Correction 14.37
On a $X \sim \mathcal{G}(p)$ i.e. $\begin{cases} X(\Omega) = \mathbb{N}^* \\ \forall k \in \mathbb{N}^*, \mathbb{P}(X = k) = (1-p)^{k-1} p \end{cases}$

D'après le théorème de transfert, on a

$$\begin{aligned} \mathbb{E}(e^{-X}) &= \sum_{k=1}^{+\infty} e^{-k} (1-p)^{k-1} p \\ &= \frac{p}{1-p} \sum_{k=1}^{+\infty} ((1-p)e^{-1})^k \\ &= \frac{p}{1-p} \times \frac{e^{-1}(1-p)}{1 - e^{-1}(1-p)} \\ &= \frac{p}{e(1 - (1-p)/e)} \\ &= \frac{p}{e - 1 + p}. \end{aligned}$$

Enfin, un petit résultat utile.

Proposition 14.38

Soit X une variable aléatoire discrète à valeurs dans \mathbb{N} .

$$\text{On a } \mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n).$$

Démonstration 14.39

Pour $n \in \mathbb{N}$, on a $(X \geq n) = (X = n) \sqcup (X \geq n+1)$ donc

$$\mathbb{P}(X \geq n) = \mathbb{P}(X = n) + \mathbb{P}(X \geq n+1).$$

Donc si X a une espérance finie alors

$$\begin{aligned} \mathbb{E}(X) &= \sum_{n=0}^{+\infty} n \mathbb{P}(X = n) \\ &= \sum_{n=1}^{+\infty} n (\mathbb{P}(X \geq n) - \mathbb{P}(X \geq n+1)). \end{aligned}$$

Or pour $N \in \mathbb{N}^*$, on a

$$\begin{aligned}
\sum_{n=1}^N n (\mathbb{P}(X \geq n) - \mathbb{P}(X \geq n+1)) &= \sum_{n=1}^N n \mathbb{P}(X \geq n) - \sum_{n=1}^N n \mathbb{P}(X \geq n+1) \\
&= \sum_{n=1}^N n \mathbb{P}(X \geq n) - \sum_{n=2}^{N+1} (n-1) \mathbb{P}(X \geq n) \\
&= \mathbb{P}(X \geq 1) - N \mathbb{P}(X \geq N+1) + \sum_{n=2}^N (n - (n-1)) \mathbb{P}(X \geq n) \\
&= \mathbb{P}(X \geq 1) + \sum_{n=2}^N \mathbb{P}(X \geq n) - N \mathbb{P}(X \geq N+1).
\end{aligned}$$

Or

$$\sum_{n=N+1}^{+\infty} n \mathbb{P}(X = n) \geq (N+1) \underbrace{\sum_{n=N+1}^{+\infty} \mathbb{P}(X = n)}_{=\mathbb{P}(X \geq N+1)} \geq 0.$$

Donc

$$0 \leq N \mathbb{P}(X \geq N+1) \leq \underbrace{\sum_{n=N+1}^{+\infty} n \mathbb{P}(X = n)}_{\text{reste d'une s\u00e9rie convergente}}.$$

Donc, par encadrement, $N \mathbb{P}(X \geq N+1) \xrightarrow[N \rightarrow +\infty]{} 0$.

Donc

$$\lim_{N \rightarrow +\infty} \left[\sum_{n=1}^N \mathbb{P}(X \geq n) - N \mathbb{P}(X \geq N+1) \right] = \mathbb{E}(X)$$

et

$$\lim_{N \rightarrow +\infty} N \mathbb{P}(X \geq N+1) = 0.$$

Donc, par op\u00e9rations sur les limites, on a

$$\mathbb{E}(X) = \lim_{N \rightarrow +\infty} \sum_{n=1}^N \mathbb{P}(X \geq n) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n).$$

■

14.2.4 In\u00e9galit\u00e9 de Markov

Proposition 14.40

Soit X une variable al\u00e9atoire discr\u00e8te r\u00e9elle positive admettant une esp\u00e9rance finie.

On a

$$\forall a > 0, \quad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Démonstration 14.41

Soit $a > 0$.

On a

$$\mathbb{E}(X) = \sum_{\substack{x \in X(\Omega) \\ x < a}} x \mathbb{P}(X = x) + \sum_{\substack{x \in X(\Omega) \\ x \geq a}} x \mathbb{P}(X = x)$$

donc

$$\sum_{\substack{x \in X(\Omega) \\ x \geq a}} x \mathbb{P}(X = x) \leq \mathbb{E}(X).$$

$$\text{Or } a \sum_{\substack{x \in X(\Omega) \\ x \geq a}} \mathbb{P}(X = x) \leq \sum_{\substack{x \in X(\Omega) \\ x \geq a}} x \mathbb{P}(X = x).$$

De plus, $(X \geq a) = \bigsqcup_{\substack{x \in X(\Omega) \\ x \geq a}} (X = x)$ (union dénombrable) donc

$$\sum_{\substack{x \in X(\Omega) \\ x \geq a}} \mathbb{P}(X = x) = \mathbb{P}(X \geq a).$$

Finalement, on a

$$a \mathbb{P}(X \geq a) \leq \mathbb{E}(X)$$

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

■

Cette inégalité n'a d'intérêt que pour $a > \mathbb{E}(X)$, sinon on majore une probabilité par un nombre plus grand que 1.

Il arrive souvent qu'on ne sache pas calculer explicitement la loi d'une variable aléatoire, mais qu'on arrive à calculer son espérance (par exemple comme somme de variables aléatoires simples). Les inégalités de ce type permettent quand même parfois d'obtenir des résultats à propos des probabilités de certains événements.

Corollaire 14.42

Soit X une variable aléatoire discrète complexe admettant une espérance finie.

On a

$$\forall a > 0, \quad \mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|)}{a}.$$

14.3 Variance d'une variable réelle

14.3.1 Moments d'ordre 2

Proposition 14.43

Soit X une variable aléatoire discrète réelle.

Si X^2 possède une espérance finie (on dit aussi que X possède un moment d'ordre 2), alors X possède une espérance finie.

Dans ce cas, pour tout $a \in \mathbb{R}$, $(X + a)^2$ possède une espérance finie.

Démonstration 14.44

On suppose que X^2 est L^1 .

On a $|X| \leq \frac{1+X^2}{2}$ et $\mathbb{E}\left(\frac{1+X^2}{2}\right) = \frac{1}{2} + \frac{1}{2}\mathbb{E}(X^2) < +\infty$.

Donc, par comparaison, $|X|$ est L^1 et donc X est L^1 .

De plus, pour $a \in \mathbb{R}$, $(X + a)^2 = X^2 + 2aX + a^2$ donc $(X + a)^2$ est une somme de variables aléatoires L^1 et est donc L^1 . ■

On note L^2 l'ensemble des variables aléatoires discrètes réelles X telles que X^2 est d'espérance finie, i.e. $\mathbb{E}(X^2) < +\infty$.

Proposition 14.45

Soient $X, Y \in L^2$.

On a $XY \in L^1$ et $\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$.

Démonstration 14.46

Comme $|XY| \leq \frac{X^2 + Y^2}{2}$, par comparaison, $XY \in L^1$.

Pour $t \in \mathbb{R}$, on pose $f(t) = \mathbb{E}\left((tX + Y)^2\right)$.

On a $f \geq 0$ et $f(t) = t^2\mathbb{E}(X^2) + 2t\mathbb{E}(XY) + \mathbb{E}(Y^2)$.

Si $\mathbb{E}(X^2) = 0$ alors, comme $X \geq 0$, X est presque sûrement nulle (Proposition 14.31), donc XY est presque sûrement nulle et donc

$$\mathbb{E}(XY)^2 = 0 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2) = 0.$$

Si $\mathbb{E}(X^2) \neq 0$ alors f est un trinôme du second degré de signe constant donc son déterminant est négatif ou nul et on a

$$\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2). \quad \blacksquare$$

Il s'agit bien sûr de l'inégalité de Cauchy-Schwarz.

14.3.2 Variance et écart-type

Définition 14.47

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$.

On appelle variance de X le nombre $\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$.

On appelle écart-type de X le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

La variance (ou l'écart-type) mesure la dispersion de X autour de sa moyenne. Une variable de variance 1 est dite réduite.

Proposition 14.48

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$.

On a $\mathbb{V}(X) = 0$ ssi X est presque sûrement constante.

En général, on calcule la variance par la formule suivante.

Proposition 14.49 (Formule de Huyghens)

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$. On a

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Exercice 14.50

Calculez la variance d'une variable suivant une loi géométrique de paramètre $p \in]0 ; 1[$.

Correction 14.51

Soit $X \sim \mathcal{G}(p)$.

On a $\mathbb{E}(X) = \frac{1}{p}$ et $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

D'après le théorème de transfert et en posant $q = 1 - p$, on a

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{k=1}^{+\infty} k^2 \mathbb{P}(X = k) \\ &= \sum_{k=1}^{+\infty} k^2 (1-p)^{k-1} p \\ &= \sum_{k=1}^{+\infty} (k(k-1) + k) q^{k-1} p \\ &= pq \sum_{k=1}^{+\infty} k(k-1) q^{k-2} + \underbrace{\sum_{k=1}^{+\infty} k q^{k-1} p}_{=\mathbb{E}(X)}.\end{aligned}$$

Sur $] -1 ; 1[$, la somme de la série entière $\sum_k x^k$ est \mathcal{C}^∞ et dérivable terme à terme.

En particulier :

$$\sum_{k=0}^{+\infty} k(k-1)x^{k-2} = \frac{d^2}{dx^2} \left(\frac{1}{1-x} \right) = \frac{2}{(1-x)^3}.$$

Donc

$$\mathbb{E}(X^2) = pq \frac{2}{(1-q)^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{1}{p}.$$

Donc

$$\mathbb{V}(X) = \frac{2q}{p} + \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}.$$

Proposition 14.52

Soient X une variable aléatoire discrète réelle telle que $X \in L^2$ et $(a, b) \in \mathbb{R}^2$.

On a $aX + b \in L^2$ et $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$.

Remarque 14.53

Si X est une variable aléatoire discrète réelle d'espérance m et de variance v , on pose $X' = \frac{1}{\sqrt{v}}(X - m)$.

X' est alors une variable aléatoire discrète réelle centrée dont la variance vaut 1 : on l'appelle la variable aléatoire discrète réelle centrée réduite associée à X .

14.3.3 Inégalité de Bienaymé-Tchebychev

Proposition 14.54

Soit X une variable aléatoire discrète réelle possédant un moment d'ordre 2. On a

$$\forall \varepsilon > 0, \quad \mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Démonstration 14.55

Soit $\varepsilon > 0$.

On a $|X - \mathbb{E}(X)| \geq \varepsilon \iff (X - \mathbb{E}(X))^2 \geq \varepsilon^2$.

Or $(X - \mathbb{E}(X))^2$ est une variable aléatoire discrète positive donc d'après l'inégalité de Markov, on a

$$\mathbb{P}\left((X - \mathbb{E}(X))^2 \geq \varepsilon^2\right) \leq \frac{\mathbb{E}\left((X - \mathbb{E}(X))^2\right)}{\varepsilon^2}$$

i.e.

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

■

La même remarque que pour l'inégalité de Markov s'applique : cette inégalité n'a d'intérêt que pour des valeurs assez grandes de ε , sinon on majore une probabilité par 1.

Exercice 14.56

On lance n fois un dé à six faces.

Comment choisir n pour que la probabilité d'obtenir moins d'une fois sur deux le nombre 6 soit au moins égale à $\frac{3}{4}$?

Correction 14.57

On pose X_n la variable aléatoire égale au nombre de 6 obtenus lors des n lancers.

On a $X_n \sim \mathcal{B}\left(n, \frac{1}{6}\right)$.

On cherche $n \in \mathbb{N}^*$ tel que $\mathbb{P}\left(X_n \leq \frac{n}{2}\right) \geq \frac{3}{4}$.

Or $\mathbb{P}\left(X_n \leq \frac{n}{2}\right) = 1 - \mathbb{P}\left(X_n > \frac{n}{2}\right) \geq \frac{3}{4}$ ssi $\mathbb{P}\left(X_n > \frac{n}{2}\right) \leq \frac{1}{4}$.

On a $\mathbb{E}(X_n) = \frac{n}{6}$.

D'après l'inégalité de Markov, on a

$$\mathbb{P}\left(X_n \geq \frac{n}{2}\right) \leq \frac{\mathbb{E}(X_n)}{n/2} = \frac{1}{3}$$

ce qui n'est pas suffisant.

On a $\mathbb{V}(X_n) = n \times \frac{1}{6} \times \frac{5}{6} = \frac{5n}{36}$.

D'après l'inégalité de Bienaymé-Tchebychev,

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|X_n - \frac{n}{6}\right| \geq \varepsilon\right) \leq \frac{5n/36}{\varepsilon^2}.$$

Or $\frac{n}{2} = \frac{n}{6} + \frac{n}{3}$ donc

$$\left|X_n - \frac{n}{6}\right| \geq \frac{n}{3} \iff \begin{cases} X_n - \frac{n}{6} \geq \frac{n}{3} \\ \text{ou} \\ X_n - \frac{n}{6} \leq \frac{-n}{3} \end{cases}$$

Donc

$$\left(\left|X_n - \frac{n}{6}\right| \geq \frac{n}{3}\right) = \left(X_n - \frac{n}{6} \geq \frac{n}{3}\right) \sqcup \left(X_n - \frac{n}{6} \leq \frac{-n}{3}\right).$$

Donc

$$\left(X_n \geq \frac{n}{2}\right) \subseteq \left(\left|X_n - \frac{n}{6}\right| \geq \frac{n}{3}\right).$$

Donc

$$\mathbb{P}\left(X_n \geq \frac{n}{2}\right) \leq \mathbb{P}\left(\left|X_n - \frac{n}{6}\right| \geq \frac{n}{3}\right) \leq \frac{5n/36}{(n/3)^2} = \frac{5}{4n}.$$

Or $\left(X_n > \frac{n}{2}\right) \subseteq \left(X_n \geq \frac{n}{2}\right)$ donc

$$\mathbb{P}\left(X_n > \frac{n}{2}\right) \leq \mathbb{P}\left(X_n \geq \frac{n}{2}\right) \leq \frac{5}{4n}.$$

Donc, pour que $\mathbb{P}\left(X_n > \frac{n}{2}\right) \leq \frac{1}{4}$, il suffit que $\frac{5}{4n} \leq \frac{1}{4}$ i.e. $n \geq 5$.

14.3.4 Généralisation

Définition 14.58

Soient X une variable aléatoire discrète réelle et $k \in \mathbb{N}^*$.

On dit que X admet un moment d'ordre k quand X^k a une espérance finie.

Proposition 14.59

Soient X une variable aléatoire discrète réelle et $k \in \mathbb{N}^*$.

Si X possède un moment d'ordre k , alors pour tout $\ell \in \llbracket 1 ; k \rrbracket$, X possède un moment d'ordre ℓ .

Démonstration 14.60

Si X a un moment d'ordre k , alors la famille $\left(x^k \mathbb{P}(X = x)\right)_{x \in X(\Omega)}$ est sommable.

Soit $\ell \in \llbracket 1 ; k - 1 \rrbracket$.

Pour $x \in X(\Omega)$,

- ▷ si $|x| \geq 1$, alors $|x|^\ell \leq |x|^k$ donc $|x|^\ell \mathbb{P}(X = x) \leq |x|^k \mathbb{P}(X = x)$;
- ▷ si $|x| < 1$, alors $|x|^\ell \mathbb{P}(X = x) \leq \mathbb{P}(X = x)$.

Donc

$$\begin{aligned} \sum_{x \in X(\Omega)} |x|^\ell \mathbb{P}(X = x) &\leq \sum_{\substack{x \in X(\Omega) \\ |x| \geq 1}} |x|^k \mathbb{P}(X = x) + \sum_{\substack{x \in X(\Omega) \\ |x| < 1}} \mathbb{P}(X = x) \\ &\leq \sum_{x \in X(\Omega)} |x|^k \mathbb{P}(X = x) + \sum_{x \in X(\Omega)} \mathbb{P}(X = x) \\ &= 1 + \mathbb{E}\left(|X|^k\right) < +\infty. \end{aligned}$$

Donc la famille $\left(x^\ell \mathbb{P}(X = x)\right)_{x \in X(\Omega)}$ est sommable, donc X possède un moment d'ordre ℓ . ■

14.4 Lois classiques

14.4.1 Loi uniforme

Définition 14.61

Soit L un ensemble fini non-vide.

On dit qu'une variable aléatoire X suit la loi uniforme sur L quand $X(\Omega) = L$ et \mathbb{P}_X est la probabilité uniforme sur L , autrement dit si pour tout $x \in L$, $\mathbb{P}(X = x) = \frac{1}{\text{Card } L}$.

On note alors $X \sim \mathcal{U}(L)$.

Le cas le plus courant est celui des variables à valeurs entières.

Définition 14.62

Soit $(a, b) \in \mathbb{N}^2$ tel que $a \leq b$.

On dit qu'une variable aléatoire suit la loi uniforme sur $\llbracket a ; b \rrbracket$ quand $X(\Omega) = \llbracket a ; b \rrbracket$ et \mathbb{P}_X est la probabilité uniforme sur $\llbracket a ; b \rrbracket$, i.e. si pour tout $k \in \llbracket a ; b \rrbracket$, $\mathbb{P}(X = k) = \frac{1}{b - a + 1}$.

On note alors $X \sim \mathcal{U}(\llbracket a ; b \rrbracket)$.

Dans ce cas, on a $\mathbb{E}(X) = \frac{a+b}{2}$ et $\mathbb{V}(X) = \frac{(b-a)(b-a+1)}{12}$.

Souvent, on a $a = 1$ et $b = n$ donc, dans ce cas, on a $\mathbb{E}(X) = \frac{n+1}{2}$ et $\mathbb{V}(X) = \frac{n^2-1}{12}$.

Exemple 14.63

Si on note X le nombre obtenu après un lancer d'un dé non-truqué, X suit la loi uniforme sur $\llbracket 1 ; 6 \rrbracket$.

14.4.2 Loi de Bernoulli

Définition 14.64

Soit $p \in [0 ; 1]$.

On dit qu'une variable aléatoire X suit la loi de Bernoulli de paramètre p quand $X(\Omega) \subseteq \{0, 1\}$ et $\mathbb{P}(X = 1) = p$.

On note alors $X \sim \mathcal{B}(p)$.

Dans ce cas, on note souvent $q = 1 - p$ et on a $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = pq$.

Exemple 14.65

- Toute expérience aléatoire à deux issues peut être représentée par une variable de Bernoulli en notant 0 et 1 les deux issues. Le cas typique est le lancer d'une pièce (équilibrée si $p = 1/2$, non-équilibrée sinon).
- En particulier, toute expérience dont seul la réussite ou l'échec importe peut être représentée par une variable de Bernoulli : 1 représente la réussite, 0 l'échec.

14.4.3 Loi binomiale

Définition 14.66

Soit $p \in [0 ; 1]$.

On dit qu'une variable aléatoire X suit la loi binomiale de paramètre (n, p) quand $X(\Omega) \subseteq \llbracket 0 ; n \rrbracket$ et pour tout $k \in \llbracket 0 ; n \rrbracket$, $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

On note alors $X \sim \mathcal{B}(n, p)$.

En notant $q = 1 - p$, on a $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = npq$.

Proposition 14.67

On considère une suite de n expériences aléatoires indépendantes qui suivent une loi de Bernoulli de même paramètre p .

On note X le nombre de réussites dans cette répétition d'expériences, appelée schéma de Bernoulli.

Alors X suit la loi binomiale de paramètre (n, p) .

Exemple 14.68

- On lance n fois une pièce dont la probabilité de tomber sur pile vaut p . Alors si X est le nombre de fois où on tombe sur pile, on a $X \sim \mathcal{B}(n, p)$.
- On fait n tirages successifs avec remise dans une urne contenant une proportion p de boules blanches. Le nombre de boules blanches tirées suit la loi binomiale de paramètre (n, p) .

14.4.4 Loi géométrique

Définition 14.69

Soit $p \in]0 ; 1[$.

On dit qu'une variable aléatoire X suit la loi géométrique de paramètre p quand $X(\Omega) \subseteq \mathbb{N}^* \cup \{+\infty\}$ et pour tout $k \in \mathbb{N}^* \cup \{+\infty\}$, $\mathbb{P}(X = k) = (1 - p)^{k-1} p$.

On note alors $X \sim \mathcal{G}(p)$.

En notant $q = 1 - p$, on a $\mathbb{P}(X = +\infty) = 0$, $\mathbb{E}(X) = \frac{1}{p}$ et $\mathbb{V}(X) = \frac{q}{p^2}$.

Proposition 14.70

La loi géométrique de paramètre p est la loi du rang du premier succès dans un schéma de Bernoulli infini de paramètre p .

Exemple 14.71

- On lance une pièce dont la probabilité de tomber sur pile vaut p . Alors si X est le premier rang pour lequel le lancer donne pile, on a $X \sim \mathcal{G}(p)$.
- On fait des tirages successifs avec remise dans une urne contenant une proportion p de boules blanches. Le premier rang où on tire une boule blanche suit la loi géométrique de paramètre p .

14.4.5 Loi de Poisson

Définition 14.72

Soit $\lambda > 0$.

On dit qu'une variable aléatoire X suit la loi de Poisson de paramètre λ quand $X(\Omega) = \mathbb{N}$ et pour tout $k \in \mathbb{N}$, $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$.

On note alors $X \sim \mathcal{P}(\lambda)$.

On a $\mathbb{E}(X) = \lambda$ et $\mathbb{V}(X) = \lambda$.

La loi de Poisson est la loi qui compte le nombre d'événements dans un intervalle de temps fixé, quand ces événements se produisent à une certaine fréquence connue et indépendamment du temps écoulé depuis le précédent. Le paramètre λ est alors le nombre moyen d'événements se produisant durant la durée fixée et X compte le nombre d'événements qui se sont effectivement produits durant la durée fixée.

Par exemple, si des statistiques montrent que le passage à une borne de péage se fait à une fréquence de f véhicules par heure, alors on choisit souvent la loi de Poisson $\mathcal{P}(f \times h)$ pour compter le nombre de véhicules qui passent dans une durée de h heures : $\mathbb{P}(X = k) = e^{-fh} \frac{(fh)^k}{k!}$ est la probabilité pour que sur un intervalle de temps de h heures, k voitures soient effectivement passées à la borne.

Souvent, la fréquence f est faible : on utilise cette loi pour l'étude des événements dits « rares ».

14.5 Couples de variables aléatoires

14.5.1 Généralités

Définition 14.73

Soient X et Y deux variables aléatoires discrètes.

La fonction $\omega \mapsto (X(\omega), Y(\omega))$ est appelée couple de variables aléatoires.

Si X et Y sont réelles, le couple est dit couple de variables aléatoires discrètes réelles.

On peut reprendre le même schéma de présentation que pour une seule variable aléatoire discrète.

Proposition 14.74

Soit (X, Y) un couple de variables aléatoires discrètes.

La famille d'événements $(\{X = x\} \cap \{Y = y\})_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est un système quasi-complet d'événements appelé système quasi-complet d'événements associé au couple (X, Y) .

La probabilité $\mathbb{P}(\{X = x\} \cap \{Y = y\})$ est souvent notée $\mathbb{P}(X = x, Y = y)$.

Définition 14.75

Soient X et Y deux variables aléatoires discrètes.

La loi conjointe de X et Y est la loi du couple (X, Y) .

Dans le cas d'un univers fini, on la représente souvent par un tableau à double entrée : si $X(\Omega) = \{x_1, \dots, x_m\}$ et $Y(\Omega) = \{y_1, \dots, y_n\}$, alors on place $\mathbb{P}(X = x_i, Y = y_j)$ sur la i -ème ligne et la j -ème colonne.

14.5.2 Lois marginales

Définition 14.76

Soient X et Y deux variables aléatoires discrètes.

Les lois marginales du couple (X, Y) sont les lois de X et de Y .

Proposition 14.77

Soient X et Y deux variables aléatoires discrètes. On a :

$$\triangleright \text{ pour tout } x \in X(\Omega), \quad \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y)$$

$$\triangleright \text{ pour tout } y \in Y(\Omega), \quad \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x, Y = y).$$

Si on a représenté la loi conjointe de (X, Y) sous forme d'un tableau, on obtient les lois marginales en additionnant les probabilités sur chaque ligne ou chaque colonne (d'où le nom « lois marginales » : celles qu'on note en marge du tableau).

Exercice 14.78

Une urne contient n boules numérotées de 1 à n .

On tire une boule avec remise et on note X le numéro de la boule.

Puis, on tire X boules et on note Y le maximum des numéros des boules tirées.

Déterminez la loi de X et la loi conjointe de (X, Y) , déduisez-en la loi de Y .

Correction 14.79

On a $X \sim \mathcal{U}(\llbracket 1 ; n \rrbracket)$ et (X, Y) est à valeurs dans $\llbracket 1 ; n \rrbracket^2$.

Soit $(k, \ell) \in \llbracket 1 ; n \rrbracket^2$.

On cherche $\mathbb{P}(X = k, Y = \ell)$.

On remarque $(X = k) \cap (Y \leq \ell) = [(X = k) \cap (Y = \ell)] \sqcup [(X = k) \cap (Y \leq \ell - 1)]$ donc

$$\mathbb{P}(X = k, Y = \ell) = \mathbb{P}(X = k, Y \leq \ell) - \mathbb{P}(X = k, Y \leq \ell - 1).$$

On cherche donc $\mathbb{P}(X = k, Y \leq \ell)$.

D'après la formule des probabilités composées, on a

$$\begin{aligned} \mathbb{P}(X = k, Y \leq \ell) &= \mathbb{P}(X = k) \mathbb{P}_{(X=k)}(Y \leq \ell) \\ &= \frac{1}{n} \mathbb{P}_{(X=k)}(Y \leq \ell). \end{aligned}$$

Dans l'expérience qui consiste à tirer simultanément k boules de l'urne, on suppose un tirage équiprobable des ensembles de k boules *i.e.*

$$\mathbb{P}_{(X=k)}(Y \leq \ell) = \frac{\binom{\ell}{k}}{\binom{n}{k}}.$$

$$\text{Donc } \mathbb{P}(X = k, Y \leq \ell) = \frac{1}{n} \frac{\binom{\ell}{k}}{\binom{n}{k}}.$$

$$\text{Donc } \mathbb{P}(X = k, Y = \ell) = \frac{1}{n \binom{n}{k}} \left(\binom{\ell}{k} - \binom{\ell-1}{k} \right).$$

Cas particulier : si $k > \ell$ alors $\mathbb{P}(X = k, Y = \ell) = 0$.

La famille $((X = k))_{k \in \llbracket 1; n \rrbracket}$ est un système complet d'événements donc, d'après la formule des probabilités totales, on a

$$\begin{aligned} \mathbb{P}(Y = \ell) &= \sum_{k=1}^n \mathbb{P}(X = k, Y = \ell) \\ &= \sum_{k=1}^n \underbrace{\frac{1}{n \binom{n}{k}} \left(\binom{\ell}{k} - \binom{\ell-1}{k} \right)}_{=0 \text{ si } \ell < k} \\ &= \sum_{k=1}^{\ell} \frac{1}{n \binom{n}{k}} \left(\binom{\ell}{k} - \binom{\ell-1}{k} \right). \end{aligned}$$

Or $\binom{\ell}{k} - \binom{\ell-1}{k} = \binom{\ell-1}{k-1}$ d'après la formule de Pascal.

Donc

$$\begin{aligned} \mathbb{P}(Y = \ell) &= \sum_{k=1}^{\ell} \frac{1}{n \binom{n}{k}} \binom{\ell-1}{k-1} \\ &= \dots \\ &= \frac{1}{n^2 \binom{n-1}{\ell-1}} \sum_{k=1}^{\ell} k \binom{n-k}{\ell-k}. \end{aligned}$$

Remarque 14.80

La loi d'une variable aléatoire ne peut pas dépendre d'une autre variable aléatoire !

Écrire par exemple que $Y(\Omega) = \llbracket X ; n \rrbracket$ dans l'exercice précédent est un non-sens car X n'est pas un nombre mais une variable aléatoire (*i.e.* une fonction!).

Écrire $Y(\Omega) = \llbracket i ; n \rrbracket$ où i est la valeur de X après le premier tirage est tout aussi absurde !

Il faut imaginer que l'expérience a lieu dans un lieu qui vous est caché et que l'opérateur ne vous donne que la valeur de Y . On laisse l'opérateur répéter son expérience un grand nombre de fois et on attend qu'il vous donne les valeurs de Y pour chacune des expériences. Qu'allez-vous obtenir ? Toutes les valeurs possibles de Y , qui peut varier de 1 à n . Certainement pas quelque chose qui dépend de X , puisque vous ne savez même pas quelle expérience réelle a lieu dans le lieu caché : pour vous, observateur extérieur, X n'existe pas.

14.5.3 Lois conditionnelles

Définition 14.81

Soient X et Y deux variables aléatoires discrètes.

Pour tout $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$, la loi de X sachant $(Y = y)$ est la loi de X dans l'espace probabilisé $(\Omega, \mathbb{P}_{(Y=y)})$:

$$\forall x \in X(\Omega), \quad \mathbb{P}_{(Y=y)}(X = x) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}.$$

On déduit de toutes les définitions précédentes les relations suivantes.

Proposition 14.82

Soient X et Y deux variables aléatoires discrètes.

On suppose que pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, $\mathbb{P}(X = x) \neq 0$ et $\mathbb{P}(Y = y) \neq 0$.

Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a :

$$\triangleright \mathbb{P}(X = x, Y = y) = \mathbb{P}(Y = y) \times \mathbb{P}_{(Y=y)}(X = x) = \mathbb{P}(X = x) \times \mathbb{P}_{(X=x)}(Y = y)$$

$$\triangleright \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(Y = y) \times \mathbb{P}_{(Y=y)}(X = x)$$

$$\triangleright \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x) \times \mathbb{P}_{(X=x)}(Y = y).$$

14.5.4 Covariance

Définition 14.83

Soient X et Y deux variables aléatoires discrètes réelles ayant une variance finie.

On appelle covariance du couple (X, Y) le nombre

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))).$$

Si, de plus, $\mathbb{V}(X)$ et $\mathbb{V}(Y)$ sont non-nulles, on appelle coefficient de corrélation du couple (X, Y) le nombre

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X) \sigma(Y)}.$$

Comme pour la variance, on a une expression plus simple.

Proposition 14.84

Avec les mêmes hypothèses :

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

L'application Cov a des propriétés classiques de bilinéarité.

Proposition 14.85

Cov est une application bilinéaire.

De plus, pour tout couple de variables aléatoires discrètes réelles (X, Y) ayant une variance finie, on a

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2 \text{Cov}(X, Y).$$

Démonstration 14.86

D'après la formule de Huygens, on a

$$\begin{aligned} \mathbb{V}(X + Y) &= \mathbb{E}\left((X + Y)^2\right) - \mathbb{E}(X + Y)^2 \\ &= \mathbb{E}(X^2 + Y^2 + 2XY) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY) - \left(\mathbb{E}(X)^2 + \mathbb{E}(Y)^2 + 2\mathbb{E}(X)\mathbb{E}(Y)\right) \\ &= \mathbb{V}(X) + \mathbb{V}(Y) + 2 \text{Cov}(X, Y). \end{aligned}$$

On généralise par récurrence :

$$\mathbb{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j).$$

■

Proposition 14.87 (Inégalité de Cauchy-Schwarz)

Pour tout couple de variables aléatoires discrètes réelles (X, Y) ayant une variance finie, on a

$$|\text{Cov}(X, Y)| \leq \sigma(X) \sigma(Y).$$

Si, de plus, $\sigma(X)$ et $\sigma(Y)$ sont non-nuls, on a $\rho(X, Y) \in [-1; 1]$ et $|\rho(X, Y)| = 1$ ssi il existe $(a, b) \in \mathbb{R}^2$ tel que $Y = aX + b$ presque sûrement.

14.6 Indépendance de variables aléatoires

14.6.1 Généralités

Définition 14.88

Soient X et Y deux variables aléatoires discrètes.

On dit que X et Y sont indépendantes quand pour toute partie A de $X(\Omega)$ et toute partie B de $Y(\Omega)$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants, *i.e.* quand

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B).$$

Dans ce cas, on note $X \perp Y$ pour signaler que X et Y sont indépendantes.

On peut se restreindre aux événements des systèmes quasi-complets d'événements associés à X et Y (résultat admis).

Proposition 14.89

Soient X et Y deux variables aléatoires discrètes.

X et Y sont indépendantes ssi pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants, *i.e.* si

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y).$$

Remarque 14.90

En général, la connaissance des lois marginales de (X, Y) ne permet pas de retrouver la loi conjointe de (X, Y) .

Dans le cas où les variables aléatoires X et Y sont indépendantes, c'est possible.

Donc pour montrer que deux variables X et Y ne sont pas indépendantes, il suffit de trouver deux valeurs x, y de X, Y respectivement telles que $\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x) \times \mathbb{P}(Y = y)$.

On a un résultat sur les composées de variables indépendantes.

Proposition 14.91

Soient X et Y deux variables aléatoires discrètes.

Si X et Y sont indépendantes, alors pour toute fonction f définie sur $X(\Omega)$ et toute fonction g définie sur $Y(\Omega)$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

14.6.2 Espérance et indépendance

On a un résultat remarquable sur les espérances de variables aléatoires discrètes indépendantes (résultat admis).

Proposition 14.92

Soient X et Y deux variables aléatoires discrètes.

Si X et Y sont indépendantes, alors on a :

$$\triangleright \mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$$

et dans le cas de variables réelles :

$$\triangleright \text{Cov}(X, Y) = 0$$

$$\triangleright \mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y).$$

Remarque 14.93

La réciproque est fausse.

Deux variables aléatoires discrètes réelles de covariance nulle sont dites non-corrélées, mais c'est un renseignement très faible sur les variables, à la différence de l'indépendance, qui est une contrainte extrêmement forte.

14.6.3 Généralisation

Tous les résultats sont admis.

Définition 14.94

Soit $(X_i)_{i \in I}$ une famille finie ou dénombrable de variables aléatoires discrètes.

On dit que les variables aléatoires $(X_i)_{i \in I}$ sont mutuellement indépendantes quand pour toute partie finie $J \subseteq I$, pour toute famille de parties $(A_j)_{j \in J}$ de $\prod_{j \in J} \mathcal{P}(X_j(\Omega))$, les événements $(\{X_j \in A_j\})_{j \in J}$ sont indépendants.

On retrouve la même caractérisation à l'aide des événements des systèmes quasi-complets d'événements associés aux différentes variables aléatoires et on peut même faire plus simple.

Proposition 14.95

Soit $(X_i)_{i \in I}$ une famille finie ou dénombrable de variables aléatoires discrètes.

Les variables aléatoires $(X_i)_{i \in I}$ sont mutuellement indépendantes quand pour toute partie finie $J \subseteq I$, pour toute famille $(x_j)_{j \in J}$ de $\prod_{j \in J} X_j(\Omega)$, on a

$$\mathbb{P} \left(\bigcap_{j \in J} \{X_j = x_j\} \right) = \prod_{j \in J} \mathbb{P}(X_j = x_j).$$

Exercice 14.96

Soient (X_1, X_2) et (Y_1, Y_2) deux couples de variables indépendantes telles que pour tout $i \in \{1, 2\}$, $X_i \sim Y_i$.

Montrez que $X_1 + X_2 \sim Y_1 + Y_2$.

Le résultat est-il encore valable en supprimant l'hypothèse d'indépendance ?

Correction 14.97

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

On veut montrer $\mathbb{P}(X_1 + X_2 = x) = \mathbb{P}(Y_1 + Y_2 = x)$.

Comme $(\{X_1 = x_1\})_{x_1 \in X_1(\Omega)}$ est un système complet d'événements, on a d'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}(X_1 + X_2 = x) &= \sum_{x_1 \in X_1(\Omega)} \mathbb{P}(X_1 + X_2 = x, X_1 = x_1) \\ &= \sum_{x_1 \in X_1(\Omega)} \mathbb{P}(X_2 = x - x_1, X_1 = x_1) \\ &= \sum_{x_1 \in X_1(\Omega)} \mathbb{P}(X_2 = x - x_1) \mathbb{P}(X_1 = x_1) \quad \left. \begin{array}{l} \\ \end{array} \right\} \begin{array}{l} X_1 \text{ et } X_2 \text{ indépendantes} \\ \end{array} \\ &= \sum_{x_1 \in X_1(\Omega)} \mathbb{P}(Y_2 = x - x_1) \mathbb{P}(Y_1 = x_1) \\ &= \mathbb{P}(Y_1 + Y_2 = x). \end{aligned}$$

Contre-exemple dans le cas où X_1 et X_2 ne sont pas indépendantes :

On prend $X \sim \mathcal{B}\left(\frac{1}{2}\right)$ et on pose $X_1 = X$ et $X_2 = X$.

On a alors $X_1 + X_2 \sim \mathcal{U}(\{0, 2\})$.

On prend $Y_1 = X$ et $Y_2 = 1 - X$.

Alors $X_1 \sim Y_1$ et $X_2 \sim Y_2$ mais $X_1 + X_2 = 2X \not\sim 1 = Y_1 + Y_2$.

Enfin, un petit lemme classique : le lemme des coalitions.

Proposition 14.98

Soit $(X_i)_{i \in I}$ une famille finie de variables aléatoires discrètes réelles indépendantes.

Soient J, K deux parties de I non-vides et disjointes, f une fonction de $\mathbb{R}^{\text{Card } J}$ dans \mathbb{R} et g une fonction de $\mathbb{R}^{\text{Card } K}$ dans \mathbb{R} .

Alors les deux variables aléatoires discrètes réelles $f(X_j, j \in J)$ et $g(X_k, k \in K)$ sont indépendantes.

En clair, si on partage des variables aléatoires discrètes réelles indépendantes en deux sous-ensembles n'ayant aucune variable en commun, alors n'importe quoi qui dépend uniquement des variables du premier ensemble est indépendant de n'importe quoi qui dépend uniquement des variables du deuxième ensemble.

Le résultat se généralise à toute famille dénombrable de variables aléatoires discrètes réelles indépendantes, puisque la définition même de l'indépendance impose de ne considérer que des sous-familles finies.

Exemple 14.99

Un cas très courant : si $(X_n)_{n \in \mathbb{N}}$ est une famille de variables aléatoires discrètes réelles indépendantes, alors :

- ▷ pour tout $n \in \mathbb{N}$, $\sum_{k=0}^n X_k$ est indépendante de X_{n+1}
- ▷ pour tout $(n, p) \in \mathbb{N}^2$ tel que $n < p$, $\sum_{k=0}^n X_k$ est indépendante de $\sum_{k=n+1}^p X_k$.

On en déduit la généralisation du résultat sur les variances.

Proposition 14.100

Soient X_1, \dots, X_n des variables aléatoires discrètes réelles indépendantes.

On a $\mathbb{V}(X_1 + \dots + X_n) = \mathbb{V}(X_1) + \dots + \mathbb{V}(X_n)$.

14.6.4 Théorème de réalisation

Le théorème de Kolmogorov est un théorème de réalisation sur les familles de variables aléatoires indépendantes (admis et jamais utilisé en pratique).

Théorème 14.101

Soit $(\mathcal{L}_i)_{i \in I}$ une famille finie ou dénombrable de lois de probabilités.

Il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et des variables aléatoires $(X_i)_{i \in I}$ indépendantes telles que pour tout $i \in I$, $X_i \sim \mathcal{L}_i$.

Un cas particulier très courant.

Définition 14.102

On dit qu'une famille finie ou dénombrable de variables aléatoires discrètes est identiquement distribuée quand toutes les variables aléatoires suivent la même loi.

Théorème 14.103

Soit \mathcal{L} une loi de probabilités.

Il existe un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et des variables aléatoires $(X_i)_{i \in I}$ indépendantes et identiquement distribuées qui suivent la loi \mathcal{L} .

Dans ce cas, on voit parfois écrit que la famille de variables aléatoires discrètes est une famille de variables indépendantes identiquement distribuées.

L'exemple typique est l'étude des suites infinies de lancers d'une pièce : on modélise cette expérience aléatoire par une suite de variables de Bernoulli indépendantes identiquement distribuées de même paramètre p .

14.6.5 Somme de variables indépendantes identiquement distribuées

Plusieurs résultats classiques.

Proposition 14.104

Soient X_1, \dots, X_n des variables aléatoires réelles.

Si ces n variables aléatoires sont mutuellement indépendantes et suivent la même loi de Bernoulli de paramètre p , alors la somme $X_1 + \dots + X_n$ suit la loi binomiale de paramètre (n, p) .

Proposition 14.105

Soient X, Y deux variables aléatoires réelles indépendantes suivant une loi de Poisson de paramètres $\lambda > 0$ et $\mu > 0$.

Alors $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.

Démonstration 14.106

On a $X(\Omega) = Y(\Omega) = \mathbb{N}$ donc $(X + Y)(\Omega) = \mathbb{N}$.

Soit $n \in \mathbb{N}$.

Comme $(\{X = k\})_{k \in \mathbb{N}}$ est un système complet d'événements, on a

$$\mathbb{P}(X + Y = n) = \sum_{k=0}^{+\infty} \mathbb{P}(X + Y = n, X = k).$$

Or $(X + Y = n, X = k) = (Y = n - k, X = k)$ donc par indépendance de X et Y , on a

$$\mathbb{P}(X + Y = n, X = k) = \mathbb{P}(X = k) \mathbb{P}(Y = n - k) = \begin{cases} 0 & \text{si } k > n \\ e^{-\lambda} \frac{\lambda^k}{k!} \times e^{-\mu} \frac{\mu^{n-k}}{(n-k)!} & \text{sinon} \end{cases}$$

Donc

$$\begin{aligned} \mathbb{P}(X + Y = n) &= \sum_{k=0}^n e^{-\lambda} \frac{\lambda^k}{k!} \times e^{-\mu} \frac{\mu^{n-k}}{(n-k)!} \\ &= \frac{e^{-(\lambda+\mu)}}{n!} \sum_{k=0}^n \binom{n}{k} \lambda^k \mu^{n-k} \\ &= e^{-(\lambda+\mu)} \frac{(\lambda + \mu)^n}{n!}. \end{aligned}$$

Donc $X + Y \sim \mathcal{P}(\lambda + \mu)$. ■

14.7 Loi faible des grands nombres

Proposition 14.107

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires discrètes réelles indépendantes identiquement distribuées admettant une variance finie.

En posant $S_n = \sum_{k=1}^n X_k$, $m = \mathbb{E}(X_1)$ et $\sigma^2 = \mathbb{V}(X_1)$, on a

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0.$$

Plus précisément, on a

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Démonstration 14.108

$$\text{On a } \mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n} \sum_{i=1}^n \underbrace{\mathbb{E}(X_i)}_{=m} = \frac{mn}{n} = m.$$

De plus, comme les X_i sont indépendantes, on a $\mathbb{V}\left(\frac{S_n}{n}\right) = \frac{1}{n^2} \mathbb{V}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \times n\sigma^2 = \frac{\sigma^2}{n}$.

Donc, d'après l'inégalité de Bienaymé-Tchebychev, on a

$$\forall \varepsilon > 0, \quad \mathbb{P} \left(\left| \frac{S_n}{n} - \mathbb{E} \left(\frac{S_n}{n} \right) \right| \geq \varepsilon \right) \leq \frac{\mathbb{V} \left(\frac{S_n}{n} \right)}{\varepsilon^2}$$

i.e.

$$\forall \varepsilon > 0, \quad \mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

■

14.8 Fonction génératrice d'une variable aléatoire à valeurs entières

14.8.1 Généralités

Définition 14.109

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

On appelle série génératrice de X la série entière $\sum_{n \geq 0} \mathbb{P}(X = n) z^n$.

Quand cette série converge, on appelle $G_X(z)$ sa somme : la fonction $z \mapsto G_X(z)$ est appelée fonction génératrice de X .

On peut écrire sa somme comme une espérance : $G_X(z) = \sum_{n=0}^{+\infty} z^n \mathbb{P}(X = n) = \mathbb{E}(z^X)$, à condition que la série $\sum_{n \geq 0} \mathbb{P}(X = n) z^n$ converge absolument.

Proposition 14.110

La série génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N} \cup \{+\infty\}$ est de rayon de convergence au moins 1 et converge normalement sur $[0 ; 1]$.

La fonction G_X est donc continue sur $[0 ; 1]$.

De plus, on a $G_X(1) = \mathbb{P}(X \neq +\infty) = 1 - \mathbb{P}(X = +\infty)$.

Démonstration 14.111

La suite $(\mathbb{P}(X = n) \times 1^n)_{n \in \mathbb{N}}$ est bornée donc, d'après le lemme d'Abel, la série entière $\sum_{n \geq 0} \mathbb{P}(X = n) z^n$ converge absolument pour tout $z \in \mathbb{C}$ tel que $|z| < 1$.

De plus, $\sum_{n \geq 0} \mathbb{P}(X = n)$ converge car $(\mathbb{P}(X = n))_{n \in \mathbb{N}}$ est un sous-système d'un système complet d'événements, donc la série $\sum_{n \geq 0} \mathbb{P}(X = n) z^n$ converge absolument pour tout $z \in \mathbb{C}$ tel que $|z| \leq 1$.

Donc le rayon de convergence est au moins 1.

Pour $x \in [0 ; 1]$, $|\mathbb{P}(X = n) x^n| \leq \mathbb{P}(X = n)$. Or $\mathbb{P}(X = n)$ est le terme général d'une série convergente indépendant de x .

Donc la série de fonctions $\sum_{n \geq 0} (x \mapsto \mathbb{P}(X = n) x^n)$ converge normalement sur $[0 ; 1]$. ■

Exercice 14.112

Déterminez la fonction génératrice associée à une variable de Bernoulli.

Correction 14.113

Soit $X \sim \mathcal{B}(p)$. On pose $q = 1 - p$.

On a

$$G_X(z) = \mathbb{P}(X = 0) z^0 + \mathbb{P}(X = 1) z^1 = q + pz.$$

Exercice 14.114

Même question avec une variable qui suit une loi binomiale.

Correction 14.115

Soit $X \sim \mathcal{B}(n, p)$. On pose $q = 1 - p$.

On a

$$\begin{aligned} G_X(z) &= \sum_{k=0}^n \mathbb{P}(X = k) z^k \\ &= \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} z^k \\ &= (q + pz)^n. \end{aligned}$$

Exercice 14.116

Même question avec une variable qui suit une loi uniforme sur $\llbracket 1 ; n \rrbracket$.

Correction 14.117

Soit $X \sim \mathcal{U}(\llbracket 1 ; n \rrbracket)$.

On a

$$G_X(z) = \sum_{k=0}^n \frac{z^k}{n+1} = \begin{cases} \frac{1}{n+1} \frac{1 - z^{n+1}}{1 - z} & \text{si } z \neq 1 \\ 1 & \text{sinon} \end{cases}$$

Exercice 14.118

Même question avec une variable qui suit une loi géométrique.

Correction 14.119

Soit $X \sim \mathcal{G}(p)$. On pose $q = 1 - p$.

On a

$$\begin{aligned} G_X(z) &= \sum_{k=1}^{+\infty} q^{k-1} p z^k \\ &= \frac{p}{q} \sum_{k=1}^{+\infty} (qz)^k \\ &= \frac{pqz}{1-p} \frac{1}{1-qz} \quad \left. \begin{array}{l} \text{rayon de convergence } \frac{1}{q} \end{array} \right\} \\ &= \frac{pz}{1-qz}. \end{aligned}$$

Exercice 14.120

Même question avec une variable qui suit une loi de Poisson.

Correction 14.121

Soit $X \sim \mathcal{P}(\lambda)$.

On a

$$\begin{aligned} G_X(z) &= \sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} z^k \\ &= e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda z)^k}{k!} \quad \left. \begin{array}{l} \text{rayon de convergence } +\infty \end{array} \right\} \\ &= e^{-\lambda} e^{\lambda z} \\ &= e^{\lambda(z-1)}. \end{aligned}$$

La fonction génératrice caractérise parfaitement la variable aléatoire.

Proposition 14.122

Soient X, Y deux variables aléatoires discrètes à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

Si $G_X = G_Y$ sur un intervalle $[0 ; \alpha]$ où $\alpha > 0$, alors $X \sim Y$.

14.8.2 Lien entre espérance et fonction génératrice

Quand $\mathbb{P}(X = +\infty) = 0$, on dit que la variable est presque sûrement à valeurs dans \mathbb{N} (exemple typique : les variables suivant des lois géométriques). C'est évidemment le cas quand X est à valeurs dans \mathbb{N} .

Théorème 14.123

Soit X une variable aléatoire discrète presque sûrement à valeurs dans \mathbb{N} .

X admet une espérance finie ssi G_X est dérivable en 1 et, dans ce cas, $\mathbb{E}(X) = G'_X(1)$.

X admet une variance finie ssi G_X est deux fois dérivable en 1 et, dans ce cas, $\mathbb{E}(X(X-1)) = G''_X(1)$.

Démonstration 14.124

G_X est définie et continue sur $[0; 1]$.

Pour $x \in [0; 1]$, on a $G_X(x) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n) x^n$.

Le rayon de convergence R de cette série étant au moins égal à 1, G_X est de classe \mathcal{C}^∞ sur $[0; 1[$, et si $R > 1$, on a même G_X de classe \mathcal{C}^∞ sur $[0; 1] \subseteq [0; R[$.

Si $R > 1$, on a $\forall x \in [0; 1]$, $G'_X(x) = \sum_{n=0}^{+\infty} n \mathbb{P}(X = n) x^{n-1}$, et donc

$$G'_X(1) = \sum_{n=0}^{+\infty} n \mathbb{P}(X = n) = \mathbb{E}(X).$$

Si $R = 1$, on a seulement $\forall x \in [0; 1[$, $G'_X(x) = \sum_{n=0}^{+\infty} n \mathbb{P}(X = n) x^{n-1}$. Si $X \in L^1$, la série $\sum_{n=0}^{+\infty} n \mathbb{P}(X = n)$ converge (absolument) donc d'après le théorème de convergence radiale, G'_X est définie et continue sur $[0; 1]$ et on a encore $G'_X(1) = \mathbb{E}(X)$.

Réciproquement, si G_X est dérivable en 1 à gauche :

On veut montrer que la série $\sum_{n=0}^{+\infty} n \mathbb{P}(X = n)$ converge.

On remarque $\forall n \in \mathbb{N}^*$, $x \mapsto n \mathbb{P}(X = n) x^{n-1}$ est croissante sur $[0; 1[$.

Donc G'_X est croissante sur $[0; 1[$.

Donc G'_X possède une limite $\ell \in \mathbb{R} \cup \{+\infty\}$ en 1^- .

On a G_X continue sur $[0; 1]$, dérivable sur $[0; 1[$ et $G'_X \xrightarrow{1^-} \ell$ donc d'après le théorème de la limite de la dérivée, si $\ell \in \mathbb{R}$, G_X est dérivable en 1^- et si $\ell = +\infty$, G_X n'est pas dérivable en 1^- .

Or G_X est dérivable en 1 par hypothèse donc $\ell \in \mathbb{R}$.

On en déduit $G'_X(1) = \ell$ et on a

$$\begin{aligned} \sum_{n=0}^N n \mathbb{P}(X = n) x^{n-1} &\leq \sum_{n=0}^{+\infty} n \mathbb{P}(X = n) x^{n-1} \\ &= G'_X(x) \\ &\leq \ell \in \mathbb{R}. \end{aligned}$$

Donc par passage à la limite quand $x \rightarrow 1^-$, on a

$$\sum_{n=0}^N n \mathbb{P}(X = n) x^{n-1} \leq \ell.$$

Les sommes partielles de la série à termes positifs $\sum_{n \geq 0} n \mathbb{P}(X = n)$ sont donc majorées par ℓ donc la série converge.

Ainsi, $X \in L^1$ et $\mathbb{E}(X) = G'_X(1)$.

Le raisonnement est similaire pour la variance et on a

$$\mathbb{V}(X) = \mathbb{E}(X(X-1)) - \mathbb{E}(X)(\mathbb{E}(X)-1) = G''_X(1) - G'_X(1)(G'_X(1)-1). \quad \blacksquare$$

Quand $\mathbb{P}(X = +\infty) > 0$, X n'est pas d'espérance finie (on a $\mathbb{E}(X) = +\infty$ dans ce cas).

14.8.3 Fonction génératrice d'une somme de variables indépendantes

Proposition 14.125

Soient X, Y deux variables aléatoires discrètes à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

Si X et Y sont indépendantes, alors $G_{X+Y} = G_X G_Y$.

Démonstration 14.126

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

$$\begin{aligned} \mathbb{P}(X+Y = n) &= \sum_{k=0}^n \mathbb{P}(X+Y = n, X = k) \\ &= \sum_{k=0}^n \mathbb{P}(Y = n-k, X = k) \\ &= \sum_{k=0}^n \mathbb{P}(Y = n-k) \mathbb{P}(X = k). \end{aligned}$$

On reconnaît un produit de Cauchy des séries absolument convergentes $\sum_{k \geq 0} \mathbb{P}(X = k) x^k$ et $\sum_{k \geq 0} \mathbb{P}(Y = k) x^k$ pour $x \in [0; 1]$.

Donc $G_{X+Y}(x) = G_X(x) G_Y(x)$ pour tout $x \in [0; 1]$. ■

Plus généralement, si X_1, \dots, X_n sont indépendantes, alors $G_{X_1+\dots+X_n} = \prod_{k=1}^n G_{X_k}$.

Exemple 14.127

Si X et Y sont indépendantes et suivent des lois de Poisson de paramètres $\lambda > 0$ et $\mu > 0$, alors $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$.

Chapitre 15

Équations différentielles linéaires

Sommaire

15.1	Équations et systèmes différentiels linéaires du premier ordre . . .	457
15.1.1	Généralités	457
15.1.2	Problème de Cauchy	459
15.1.3	Équation différentielle linéaire homogène	460
15.1.4	Cas général	463
15.2	Exponentielle d'un endomorphisme, d'une matrice	464
15.2.1	Définition	464
15.2.2	Propriétés algébriques	464
15.2.3	Propriétés fonctionnelles	465
15.3	Équations linéaires homogènes à coefficients constants.	467
15.3.1	Forme générale de la solution	467
15.3.2	Cas praticable	467
15.4	Équations différentielles linéaires scalaires d'ordre n.	471
15.4.1	Généralités	471
15.4.2	Représentation matricielle	472
15.4.3	Équation différentielle linéaire homogène	473
15.4.4	Cas général	473
15.4.5	Diverses idées pour résoudre une équation différentielle linéaire d'ordre n . .	474

Dans tout le cours, I désigne un intervalle de \mathbb{R} contenant au moins deux points, n désigne un entier naturel non-nul, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} , et E est un espace vectoriel normé de dimension finie n .

Si f est un endomorphisme de E et v un vecteur de E , on note $f.v$ l'image de v par f , au lieu de la traditionnelle notation $f(v)$. Sans cette notation, on rencontrerait des notations lourdes comme $f(t)(v)$ ou $f(t)(v(t))$, qui seront donc avantageusement remplacées par $f(t).v$ ou $f(t).v(t)$.

15.1 Équations et systèmes différentiels linéaires du premier ordre

15.1.1 Généralités

Définition 15.1

On appelle équation différentielle linéaire (du premier ordre) toute équation différentielle de la forme

$$x' = a(t) \cdot x + b(t)$$

où

- ▷ a est une fonction continue de I dans $\mathcal{L}(E)$
- ▷ b est une fonction continue de I dans E (appelée second membre)
- ▷ x est la fonction inconnue dérivable de I dans E et x' sa dérivée.

Définition 15.2

Avec les notations précédentes, on appelle solution (sur I) de l'équation différentielle toute application x dérivable sur I qui vérifie $\forall t \in I, x'(t) = a(t) \cdot x(t) + b(t)$.

Remarque 15.3

Si le second membre b peut s'écrire sous la forme $b = b_1 + b_2$, alors en additionnant les solutions de l'équation $x' = a(t) \cdot x + b_1(t)$ avec celles de l'équation $x' = a(t) \cdot x + b_2(t)$, on obtient des solutions de $x' = a(t) \cdot x + b(t)$.

On appelle cette idée le principe de superposition.

Comme E est de dimension finie n , si on choisit une base \mathcal{B} , alors en exprimant les coordonnées des vecteurs et les matrices des endomorphismes dans la base \mathcal{B} , on obtient une traduction matricielle de l'équation différentielle précédente.

Définition 15.4

On appelle système différentiel linéaire (du premier ordre) toute équation différentielle de la forme

$$X' = A(t) X + B(t)$$

où

- ▷ A est une fonction continue de I dans $\mathcal{M}_n(\mathbb{K})$
- ▷ B est une fonction continue de I dans $\mathcal{M}_{n,1}(\mathbb{K}) = \mathbb{K}^n$
- ▷ X est la fonction inconnue dérivable de I dans $\mathcal{M}_{n,1}(\mathbb{K})$ et X' sa dérivée.

Exemple 15.5

► Le système $\begin{cases} x' = 2x + 3y \\ y' = -x + 4y \end{cases}$ est un système différentiel linéaire où A est la fonction constante égale à $\begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix}$ et $B = 0$.

► Le système $\begin{cases} x' = t^2x + e^t y + \sin t \\ y' = -tx + \cos(t)y - 1 \end{cases}$ est un système différentiel linéaire où A est la fonction $t \mapsto \begin{pmatrix} t^2 & e^t \\ -t & \cos t \end{pmatrix}$ et B est la fonction $t \mapsto \begin{pmatrix} \sin t \\ -1 \end{pmatrix}$.

Définition 15.6

Avec les notations précédentes, on appelle solution (sur I) du système différentiel toute application X dérivable sur I qui vérifie $\forall t \in I, X'(t) = A(t)X(t) + B(t)$.

Proposition 15.7

Les solutions d'une telle équation différentielle ou d'un tel système différentiel linéaire sont au moins de classe \mathcal{C}^1 sur I .

Démonstration 15.8

L'application $\Phi : \mathcal{L}(E) \times E \longrightarrow E$ est bilinéaire.

$$(f, v) \longmapsto f.v$$

E étant de dimension finie, $\mathcal{L}(E)$ aussi, donc Φ est continue.

Donc si x est continue de I dans E et a est continue de I dans $\mathcal{L}(E)$, alors $t \mapsto a(t).x(t)$ est continue sur I par composition de fonctions continues.

Par conséquent, si x est solution de l'équation différentielle linéaire $x' = a(t).x + b(t)$, alors x' est continue sur I , donc x est de classe \mathcal{C}^1 sur I . ■

À partir de maintenant, le cours portera sur les équations différentielles, mais tout ce qui est énoncé à leur propos est valable pour les systèmes différentiels.

15.1.2 Problème de Cauchy

Définition 15.9

Soit $(*) : x' = a(t).x + b(t)$ une équation différentielle linéaire qui respecte les conditions ci-dessus. On choisit un réel $t_0 \in I$ et un vecteur $v_0 \in E$.

Le problème d'existence et d'unicité d'une solution de $(*)$ qui satisfait la condition initiale $x(t_0) = v_0$ est appelé un problème de Cauchy.

Le problème de Cauchy

$$(1) : \begin{cases} x' = a(t) \cdot x + b(t) \\ x(t_0) = v_0 \end{cases}$$

est équivalent au problème intégral

$$(2) : x(t) = \int_{t_0}^t (a(\tau) \cdot x(\tau) + b(\tau)) d\tau + v_0$$

i.e. x est solution de (1) ssi x est solution de (2).

15.1.3 Équation différentielle linéaire homogène

Définition 15.10

On appelle équation différentielle linéaire homogène (du premier ordre) toute équation différentielle de la forme

$$x' = a(t) \cdot x$$

où a est une fonction continue de I dans $\mathcal{L}(E)$.

Il y a un seul résultat à connaître sur de telles équations : le théorème de Cauchy-Lipschitz (ou théorème de Cauchy linéaire).

Théorème 15.11

L'ensemble des solutions d'une équation différentielle linéaire homogène $x' = a(t) \cdot x$ est un sous-espace vectoriel de dimension n de $\mathcal{F}(I, \mathbb{K})$.

Plus précisément, en notant S l'ensemble des solutions, pour tout $t_0 \in I$, l'application

$$\begin{array}{ccc} S & \longrightarrow & E \\ x & \longmapsto & x(t_0) \end{array}$$

est un isomorphisme.

Autrement dit, pour tout $t_0 \in I$ et $v_0 \in E$ (appelé vecteur des conditions initiales), il existe une unique solution telle que $x(t_0) = v_0$.

Démonstration 15.12 (Cas particulier : $I = [c ; b]$ est un compact)

On choisit $\alpha > 0$ à ajuster ultérieurement et on travaille sur $[t_0 ; t_0 + \alpha] = J$.

On choisit une norme $\|\cdot\|$ sur E à laquelle on associe $\|\|\cdot\|\|$ sur $\mathcal{L}(E)$.

a étant continue sur $[c ; b]$, $t \mapsto \|\|a(t)\|\|$ est continue sur le segment $[c ; b]$.

D'après le théorème des bornes atteintes, $M = \sup_{t \in I} \|\|a(t)\|\|$ existe dans \mathbb{R} .

On pose $x_0 : t \mapsto v_0$ et pour tout $k \in \mathbb{N}$

$$x_{k+1} : t \mapsto \int_{t_0}^t a(\tau) \cdot x_k(\tau) d\tau + v_0.$$

On veut montrer que la suite de fonctions (x_k) converge uniformément vers une fonction x .

Par récurrence immédiate, les fonctions (x_k) sont \mathcal{C}^1 sur I .

Pour $k \in \mathbb{N}^*$, par linéarité de l'intégrale et de a , on a

$$x_{k+1}(t) - x_k(t) = \int_{t_0}^t a(\tau) \cdot (x_k(\tau) - x_{k-1}(\tau)) d\tau.$$

Pour tout $t \in J$, on a donc

$$\begin{aligned} \|x_{k+1}(t) - x_k(t)\| &\leq \int_{t_0}^t \|a(\tau) \cdot (x_k(\tau) - x_{k-1}(\tau))\| d\tau \\ &\leq \int_{t_0}^t \|a(\tau)\| \|x_k(\tau) - x_{k-1}(\tau)\| d\tau \\ &\leq M \int_{t_0}^t \|x_k(\tau) - x_{k-1}(\tau)\| d\tau \\ &\leq M \|x_k - x_{k-1}\|_\infty^J (t - t_0) \\ &\leq M\alpha \|x_k - x_{k-1}\|_\infty^J. \end{aligned}$$

Donc $\|x_{k+1} - x_k\|_\infty^J \leq M\alpha \|x_k - x_{k-1}\|_\infty^J$.

Maintenant, on choisit α tel que $M\alpha < 1$.

Par récurrence immédiate, on en déduit que

$$\forall k \in \mathbb{N}, \|x_{k+1} - x_k\|_\infty^J \leq \underbrace{(M\alpha)^k \|x_k - x_{k-1}\|_\infty^J}_{\text{terme général d'une série convergente}}.$$

Donc la série de fonctions $\sum_{k \geq 0} (x_{k+1} - x_k)$ converge normalement et donc uniformément sur J i.e. la

suite (x_k) converge uniformément sur J vers la fonction $x = \sum_{k=0}^{+\infty} (x_{k+1} - x_k) + x_0$.

Pour tout $k \in \mathbb{N}$ et $t \in J$, on a

$$x_{k+1}(t) = \int_{t_0}^t a(\tau) \cdot x_k(\tau) d\tau + v_0.$$

Or pour tout $\tau \in J$, on a

$$\begin{aligned} \|a(\tau) \cdot (x_k(\tau) - x(\tau))\| &\leq M \|x_k(\tau) - x(\tau)\| \\ &\leq M \|x_k - x\|_\infty^J \\ &\xrightarrow[k \rightarrow +\infty]{} 0. \end{aligned}$$

Donc $(t \mapsto a(t) \cdot x_k(t))_{k \in \mathbb{N}}$ converge uniformément vers $t \mapsto a(t) \cdot x(t)$ sur J .

Donc par interversion limite-intégrale, on a

$$\forall t \in J, \quad x(t) = \int_{t_0}^t a(\tau) x(\tau) d\tau + v_0.$$

On remarque que ce procédé peut être réalisé sur tout intervalle de longueur α et indépendamment de v_0 .

On réitère sur $[t_0 + \alpha ; t_0 + 2\alpha]$ avec la condition initiale $x(t_0 + \alpha) = x(t_0 + \alpha)$, etc.

On peut recouvrir I par un nombre fini d'intervalles $[t_0 + j\alpha ; t_0 + (j+1)\alpha]$ où $j \in \mathbb{Z}$.

En recollant les solutions locales, on obtient une solution globale sur I (et unique par unicité locale). ■

Démonstration 15.13 (Cas général)

On recolle sur I les solutions sur chaque segment inclus dans I .

Cela justifie le théorème car la linéarité est triviale et car on a unicité donc isomorphisme. ■

La conjonction d'une équation différentielle linéaire homogène et d'une condition initiale s'appelle un problème de Cauchy linéaire.

Ce qui précède signifie que n'importe quel problème de Cauchy linéaire possède une unique solution.

Une base de l'ensemble des solutions est appelé un système fondamental de solutions.

Exercice 15.14

Montrez que les fonctions $t \mapsto \begin{pmatrix} t \cos t \\ t \sin t \end{pmatrix}$ et $t \mapsto \begin{pmatrix} -t \sin t \\ t \cos t \end{pmatrix}$ forment un système fondamental de solutions du système différentiel sur \mathbb{R}_+^*
$$\begin{cases} x' = \frac{1}{t}x - y \\ y' = x + \frac{1}{t}y \end{cases}$$

Correction 15.15

On pose $X = \begin{pmatrix} x \\ y \end{pmatrix}$. Le système s'écrit $X'(t) = A(t) X(t)$ où $A(t) = \begin{pmatrix} 1/t & -1 \\ 1 & 1/t \end{pmatrix}$.

D'après le théorème de Cauchy-Lipschitz, l'ensemble des solutions est un espace vectoriel de dimension 2.

Les fonctions $t \mapsto \begin{pmatrix} t \cos t \\ t \sin t \end{pmatrix}$ et $t \mapsto \begin{pmatrix} -t \sin t \\ t \cos t \end{pmatrix}$ sont solutions et non-colinéaires, donc elles forment une base \mathcal{S} de l'ensemble des solutions *i.e.* un système fondamental de solutions.

On a $\mathcal{S} = \left\{ t \mapsto \begin{pmatrix} at \cos t - bt \sin t \\ at \sin t + bt \cos t \end{pmatrix} \mid (a, b) \in \mathbb{R}^2 \right\}$.

Exercice 15.16

Déterminez un système fondamental de solutions polynomiales du système

$$\text{différentiel } \begin{cases} x' = \frac{t}{1+t^2}x + \frac{1}{1+t^2}y \\ y' = \frac{-1}{1+t^2}x + \frac{t}{1+t^2}y \end{cases}$$

Correction 15.17

Si $\begin{pmatrix} x \\ y \end{pmatrix}$ est solution avec $x = \lambda$ constante, la première équation donne $y(t) = -\lambda t$ et donc

$$y'(t) = -\lambda = \frac{-1}{1+t^2}\lambda + \frac{t}{1+t^2}(-\lambda)t$$

est vraie.

Donc $t \mapsto \begin{pmatrix} -1 \\ t \end{pmatrix}$ est solution.

Idem, $t \mapsto \begin{pmatrix} t \\ 1 \end{pmatrix}$ est solution.

D'après le théorème de Cauchy-Lipschitz, $\dim \mathcal{S} = 2$ donc

$$\mathcal{S} = \left\{ t \mapsto \begin{pmatrix} -a + bt \\ at + b \end{pmatrix} \mid (a, b) \in \mathbb{R}^2 \right\}.$$

15.1.4 Cas général

Proposition 15.18

Soit $x' = a(t)x + b(t)$ une équation différentielle linéaire sur I .

Si on connaît une solution p dite particulière de l'équation, alors on connaît la forme générale des solutions : elles sont toutes de la forme

$$t \mapsto g(t) + p(t)$$

où g est la solution générale de l'équation différentielle linéaire homogène associée.

Autrement dit, l'ensemble des solutions est un sous-espace affine de $\mathcal{F}(I, \mathbb{K}) : p + S$ où S est l'ensemble des solutions de l'équation homogène associée.

Le théorème de Cauchy-Lipschitz est encore valable sous une forme amenuisée.

Théorème 15.19

Pour tout $t_0 \in I$ et $v_0 \in E$, il existe une unique solution telle que $x(t_0) = v_0$.

Remarque 15.20

Il n'existe pas de méthode simple pour résoudre concrètement de telles équations différentielles linéaires, à part dans le cas où a est une constante.

15.2 Exponentielle d'un endomorphisme, d'une matrice

15.2.1 Définition

Proposition 15.21

Soit $u \in \mathcal{L}(E)$. La série $\sum_{k \geq 0} \frac{u^k}{k!}$ converge absolument dans $\mathcal{L}(E)$.

Soit $A \in \mathcal{M}_n(\mathbb{K})$. La série $\sum_{k \geq 0} \frac{A^k}{k!}$ converge absolument dans $\mathcal{M}_n(\mathbb{K})$.

Définition 15.22

Avec les mêmes hypothèses, on pose

$$\exp(u) = e^u = \sum_{k=0}^{+\infty} \frac{u^k}{k!} \quad \text{et} \quad \exp(A) = e^A = \sum_{k=0}^{+\infty} \frac{A^k}{k!}.$$

Il est bien sûr évident que si u a pour matrice A dans une base \mathcal{B} , alors e^u a pour matrice e^A dans cette même base.

15.2.2 Propriétés algébriques

L'exponentielle d'endomorphismes ou de matrices a de nombreuses propriétés communes avec l'exponentielle de nombres, mais pas toutes, car les algèbres $\mathcal{L}(E)$ et $\mathcal{M}_n(\mathbb{K})$ ne sont pas commutatives.

Proposition 15.23

Pour tout $u \in \mathcal{L}(E)$, u et e^u commutent. De même avec les matrices.

Proposition 15.24

Soient u, v deux endomorphismes de E qui commutent. On a $e^{u+v} = e^u e^v = e^v e^u$.

De même avec des matrices de $\mathcal{M}_n(\mathbb{K})$ qui commutent.

Attention ! Sans hypothèse de commutation, ce résultat est faux en général.

Une conséquence immédiate est que pour tout $u \in \mathcal{L}(E)$, e^u est inversible (donc non nul) et son inverse est e^{-u} . De même pour les matrices.

Si A et B sont semblables, alors e^A et e^B le sont aussi avec la même matrice de passage : si $B = PAP^{-1}$ alors $e^B = Pe^AP^{-1}$.

En général, il est difficile de calculer une exponentielle d'endomorphisme ou de matrice. Deux cas simples :

Dans le cas diagonalisable, c'est faisable, car l'exponentielle d'une matrice diagonale est très simple.

Proposition 15.25

Soit $A = \text{diag}(\lambda_1, \dots, \lambda_n) \in \mathcal{M}_n(\mathbb{K})$.

On a $e^A = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$.

Démonstration 15.26

On a $\forall k \in \mathbb{N}$, $A^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$.

$$\text{Donc } \sum_{k=0}^N \frac{A^k}{k!} = \sum_{k=0}^N \frac{\text{diag}(\lambda_1^k, \dots, \lambda_n^k)}{k!} = \begin{pmatrix} \sum_{k=0}^N \frac{\lambda_1^k}{k!} & 0 & \dots & 0 \\ 0 & \sum_{k=0}^N \frac{\lambda_2^k}{k!} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sum_{k=0}^N \frac{\lambda_n^k}{k!} \end{pmatrix}.$$

Par passage à la limite, on obtient

$$e^A = e^D = \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n} \end{pmatrix}. \quad \blacksquare$$

Si une matrice A est nilpotente, alors son exponentielle est en fait un polynôme en A .

Dans le cas général, le calcul est souvent pénible. Néanmoins, on a quelques résultats généraux simples.

Proposition 15.27

Soit $A \in \mathcal{M}_n(\mathbb{K})$ scindée.

On a $\text{Sp}(e^A) = e^{\text{Sp}(A)}$.

En particulier, on a $\det e^A = e^{\text{tr} A}$.

15.2.3 Propriétés fonctionnelles

La convergence absolue de la série définissant l'exponentielle entraîne quelques résultats simples à propos de la continuité et de la dérivabilité.

Proposition 15.28

Les fonctions $u \mapsto e^u$ et $A \mapsto e^A$ sont continues sur $\mathcal{L}(E)$ et $\mathcal{M}_n(\mathbb{K})$ respectivement.

Démonstration 15.29 (Continuité de $u \mapsto e^u$)

On choisit une norme $\|\cdot\|$ sous-multiplicative sur $\mathcal{L}(E)$.

Soit $R > 0$.

On a $\forall u \in \overline{B}(0, R), \forall k \in \mathbb{N}, \|u^k\| \leq \|u\|^k \leq R^k$.

Donc $\forall k \in \mathbb{N}, \left\| \frac{u^k}{k!} \right\| \leq \frac{R^k}{k!}$.

Donc la série $\sum_{k=0}^{+\infty} \frac{u^k}{k!}$ converge normalement sur $\overline{B}(0, R)$.

Or les fonctions $u \mapsto u^k$ sont continues.

Donc $u \mapsto e^u = \sum_{k=0}^{+\infty} \frac{u^k}{k!}$ est continue sur $\overline{B}(0, R)$.

Donc par réunion des boules $\overline{B}(0, R)$, $u \mapsto e^u$ est continue sur $\mathcal{L}(E) = \bigcup_{R>0} \overline{B}(0, R)$. ■

Proposition 15.30

Soit $u \in \mathcal{L}(E)$.

La fonction de variable réelle $t \mapsto e^{tu}$ est dérivable et sa dérivée est $t \mapsto ue^{tu} = ue^{tu}$.

Démonstration 15.31

On a $e^{tu} = \sum_{k=0}^{+\infty} t^k \frac{u^k}{k!}$.

On pose $f_k : t \mapsto t^k \frac{u^k}{k!}$ de classe \mathcal{C}^1 sur \mathbb{R} et, pour $k \geq 1$, on a

$$\forall t \in \mathbb{R}, f'_k(t) = kt^{k-1} \frac{u^k}{k!} = t^{k-1} \frac{u^k}{(k-1)!}.$$

Soit $a > 0$. Pour $t \in [-a; a]$, on a

$$\|f'_k(t)\| = |t^{k-1}| \left\| \frac{u^k}{(k-1)!} \right\| \leq a^{k-1} \frac{\|u\|^k}{(k-1)!}.$$

Donc la série de fonctions $\sum_{k=1}^{+\infty} f'_k$ converge normalement sur $[-a ; a]$.

D'après le théorème de dérivation sous le signe somme, la fonction $t \mapsto e^{tu}$ est de classe \mathcal{C}^1 sur $[-a ; a]$.

Par réunion d'intervalles, elle est de classe \mathcal{C}^1 sur \mathbb{R} et

$$\frac{d}{dt} (e^{tu}) = \left(\sum_{k=1}^{+\infty} t^{k-1} \frac{u^{k-1}}{(k-1)!} \right) \circ u = e^{tu} \circ u = ue^{tu}.$$

■

15.3 Équations linéaires homogènes à coefficients constants

15.3.1 Forme générale de la solution

Proposition 15.32

Soient $a \in \mathcal{L}(E)$, $t_0 \in I$ et $v_0 \in E$.

L'unique solution au problème de Cauchy $\begin{cases} x' = a.x \\ x(t_0) = v_0 \end{cases}$ est la fonction $t \mapsto e^{(t-t_0)a}.v_0$.

On a bien sûr l'équivalent matriciel.

Proposition 15.33

Soient $A \in \mathcal{M}_n(\mathbb{K})$, $t_0 \in I$ et $V_0 \in \mathbb{K}^n$.

L'unique solution au problème de Cauchy $\begin{cases} X' = AX \\ X(t_0) = V_0 \end{cases}$ est la fonction $t \mapsto e^{(t-t_0)A}V_0$.

15.3.2 Cas praticable

Proposition 15.34

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

On suppose que A est diagonalisable : on note a_1, \dots, a_n les valeurs propres non-distinctes de A et (V_1, \dots, V_n) une base de l'espace $\mathcal{M}_{n1}(\mathbb{K})$ constituée de vecteurs propres associés.

Les solutions du système différentiel linéaire homogène $X' = AX$ sont les fonctions

$$t \mapsto \sum_{i=1}^n \lambda_i e^{a_i t} V_i$$

où $\lambda_1, \dots, \lambda_n$ sont des constantes dites d'intégration.

En pratique, on diagonalise la matrice A en donnant une matrice de passage P et une matrice diagonale D telles que $A = PDP^{-1}$, on pose $X = PY$ et on résout le système différentiel $Y' = DY$, ce qui donne la forme générale de Y puis celle de $X = PY$. Ce qui est remarquable, c'est qu'il est inutile de calculer P^{-1} .

Quand la matrice A n'est pas diagonalisable mais seulement trigonalisable, alors en notant T une matrice triangulaire semblable à A , la même stratégie peut être appliquée, moyennant une résolution un peu plus difficile du système différentiel $Y' = TY$.

Exercice 15.35
Résoudre le système différentiel
$$\begin{cases} x' = 2x + y - z \\ y' = x + 2y - z \\ z' = -x - y + 2z \end{cases}$$

Correction 15.36

On a $A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \in \mathcal{S}_3(\mathbb{R})$ donc diagonalisable.

On a $A - I_3 = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$ de rang 1.

Donc 1 est valeur propre et $\dim \text{sep}(A, 1) = 3 - 1 = 2$.

A est diagonalisable donc elle a trois valeurs propres : 1, 1 et λ .

Or $\text{tr } A = 6$ donc $1 + 1 + \lambda = 6$ i.e. $\lambda = 4$.

Avec $U = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, on a $AU = U \iff a + b - c = 0$.

Donc $\text{sep}(A, 1) = \text{Vect}(V_1, V_2)$ où $V_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ et $V_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

D'après le théorème spectral, on a

$$\text{sep}(A, 4) = \text{sep}(A, 1)^\perp = \text{Vect}(V_1 \wedge V_2) = \text{Vect}(V_3)$$

où $V_3 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$.

On pose $P = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$ et on a alors $A = PDP^{-1}$ où $D = \text{diag}(1, 1, 4)$.

On a

$$\begin{aligned} X' = AX &\iff X' = PDP^{-1}X \\ &\iff P^{-1}X' = DP^{-1}X. \end{aligned}$$

En posant $Y = P^{-1}X$, on a l'équivalence $X' = AX \iff Y' = DY$.

Avec $Y = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$, on a

$$Y' = DY \iff \begin{cases} u' = u \\ v' = v \\ w' = 4w \end{cases}$$

Donc X est solution ssi il existe $(\alpha, \beta, \gamma) \in \mathbb{R}^3$ tels que

$$\begin{cases} u(t) = \alpha e^t \\ v(t) = \beta e^t \\ w(t) = \gamma e^{4t} \end{cases}$$

i.e.

$$X = PY = \begin{pmatrix} \alpha e^t - \gamma e^{4t} \\ \beta e^t - \gamma e^{4t} \\ (\alpha + \beta) e^t + \gamma e^{4t} \end{pmatrix}.$$

Exercice 15.37

Résoudre le système différentiel $\begin{cases} x' = -2x - y + 7z \\ y' = 5x + 4y - 8z \\ z' = x + y + z \end{cases}$

Correction 15.38

On pose $A = \begin{pmatrix} -2 & -1 & 7 \\ 5 & 4 & -8 \\ 1 & 1 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$.

On a

$$\chi_A = \begin{vmatrix} X+2 & 1 & -7 \\ -5 & X-4 & 8 \\ -1 & -1 & X-1 \end{vmatrix} = (X+1)(X-2)^2.$$

De plus, on a

$$AU = U \iff \begin{cases} -x - y + 7z = 0 \\ 5x + 5y - 8z = 0 \\ x + y + 2z = 0 \end{cases}$$

$$\iff \begin{cases} 9z = 0 \\ -6z = 0 \\ x + y = 0 \end{cases}$$

$$\iff U \in \text{Vect}(V_1) \text{ où } V_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Idem, $\text{sep}(A, 2) = \text{Vect}(V_2)$ où $V_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$.

$\dim \text{sep}(A, 2) = 1 \neq 2$ donc A n'est pas diagonalisable.

Or A est scindée donc trigonalisable : il existe $P \in \text{GL}_3(\mathbb{R})$ et $T \in \mathcal{T}_3^+(\mathbb{R})$ telles que $A = PTP^{-1}$.

On cherche T sous la forme $T = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ *i.e.* on cherche $V \notin \text{Vect}(V_1, V_2)$ tel que $AV = 2V + V_2$.

Avec $V = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, on a

$$AV = 2V + V_2 \iff \begin{cases} -2a - b + 7c = 2a + 2 \\ 5a + 4b - 8c = 2b - 1 \\ a + b + c = 2c + 1 \end{cases}$$

$$\iff \begin{cases} a + b - c = 1 \\ -4a - b + 7c = 2 \\ 5a + 2b - 8c = -1 \end{cases}$$

$$\iff \begin{cases} a + b - c = 1 \\ 3b + 3c = 6 \\ -3b - 3c = -6 \end{cases}$$

$$\iff \begin{cases} a + b + c = 1 \\ b + c = 2 \end{cases}$$

$$\iff \begin{cases} a = -1 + 2c \\ b = 2 - c \end{cases}$$

Par exemple, $V = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$ convient.

On pose $P = \begin{pmatrix} -1 & 2 & -1 \\ 1 & -1 & 2 \\ 0 & 1 & 0 \end{pmatrix}$.

On a

$$\begin{aligned} X' = AX &\iff P^{-1}X = TP^{-1}X \\ &\iff Y' = TY \end{aligned}$$

où $Y = P^{-1}X = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$.

$$\text{On a } \begin{cases} u' = -u \\ v' = 2v + w \\ w' = 2w \end{cases}$$

Or

$$\begin{aligned} \triangleright u' = -u &\iff u(t) = \alpha e^{-t}; \\ \triangleright w' = 2w &\iff w(t) = \gamma e^{2t}; \\ \triangleright v' = 2v + \gamma e^{2t} &\iff v' - 2v = \gamma e^{2t} \iff v(t) = \beta e^{2t} + \gamma t e^{2t}. \end{aligned}$$

Donc

$$X = PY = \begin{pmatrix} -\alpha e^{-t} + 2e^{2t}(\gamma t + \beta) - \gamma e^{2t} \\ \alpha e^{-t} - e^{2t}(\gamma t + \beta) + 2\gamma e^{2t} \\ e^{2t}(\gamma t + \beta) \end{pmatrix} = \alpha \begin{pmatrix} -e^{-t} \\ e^{-t} \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 2e^{2t} \\ -e^{2t} \\ e^{2t} \end{pmatrix} + \gamma \begin{pmatrix} (2t-1)e^{2t} \\ (2-t)e^{2t} \\ te^{2t} \end{pmatrix}.$$

Dans la suite, on s'intéresse à d'autres types d'équations différentielles linéaires, où les inconnues sont à valeurs réelles ou complexes.

15.4 Équations différentielles linéaires scalaires d'ordre n

15.4.1 Généralités

Définition 15.39

On appelle équation différentielle linéaire scalaire d'ordre n toute équation différentielle de la forme

$$y^{(n)} = a_{n-1}(t) y^{(n-1)} + \dots + a_1(t) y' + a_0(t) y + b(t)$$

où a_0, \dots, a_{n-1}, b sont $n+1$ fonctions continues de I dans \mathbb{K} et y est la fonction inconnue de I dans \mathbb{K} .

Définition 15.40

Avec les notations précédentes, on appelle solution (sur I) de l'équation différentielle toute application f n fois dérivable sur I et telle que

$$\forall t \in I, \quad f^{(n)}(t) = a_{n-1}(t) f^{(n-1)}(t) + \dots + a_1(t) f'(t) + a_0(t) f(t) + b(t).$$

Exemple 15.41

$\triangleright y'' - (t-1)y' + ty = 0$ a pour solution sur \mathbb{R} la fonction exp.

$\triangleright y'' + \frac{1}{t}y' + 4t^2y = 1 + t^4$ a pour solution sur \mathbb{R}^* la fonction $t \mapsto \cos(t^2) + \frac{t^2}{4}$.

Proposition 15.42

Les solutions d'une telle équation différentielle linéaire sont au moins de classe \mathcal{C}^n sur I .

15.4.2 Représentation matricielle

On considère une équation différentielle linéaire scalaire d'ordre n

$$(1) : y^{(n)} = a_{n-1}(t) y^{(n-1)} + \cdots + a_1(t) y' + a_0(t) y + b(t)$$

On pose $X = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$, $A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}$ et $B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b \end{pmatrix}$.

On définit ainsi trois fonctions vectorielles (matricielles) $X \in \mathcal{C}^1(I, \mathbb{K}^n)$, $A \in \mathcal{C}^0(I, \mathcal{M}_n(\mathbb{K}))$ et $B \in \mathcal{C}^0(I, \mathbb{K}^n)$.

L'équation (1) est alors équivalente au système différentiel $X' = AX + B$.

Par conséquent, tous les résultats précédents sont applicables dans ce cas particulier.

Démonstration 15.43

On a l'équivalence

$$y \text{ solution de (1)} \iff y^{(n)} = \sum_{i=0}^{n-1} a_i y^{(i)} + b.$$

On pose $U = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ où x_1, \dots, x_n sont n fonctions scalaires dérivables.

On a

$$U' = AU + B \iff \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b \end{pmatrix}$$

$$\iff \begin{cases} x_1' = x_2 \\ \vdots \\ x_{n-1}' = x_n \\ x_n' = \sum_{i=0}^{n-1} a_i x_{i+1} + b \end{cases}$$

$$\iff \begin{cases} x_1' = x_2 \\ \vdots \\ x_{n-1}' = x_n \\ x_n' = \sum_{i=0}^{n-1} a_i x_1^{(i)} + b \end{cases}$$

D'où (1) $\iff X' = AX + B$. ■

15.4.3 Équation différentielle linéaire homogène

Définition 15.44

On appelle équation différentielle linéaire scalaire homogène d'ordre n toute équation différentielle de la forme

$$y^{(n)} = a_{n-1}(t) y^{(n-1)} + \cdots + a_1(t) y' + a_0(t) y$$

où a_0, \dots, a_{n-1} sont n fonctions continues sur I à valeurs dans \mathbb{K} et y est la fonction inconnue de I dans \mathbb{K} .

Là encore, il y a un seul résultat à connaître sur de telles équations : le théorème de Cauchy-Lipschitz.

Théorème 15.45

L'ensemble des solutions d'une équation différentielle linéaire scalaire homogène d'ordre n est un \mathbb{K} -espace vectoriel de dimension n .

Plus précisément, en notant S l'ensemble des solutions, pour tout $t_0 \in I$, l'application

$$\begin{aligned} S &\longrightarrow \mathbb{K}^n \\ f &\longmapsto (f(t_0), f'(t_0), \dots, f^{(n-1)}(t_0)) \end{aligned}$$

est un isomorphisme.

Autrement dit, pour tout $t_0 \in I$ et $(v_0, v_1, \dots, v_{n-1}) \in \mathbb{K}^n$ (appelé vecteur des conditions initiales), il existe une unique solution telle que pour tout $k \in \llbracket 0 ; n-1 \rrbracket$, $f^{(k)}(t_0) = v_k$.

Là encore, toute base de l'ensemble des solutions est appelée système fondamental de solutions. Il n'y a hélas pas de méthode générale pour en calculer concrètement.

15.4.4 Cas général

Proposition 15.46

Soit

$$(1) : y^{(n)} = a_{n-1}(t) y^{(n-1)} + \cdots + a_1(t) y' + a_0(t) y + b(t)$$

une équation différentielle linéaire scalaire d'ordre n sur l'intervalle I .

Si on connaît une solution p dite particulière de l'équation, alors on connaît la forme générale des solutions : elles sont de la forme

$$t \longmapsto g(t) + p(t)$$

où g est la solution générale de l'équation différentielle linéaire homogène associée.

Le théorème de Cauchy-Lipschitz est encore valable sous une forme amenuisée.

Théorème 15.47

Pour tout choix des conditions initiales $(v_0, \dots, v_{n-1}) \in \mathbb{K}^n$, le problème de Cauchy

$$\begin{cases} y^{(n)} = a_{n-1}(t) y^{(n-1)} + \dots + a_1(t) y' + a_0(t) y + b(t) \\ y(t_0) = v_0 \\ y'(t_0) = v_1 \\ \vdots \\ y^{(n-1)}(t_0) = v_{n-1} \end{cases}$$

possède toujours une unique solution.

Remarque 15.48

La recherche d'une solution particulière est un problème difficile. Sauf cas évidents, des indications seront données.

15.4.5 Diverses idées pour résoudre une équation différentielle linéaire d'ordre n

En pratique, on a souvent $n = 2$.

- ▶ Chercher des solutions sous une forme *a priori* raisonnable (polynomiale, exponentielle, etc.).
- ▶ Chercher des solutions développables en série entière.
- ▶ Faire un changement de variable ou un changement de fonction inconnue en suivant les indications de l'énoncé.
- ▶ Quand on connaît une solution non-nulle z , en chercher une autre sous la forme uz où u est une fonction inconnue à déterminer (c'est la méthode de variation de la constante!), qui transforme le problème différentiel d'ordre n en un problème d'ordre $n - 1$ plus une intégration.

Exercice 15.49

Résolvez l'équation différentielle suivante sur $\left] -\frac{1}{2} ; +\infty \right[: (2t + 1) y''(t) + (4t - 2) y'(t) - 8y(t) = 0$ en cherchant une solution polynomiale et une solution exponentielle.

Correction 15.50

Si f est une solution polynomiale de degré n , on a

$$f(t) = t^n + \dots \quad f'(t) = nt^{n-1} + \dots \quad f''(t) = n(n-1)t^{n-2} + \dots$$

Alors $(2t+1)f''(t) + (4t-2)f'(t) - 8f(t) = 0$ donne

$$(2n(n-1)t^{n-1} + \dots) + (4nt^n + \dots) + (-8t^n + \dots) = 0.$$

Donc $n = 2$ *i.e.*

$$f(t) = t^2 + at + b \quad f'(t) = 2t + a \quad f''(t) = 2.$$

Donc

$$(2t+1)2 + (4t-2)(2t+a) - 8(t^2 + at + b) = 0 \iff -4at + 2 - 2a - 8b = 0$$

$$\iff \begin{cases} -4a = 0 \\ 2 - 2a - 8b = 0 \end{cases}$$

$$\iff \begin{cases} a = 0 \\ b = \frac{1}{4} \end{cases}$$

Donc $f : t \mapsto t^2 + \frac{1}{4}$ est solution.

De plus, si $f : t \mapsto e^{kt}$ est solution, alors

$$(2t+1)k^2e^{kt} + (4t-2)ke^{kt} - 8e^{kt} = 0 \iff k^2(2t+1) + k(4t-2) - 8 = 0$$

$$\iff \begin{cases} 2k^2 + 4k = 0 \\ k^2 - 2k - 8 = 0 \end{cases}$$

$$\iff k = -2.$$

Donc la famille $\left(t \mapsto t^2 + \frac{1}{4}, t \mapsto e^{-2t}\right)$ est une famille libre de solutions.

Or d'après le théorème de Cauchy-Lipschitz, l'ensemble des solutions est un \mathbb{R} -espace vectoriel de dimension 2 *i.e.*

$$\mathcal{S} = \left\{ t \mapsto \lambda \left(t^2 + \frac{1}{4} \right) + \mu e^{-2t} \mid (\lambda, \mu) \in \mathbb{R}^2 \right\}.$$

Exercice 15.51

Même exercice sur \mathbb{R}_+^* avec : $y''(t) + 2t(1+t)y'(t) - 2(1+t)y(t) = 0$ en cherchant une première solution de la forme $t \mapsto t^\alpha$ où $\alpha \in \mathbb{R}$.

Exercice 15.53

Même exercice sur \mathbb{R} avec : $(1+t^2)y''(t)+4ty'(t)+2y(t)=1$ en cherchant des solutions développables en série entière.

Correction 15.54

Avec $f : t \mapsto \sum_{n=0}^{+\infty} a_n t^n$ on a

$$f'(t) = \sum_{n=0}^{+\infty} n a_n t^{n-1} \quad \text{et} \quad f''(t) = \sum_{n=0}^{+\infty} n(n-1) a_n t^{n-2}.$$

Alors

$$\begin{aligned} f \text{ solution} &\iff \sum_{n=2}^{+\infty} n(n-1) a_n t^{n-2} + \sum_{n=0}^{+\infty} n(n-1) a_n t^n + \sum_{n=0}^{+\infty} 4n a_n t^n + \sum_{n=0}^{+\infty} 2a_n t^n = 1 \\ &\iff \sum_{n=0}^{+\infty} (n+2)(n+1) a_{n+2} t^n + \sum_{n=0}^{+\infty} (n(n-1) + 4n + 2) a_n t^n = 1 \\ &\iff \begin{cases} 2a_2 + 2a_0 = 1 \\ \forall n \geq 1, (n+2)(n+1) a_{n+2} + (n^2 + 3n + 2) a_n = 0 \end{cases} \\ &\iff \begin{cases} a_2 = \frac{1}{2} - a_0 \\ \forall n \geq 1, a_{n+2} = -a_n \end{cases} \\ &\iff \begin{cases} a_0 \text{ quelconque} \\ a_1 \text{ quelconque} \\ a_2 = \frac{1}{2} - a_0 \\ \forall n \in \mathbb{N}^*, \begin{cases} a_{2n} = (-1)^{n-1} a_2 \\ a_{2n+1} = (-1)^n a_n \end{cases} \end{cases} \end{aligned}$$

Donc, si f est développable en série entière au voisinage de 0, alors

$$f(t) = a_1 \sum_{n=0}^{+\infty} (-1)^n t^{2n+1} + a_0 + \left(\frac{1}{2} - a_0\right) \sum_{n=1}^{+\infty} (-1)^{n-1} t^{2n}.$$

Pour $t \in]-1 ; 1[$, on pose

$$\varphi(t) = \sum_{n=0}^{+\infty} (-1)^n t^{2n+1} \quad \text{et} \quad \psi(t) = \sum_{n=1}^{+\infty} (-1)^{n-1} t^{2n}.$$

Alors si f est une solution développable en série entière, on a

$$f(t) = a_1 \varphi(t) + \left(\frac{1}{2} - a_0\right) (\psi(t) - 1) + \frac{1}{2}.$$

Réciproquement, si f est de cette forme, alors f est développable en série entière sur $]-1 ; 1[$ et est solution de l'équation différentielle.

Or pour $t \in]-1 ; 1[$, on a

$$\varphi(t) = t \sum_{n=0}^{+\infty} (-t^2)^n = \frac{t}{1+t^2}$$

et

$$\psi(t) = - \sum_{n=1}^{+\infty} (-t^2)^n = \frac{t^2}{1+t^2}.$$

On vérifie que $t \mapsto \frac{t}{1+t^2}$ et $t \mapsto \frac{t^2}{1+t^2} - 1$ sont solutions sur \mathbb{R} et elles sont non-colinéaires donc d'après le théorème de Cauchy-Lipschitz :

$$\mathcal{S} = \left\{ t \mapsto \lambda \frac{t}{1+t^2} + \mu \left(\frac{t^2}{1+t^2} - 1 \right) + \frac{1}{2} \mid (\lambda, \mu) \in \mathbb{R}^2 \right\}.$$

Exercice 15.55

Même exercice sur \mathbb{R}_+^* avec : $t^2 y''(t) + 4ty'(t) - (t^2 - 2)y(t) = e^t$ en effectuant le changement de fonction inconnue $z(t) = t^2 y(t)$.

Correction 15.56

Avec $z : t \mapsto t^2 y(t)$, on a

$$z'(t) = 2ty(t) + t^2 y'(t)$$

et

$$z''(t) = t^2 y''(t) + 4ty'(t) + 2y(t).$$

On a donc

$$t^2 y'' + 4ty' - (t^2 - 2)y = z'' - z = e^t.$$

Donc z est de la forme $t \mapsto \alpha e^t + \beta e^{-t} + \frac{1}{2} t e^t$ avec $(\alpha, \beta) \in \mathbb{R}^2$.

Sur \mathbb{R}_+^* , on a donc

$$\mathcal{S} = \left\{ t \mapsto \alpha \frac{e^t}{t^2} + \beta \frac{e^{-t}}{t^2} + \frac{e^t}{2t} \mid (\alpha, \beta) \in \mathbb{R}^2 \right\}.$$

Dans le cas $n = 2$, pour trouver une solution particulière de l'équation $y'' = a(t)y' + b(t)y + c(t)$, si on connaît un système fondamental de solutions de l'équation homogène associée $y'' = a(t)y' + b(t)y$, alors on peut parfois réussir à trouver une solution particulière par variations des constantes : en notant (s_1, s_2) un système fondamental de solutions de l'équation homogène associée, on pose $p = u_1 s_1 + u_2 s_2$ où u_1 et u_2 sont deux fonctions inconnues deux fois dérivables. Comme on a deux inconnues pour une seule équation, on impose une condition supplémentaire : $u_1' s_1 + u_2' s_2 = 0$, afin que la recherche de solutions particulières aboutisse au système suivant :

$$\begin{cases} u_1' s_1 + u_2' s_2 = 0 \\ u_1' s_1' + u_2' s_2' = c \end{cases}$$

Comme dans le cas de la méthode de variation de la constante vue en première année, la détermination explicite de u_1 et u_2 est parfois empêchée par les calculs de primitives non-explicitables.

Démonstration 15.57

Avec $p = u_1 s_1 + u_2 s_2$, on a

$$p' = u_1' s_1 + u_2' s_2 + u_1 s_1' + u_2 s_2'$$

et

$$p'' = u_1'' s_1 + 2u_1' s_1' + u_1 s_1'' + u_2'' s_2 + 2u_2' s_2' + u_2 s_2''.$$

Alors

$$\begin{aligned} p \text{ solution} &\iff p'' = ap' + bp + c \\ &\iff u_1'' s_1 + 2u_1' s_1' + \cancel{u_1 s_1''} + u_2'' s_2 + 2u_2' s_2' + \cancel{u_2 s_2''} = au_1' s_1 + au_2' s_2 + \cancel{au_1 s_1'} + \cancel{au_2 s_2'} \\ &\quad + \cancel{bu_1 s_1} + \cancel{bu_2 s_2} + c \\ &\iff (u_1' s_1 + u_2' s_2)' + u_1' s_1' + u_2' s_2' = \underbrace{au_1' s_1 + au_2' s_2 + c}_{=a(u_1' s_1 + u_2' s_2)}. \end{aligned}$$

En effet, on a

$$u_1'' s_1 + 2u_1' s_1' = u_1'' s_1 + u_1' s_1' + u_1' s_1' = (u_1' + s_1)' + u_1' s_1'$$

et on justifie les annulations par $s_1'' = as_1' + bs_1$ et $s_2'' = as_2' + bs_2$.

Si on choisit $u_1' s_1 + u_2' s_2 = 0$ alors

$$p \text{ solution} \iff u_1' s_1' + u_2' s_2' = c. \quad \blacksquare$$

Exercice 15.58

Résolvez l'équation $y'' + 3y' + 2y = \frac{t-1}{t^2} e^{-t}$ sur $]0; +\infty[$.

Correction 15.59

Les solutions homogènes sont $t \mapsto \lambda e^{-t} + \mu e^{-2t}$ avec $(\lambda, \mu) \in \mathbb{R}^2$.

On pose

$$p(t) = u_1(t) e^{-t} + u_2(t) e^{-2t}$$

et on impose

$$u_1'(t) e^{-t} + u_2'(t) e^{-2t} = 0.$$

Alors

$$\begin{aligned} p \text{ solution} &\iff -u_1'(t) e^{-t} - 2u_2'(t) e^{-2t} = \frac{t-1}{t^2} e^{-t} \\ &\iff \begin{cases} u_1'(t) + e^{-t} u_2'(t) = 0 \\ -u_1'(t) - 2e^{-t} u_2'(t) = \frac{t-1}{t^2} \end{cases} \\ &\iff \begin{cases} u_1'(t) = \frac{t-1}{t^2} \\ u_2'(t) = \frac{1-t}{t^2} e^t \end{cases} \end{aligned}$$

On choisit par exemple

$$u_1(t) = \ln t + \frac{1}{t}$$

et

$$u_2(t) = -\frac{e^t}{t}.$$

Conclusion :

$$\mathcal{S} = \{t \mapsto \lambda e^{-t} + \mu e^{-2t} + \ln(t) e^{-t} \mid (\lambda, \mu) \in \mathbb{R}^2\}.$$

Exercice 15.60

Résolvez l'équation $y'' + 4y = \tan t$ sur $\left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$.

Correction 15.61

Les solutions homogènes sont $t \mapsto \lambda \cos(2t) + \mu \sin(2t)$ avec $(\lambda, \mu) \in \mathbb{R}^2$.

On pose

$$p(t) = u_1(t) \cos(2t) + u_2(t) \sin(2t)$$

et on impose

$$u_1'(t) \cos(2t) + u_2'(t) \sin(2t) = 0.$$

Alors

$$p \text{ solution} \iff \begin{cases} u_1'(t) \cos(2t) + u_2'(t) \sin(2t) = 0 \\ -2u_1'(t) \sin(2t) + 2u_2'(t) \cos(2t) = \tan t \end{cases}$$

Or $\begin{vmatrix} \cos(2t) & \sin(2t) \\ -2\sin(2t) & 2\cos(2t) \end{vmatrix} = 2$ donc

$$\begin{cases} u_1'(t) = \frac{\begin{vmatrix} 0 & \sin(2t) \\ \tan t & 2\cos(2t) \end{vmatrix}}{2} = \frac{-\tan(t)\sin(2t)}{2} = -\sin^2 t = \frac{\cos(2t) - 1}{2} \\ u_2'(t) = \frac{\begin{vmatrix} \cos(2t) & 0 \\ -2\sin(2t) & \tan t \end{vmatrix}}{2} = \frac{\cos(2t)\tan(t)}{2} = \frac{(2\cos^2 t - 1)\frac{\sin t}{\cos t}}{2} = \frac{\sin(2t)}{2} - \frac{1}{2} \frac{\sin t}{\cos t} \end{cases}$$

On choisit par exemple

$$u_1(t) = \frac{1}{4} \sin(2t) - \frac{t}{2}$$

et

$$u_2(t) = -\frac{1}{4} \cos(2t) + \frac{1}{2} \ln |\cos t|.$$

Conclusion :

$$\mathcal{S} = \left\{ t \mapsto \lambda \cos(2t) + \mu \sin(2t) - \frac{t}{2} \cos(2t) + \frac{1}{2} \sin(2t) \ln |\cos t| \mid (\lambda, \mu) \in \mathbb{R}^2 \right\}.$$

Chapitre 16

Calcul différentiel

Sommaire

16.1	Dérivées partielles482
16.1.1	Dérivée selon un vecteur	482
16.1.2	Dérivées partielles dans une base	482
16.1.3	Absence de lien entre la continuité et l'existence de dérivées selon tout vecteur	483
16.2	Différentielle.486
16.2.1	Application différentiable	486
16.2.2	Différentielle	492
16.2.3	Différentiabilité sur un ouvert	493
16.2.4	Lien avec les dérivées partielles	493
16.2.5	Caractérisation des fonctions à dérivée partielle nulle	495
16.2.6	Matrice jacobienne	497
16.2.7	Cas particulier où $F = \mathbb{R}$	498
16.3	Opérations sur les fonctions différentiables499
16.3.1	Combinaison linéaire	499
16.3.2	Composition par une application linéaire	499
16.3.3	Composition par une application k -linéaire	500
16.3.4	Composition d'applications différentiables	500
16.3.5	Dérivation le long d'un chemin	504
16.4	Fonctions de classe \mathcal{C}^1505
16.4.1	Définition	505
16.4.2	Caractérisation	505
16.4.3	Opérations sur les fonctions de classe \mathcal{C}^1	507
16.4.4	Caractérisation des fonctions constantes parmi les \mathcal{C}^1	510
16.5	Vecteurs tangents à une partie510
16.6	Optimisation au premier ordre513
16.6.1	Vocabulaire	513
16.6.2	Points critiques, extrema locaux d'une fonction sur un ouvert	514
16.6.3	Extrema locaux d'une fonction sur une partie	517
16.7	Fonctions de classe \mathcal{C}^k522
16.7.1	Dérivées partielles d'ordre supérieur	522
16.7.2	Fonctions de classe \mathcal{C}^k	522
16.7.3	Théorème de Schwarz	522
16.8	Optimisation au second ordre523
16.8.1	Hessienne	523

16.8.2 Développement limité à l'ordre 2	523
16.8.3 Application à l'étude des points critiques	524

Dans tout le cours, E, F (et éventuellement G) sont des \mathbb{R} -espaces vectoriels normés de dimensions finies p, n respectivement (éventuellement q). Comme dans le chapitre précédent, si $\varphi \in \mathcal{L}(E, F)$ et $v \in E$, on note $\varphi.v$ plutôt que $\varphi(v)$ l'image de v par l'application linéaire φ .

16.1 Dérivées partielles

16.1.1 Dérivée selon un vecteur

Définition 16.1

Soient U un ouvert de E , $f : U \longrightarrow F$, $a \in U$, $v \in E \setminus \{0\}$ et $t \in \mathbb{R}$.

On dit que f possède une dérivée en a selon le vecteur v quand la fonction $t \longmapsto f(a + tv)$ est dérivable en 0, *i.e.* quand $\frac{f(a + tv) - f(a)}{t}$ a une limite finie dans F quand t tend vers 0.

Dans ce cas, on pose $D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}$.

16.1.2 Dérivées partielles dans une base

Définition 16.2

Soient $\mathcal{B} = (e_1, \dots, e_p)$ une base de E , U un ouvert de E , $f : U \longrightarrow F$ et $a \in U$.

On dit que f possède des dérivées partielles en a dans la base \mathcal{B} quand pour tout $j \in \llbracket 1 ; p \rrbracket$, f possède une dérivée en a selon le vecteur e_j .

Dans ce cas, on note $\partial_j f(a) = D_{e_j} f(a) = \lim_{t \rightarrow 0} \frac{f(a + te_j) - f(a)}{t}$.

Avec les mêmes notations, en notant $x = \sum_{j=1}^p x_j e_j$ un vecteur générique de E , on identifie en général les notations $f(x)$ et $f(x_1, \dots, x_p)$: moyennant le choix d'une base de E , toute fonction de E dans F peut être vue comme une fonction de \mathbb{R}^p dans F , donc comme une fonction de p variables réelles et à valeurs dans F . On note alors aussi $\frac{\partial f}{\partial x_j}(a) = \partial_j f(a)$.

Autrement dit, f possède une j -ème dérivée partielle en $a = (a_1, \dots, a_p)$ quand la fonction $\varphi_j : x_j \longmapsto f(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_p)$ est dérivable en a_j . Dans ce cas, $\partial_j f(a) = \varphi'_j(a_j)$.

16.1.3 Absence de lien entre la continuité et l'existence de dérivées selon tout vecteur

Pour une fonction d'une seule variable, l'existence d'une dérivée en un point implique la continuité en ce point. Dès que $p \geq 2$, ce lien est faux : une fonction peut très bien avoir des dérivées selon tout vecteur mais n'être pas continue.

Exercice 16.3

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par : si $(x, y) \neq 0$, $f(x, y) = \frac{\sin(x) \sin^2(y)}{x^2 + y^2}$ et $f(0, 0) = 0$.

Vérifiez que f a des dérivées selon tout vecteur en $(0, 0)$ et qu'elle est continue en ce point.

Correction 16.4

Soit $v = (h, k) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

On a

$$\begin{aligned} \frac{f((0, 0) + t(h, k)) - f((0, 0))}{t} &= \frac{f(th, tk)}{t} \\ &= \frac{\frac{\sin(th) \sin^2(tk)}{(th)^2 + (tk)^2}}{t} \\ &= \frac{\sin(th) \sin^2(tk)}{t^3(h^2 + k^2)} \\ &\underset{t \rightarrow 0}{\sim} \frac{hk^2}{h^2 + k^2} \\ &\xrightarrow{t \rightarrow 0} \frac{hk^2}{h^2 + k^2}. \end{aligned}$$

f a donc une dérivée en $(0, 0)$ selon le vecteur v et

$$D_v f(0, 0) = \frac{hk^2}{h^2 + k^2}.$$

On veut montrer $\lim_{(x, y) \rightarrow (0, 0)} f(x, y) = f(0, 0) = 0$.

On a $\left| \frac{xy^2}{x^2 + y^2} \right| = |x| \frac{y^2}{x^2 + y^2} \leq |x|$ donc par encadrement

$$\frac{xy^2}{x^2 + y^2} \xrightarrow{(x, y) \rightarrow (0, 0)} 0.$$

De plus, $|f(x, y)| = \frac{|\sin x| |\sin y|^2}{x^2 + y^2} \leq \frac{|x| |y|^2}{x^2 + y^2} \leq |x|$ donc par encadrement

$$\lim_{(x, y) \rightarrow (0, 0)} f(x, y) = 0.$$

Exercice 16.5

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par : si $y \neq 0$, $f(x, y) = \frac{x^2}{y}$ et si $y = 0$, $f(x, 0) = 0$.

La fonction f est-elle continue en $(0, 0)$? A-t-elle une dérivée selon un vecteur v en $(0, 0)$?

Correction 16.6

On remarque

$$f(t, t) = \frac{t^2}{t} = t \xrightarrow[t \rightarrow 0]{} 0$$

et

$$f(t, t^2) = \frac{t^2}{t^2} = 1 \xrightarrow[t \rightarrow 0]{} 1.$$

Si f est continue en $(0, 0)$, alors par composition des limites, pour toutes fonctions $\varphi, \psi : \mathbb{R} \longrightarrow \mathbb{R}$ telles que $\varphi(0) = \psi(0) = 0$ et φ, ψ continues en 0, la composée $t \longmapsto f(\varphi(t), \psi(t))$ est continue en 0 *i.e.*

$$\lim_{t \rightarrow 0} f(\varphi(t), \psi(t)) = f(0, 0).$$

Avec $\varphi : t \longmapsto t$ et $\psi : t \longmapsto t^2$ on a $f(t, t^2) = 1 \not\xrightarrow[t \rightarrow 0]{} f(0, 0)$: contradiction.

Soit $v = (h, k) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

On a

$$\frac{f(th, tk)}{t} = \begin{cases} 0 \xrightarrow[t \rightarrow 0]{} 0 & \text{si } k = 0 \\ \frac{h^2}{k} \xrightarrow[t \rightarrow 0]{} \frac{h^2}{k} & \text{si } k \neq 0 \end{cases}$$

Finalement, $D_v f(0, 0) = f(v)$.

Exercice 16.7

Soit $f : (x, y) \longmapsto \frac{x^3 + y^3}{x^2 + y^2}$ prolongée en $(0, 0)$ par 0.

Montrez que f possède des dérivées partielles en tout point de \mathbb{R}^2 . Ces dérivées partielles sont-elles continues ?

Correction 16.8

Sur $\mathbb{R}^2 \setminus \{(0, 0)\}$, f admet des dérivées partielles en tout point car $x \longmapsto f(x, y_0)$ et $y \longmapsto f(x_0, y)$ sont dérivables en x_0, y_0 respectivement si $(x_0, y_0) \neq (0, 0)$.

On peut donc définir $\partial_1 f(x_0, y_0)$ et $\partial_2 f(x_0, y_0)$.

En $(0, 0)$, soit $e_1 = (1, 0)$ et $e_2 = (0, 1)$.

On a

$$\frac{f((0, 0) + t(1, 0)) - f(0, 0)}{t} = \frac{f(t, 0)}{t} = 1 \xrightarrow[t \rightarrow 0]{} 1$$

donc $\partial_1 f(0, 0) = 1$ et

$$\frac{f(0, t) - f(0, 0)}{t} = 1 \xrightarrow{t \rightarrow 0} 1$$

donc $\partial_2 f(0, 0) = 1$.

Donc $\partial_1 f$ et $\partial_2 f$ sont définies sur \mathbb{R}^2 . Y sont-elles continues ?

Pour $(x, y) \neq (0, 0)$, on a

$$\partial_1 f(x, y) = \frac{3x^2(x^2 + y^2) - (x^3 + y^3)2x}{(x^2 + y^2)^2} = \frac{x^4 + 3x^2y^2 - 2xy^3}{(x^2 + y^2)^2}.$$

Donc $\partial_1 f$ est continue sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ par opérations sur les fonctions continues.

De plus,

$$\partial_1 f(t, t) = \frac{2t^4}{4t^4} = \frac{1}{2} \xrightarrow{t \rightarrow 0} \frac{1}{2} \neq 1.$$

Donc $\partial_1 f$ n'est pas continue en $(0, 0)$.

De même, $\partial_2 f$ n'est pas continue en $(0, 0)$.

Exercice 16.9

Soit $f : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}$ définie par : $f(A) = \text{tr } A^2$.

Montrez que f possède en tout point des dérivées selon tout vecteur.

En choisissant comme base de $\mathcal{M}_n(\mathbb{R})$ la base canonique et en notant x_{ij} les coordonnées dans cette base, calculez $\frac{\partial f}{\partial x_{ij}}(A)$.

Correction 16.10

Soient $A \in \mathcal{M}_n(\mathbb{R})$ et $V \in \mathcal{M}_n(\mathbb{R}) \setminus \{0\}$.

On a

$$\begin{aligned} \frac{f(A + tV) - f(A)}{t} &= \frac{\text{tr}(A^2) + t \text{tr}(AV) + t \text{tr}(VA) + t^2 \text{tr}(V^2) - \text{tr}(A^2)}{t} \\ &= \text{tr}(AV) + \text{tr}(VA) + t \text{tr}(V^2) \\ &\xrightarrow{t \rightarrow 0} 2 \text{tr}(AV). \end{aligned}$$

Donc $D_V f(A) = 2 \text{tr}(AV)$.

Alors

$$\frac{\partial f}{\partial x_{ij}}(A) = D_{E_{ij}} f(A) = 2 \text{tr}(AE_{ij})$$

et si $A = (a_{ij})$ alors

$$\frac{\partial f}{\partial x_{ij}}(A) = 2a_{ji}.$$

16.2 Différentielle

16.2.1 Application différentiable

Définition 16.11

Soient U un ouvert de E , $f : U \rightarrow F$ et $a \in U$.

On dit que f est différentiable en a quand il existe une application linéaire $L : E \rightarrow F$ continue, un réel $r > 0$ et une application $\varepsilon : B(0, r) \rightarrow F$ telles que

$$\begin{cases} \forall v \in B(0, r), f(a + v) = f(a) + L.v + \|v\| \varepsilon(v) \\ \lim_{v \rightarrow 0} \varepsilon(v) = 0 \end{cases}$$

On note classiquement $o(\|v\|)$ toute expression du type $\|v\| \varepsilon(v)$ où $\lim_{v \rightarrow 0} \varepsilon(v) = 0$.

La définition précédente affirme donc l'existence d'un développement limité à l'ordre 1 : $f(a + v) = f(a) + L.v + o(\|v\|)$.

Remarque 16.12

Dans la définition précédente, l'hypothèse de continuité de L est superflue car E est de dimension finie donc toutes les applications linéaires de E dans F sont continues. Mais, au cas où un sujet hors-programme vous placerait dans un espace E de dimension infinie, vous avez la définition complète.

Exercice 16.13

Soit $f : (x, y) \mapsto y \ln x + e^{xy}$.

Montrez que f est différentiable en $(1, 1)$.

Correction 16.14

On pose $v = (h, k) \in \mathbb{R}^2$ et $a = (1, 1)$.

On a $f(a) = e$ et $\|v\| = \max(|h|, |k|)$.

On a

$$\begin{aligned} f(a + v) &= f(1 + h, 1 + k) \\ &= (1 + k) \ln(1 + h) + e^{(1+h)(1+k)} \\ &= (1 + k) \ln(1 + h) + e e^{h+k+hk} \\ &= (1 + k) (h + h\alpha(h)) + e (1 + (h + k + hk) + (h + k + hk) \beta(h + k + hk)) \\ &= f(a) + h + h\alpha(h) + hk + hk\alpha(h) + eh + ek + ehk + e(h + k + hk) \beta(h + k + hk) \\ &= f(a) + \underbrace{(e + 1)h + ek}_{=L.(h,k)} + \|v\| \underbrace{\frac{h\alpha(h) + hk + hk\alpha(h) + ehk + e(h + k + hk) \beta(h + k + hk)}{\|v\|}}_{=\varepsilon(v)}. \end{aligned}$$

$\lim_{\alpha \rightarrow 0} \alpha = 0$ et $\lim_{\beta \rightarrow 0} \beta = 0$

Il reste à montrer $\lim_{v \rightarrow (0,0)} \varepsilon(v) = 0$.

On a

$$|\varepsilon(v)| \leq \frac{1}{\|v\|} (|h| |\alpha(h)| + |h| |k| + |h| |k| |\alpha(h)| + e |h| |k| + e |h+k+hk| |\beta(h+k+hk)|).$$

Or $|h| \leq \|v\|$ et $|k| \leq \|v\|$ donc $\frac{|h|}{\|v\|} \leq 1$ et $\frac{|k|}{\|v\|} \leq 1$.

Donc

$$|\varepsilon(v)| \leq |\alpha(h)| + |k| + |k| |\alpha(h)| + e |h| + e (2 + |k|) |\beta(h+k+hk)|.$$

Par encadrement, on a

$$\lim_{v \rightarrow (0,0)} \varepsilon(v) = 0.$$

Donc f est différentiable en $(1, 1)$.

Exercice 16.15

Soit $C : v \mapsto \|v\|^2$ (où la norme est ici la norme euclidienne).

Montrez que C est différentiable en tout point.

En est-il de même pour l'application norme elle-même ?

Correction 16.16

Soient $a, v \in E$.

On a

$$\begin{aligned} C(a+v) &= \|a+v\|^2 \\ &= \|a\|^2 + \|v\|^2 + 2 \langle a | v \rangle \\ &= C(a) + L.v + \|v\| \varepsilon(v) \end{aligned}$$

où $L.v = 2 \langle a | v \rangle$ (L est linéaire) et $\varepsilon(v) = \|v\| \xrightarrow{v \rightarrow 0} 0$.

Donc C est différentiable en tout point de E .

On pose $N(a) = \|a\|$.

Si $a \neq 0$, on a

$$\begin{aligned} N(a+v) &= N(a) \sqrt{1 + 2 \frac{\langle a | v \rangle}{C(a)} + \frac{\|v\|^2}{C(a)}} \\ &= N(a) \left(1 + \frac{\langle a | v \rangle}{C(a)} + \frac{1}{2} \frac{\|v\|^2}{C(a)} + \underbrace{\left(2 \frac{\langle a | v \rangle}{C(a)} + \frac{\|v\|^2}{C(a)} \right)}_{=u} \alpha(u) \right) \quad \left. \vphantom{\frac{\|v\|^2}{C(a)}} \right\} \lim_0 \alpha = 0 \\ &= N(a) + \frac{1}{N(a)} \langle a | v \rangle + \|v\| \varepsilon(v) \end{aligned}$$

où

$$\varepsilon(v) = \frac{\|v\|}{2C(a)} + \left(\frac{2\langle a | v \rangle}{C(a)\|v\|} + \frac{\|v\|}{C(a)} \right) \alpha \left(2\frac{\langle a | v \rangle}{C(a)} + \frac{\|v\|^2}{C(a)} \right).$$

Donc

$$|\varepsilon(v)| \leq \underbrace{\frac{\|v\|}{2C(a)} + \frac{2}{C(a)} \underbrace{\left| \frac{\langle a | v \rangle}{\|v\|} \right|}_{\substack{\leq \|a\| \\ \text{d'après} \\ \text{Cauchy-Schwarz}}} + \frac{\|v\|}{C(a)}}_{\text{bornée}} \underbrace{\alpha \left(2\frac{\langle a | v \rangle}{C(a)} + \frac{\|v\|^2}{C(a)} \right)}_{\xrightarrow{v \rightarrow 0} 0}.$$

Donc $\varepsilon(v) \xrightarrow{v \rightarrow 0} 0$.

Donc N est différentiable sur $E \setminus \{0\}$.

Si $\|\cdot\|$ est différentiable en 0, alors il existe L linéaire telle que

$$\forall v \in E, \quad \|v\| = \|0\| + L.v + \|v\| \alpha(v)$$

où $\alpha(t) \xrightarrow{t \rightarrow 0} 0$.

On a aussi

$$\begin{aligned} \forall v \in E, \quad \|-v\| &= \|v\| = L.(-v) + \|v\| \alpha(-v) \\ &= -L.v + \|v\| \alpha(-v). \end{aligned}$$

Donc, en additionnant les égalités, on a

$$\begin{aligned} 2\|v\| &= \|v\| (\alpha(v) + \alpha(-v)) \\ 2 &= \alpha(v) + \alpha(-v). \end{aligned}$$

Or $\alpha(v) + \alpha(-v) \xrightarrow{v \rightarrow 0} 0 \neq 2$.

Donc $\|\cdot\|$ n'est pas différentiable en 0.

Exercice 16.17

Soit $f : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathcal{M}_n(\mathbb{R})$ définie par : $f(A) = A^2$.

Montrez que f est différentiable en tout point de $\mathcal{M}_n(\mathbb{R})$.

Correction 16.18

Soient $A, V \in \mathcal{M}_n(\mathbb{R})$.

On choisit une norme $\|\cdot\|$ sous-multiplicative sur $\mathcal{M}_n(\mathbb{R})$.

On a

$$\begin{aligned} f(A+V) &= (A+V)^2 \\ &= A^2 + AV + VA + V^2 \\ &= f(A) + L.V + \underbrace{\|V\| \frac{V^2}{\|V\|}}_{=\varepsilon(V)} \end{aligned}$$

où $L : V \mapsto AV + VA$ est linéaire de $\mathcal{M}_n(\mathbb{R})$ dans $\mathcal{M}_n(\mathbb{R})$.

On pose aussi $\varepsilon(0) = 0$ et on vérifie $\varepsilon(V) \xrightarrow{V \rightarrow 0} 0$:

$$\|\varepsilon(V)\| = \left\| \frac{V^2}{\|V\|} \right\| = \frac{1}{\|V\|} \|V^2\| \leq \frac{\|V\|^2}{\|V\|} = \|V\|$$

donc par encadrement $\varepsilon(V) \xrightarrow{V \rightarrow 0} 0$.

Donc f est différentiable en A .

Exercice 16.19

Soit $g : \text{GL}_n(\mathbb{R}) \longrightarrow \text{GL}_n(\mathbb{R})$ définie par : $g(A) = A^{-1}$.

Rappelez pourquoi $\text{GL}_n(\mathbb{R})$ est un ouvert de $\mathcal{M}_n(\mathbb{R})$.

Montrez qu'il existe $r > 0$ tel que pour tout $V \in B(0, r)$, $\sum_{k \geq 0} (-V)^k$ converge.

Dans ce cas, que vaut sa somme ?

Montrez que g est différentiable en I_n , puis en tout point de $\text{GL}_n(\mathbb{R})$.

Correction 16.20

On a $\text{GL}_n(\mathbb{R}) = \det^{-1}(\mathbb{R}^*)$. Or \det est continue et \mathbb{R}^* est un ouvert donc $\text{GL}_n(\mathbb{R})$ aussi.

On choisit une norme $\|\cdot\|$ sous-multiplicative sur $\mathcal{M}_n(\mathbb{R})$.

Soit $V \in B(0, 1)$.

Pour tout $k \in \mathbb{N}$, $\|(-V)^k\| \leq \|V\|^k$.

Or $\|V\| < 1$ donc la série géométrique $\sum \|V\|^k$ converge et par théorème de comparaison des séries à termes positifs, $\sum \|(-V)^k\|$ converge i.e. $\sum (-V)^k$ converge absolument et donc converge car $\mathcal{M}_n(\mathbb{R})$ est de dimension finie.

A-t-on $\sum_{k=0}^{+\infty} (-V)^k = (I_n + V)^{-1}$?

On a

$$\begin{aligned}
 (I_n + V) \sum_{k=0}^{+\infty} (-V)^k &= \sum_{k=0}^{+\infty} (-V)^k + V \sum_{k=0}^{+\infty} (-V)^k \\
 &= \sum_{k=0}^{+\infty} (-V)^k - (-V) \sum_{k=0}^{+\infty} (-V)^k \\
 &= \sum_{k=0}^{+\infty} (-V)^k - \sum_{k=0}^{+\infty} (-V)^{k+1} \\
 &= \sum_{k=0}^{+\infty} (-V)^k - \sum_{k=1}^{+\infty} (-V)^k \\
 &= I_n.
 \end{aligned}
 \quad \left. \begin{array}{l} \text{linéarité à droite} \\ \text{et continuité du} \\ \text{produit matriciel} \end{array} \right\}$$

Ceci prouve que $I_n + V \in \text{GL}_n(\mathbb{R})$ et $(I_n + V)^{-1} = \sum_{k=0}^{+\infty} (-V)^k$.

Si $V \in B(0, 1)$, on a

$$\begin{aligned}
 g(I_n + V) &= (I_n + V)^{-1} \\
 &= I_n - V + \sum_{k=2}^{+\infty} (-V)^k \\
 &= g(I_n) + L.V + \sum_{k=2}^{+\infty} (-V)^k
 \end{aligned}$$

où $L : V \mapsto -V$ est un endomorphisme de $\mathcal{M}_n(\mathbb{R})$.

Il reste à vérifier $\sum_{k=2}^{+\infty} (-V)^k \underset{V \rightarrow 0}{\underset{\sim}{=}} o(\|V\|)$ i.e. $\frac{\sum_{k=2}^{+\infty} (-V)^k}{\|V\|} \xrightarrow{V \rightarrow 0} 0$.

On a

$$\begin{aligned}
 \left\| \frac{\sum_{k=2}^{+\infty} (-V)^k}{\|V\|} \right\| &= \frac{1}{\|V\|} \left\| \sum_{k=2}^{+\infty} (-V)^k \right\| \\
 &\leq \frac{1}{\|V\|} \sum_{k=2}^{+\infty} \|V\|^k \\
 &= \frac{1}{\|V\|} \frac{\|V\|^2}{1 - \|V\|} \\
 &= \frac{\|V\|}{1 - \|V\|} \\
 &\xrightarrow{V \rightarrow 0} 0.
 \end{aligned}$$

Donc $\sum_{k=2}^{+\infty} (-V)^k \underset{V \rightarrow 0}{\underset{\sim}{=}} o(\|V\|)$ par encadrement.

Donc g est différentiable en I_n .

Soient $A \in \text{GL}_n(\mathbb{R})$ et $V \in \mathcal{M}_n(\mathbb{R})$.

On a $A + V = A(I_n + A^{-1}V)$ donc

$$A + V \in \text{GL}_n(\mathbb{R}) \iff I_n + A^{-1}V \in \text{GL}_n(\mathbb{R}).$$

Donc il suffit que $\|A^{-1}V\| < 1$.

Or $\|A^{-1}V\| \leq \|A^{-1}\| \|V\|$.

Donc il suffit que $\|A^{-1}\| \|V\| < 1$ i.e. $\|V\| < \frac{1}{\|A^{-1}\|}$.

Si $V \in B\left(0, \frac{1}{\|A^{-1}\|}\right)$, on a

$$\begin{aligned} g(A + V) &= (A + V)^{-1} \\ &= (A(I_n + A^{-1}V))^{-1} \\ &= (I_n + A^{-1}V)^{-1} A^{-1} \\ &= g(I_n + A^{-1}V) A^{-1} \\ &= (g(I_n) + L.(A^{-1}V) + o(\|A^{-1}V\|)) A^{-1} \\ &= (I_n - A^{-1}V + o(\|A^{-1}V\|)) A^{-1} \\ &= A^{-1} - A^{-1}VA^{-1} + o(\|A^{-1}V\|). \end{aligned}$$

On pose $L_A : V \mapsto -A^{-1}VA^{-1}$ endomorphisme de $\mathcal{M}_n(\mathbb{R})$.

Comme on a $\|A^{-1}V\| \leq \|A^{-1}\| \|V\|$, on a finalement

$$g(A + V) = g(A) + L_A.V + o(\|V\|).$$

Donc g est différentiable sur $\text{GL}_n(\mathbb{R})$.

Proposition 16.21

Avec les mêmes notations, si f est différentiable en a , alors

- ▷ f est continue en a
- ▷ f admet des dérivées selon tout vecteur en a .

Démonstration 16.22

On suppose f différentiable en a .

Il existe $L : E \longrightarrow F$ linéaire et ε telle que $\varepsilon(v) \xrightarrow[v \rightarrow 0]{} 0$ tels que

$$f(a+v) = f(a) + L.v + \|v\| \varepsilon(v).$$

L étant linéaire, elle est continue (dimension finie) donc

$$L.v \xrightarrow[v \rightarrow 0]{} L.0 = 0.$$

Donc par opérations sur les limites, $f(a+v) \xrightarrow[v \rightarrow 0]{} f(a)$ i.e. f est continue en a .

De plus, si $v \neq 0$, on a

$$\begin{aligned} \frac{f(a+tv) - f(a)}{t} &= \frac{f(a) + tL.v + |t| \|v\| \varepsilon(tv) - f(a)}{t} \\ &= L.v + \underbrace{\frac{|t|}{t}}_{\text{bornée}} \|v\| \underbrace{\varepsilon(tv)}_{\xrightarrow[t \rightarrow 0]{} 0} \\ &\xrightarrow[t \rightarrow 0]{} L.v. \end{aligned}$$

Donc $D_v f(a) = L.v$. ■

Bien évidemment, la réciproque est fausse.

16.2.2 Différentielle

Proposition 16.23

Avec les mêmes notations, l'application L est unique.

Démonstration 16.24

On a

$$L.v = D_v f(a) = \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t}$$

qui est unique. ■

Dans le cas où f est différentiable en a , l'application L s'appelle la différentielle de f en a ou l'application linéaire tangente en a . Elle est notée $df(a)$.

Le développement limité en a à l'ordre 1 est donc

$$f(a+v) = f(a) + df(a).v + o(\|v\|)$$

et la dérivée selon le vecteur v en a est

$$D_v f(a) = df(a).v$$

16.2.3 Différentiabilité sur un ouvert

Définition 16.25

Soient U un ouvert de E et $f : U \longrightarrow F$.

On dit que f est différentiable sur U quand f est différentiable en tout point de U . On lui associe donc une unique application $df : U \longrightarrow \mathcal{L}(E, F)$, appelée la différentielle de f sur U .

Remarque 16.26

En fait, on devrait noter $df : U \longrightarrow \mathcal{L}_c(E, F)$.

Deux cas particuliers :

- si f est constante sur U , alors $df = 0$
- si f est linéaire, alors $df = f$.

Démonstration 16.27

Si f est linéaire, on a

$$\begin{aligned} f(a + v) &= f(a) + f(v) \\ &= f(a) + f(v) + \|v\| 0. \end{aligned}$$

Comme f est linéaire et $0 \xrightarrow[v \rightarrow 0]{} 0$, f est différentiable en a et $df(a) = f$. ■

16.2.4 Lien avec les dérivées partielles

Proposition 16.28

Soient U un ouvert de E , $f : U \longrightarrow F$, $a \in U$ et \mathcal{B} une base de E .

Si f est différentiable en a , alors f possède des dérivées partielles en a dans la base \mathcal{B} et pour tout $v \in$

E de coordonnées (h_1, \dots, h_p) , $df(a) \cdot v = \sum_{j=1}^p \partial_j f(a) h_j = \sum_{j=1}^p \frac{\partial f}{\partial x_j}(a) h_j$.

Démonstration 16.29

On pose $\mathcal{B} = (e_1, \dots, e_p)$ et on a $v = \sum_{i=1}^p h_i e_i$.

On a

$$\begin{aligned} df(a) \cdot v &= df(a) \cdot \left(\sum_{i=1}^p h_i e_i \right) \\ &= \sum_{i=1}^p h_i df(a) \cdot e_i. \end{aligned} \quad \left. \begin{array}{c} \searrow \\ \downarrow \end{array} \right\} \text{linéarité de } df(a)$$

Or $df(a) \cdot e_i = D_{e_i} f(a) = \partial_i f(a)$.

Donc $df(a) \cdot v = \sum_{i=1}^p h_i \partial_i f(a) \in F$. ■

Remarque 16.30

Dans la notation précédente, on a exceptionnellement fait une entorse à la convention habituelle qui consiste à écrire les produits externes scalaire-vecteur dans le sens λv .

Ici, les dérivées partielles $\partial_j f(a)$ sont des vecteurs de F et les h_j sont des scalaires, on devrait donc noter les produits externes $h_j \partial_j f(a)$. Mais l'usage veut que dans le cas des différentielles, on respecte l'ordre des objets plutôt que la convention du produit externe.

Un cas particulier : si $p = 1$, alors f est différentiable en a ssi f est dérivable en a et dans ce cas, pour tout $h \in \mathbb{R}$, $df(a) \cdot h = f'(a) h$.

Exercice 16.31

Soit $f : (x, y) \mapsto y \ln x + e^{xy}$.

En admettant momentanément que f est différentiable sur $\mathbb{R}_+^* \times \mathbb{R}$, calculez sa différentielle en tout point.

Correction 16.32

On pose $a = (a_1, a_2)$ et $v = (h_1, h_2)$.

On a $df(a) \cdot v = \sum_{i=1}^2 h_i \partial_i f(a)$.

Or

$$\partial_1 f(x, y) = \frac{y}{x} + ye^{xy} \quad \text{et} \quad \partial_2 f(x, y) = \ln x + xe^{xy}.$$

Donc

$$df(a) \cdot v = h_1 \left(\frac{a_2}{a_1} + a_2 e^{a_1 a_2} \right) + h_2 (\ln a_1 + a_1 e^{a_1 a_2}).$$

Exercice 16.33

Soit $g : \mathcal{M}_2(\mathbb{R}) \longrightarrow \mathcal{M}_2(\mathbb{R})$ définie par : $g(A) = A^\top A$.

En notant $A = \begin{pmatrix} a_1 & a_3 \\ a_2 & a_4 \end{pmatrix}$, calculez la différentielle de g en la matrice $J = (i + j)_{1 \leq i, j \leq 2}$.

Correction 16.34

On a $J = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$ et $A = \begin{pmatrix} a_1 & a_3 \\ a_2 & a_4 \end{pmatrix}$ et on pose $V = \begin{pmatrix} h_1 & h_3 \\ h_2 & h_4 \end{pmatrix}$.

On a $dg(J) \cdot V = \sum_{i=1}^4 h_i \partial_i g(J)$.

Or $g(A) = \begin{pmatrix} a_1^2 + a_2^2 & a_1 a_3 + a_2 a_4 \\ a_1 a_3 + a_2 a_4 & a_3^2 + a_4^2 \end{pmatrix}$ donc

$$\partial_1 g(A) = \begin{pmatrix} 2a_1 & a_3 \\ a_3 & 0 \end{pmatrix} \quad \partial_2 g(A) = \begin{pmatrix} 2a_2 & a_4 \\ a_4 & 0 \end{pmatrix} \quad \partial_3 g(A) = \begin{pmatrix} 0 & a_1 \\ a_1 & 2a_3 \end{pmatrix} \quad \partial_4 g(A) = \begin{pmatrix} 0 & a_2 \\ a_2 & 2a_4 \end{pmatrix}.$$

Donc

$$dg(J) \cdot V = h_1 \begin{pmatrix} 4 & 3 \\ 3 & 0 \end{pmatrix} + h_2 \begin{pmatrix} 6 & 4 \\ 4 & 0 \end{pmatrix} + h_3 \begin{pmatrix} 0 & 2 \\ 2 & 6 \end{pmatrix} + h_4 \begin{pmatrix} 0 & 3 \\ 3 & 8 \end{pmatrix}.$$

16.2.5 Caractérisation des fonctions à dérivée partielle nulle

Proposition 16.35

Soient $D = I \times J$ où I, J sont deux intervalles ouverts de \mathbb{R} et f une fonction différentiable sur D .

► Si pour tout $a \in D$, $\frac{\partial f}{\partial x}(a) = 0$, alors il existe une fonction v dérivable sur J telle que

$$\text{pour tout } (x, y) \in D, \quad f(x, y) = v(y).$$

► Si pour tout $a \in D$, $\frac{\partial f}{\partial y}(a) = 0$, alors il existe une fonction u dérivable sur I telle que

$$\text{pour tout } (x, y) \in D, \quad f(x, y) = u(x).$$

Démonstration 16.36

On suppose $\forall a \in D, \frac{\partial f}{\partial x}(a) = \partial_1 f(a) = 0$.

Soit $a = (a_1, a_2) \in D$.

Sur I , la fonction $\varphi : t \mapsto f(t, a_2)$ est constante car sa dérivée est nulle sur I .

Donc il existe $\lambda \in F$ tel que $\forall t \in I, f(t, a_2) = \lambda$.

On a donc montré :

$$\forall a_2 \in J, \exists ! \lambda \in F, \forall t \in I, f(t, a_2) = \lambda.$$

On construit ainsi une fonction $v : J \longrightarrow F$.

$$a_2 \longmapsto \lambda$$

Avec cette notation, on a

$$\forall a_2 \in J, \forall t \in I, f(t, a_2) = v(a_2).$$

Comme f est différentiable en tout point de D , il faut que v soit dérivable sur J .

Idem en échangeant x et y . ■

Exprimé de façon plus grossière, si la dérivée partielle par rapport à une variable est constamment nulle, alors la fonction ne dépend pas de cette variable. Bien sûr, ce résultat s'étend à des fonctions à plus de deux variables.

Corollaire 16.37

Avec les mêmes hypothèses, si pour tout $a \in D, \frac{\partial f}{\partial x}(a) = \frac{\partial f}{\partial y}(a) = 0$, alors f est constante sur D .

Remarque 16.38

Ce résultat reste valable sur des ouverts de forme plus générale (voir plus loin) ou en adaptant légèrement l'énoncé, avec plus de trois variables.

Exercice 16.39

Déterminez les fonctions f différentiables sur \mathbb{R}^2 telle que pour tout $(x, y) \in \mathbb{R}^2, \frac{\partial f}{\partial x}(x, y) = \lambda$, où λ est une constante.

Correction 16.40

Soient $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ différentiable et $\lambda \in \mathbb{R}$.

En posant $g : (x, y) \mapsto f(x, y) - \lambda x$, on a

$$\forall (x, y) \in \mathbb{R}^2, \frac{\partial f}{\partial x}(x, y) = \lambda \iff \forall (x, y) \in \mathbb{R}^2, \frac{\partial g}{\partial x}(x, y) = 0$$

$$\iff \exists v : \mathbb{R} \longrightarrow \mathbb{R} \text{ dérivable, } \forall (x, y) \in \mathbb{R}^2, g(x, y) = v(y)$$

$$\iff \exists v : \mathbb{R} \longrightarrow \mathbb{R} \text{ dérivable, } \forall (x, y) \in \mathbb{R}^2, f(x, y) = \lambda x + v(y).$$

Exercice 16.41

Faites de même avec la condition $\frac{\partial f}{\partial x}(x, y) = \cos x$.

Correction 16.42

En posant $g : (x, y) \mapsto f(x, y) - \sin x$, on a

$$\begin{aligned} \forall (x, y) \in \mathbb{R}^2, \quad \frac{\partial f}{\partial x}(x, y) = \cos x &\iff \forall (x, y) \in \mathbb{R}^2, \quad \frac{\partial g}{\partial x}(x, y) = 0 \\ &\iff \exists v : \mathbb{R} \longrightarrow \mathbb{R} \text{ dérivable, } \forall (x, y) \in \mathbb{R}^2, \quad f(x, y) = \sin x + v(y). \end{aligned}$$

Exercice 16.43

Faites de même avec la condition $\frac{\partial f}{\partial x}(x, y) = f(x, y)$.

Correction 16.44

En posant $g : (x, y) \mapsto f(x, y) e^{-x}$, on a

$$\begin{aligned} \forall (x, y) \in \mathbb{R}^2, \quad \frac{\partial f}{\partial x}(x, y) = f(x, y) &\iff \forall (x, y) \in \mathbb{R}^2, \quad \frac{\partial g}{\partial x}(x, y) = 0 \\ &\iff \exists v : \mathbb{R} \longrightarrow \mathbb{R} \text{ dérivable, } \forall (x, y) \in \mathbb{R}^2, \quad f(x, y) = e^x v(y). \end{aligned}$$

Ce genre de problème s'appelle des équations aux dérivées partielles.

16.2.6 Matrice jacobienne

Définition 16.45

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}^n$ et $a \in U$.

Si f est différentiable en a , alors on appelle (matrice) jacobienne de f en a la matrice dans les bases canoniques de $df(a)$, souvent notée $Jf(a)$. C'est une matrice de $\mathcal{M}_{np}(\mathbb{R})$.

On note $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$ où f_1, \dots, f_n sont n fonctions de U dans \mathbb{R} .

Proposition 16.46

Avec les mêmes notations, f est différentiable en a ssi pour tout $i \in \llbracket 1 ; p \rrbracket$, f_i est différentiable en a .

Dans ce cas, on a

$$J f(a) = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \dots & \partial_p f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \dots & \partial_p f_2(a) \\ \vdots & \vdots & & \vdots \\ \partial_1 f_n(a) & \partial_2 f_n(a) & \dots & \partial_p f_n(a) \end{pmatrix} = (\partial_j f_i(a))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$$

Exercice 16.47

Calculez la jacobienne du changement de variables en polaires $(r, \theta) \mapsto (r \cos \theta, r \sin \theta)$.

Correction 16.48

On pose $f : (r, \theta) \mapsto \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$.

On a

$$J f(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

avec, par ligne, les dérivées de $r \cos \theta$ et $r \sin \theta$ par rapport à r et θ , par colonne.

16.2.7 Cas particulier où $F = \mathbb{R}$

Si $n = 1$, alors $f : U \rightarrow \mathbb{R}$, donc si f est différentiable en a , alors $df(a)$ est une forme linéaire de E .

Si on a muni E d'un produit scalaire \cdot , on appelle alors gradient de f en a , noté $\nabla f(a)$, l'unique vecteur de E tel que

$$\forall v \in E, \quad df(a) \cdot v = \nabla f(a) \cdot v.$$

Le développement limité à l'ordre 1 devient donc dans ce cas

$$f(a + v) = f(a) + \nabla f(a) \cdot v + o(\|v\|)$$

et la dérivée selon le vecteur v en a est donc

$$D_v f(a) = \nabla f(a) \cdot v.$$

Dans une base orthonormée $\mathcal{B} = (e_1, \dots, e_p)$, le vecteur $\nabla f(a)$ a pour coordonnées $\begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_p f(a) \end{pmatrix}$, tandis

que la matrice jacobienne est donc une matrice-ligne :

$$J f(a) = (\partial_1 f(a) \quad \partial_2 f(a) \quad \dots \quad \partial_p f(a)).$$

Remarque 16.49

Si $\nabla f(a) \neq 0$, alors $\nabla f(a)$ est positivement colinéaire au vecteur unitaire selon lequel la dérivée de f en a est maximale : l'application $v \mapsto \nabla f(a) \cdot v = D_v f(a)$, définie sur la sphère-unité, est maximale en le vecteur $v = \frac{1}{\|\nabla f(a)\|} \nabla f(a)$.

16.3 Opérations sur les fonctions différentiables

16.3.1 Combinaison linéaire

Proposition 16.50

Soient U un ouvert de E , $f, g : U \longrightarrow F$ et $a \in U$.

Si f et g sont différentiables en a , alors $f + g$ l'est aussi et $d(f + g)(a) = df(a) + dg(a)$.

Si f est différentiable en a , alors pour tout $\lambda \in \mathbb{R}$, λf l'est aussi et $d(\lambda f)(a) = \lambda df(a)$.

Autrement dit, l'application $f \mapsto df(a)$ est linéaire.

16.3.2 Composition par une application linéaire

Proposition 16.51

Soient U un ouvert de E , $f : U \longrightarrow F$, $a \in U$ et $L \in \mathcal{L}(F, G)$.

Si f est différentiable en a , alors $L \circ f$ l'est aussi et $d(L \circ f)(a) = L \circ df(a)$.

Démonstration 16.52

f est différentiable en a donc pour v au voisinage de 0, on a

$$f(a + v) = f(a) + df(a) \cdot v + \|v\| \varepsilon(v).$$

L est linéaire donc

$$L(f(a + v)) = L(f(a)) + L(df(a) \cdot v) + \|v\| L(\varepsilon(v))$$

i.e.

$$L \circ f(a + v) = L \circ f(a) + \underbrace{(L \circ df(a))}_{\text{linéaire}} \cdot v + \|v\| \underbrace{L(\varepsilon(v))}_{\xrightarrow{v \rightarrow 0} L(0)=0}.$$

Donc $L \circ f$ est différentiable en a et

$$d(L \circ f)(a) = L \circ df(a).$$

■

Autrement dit, pour tout vecteur $v \in E$,

$$d(L \circ f)(a) \cdot v = L(df(a) \cdot v).$$

16.3.3 Composition par une application k -linéaire

Proposition 16.53

Soient U un ouvert de E , $f_1, \dots, f_k : U \longrightarrow F$, $a \in U$ et $M : F^k \longrightarrow G$ k -linéaire.

Si f_1, \dots, f_k sont différentiable en a , alors $M(f_1, \dots, f_k)$ l'est aussi et

$$d(M(f_1, \dots, f_k))(a) = \sum_{i=1}^k M(f_1(a), \dots, f_{i-1}(a), df_i(a), f_{i+1}(a), \dots, f_k(a)).$$

Autrement dit, pour tout vecteur $v \in E$,

$$d(M(f_1, \dots, f_k))(a) \cdot v = \sum_{i=1}^k M(f_1(a), \dots, f_{i-1}(a), df_i(a) \cdot v, f_{i+1}(a), \dots, f_k(a)).$$

Un cas particulier important : le produit externe.

Proposition 16.54

Soient U un ouvert de E , $\lambda : U \longrightarrow \mathbb{R}$, $g : U \longrightarrow F$ et $a \in U$.

Si λ et g sont différentiables en a , alors

$$d(\lambda g)(a) = d\lambda(a) g(a) + \lambda(a) dg(a).$$

D'une manière générale, tout produit vérifie le même genre de relation (produit de deux réels, de deux matrices, de deux polynômes, composée de deux endomorphismes, etc).

16.3.4 Composition d'applications différentiables

Ce résultat est souvent appelé règle de la chaîne.

Proposition 16.55

Soient U un ouvert de E , V un ouvert de F , $f : U \longrightarrow V$, $g : V \longrightarrow G$ et $a \in U$.

Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et

$$d(g \circ f)(a) = dg(f(a)) \circ df(a).$$

Démonstration 16.56

f est différentiable en a donc pour v au voisinage de 0, on a

$$f(a + v) = f(a) + df(a) \cdot v + \|v\| \alpha(v).$$

g est différentiable en $f(a)$ donc pour w au voisinage de 0, on a

$$g(f(a) + w) = g(f(a)) + dg(f(a)) \cdot w + \|w\| \beta(w).$$

En posant $w = df(a) \cdot v + \|v\| \alpha(v) \xrightarrow{v \rightarrow 0} 0$, on a

$$\begin{aligned} g \circ f(a + v) &= g(f(a) + w) \\ &= g \circ f(a) + dg(f(a)) \cdot df(a) \cdot v + \|v\| dg(f(a)) \cdot \alpha(v) + \|df(a) \cdot v + \|v\| \alpha(v)\| \beta(w). \end{aligned}$$

On pose $\varepsilon(v) = dg(f(a)) \cdot \alpha(v) + \frac{\|df(a) \cdot v + \|v\| \alpha(v)\|}{\|v\|} \beta(w)$.

Par continuité de $dg(f(a))$, on a

$$dg(f(a)) \cdot \alpha(v) \xrightarrow{v \rightarrow 0} 0$$

et, par compositions de limites, on a

$$\beta(w) \xrightarrow{v \rightarrow 0} 0.$$

De plus, on a $\|df(a) \cdot v + \|v\| \alpha(v)\| \leq \|df(a) \cdot v\| + \|v\| \|\alpha(v)\|$.

Or $df(a)$ est linéaire en dimension finie donc il existe $K > 0$ tel que $\forall u \in E, \|df(a) \cdot u\| \leq K \|u\|$.

Donc $\left\| \frac{df(a) \cdot v + \|v\| \alpha(v)}{\|v\|} \right\| \leq K + \|\alpha(v)\|$.

Or $\alpha(v) \xrightarrow{v \rightarrow 0} 0$ donc α est bornée au voisinage de 0.

Donc pour v dans ce voisinage, $\left\| \frac{df(a) \cdot v + \|v\| \alpha(v)}{\|v\|} \right\|$ est bornée et donc

$$\varepsilon(v) \xrightarrow{v \rightarrow 0} 0.$$

Donc $g \circ f$ est différentiable en a et

$$d(g \circ f)(a) = dg(f(a)) \circ df(a). \quad \blacksquare$$

Autrement dit, pour tout vecteur $v \in E$,

$$d(g \circ f)(a) \cdot v = dg(f(a)) \cdot (df(a) \cdot v)$$

ce qui est conventionnellement noté $dg(f(a)) \cdot df(a) \cdot v$.

Proposition 16.57

Si E est identifié à \mathbb{R}^p , F à \mathbb{R}^n et G à \mathbb{R}^q par choix de bases, alors cela se traduit sur les matrices jacobiniennes :

$$J(g \circ f)(a) = Jg(f(a)) \times Jf(a).$$

On retrouve la règle de composition des dérivées partielles de première année. En particulier, les changements de variables entrent dans cette catégorie des composées de fonctions différentiables.

Exercice 16.58

Soient U, V deux ouverts et $f : U \rightarrow V$ une bijection différentiable sur U telle que f^{-1} le soit aussi.

Que dire de la matrice jacobienne de f en tout point de U ?

Exemple 16.59

Soient f une fonction différentiable de \mathbb{R}^2 dans \mathbb{R} définie sur un ouvert D et x, y deux fonctions différentiables sur un ouvert U de \mathbb{R}^2 telles que pour tout $(u, v) \in U$, $\Phi(u, v) = (x(u, v), y(u, v)) \in D$.

Alors la fonction composée $g : (u, v) \mapsto f(\Phi(u, v))$ est différentiable sur U et on a l'égalité matricielle

$$(\partial_1 g(u, v) \quad \partial_2 g(u, v)) = (\partial_1 f(\Phi(u, v)) \quad \partial_2 f(\Phi(u, v))) \times \begin{pmatrix} \partial_1 x(u, v) & \partial_2 x(u, v) \\ \partial_1 y(u, v) & \partial_2 y(u, v) \end{pmatrix},$$

ce qui se traduit, en clair, par

$$\forall (u, v) \in U, \quad \begin{cases} \partial_1 g(u, v) = \partial_1 f(\Phi(u, v)) \partial_1 x(u, v) + \partial_2 f(\Phi(u, v)) \partial_1 y(u, v) \\ \partial_2 g(u, v) = \partial_1 f(\Phi(u, v)) \partial_2 x(u, v) + \partial_2 f(\Phi(u, v)) \partial_2 y(u, v) \end{cases}$$

Avec les notations des physiciens, c'est plus clair, à condition de fixer les noms des variables selon leur rang :

$$\forall (u, v) \in U, \quad \begin{cases} \frac{\partial g}{\partial u}(u, v) = \frac{\partial f}{\partial x}(\Phi(u, v)) \frac{\partial x}{\partial u}(u, v) + \frac{\partial f}{\partial y}(\Phi(u, v)) \frac{\partial y}{\partial u}(u, v) \\ \frac{\partial g}{\partial v}(u, v) = \frac{\partial f}{\partial x}(\Phi(u, v)) \frac{\partial x}{\partial v}(u, v) + \frac{\partial f}{\partial y}(\Phi(u, v)) \frac{\partial y}{\partial v}(u, v) \end{cases}$$

voire même, de façon encore plus abrégée :

$$\begin{cases} \frac{\partial g}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} \\ \frac{\partial g}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v} \end{cases}$$

Remarque 16.60

Attention ! Plus on utilise des notations abrégées, plus il y a de sous-entendus ! Donc pour comprendre correctement ces égalités, il faut les replacer dans le contexte, donc ne pas oublier ces sous-entendus.

Exemple 16.61 (Cas particulier important : le passage en coordonnées polaires)

Si f est différentiable sur \mathbb{R}^2 , alors $g : (r, \theta) \mapsto f(r \cos \theta, r \sin \theta)$ est différentiable sur \mathbb{R}^2 et

$$\forall (r, \theta) \in \mathbb{R}^2, \quad \begin{cases} \frac{\partial g}{\partial r}(r, \theta) = \frac{\partial f}{\partial x}(\Phi(r, \theta)) \cos \theta + \frac{\partial f}{\partial y}(\Phi(r, \theta)) \sin \theta \\ \frac{\partial g}{\partial \theta}(r, \theta) = -\frac{\partial f}{\partial x}(\Phi(r, \theta)) r \sin \theta + \frac{\partial f}{\partial y}(\Phi(r, \theta)) r \cos \theta \end{cases}$$

Avec ces changements de variables, on peut résoudre quelques équations aux dérivées partielles simples, la difficulté étant de trouver un bon changement de variables. En pratique, il est presque toujours donné par l'énoncé.

Exercice 16.62

Déterminez les fonctions différentiables sur $\mathbb{R}_+^* \times \mathbb{R}$ telles que $x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = 0$.

Correction 16.63

On pose $g : (r, \theta) \mapsto f(r \cos \theta, r \sin \theta)$.

On a

$$\begin{aligned} x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} = 0 &\iff r \frac{\partial g}{\partial r}(r, \theta) = 0 \\ &\iff \frac{\partial g}{\partial r}(r, \theta) = 0 \\ &\iff \exists v : \left] \frac{-\pi}{2} ; \frac{\pi}{2} \right[\longrightarrow \mathbb{R} \text{ dérivable, } \forall (r, \theta) \in \mathbb{R}_+^* \times \left] \frac{-\pi}{2} ; \frac{\pi}{2} \right[, \quad g(r, \theta) = v(\theta) \\ &\iff \exists v : \left] \frac{-\pi}{2} ; \frac{\pi}{2} \right[\longrightarrow \mathbb{R} \text{ dérivable, } \forall (x, y) \in \mathbb{R}_+^* \times \mathbb{R}, \quad f(x, y) = v \circ \text{Arctan} \left(\frac{y}{x} \right) \\ &\iff \exists w : \mathbb{R} \longrightarrow \mathbb{R} \text{ dérivable, } \forall (x, y) \in \mathbb{R}_+^* \times \mathbb{R}, \quad f(x, y) = w \left(\frac{y}{x} \right). \end{aligned}$$

En effet, on a

$$\begin{array}{ccc} \mathbb{R}_+^* \times \left] \frac{-\pi}{2} ; \frac{\pi}{2} \right[& \xrightarrow{\sim} & \mathbb{R}_+^* \times \mathbb{R} \\ (r, \theta) & \mapsto & (r \cos \theta, r \sin \theta) \\ \left(\sqrt{x^2 + y^2}, \text{Arctan} \frac{y}{x} \right) & \longleftarrow & (x, y) \end{array}$$

Exercice 16.64

Déterminez les fonctions différentiables sur \mathbb{R}^2 telles que $2 \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} = 0$, en posant $x = 2u - v$ et $y = v - u$.

Correction 16.65

On pose $g : (u, v) \mapsto f(2u - v, v - u)$.

On a

$$\begin{aligned} 2 \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} = 0 &\iff \forall (u, v) \in \mathbb{R}^2, \quad \frac{\partial g}{\partial u}(u, v) = 0 \\ &\iff \exists h : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \text{ dérivable, } \forall (u, v) \in \mathbb{R}^2, \quad g(u, v) = h(v) \\ &\iff \exists h : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \text{ dérivable, } \forall (x, y) \in \mathbb{R}^2, \quad f(x, y) = h(x + 2y). \end{aligned}$$

Exercice 16.66

Déterminez les fonctions différentiables sur $\mathbb{R}_+^* \times \mathbb{R}$ telles que $2xy \frac{\partial f}{\partial x} + (1 + y^2) \frac{\partial f}{\partial y} = 0$, en posant

$$x = \frac{u^2 + v^2}{2} \text{ et } y = \frac{u}{v}.$$

16.3.5 Dérivation le long d'un chemin

On appelle ici chemin une fonction $\gamma : I \longrightarrow E$ continue sur I , comme pour la définition de connexité par arcs (cf. Définition 1.239).

Proposition 16.67

Soient U un ouvert de E , $f : U \longrightarrow F$ et $\gamma : I \longrightarrow U$ un chemin.

Si γ est dérivable en $t \in I$ et f est différentiable en $\gamma(t)$, alors $f \circ \gamma$ est dérivable en t et

$$(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t).$$

Un cas particulier : si f est à valeurs réelles.

Proposition 16.68

Soient U un ouvert de E , $f : U \longrightarrow \mathbb{R}$ et $\gamma : I \longrightarrow U$ un chemin.

Si γ est dérivable en $t \in I$ et f est différentiable en $\gamma(t)$, alors $f \circ \gamma$ est dérivable en t et

$$(f \circ \gamma)'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t).$$

Dans le cas où E est identifié à \mathbb{R}^p par choix d'une base, en notant $\gamma(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_p(t) \end{pmatrix}$, on a

$$(f \circ \gamma)'(t) = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(\gamma(t)) u_i'(t).$$

Si la situation précédente est valable pour tout $t \in I$, alors le support Γ de γ (i.e. son image) est une courbe.

- Si pour tout $t \in I$, les vecteurs $\nabla f(\gamma(t))$ et $\gamma'(t)$ sont orthogonaux, alors la fonction f est constante sur la courbe Γ : on dit que Γ est une ligne de niveau.
- Si pour tout $t \in I$, les vecteurs $\nabla f(\gamma(t))$ et $\gamma'(t)$ sont colinéaires de même sens, alors la courbe Γ est une courbe sur laquelle quand on se déplace, les variations relatives de f sont maximales : on dit que Γ est une ligne de champ de f .
- Par conséquent, si une ligne de niveau et une ligne de champ se croisent en un point, les vecteurs tangents en ce point sont orthogonaux : on dit que les lignes de champ sont orthogonales aux lignes de niveau.

Exemple 16.69

- Les lignes de niveau dans \mathbb{R}^2 de la fonction $f : (x, y) \longmapsto x^2 + y^2$ sont les courbes inscrites dans des cercles ; les lignes de champ sont les courbes inscrites dans les droites passant par l'origine.
- Les lignes de niveau dans \mathbb{R}^2 de la fonction $f : (x, y) \longmapsto xy$ sont les courbes inscrites dans les hyperboles d'asymptotes (Ox) et (Oy) ; les lignes de champ sont les courbes inscrites dans les hyperboles d'asymptotes les deux bissectrices des axes.

16.4 Fonctions de classe \mathcal{C}^1

16.4.1 Définition

Si une fonction $f : U \longrightarrow F$ est différentiable en tout point d'un ouvert U , alors pour tout $a \in U$, $df(a)$ est une application linéaire (continue) de E dans F , donc un élément de $\mathcal{L}_c(E, F)$ (qui est égal à $\mathcal{L}(E, F)$ dans ce cours, puisque les dimensions de E et F sont finies). On peut alors définir l'application $df : U \longrightarrow \mathcal{L}(E, F)$ (remarque : $\mathcal{L}(E, F)$ est aussi de dimension finie selon les hypothèses de ce cours).

Définition 16.70

On dit que f est de classe \mathcal{C}^1 sur U quand df est une application continue de U dans $\mathcal{L}(E, F)$.

Un exemple fondamental : les applications linéaires ou k -linéaires.

16.4.2 Caractérisation

Théorème 16.71

Soient U un ouvert de E et $f : U \longrightarrow F$.

En identifiant E et \mathbb{R}^p par choix d'une base quelconque, f est de classe \mathcal{C}^1 sur U ssi f possède des dérivées partielles en tout point de U et si toutes ses dérivées partielles sont continues sur U .

Démonstration 16.72

\Rightarrow Trivial car « $df(a) = \sum_i \partial_i f$ ».

\Leftarrow

Montrons que si f a des dérivées partielles continues alors f est différentiable (et donc *a fortiori* \mathcal{C}^1).

Soient $f : U \longrightarrow F$ ayant des dérivées partielles continues en tout point, $a = (a_1, \dots, a_p) \in U$ et $v = (h_1, \dots, h_p) \in E \simeq \mathbb{R}^p$.

On a

$$f(a+v) - f(a) = \sum_{i=1}^p [f(a_1 + h_1, \dots, a_i + h_i, a_{i+1}, \dots, a_p) - f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, a_{i+1}, \dots, a_p)].$$

Pour $i \in \llbracket 1 ; p \rrbracket$, on pose

$$\varphi_i : z \longmapsto f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, z, a_{i+1}, \dots, a_p)$$

définie au voisinage de a_i .

f ayant des dérivées partielles, φ_i est dérivable au voisinage de a_i .

Donc d'après le théorème des accroissements finis, il existe $c_i \in [a_i ; a_i + h_i]$ tel que

$$\varphi_i(a_i + h_i) - \varphi_i(a_i) = \varphi'_i(c_i) h_i.$$

Donc

$$\begin{aligned} f(a + v) - f(a) &= \sum_{i=1}^p \varphi'_i(c_i) h_i \\ &= \sum_{i=1}^p \partial_i f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, c_i, a_{i+1}, \dots, a_p) h_i. \end{aligned}$$

Par encadrement, on a $c_i \xrightarrow[v \rightarrow 0]{} a_i$.

Donc, par continuité de $\partial_i f$, on a

$$\partial_i f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, c_i, a_{i+1}, \dots, a_p) \xrightarrow[v \rightarrow 0]{} \partial_i f(a_1, \dots, a_p).$$

Il existe donc $\alpha_i : v \mapsto \alpha_i(v)$ telle que $\alpha_i(v) \xrightarrow[v \rightarrow 0]{} 0$ et

$$\partial_i f(a_1 + h_1, \dots, a_{i-1} + h_{i-1}, c_i, a_{i+1}, \dots, a_p) = \partial_i f(a) + \alpha_i(v).$$

Donc

$$f(a + v) - f(a) = \sum_{i=1}^p \partial_i f(a) h_i + \sum_{i=1}^p h_i \alpha_i(v).$$

On pose $\varepsilon(v) = \frac{1}{\|v\|} \sum_{i=1}^p h_i \alpha_i(v)$.

On a $\|\varepsilon(v)\| \leq \sum_{i=1}^p \frac{|h_i|}{\|v\|} \|\alpha_i(v)\|$.

On choisit $\|v\| = \max_{1 \leq j \leq p} |h_j|$ et on a $\frac{|h_i|}{\|v\|} \leq 1$.

Donc $\|\varepsilon(v)\| \leq \sum_{i=1}^p \|\alpha_i(v)\|$.

Par encadrement, $\varepsilon(v) \xrightarrow[v \rightarrow 0]{} 0$.

Donc f est différentiable en a et

$$df(a) : v \mapsto \sum_{i=1}^p \partial_i f(a) h_i.$$

Les dérivées partielles étant continues, $df : a \mapsto df(a)$ est continue, *i.e.* f est de classe \mathcal{C}^1 . ■

16.4.3 Opérations sur les fonctions de classe \mathcal{C}^1

Grâce aux théorèmes d'opérations et de composition des fonctions continues, il devient évident que

- toute combinaison linéaire d'applications \mathcal{C}^1 est \mathcal{C}^1
- tout produit ou quotient (sous réserve de définition) de fonctions \mathcal{C}^1 est \mathcal{C}^1
- toute composée de fonctions \mathcal{C}^1 est \mathcal{C}^1 .

Exemple 16.73

- Les applications coordonnées $(x_1, \dots, x_p) \mapsto x_i$ sont de classe \mathcal{C}^1 donc toute application de U dans \mathbb{R} qui est polynomiale en ses p variables est elle-même de classe \mathcal{C}^1 : le produit scalaire, le déterminant sont des applications de classe \mathcal{C}^1 .
- On retrouve ainsi que l'application $A \mapsto A^{-1}$ est de classe \mathcal{C}^1 sur $\text{GL}_n(\mathbb{R})$.

Exercice 16.74

La fonction $f : (x, y) \mapsto \frac{y^4}{x^2 + y^2}$ prolongée en $(0, 0)$ par 0 est-elle de classe \mathcal{C}^1 sur \mathbb{R}^2 ?

Correction 16.75

Sur $\mathbb{R}^2 \setminus \{(0, 0)\}$, f est \mathcal{C}^1 comme quotient de fonctions \mathcal{C}^1 .

De plus, on a

$$\frac{f(t, 0) - f(0, 0)}{t} = 0 \xrightarrow[t \rightarrow 0]{} 0$$

donc $\partial_1 f(0, 0) = 0$ et

$$\frac{f(0, t) - f(0, 0)}{t} = t \xrightarrow[t \rightarrow 0]{} 0$$

donc $\partial_2 f(0, 0) = 0$.

Or, on a

$$|\partial_1 f(x, y)| = \left(\frac{y^2}{x^2 + y^2} \right)^2 \times 2|x| \leq 2|x|$$

et

$$\begin{aligned} |\partial_2 f(x, y)| &\leq 4 \frac{x^2}{x^2 + y^2} \times \frac{y^2}{x^2 + y^2} |y| + 2 \left(\frac{x^2}{x^2 + y^2} \right)^2 |y| \\ &\leq 6|y|. \end{aligned}$$

Par encadrement, pour $i \in \{1, 2\}$, on a

$$\partial_i f(x, y) \xrightarrow[(x, y) \rightarrow (0, 0)]{} 0 = \partial_i f(0, 0).$$

Donc les dérivées partielles de f sont continues sur \mathbb{R}^2 .

Donc d'après le Théorème 16.71, f est \mathcal{C}^1 sur \mathbb{R}^2 .

Exercice 16.76

Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ de classe \mathcal{C}^1 .

On pose $\varphi(x, y) = \frac{1}{y-x} \int_x^y f$ si $x \neq y$ et $\varphi(x, x) = f(x)$.

Montrez que φ est de classe \mathcal{C}^1 sur \mathbb{R}^2 .

Correction 16.77

f étant \mathcal{C}^1 sur \mathbb{R} , elle possède une primitive F de classe \mathcal{C}^2 sur \mathbb{R} .

Alors, pour $x \neq y$, on a $\varphi(x, y) = \frac{F(y) - F(x)}{y-x}$.

φ est donc \mathcal{C}^1 sur $\mathbb{R}^2 \setminus \Delta$ où $\Delta = \{(x, x) \mid x \in \mathbb{R}\}$ par opérations sur les fonctions \mathcal{C}^1 .

Soit $(a, a) \in \Delta$.

On a

$$\begin{aligned} \frac{\varphi(a+t, a) - \varphi(a, a)}{t} &= \frac{\frac{F(a+t) - F(a)}{t} - f(a)}{t} \\ &= \frac{F(a+t) - F(a) - tF'(a)}{t^2}. \end{aligned}$$

F étant de classe \mathcal{C}^2 , d'après Taylor-Young, on a

$$F(a+t) \underset{t \rightarrow 0}{=} F(a) + tF'(a) + \frac{t^2}{2}F''(a) + o(t^2).$$

$$\text{Donc } \frac{\varphi(a+t, a) - \varphi(a, a)}{t} \underset{t \rightarrow 0}{\longrightarrow} \frac{F''(a)}{2} = \frac{f'(a)}{2}.$$

$$\text{Donc } \partial_1 \varphi(a, a) = \frac{f'(a)}{2}.$$

$$\text{Idem, } \partial_2 \varphi(a, a) = \frac{f'(a)}{2}.$$

Pour $x \neq y$, on a

$$\partial_1 \varphi(x, y) = \frac{f(x)(x-y) - (F(x) - F(y))}{(x-y)^2}$$

et

$$\partial_2 \varphi(x, y) = \frac{f(y)(y-x) - (F(y) - F(x))}{(y-x)^2}.$$

En posant $T(x, y) = \frac{f(x)(x-y) - (F(x) - F(y))}{(x-y)^2}$, il suffit de montrer

$$T(x, y) \underset{(x,y) \rightarrow (a,a)}{\longrightarrow} \frac{f'(a)}{2}.$$

F étant \mathcal{C}^2 , d'après la formule de Taylor avec reste intégral, on a

$$F(y) = F(x) + (y-x)F'(x) + \int_x^y (y-t)F''(t) dt$$

donc

$$F(y) - F(x) - (y-x)f'(x) = \int_x^y (y-t)f''(t) dt.$$

Donc

$$\begin{aligned} T(x, y) &= \frac{1}{(y-x)^2} \int_x^y (y-t)f''(t) dt \\ &= \frac{1}{(y-x)^2} \int_x^y (y-t)[f''(t) - f''(a)] dt + \frac{f''(a)}{(y-x)^2} \int_x^y (y-t) dt \\ &= \frac{1}{(y-x)^2} \int_x^y (y-t)[f''(t) - f''(a)] dt + \frac{f''(a)}{2}. \end{aligned}$$

De plus, on a

$$\left| \underbrace{\frac{1}{(y-x)^2} \int_x^y (y-t)[f''(t) - f''(a)] dt}_{V(x,y)} \right| \leq \frac{1}{(y-x)^2} \int_x^y |y-t| |f''(t) - f''(a)| dt.$$

On remarque que $|y-t| = y-t$ dans tous les cas et on a

$$\begin{aligned} |V(x, y)| &\leq \frac{1}{(y-x)^2} \int_x^y (y-t) \underbrace{|f''(t) - f''(a)|}_{\leq \sup_{t \in [x,y]} |f''(t) - f''(a)| = M_{xy}} dt \\ &\leq \frac{M_{xy}}{(y-x)^2} \int_x^y (y-t) dt \\ &\leq \frac{M_{xy}}{2}. \end{aligned}$$

f'' étant continue en a , on a $M_{xy} \xrightarrow{(x,y) \rightarrow (a,a)} 0$.

Donc $V(x, y) \xrightarrow{(x,y) \rightarrow (a,a)} 0$.

Donc $T(x, y) \xrightarrow{(x,y) \rightarrow (a,a)} \frac{f''(a)}{2} = \partial_1^2 \varphi(a, a)$.

Idem pour $\partial_2^2 \varphi(a, a)$.

Donc φ est \mathcal{C}^2 .

16.4.4 Caractérisation des fonctions constantes parmi les \mathcal{C}^1

Proposition 16.78

Soient U un ouvert de E , $f : U \longrightarrow F$ de classe \mathcal{C}^1 et $a, b \in U$.

Pour tout chemin $\gamma : [0 ; 1] \longrightarrow U$ de classe \mathcal{C}^1 tel que $\gamma(0) = a$ et $\gamma(1) = b$, on a

$$f(b) - f(a) = \int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt.$$

Démonstration 16.79

D'après la Proposition 16.67, on a $(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t)$.

f et γ étant \mathcal{C}^1 , $f \circ \gamma$ est \mathcal{C}^1 et donc

$$\begin{aligned} [f \circ \gamma(t)]_0^1 &= \int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt \\ &= f(b) - f(a). \end{aligned}$$

■

Théorème 16.80

Si U est un ouvert connexe par arcs de E et f est de classe \mathcal{C}^1 sur U , alors f est constante sur U ssi $df = 0$ sur U .

Remarque 16.81

Attention ! Ce résultat n'est valable que sur un ouvert connexe par arcs ! Dans le cas contraire, on est dans une situation analogue à celle rencontrée sur \mathbb{R} : la fonction $x \mapsto \text{Arctan } x + \text{Arctan } \frac{1}{x}$ est de dérivée nulle sur \mathbb{R}^* mais n'est pas constante sur \mathbb{R}^* .

16.5 Vecteurs tangents à une partie

Définition 16.82

Soient A une partie de E , $a \in A$ et $v \in E$.

On dit que v est un vecteur tangent à A en a quand il existe un chemin γ défini au voisinage de 0, dérivable en 0 et à valeurs dans A tel que $\gamma(0) = a$ et $\gamma'(0) = v$.

L'ensemble des vecteurs tangents à A en a est noté $T_a(A)$. Il contient toujours le vecteur nul. Quand il s'agit d'un sous-espace vectoriel de E , on l'appelle le sous-espace tangent à A en a .

Proposition 16.83

Soient U un ouvert de E et $g : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^1 . On pose $A = g^{-1}(\{0\})$.

Pour tout $a \in A$ tel que $dg(a) \neq 0$, l'ensemble des vecteurs tangents à A en a est le sous-espace $\ker dg(a)$.

Dans le cas où $dg(a) = 0$, il n'y a pas de résultat général.

Remarque 16.84

► Dans le cas où $p = 2$, A est appelé une ligne de niveau d'équation $g(x, y) = 0$.

Si $M = (a, b) \in A$ et $dg(M) \neq 0$, alors $T_M(A)$ est la droite vectorielle d'équation $\frac{\partial g}{\partial x}(M)x + \frac{\partial g}{\partial y}(M)y = 0$ et $\nabla g(M)$ est un vecteur normal à cette droite.

La droite affine $M + T_M(A)$ d'équation $\frac{\partial g}{\partial x}(M)(x - a) + \frac{\partial g}{\partial y}(M)(y - b) = 0$ est la tangente en M .

► Dans le cas où $p = 3$, A est appelé une surface de niveau d'équation $g(x, y, z) = 0$.

Si $M = (a, b, c) \in A$ et $dg(M) \neq 0$, alors $T_M(A)$ est le plan d'équation $\frac{\partial g}{\partial x}(M)x + \frac{\partial g}{\partial y}(M)y + \frac{\partial g}{\partial z}(M)z = 0$ et $\nabla g(M)$ est un vecteur normal à ce plan.

Le plan affine $M + T_M(A)$ d'équation $\frac{\partial g}{\partial x}(M)(x - a) + \frac{\partial g}{\partial y}(M)(y - b) + \frac{\partial g}{\partial z}(M)(z - c) = 0$ est le plan tangent en M .

Les physiciens utilisent aussi le terme d'équipotentielle. Par extension, en dimension quelconque, le vecteur $\nabla g(M)$ est appelé vecteur normal à la ligne de niveau d'équation $g(M) = 0$.

Exercice 16.85

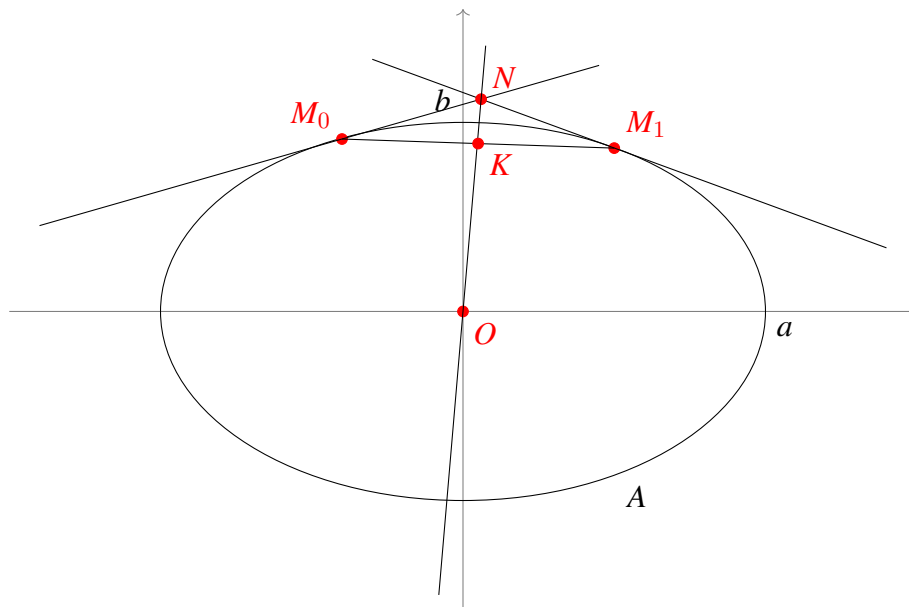
Soit A l'ellipse d'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Soient M_0 et M_1 deux points de A tels que les tangentes à A en M_0 et M_1 se coupent en un point N .

Montrez que la droite (ON) passe par le milieu de $[M_0M_1]$.

Correction 16.86

On représente le problème graphiquement :



En posant $g : (x, y) \mapsto \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$, on a

$$A = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}.$$

On a $dg(x, y) \cdot (h, k) = \frac{2x}{a^2}h + \frac{2y}{b^2}k$ donc

$$dg(x, y) = 0 \iff x = y = 0.$$

Or $(0, 0) \notin A$ donc $\forall (x, y) \in A, dg(x, y) \neq 0$.

On note $M_0 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ et $M_1 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$.

La tangente en M_0 est la droite d'équation

$$\frac{2x_0}{a^2} (x - x_0) + \frac{2y_0}{b^2} (y - y_0) = 0$$

i.e.

$$\frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1.$$

Idem, la tangente en M_1 est la droite d'équation

$$\frac{xx_1}{a^2} + \frac{yy_1}{b^2} = 1.$$

On détermine donc les coordonnées du point N en résolvant le système

$$\begin{cases} \frac{x_0}{a^2}x + \frac{y_0}{b^2}y = 1 \\ \frac{x_1}{a^2}x + \frac{y_1}{b^2}y = 1 \end{cases}$$

On pose $\Delta = \begin{vmatrix} x_0/a^2 & y_0/b^2 \\ x_1/a^2 & y_1/b^2 \end{vmatrix} = \frac{1}{a^2 b^2} (x_0 y_1 - x_1 y_0)$.

On a alors

$$x = \frac{\begin{vmatrix} 1 & y_0/b^2 \\ 1 & y_1/b^2 \end{vmatrix}}{\Delta} = \frac{a^2 (y_1 - y_0)}{x_0 y_1 - x_1 y_0} \quad \text{et} \quad y = \frac{\begin{vmatrix} x_0/a^2 & 1 \\ x_1/a^2 & 1 \end{vmatrix}}{\Delta} = \frac{b^2 (x_0 - x_1)}{x_0 y_1 - x_1 y_0}.$$

De plus, on a $K \begin{pmatrix} \frac{x_0 + x_1}{2} \\ \frac{y_0 + y_1}{2} \end{pmatrix}$.

En notant $\delta = x_0 y_1 - x_1 y_0$, on a alors

$$\begin{aligned} \det \left(\overrightarrow{ON}, \overrightarrow{OK} \right) &= \begin{vmatrix} \frac{a^2 (y_1 - y_0)}{\delta} & \frac{x_0 + x_1}{2} \\ \frac{b^2 (x_0 - x_1)}{\delta} & \frac{y_0 + y_1}{2} \end{vmatrix} \\ &= \frac{1}{2\delta} (a^2 (y_1^2 - y_0^2) - b^2 (x_0^2 - x_1^2)) \\ &= \frac{1}{2\delta} ((a^2 y_1^2 + b^2 x_1^2) - (a^2 y_0^2 + b^2 x_0^2)) \\ &= \frac{a^2 b^2}{2\delta} \left(\left(\frac{y_1^2}{b^2} + \frac{x_1^2}{a^2} \right) - \left(\frac{y_0^2}{b^2} + \frac{x_0^2}{a^2} \right) \right) \\ &= 0. \end{aligned}$$

Donc les vecteurs \overrightarrow{ON} et \overrightarrow{OK} sont colinéaires *i.e.* les points O , N et K sont alignés.

16.6 Optimisation au premier ordre

16.6.1 Vocabulaire

Définition 16.87

Soient f une fonction de E dans \mathbb{R} définie sur une partie A et $a_0 \in A$.

On dit que

- f possède un maximum local sur A en a_0 quand il existe $r > 0$ tel que

$$\forall a \in B(a_0, r) \cap A, \quad f(a) \leq f(a_0)$$

- f possède un minimum local sur A en a_0 quand il existe $r > 0$ tel que

$$\forall a \in B(a_0, r) \cap A, \quad f(a) \geq f(a_0)$$

- f possède un extremum local sur A en a_0 quand f possède un maximum ou un minimum local sur A en a_0 .

Définition 16.88

Soient f une fonction de E dans \mathbb{R} définie sur une partie A et $a_0 \in A$.

On dit que

- ▷ f possède un maximum global sur A en a_0 quand

$$\forall a \in A, \quad f(a) \leq f(a_0)$$

- ▷ f possède un minimum global sur A en a_0 quand

$$\forall a \in A, \quad f(a) \geq f(a_0)$$

- ▷ f possède un extremum global sur A en a_0 quand f possède un maximum ou un minimum global sur A en a_0 .

La recherche des points en lesquels une fonction possède des extrema (locaux ou globaux) dépend à la fois des propriétés de la fonction et de l'ensemble sur lequel la fonction est définie.

16.6.2 Points critiques, extrema locaux d'une fonction sur un ouvert

Définition 16.89

Soient U un ouvert de E et $f : U \rightarrow F$ différentiable sur U .

Un point critique de f est un point $a \in U$ tel que $df(a) = 0$.

Comme dans le cours de première année, on retrouve la condition nécessaire d'existence d'un extremum local pour les fonctions à valeurs réelles.

Proposition 16.90

Soient U un ouvert de E , $f : U \rightarrow \mathbb{R}$ différentiable sur U et $a \in U$.

Si f possède un extremum local en a , alors a est un point critique de f .

Démonstration 16.91

On pose $a = (a_1, \dots, a_p)$.

Si f possède un minimum local en a , alors il existe $r > 0$ tel que

$$\forall b \in B(a, r), \quad f(b) \geq f(a).$$

En particulier, pour tout $i \in \llbracket 1 ; p \rrbracket$, sur $[a_i - r ; a_i + r]$, la fonction $t \mapsto f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_p)$ a un minimum local en a_i donc sa dérivée en a_i est nulle i.e. $\partial_i f(a) = 0$.

Toutes les dérivées partielles en a sont nulles donc $df(a) = 0$.

Idem pour un maximum local. ■

Remarque 16.92

- La réciproque est fausse (contre-exemple : la selle de cheval $(x, y) \mapsto x^2 - y^2$).
- Ce résultat n'est valable que sur un ouvert, donc en particulier en tout point intérieur à une partie, mais pas sur les bords. En général, on distingue donc dans l'étude des extrema les points sur la frontière et ceux à l'intérieur.

Exercice 16.93

Déterminez les extrema de la fonction $(x, y) \mapsto x^3 + y^3 - 3xy$.

Correction 16.94

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

On a

$$\begin{aligned}(x, y) \text{ point critique} &\iff \begin{cases} \partial_1 f(x, y) = 0 \\ \partial_2 f(x, y) = 0 \end{cases} \\ &\iff \begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases} \\ &\iff \begin{cases} y = x^2 \\ x = y^2 \end{cases} \\ &\iff \begin{cases} x^4 = x \\ y = x^2 \end{cases} \\ &\iff \begin{cases} x = 0 \text{ ou } x = 1 \\ y = x^2 \end{cases}\end{aligned}$$

Donc f possède deux points critiques : $(0, 0)$ et $(1, 1)$.

Étude en $(0, 0)$:

On étudie le signe de $(x, y) \mapsto f(x, y) - f(0, 0)$ au voisinage de $(0, 0)$.

On a $f(x, 0) = x^3$ donc pour $x \in]0 ; +\infty[$ on a

$$f(x, 0) - f(0, 0) > 0$$

et pour $x \in]-\infty ; 0[$, on a

$$f(x, 0) - f(0, 0) < 0.$$

Donc f n'a pas d'extremum en $(0, 0)$.

Étude en $(1, 1)$:

On étudie le signe de $g : (h, k) \mapsto f(1+h, 1+k) - f(1, 1)$ au voisinage de $(0, 0)$.

On a

$$\begin{aligned}g(h, k) &= (1+h)^3 + (1+k)^3 - 3(1+h)(1+k) - f(1, 1) \\&= 3h^2 + h^3 + 3k^2 + k^3 - 3hk \\&= 3(h^2 + k^2 - hk) + h^3 + k^3 \\&= 3(h^2 + k^2 - hk) + (h+k)(h^2 - hk + k^2) \\&= (h^2 + k^2 - hk)(3 + h + k).\end{aligned}$$

$3 + h + k \xrightarrow{(h,k) \rightarrow (0,0)} 3 > 0$ donc au voisinage de $(0, 0)$, $3 + h + k \geq 0$.

De plus $h^2 - hk + k^2 = \left(h - \frac{k}{2}\right)^2 + \frac{3}{4}k^2 \geq 0$.

Donc g est à valeurs positives au voisinage de $(0, 0)$, donc f possède un minimum local en $(1, 1)$.

Exercice 16.95

Même question avec $(x, y) \mapsto x^3 + x^2 + y^2$.

Correction 16.96

On a

$$\begin{aligned}(x, y) \text{ point critique} &\iff \begin{cases} \partial_1 f(x, y) = 0 \\ \partial_2 f(x, y) = 0 \end{cases} \\&\iff \begin{cases} 3x^2 + 2x = 0 \\ 2y = 0 \end{cases} \\&\iff \begin{cases} x(3x + 2) = 0 \\ y = 0 \end{cases} \\&\iff \begin{cases} x = 0 \text{ ou } x = -\frac{2}{3} \\ y = 0 \end{cases}\end{aligned}$$

Donc f possède deux points critiques : $(0, 0)$ et $(-2/3, 0)$.

Étude en $(0, 0)$:

On a $f(x, y) = x^2(x + 1) + y^2 \geq 0$ au voisinage de $(0, 0)$ donc f possède un minimum local en $(0, 0)$.

Étude en $(-2/3, 0)$:

On pose $g : (h, k) \mapsto f(-2/3 + h, k) - f(-2/3, 0)$.

On a

$$\begin{aligned}g(h, k) &= (-2/3 + h)^3 + (-2/3 + h)^2 + k^2 - f(-2/3, 0) \\&= 3(-2/3)h^2 + h^3 + h^2 + k^2 \\&= -h^2 + k^2 + h^3.\end{aligned}$$

Donc $g(h, 0) = -h^2 + h^3 \underset{h \rightarrow 0}{\sim} -h^2$ donc $g(h, 0) \leq 0$ au voisinage de 0.

Or $g(0, k) = k^2 \geq 0$ au voisinage de 0.

Donc g n'a pas d'extremum en $(-2/3, 0)$.

Exercice 16.97

Même question avec $(x, y) \mapsto x^2 + x^2y + y^3$.

Correction 16.98

On a

$$\begin{aligned} (x, y) \text{ point critique} &\iff \begin{cases} \partial_1 f(x, y) = 0 \\ \partial_2 f(x, y) = 0 \end{cases} \\ &\iff \begin{cases} 2x(1+y) = 0 \\ x^2 + 3y^2 = 0 \end{cases} \\ &\iff \begin{cases} x = 0 \\ y = 0 \end{cases} \quad \text{ou} \quad \begin{cases} y = -1 \\ x^2 = -3 \end{cases} \\ &\iff \begin{cases} x = 0 \\ y = 0 \end{cases} \end{aligned}$$

Donc f possède un point critique en $(0, 0)$.

On remarque $f(x, y) = x^2(1+y) + y^3$ donc $f(0, y) = y^3$.

Donc f ne possède pas d'extremum.

16.6.3 Extrema locaux d'une fonction sur une partie

La recherche des extrema locaux sur une partie quelconque est souvent un problème difficile. Néanmoins, on dispose de quelques résultats.

D'abord, on généralise le théorème précédent (admis).

Proposition 16.99

Soient U un ouvert de E , $f : U \rightarrow \mathbb{R}$ différentiable sur U , A une partie de U et $a \in A$.

Si f possède un extremum local sur A en a , alors $df(a)$ s'annule sur $T_a(A)$.

En conséquence, on a un résultat dans certains cas particuliers d'équipotentielles, appelé théorème d'optimisation sous une contrainte.

Proposition 16.100

Soient U un ouvert de E et $f, g : U \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 .

Soient $A = g^{-1}(\{0\})$ et $a \in A$.

Si f possède un extremum local sur A en a et si $dg(a) \neq 0$, alors $df(a)$ est colinéaire à $dg(a)$, ce qui revient à dire que les vecteurs gradients de f et g en a sont colinéaires.

Remarque 16.101

Là encore, il s'agit de conditions nécessaires mais pas suffisantes en général. Une fois les points candidats trouvés, il faut toujours une étude locale pour les accepter ou non, ou bien utiliser un théorème d'existence comme le théorème des bornes atteintes.

Exercice 16.102

Déterminez les extrema de la fonction $(x, y) \mapsto xy$ sur la courbe d'équation $x^4 + y^4 = 1$.

Correction 16.103

On pose $f : (x, y) \mapsto xy$, $g : (x, y) \mapsto x^4 + y^4 - 1$ et $A = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}$.

On a $\nabla f(x, y) \begin{pmatrix} y \\ x \end{pmatrix}$ et $\nabla g(x, y) \begin{pmatrix} 4x^3 \\ 4y^3 \end{pmatrix}$.

Si f possède un extremum local en $(x, y) \in A$, alors

$$\begin{aligned} \begin{cases} \det(\nabla f(x, y), \nabla g(x, y)) = 0 \\ x^4 + y^4 = 1 \end{cases} &\iff \begin{cases} 4y^4 - 4x^4 = 0 \\ x^4 + y^4 = 1 \end{cases} \\ &\iff \begin{cases} y^4 = x^4 \\ 2x^4 = 1 \end{cases} \end{aligned}$$

On a donc quatre points critiques sous contrainte : $\begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$, $\begin{pmatrix} -\alpha \\ -\alpha \end{pmatrix}$, $\begin{pmatrix} \alpha \\ -\alpha \end{pmatrix}$ et $\begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$ en posant $\alpha = \frac{1}{\sqrt[4]{2}}$.

g est continue et $A = g^{-1}(\{0\})$ donc A est un fermé.

A est bornée car si $(x, y) \in A$, on a $x^4 \leq x^4 + y^4 = 1$ donc $|x| \leq 1$ et $|y| \leq 1$.

Donc A est un compact car \mathbb{R}^2 est de dimension finie.

D'après le théorème des bornes atteintes, f étant continue, elle est bornée sur A et atteint ses bornes en des points parmi les quatre précédents.

Or $f(\alpha, \alpha) = f(-\alpha, -\alpha) = \alpha^2 > -\alpha^2 = f(\alpha, -\alpha) = f(-\alpha, \alpha)$.

Donc f a des maxima globaux sur A en (α, α) et $(-\alpha, -\alpha)$ et des minima globaux sur A en $(\alpha, -\alpha)$ et $(-\alpha, \alpha)$.

Exercice 16.104

Même question avec la fonction $(x, y) \mapsto x^3 + 2y^3$ sur la courbe d'équation $x^2 - y^2 = 1$.

Correction 16.105

On pose $f : (x, y) \mapsto x^3 + 2y^3$, $g : (x, y) \mapsto x^2 - y^2 - 1$ et $A = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) = 0\}$.

On a $\nabla f(x, y) \begin{pmatrix} 3x^2 \\ 6y^2 \end{pmatrix}$ et $\nabla g(x, y) \begin{pmatrix} 2x \\ -2y \end{pmatrix}$.

Si f possède un extremum local en $(x, y) \in A$, alors

$$\begin{cases} -6x^2y - 12xy^2 = 0 \\ x^2 - y^2 = 1 \end{cases} \iff \begin{cases} -6xy(x + 2y) = 0 \\ x^2 - y^2 = 1 \end{cases}$$

Donc

$$\begin{cases} y = 0 \\ x^2 = 1 \end{cases} \quad \text{ou} \quad \begin{cases} x = -2y \\ y^2 = 1/3 \end{cases}$$

On a donc quatre points critiques sous contrainte : $C \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -2/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$ et $B \begin{pmatrix} 2/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}$.

On pose $M = (x, y)$.

On remarque

$$|f(M)| \xrightarrow[d(M,0) \rightarrow +\infty]{M \in A} +\infty.$$

Sur la demi-hyperbole A_+ de droite, on fait la conjecture suivante :

M	B	C
$f(M)$	$ \begin{array}{ccc} & 2/\sqrt{3} & \\ \nearrow & & \searrow \\ -\infty & & 1 \end{array} $	$+\infty$

On a

$$\forall K > 0, \exists r > 0, \forall M \in A, d(M, 0) > r \implies |f(M)| \geq K.$$

Avec $K = 42$, il existe $r > 0$ tel que

$$\forall M \notin \overline{B}(0, r), |f(M)| \geq 42.$$

On a $|f(C)| = 1 < 42$ et $|f(B)| = \frac{2}{\sqrt{3}} < 42$ donc $B, C \in \overline{B}(0, r)$.

$\Gamma_+ = A_+ \cap \overline{B}(0, r)$ est un fermé (intersection de deux fermés) borné et donc un compact car \mathbb{R}^2 est de dimension finie.

Sur ce compact, f possède un maximum et un minimum.

Pour $M \in A_+$, on a $x^2 = 1 + y^2$ donc $x = \sqrt{1 + y^2}$.

Donc $f(M) = (1 + y^2)^{3/2} + 2y^3 = \varphi(y)$.

On a $\varphi'(y) = 3y(2y + \sqrt{1 + y^2})$ donc

$$\varphi'(y) = 0 \iff \left[y = 0 \quad \text{ou} \quad y = \frac{-1}{\sqrt{3}} \right].$$

On a donc

y	$-\infty$	$-1/\sqrt{3}$	0	$+\infty$	
$\varphi'(y)$	$+$	0	$-$	0	$+$
φ	$-\infty$	$2/\sqrt{3}$	1	$+\infty$	

Idem sur A_- .

Exercice 16.106

Même question avec la fonction $(x, y, z) \mapsto x + y + z$ sur la surface d'équation $x^2 + y^2 - 2z^2 = 2$.

Correction 16.107

On pose $f : (x, y, z) \mapsto x + y + z$, $g : (x, y, z) \mapsto x^2 + y^2 - 2z^2 - 2$ et $A = \{(x, y, z) \in \mathbb{R}^3 \mid g(x, y, z) = 0\}$.

On a $\nabla f(x, y, z) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ et $\nabla g(x, y, z) \begin{pmatrix} 2x \\ 2y \\ -4z \end{pmatrix}$.

On a

$$\begin{aligned} \nabla f(x, y, z) \text{ et } \nabla g(x, y, z) \text{ colinéaires} &\iff \begin{cases} 2x = 2y \\ 2x = -4z \end{cases} \\ &\iff \begin{cases} x = y \\ z = \frac{-1}{2}x \end{cases} \end{aligned}$$

Donc

$$\begin{cases} y = x \\ z = \frac{-1}{2}x \\ x^2 + y^2 - 2z^2 = 2 \end{cases} \iff \begin{cases} y = x \\ z = \frac{-1}{2}x \\ x^2 = \frac{4}{3} \end{cases}$$

On a donc deux points critiques sous contraintes : $B \begin{pmatrix} 2/\sqrt{3} \\ 2/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}$ et $C \begin{pmatrix} -2/\sqrt{3} \\ -2/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$.

On pose s la symétrie centrale de centre 0 et on a

$$\forall M \in A, [s(M) \in A \quad \text{et} \quad f \circ s(M) = -f(M)].$$

Donc si f possède un maximum (respectivement minimum) local en B , f a un minimum (respectivement maximum) local en C .

Si $M = (x, y, z)$ est au voisinage de B , alors $z < 0$ donc $z = -\sqrt{\frac{x^2 + y^2 - 2}{2}}$.

On cherche le signe de $f(M) - f(B)$ au voisinage de B :

$$f(M) - f(B) = F(x, y) = x + y - \sqrt{\frac{x^2 + y^2 - 2}{2}} - f(B) = G(h, k).$$

On pose $\alpha = \frac{2}{\sqrt{3}}$ et $\begin{cases} x = \alpha + h \\ y = \alpha + k \end{cases}$

On a

$$\begin{aligned} G(h, k) &= \alpha + h + \alpha + k - \sqrt{\frac{(\alpha + h)^2 + (\alpha + k)^2}{2}} - 1 - f(B) \\ &= h + k - \sqrt{\frac{h^2 + k^2}{2} + \alpha(h + k) + \alpha^2 - 1 + 2\alpha} - f(B) \\ &= h + k - \frac{1}{\sqrt{3}} \sqrt{1 + \frac{\frac{h^2 + k^2}{2} + \alpha(h + k)}{1/3}} + 2\alpha - f(B) \\ &= \cancel{h+k} - \frac{1}{\sqrt{3}} \left[\cancel{1} + \frac{3(h^2 + k^2)}{4} + \cancel{\frac{3\alpha}{2}(h+k)} - \frac{1}{8} \left(\underbrace{\frac{3}{2}(h^2 + k^2) + 3\alpha(h + k)}_u \right)^2 + u^2 \varepsilon(u) \right] + \cancel{2\alpha - f(B)} \\ &= -\frac{\sqrt{3}}{4} (h^2 + k^2) + \frac{1}{8\sqrt{3}} (9\alpha^2 (h^2 + k^2 + 2hk)) + (h^2 + k^2) \gamma(h, k) \\ &= \frac{\sqrt{3}}{4} (h^2 + k^2 + 4hk) + (h^2 + k^2) \gamma(h, k) \\ &= \frac{\sqrt{3}}{4} ((h + 2k)^2 - 3k^2) + (h^2 + k^2) \gamma(h, k) \end{aligned}$$

16.7 Fonctions de classe \mathcal{C}^k

16.7.1 Dérivées partielles d'ordre supérieur

Pour une fonction quelconque définie sur un ouvert de \mathbb{R}^p , on peut éventuellement définir récursivement ses dérivées partielles d'ordre k comme étant les dérivées partielles des dérivées partielles d'ordre $(k-1)$ quand cela a un sens.

Évidemment, l'ordre dans lequel on effectue les dérivations a une importance.

Si $i = (i_1, \dots, i_k) \in \llbracket 1 ; p \rrbracket^k$ est un multi-indice, on note $\partial_i f$, $\partial_{i_1} \partial_{i_2} \dots \partial_{i_k} f$ ou $\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}$ la dérivée partielle $\partial_{i_1} (\partial_{i_2} \dots \partial_{i_k} f)$.

L'ordre des dérivées partielles se lit donc de la droite vers la gauche.

A priori si une fonction a p variables, elle peut posséder jusqu'à p^k dérivées partielles d'ordre k .

16.7.2 Fonctions de classe \mathcal{C}^k

La classe \mathcal{C}^0 étant celle des fonctions continues, on peut donner une définition récursive de la classe \mathcal{C}^k pour $k \in \mathbb{N}$.

Définition 16.108

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow F$ et $k \in \mathbb{N}^*$.

On dit que f est de classe \mathcal{C}^k sur U quand elle a des dérivées partielles en tout point de U et que celles-ci sont de classe \mathcal{C}^{k-1} sur U .

On dit que f est de classe \mathcal{C}^∞ sur U quand pour tout $k \in \mathbb{N}$, f est de classe \mathcal{C}^k .

Par application récursive des théorèmes d'opérations précédents, on obtient les théorèmes d'opérations sur les fonctions de classe \mathcal{C}^k sans difficulté.

16.7.3 Théorème de Schwarz

Dans le cas des fonctions de classe \mathcal{C}^k , il y a finalement bien moins de dérivées partielles que prévu.

Théorème 16.109

Soient U un ouvert de \mathbb{R}^p et $f : U \longrightarrow F$.

Si f est de classe \mathcal{C}^2 sur U , alors pour tout $(i, j) \in \llbracket 1 ; p \rrbracket^2$, $\partial_i (\partial_j f) = \partial_j (\partial_i f)$.

On en déduit par récurrence que pour toute fonction de classe \mathcal{C}^k , l'ordre des dérivations (jusqu'à k dérivations successives) n'importe pas.

16.8 Optimisation au second ordre

16.8.1 Hessienne

Définition 16.110

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $a \in U$.

On appelle hessienne de f en a la matrice $H_f(a) = (h_{ij}) \in \mathcal{M}_p(\mathbb{R})$ telle que

$$\forall (i, j) \in \llbracket 1 ; p \rrbracket^2, \quad h_{ij} = \partial_i \partial_j f(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

D'après le théorème de Schwarz, la hessienne de f est alors une matrice symétrique.

On lui associe l'application $h_f(a)$ de \mathbb{R}^p dans \mathbb{R} définie de la façon suivante :

$$\forall v = \begin{pmatrix} v_1 \\ \vdots \\ v_p \end{pmatrix} \in \mathbb{R}^p, \quad h_f(a) \cdot v = v^\top H_f(a) v.$$

C'est un produit scalaire (canonique) entre le vecteur-colonne v et $H_f(a) v$. Cette application est appelée la forme hessienne de f en a (à titre culturel, on appelle ce genre d'application des formes quadratiques).

En clair, on a

$$h_f(a) \cdot v = \sum_{1 \leq i, j \leq p} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j.$$

16.8.2 Développement limité à l'ordre 2

On admet le résultat suivant, appelé formule de Taylor-Young à l'ordre 2.

Proposition 16.111

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $a \in U$.

Pour tout vecteur v au voisinage de 0, on a

$$f(a + v) = f(a) + df(a) \cdot v + \frac{1}{2} h_f(a) \cdot v + o(\|v\|^2).$$

Ce développement limité est souvent écrit à l'aide du gradient et de la hessienne :

$$f(a + v) = f(a) + \nabla f(a) \cdot v + \frac{1}{2} (H_f(a) v) \cdot v + o(\|v\|^2).$$

16.8.3 Application à l'étude des points critiques

Proposition 16.112

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $a \in U$.

Si f possède un minimum local en a , alors a est un point critique de f et $H_f(a) \in \mathcal{S}_p^+(\mathbb{R})$.

Remarque 16.113

Attention ! La réciproque est fausse. Néanmoins, elle est « presque vraie » en ajoutant une précision.

Proposition 16.114

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $a \in U$.

Si a est un point critique de f et $H_f(a) \in \mathcal{S}_p^{++}(\mathbb{R})$, alors f possède un minimum local strict en a .

Cas particulier très courant : les applications à deux variables.

Soient f une fonction de classe \mathcal{C}^2 sur un ouvert U de \mathbb{R}^2 et $a \in U$.

On a

$$\nabla f(a) = \begin{pmatrix} \frac{\partial f}{\partial x}(a) \\ \frac{\partial f}{\partial y}(a) \end{pmatrix} \quad \text{et} \quad H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x \partial y}(a) \\ \frac{\partial^2 f}{\partial x \partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a) \end{pmatrix}.$$

Cette matrice est définie-positive ssi sa trace et son déterminant sont strictement positifs.

En remplaçant f par $-f$, on en déduit des résultats similaires à propos des maxima locaux.

Proposition 16.115

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $a \in U$.

Si f possède un maximum local en a , alors a est un point critique de f et $H_f(a) \in \mathcal{S}_p^-(\mathbb{R})$ (i.e. c'est une matrice symétrique négative, ses valeurs propres sont négatives).

Proposition 16.116

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $a \in U$.

Si a est un point critique de f et $H_f(a) \in \mathcal{S}_p^{--}(\mathbb{R})$ (i.e. c'est une matrice définie-négative, ses valeurs propres sont strictement négatives), alors f possède un maximum local strict en a .

Pour une application à deux variables, sa hessienne est définie-négative ssi sa trace est strictement négative et son déterminant strictement positif.

Proposition 16.117

Soient U un ouvert de \mathbb{R}^p , $f : U \longrightarrow \mathbb{R}$ de classe \mathcal{C}^2 et $a \in U$.

Si a est un point critique de f et $H_f(a)$ est une matrice symétrique ayant deux valeurs propres non-nulles de signes opposés, alors f ne possède pas d'extremum en a (le point a est appelé point-col ou point-selle).

Correction 16.118 (Reprise de l'Exercice 16.93 avec la hessienne)

On pose $f : (x, y) \longmapsto x^3 + y^3 - 3xy$.

$A \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ et $B \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sont des points critiques.

On a $\frac{\partial f}{\partial x}(x, y) = 3x^2 - 3y$ donc

$$\frac{\partial^2 f}{\partial x^2} = 6x \quad \frac{\partial^2 f}{\partial x \partial y} = -3 \quad \frac{\partial^2 f}{\partial y^2} = 6y.$$

Donc

$$H_f(A) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$$

a pour valeurs propres 3 et -3, donc f n'admet pas d'extremum en A .

De plus, en posant

$$M = H_f(B) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

on a $\text{tr}(M) > 0$, $\det(M) > 0$ et M est symétrique réelle donc M est définie-positive, et donc f a un minimum local en B .

Exercice 16.119

Déterminez les extrema de la fonction $(x, y) \longmapsto 2y^2 + 2x^2 - x^4$.

Correction 16.120

On pose $f : (x, y) \longmapsto 2y^2 + 2x^2 - x^4$.

On a

$$\begin{cases} 4y = 0 \\ 4x(1 - x^2) = 0 \end{cases} \iff \begin{cases} y = 0 \\ x = 0 \text{ ou } x^2 = 1 \end{cases}$$

On a donc trois points critiques : $(0, 0)$, $(1, 0)$ et $(-1, 0)$.

On a

$$\frac{\partial^2 f}{\partial x^2} = -12x^2 + 4 \quad \frac{\partial^2 f}{\partial y^2} = 4 \quad \frac{\partial^2 f}{\partial x \partial y} = 0.$$

Donc

$$H_f(x, y) = \begin{pmatrix} 4 - 12x^2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Donc

$$H_f(0, 0) = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \in \mathcal{S}_2^{++}(\mathbb{R})$$

donc f admet un minimum local en $(0, 0)$.

De plus, on a

$$H_f(1, 0) = \begin{pmatrix} -8 & 0 \\ 0 & 4 \end{pmatrix} \in \mathcal{S}_2(\mathbb{R})$$

or les valeurs propres sont non-nulles mais opposées donc f n'admet pas d'extremum en $(1, 0)$.

Idem en $(-1, 0)$.

Exercice 16.121

Même question avec $(x, y) \mapsto y(x^2 + \ln^2 y)$.

Correction 16.122

On pose $f : (x, y) \mapsto y(x^2 + \ln^2 y)$.

On a

$$\begin{cases} 2yx = 0 \\ x^2 + \ln^2 y + 2y \frac{\ln y}{y} = 0 \end{cases} \iff \begin{cases} 2xy = 0 \\ x^2 + \ln^2 y + 2 \ln y = 0 \end{cases}$$

$$\text{donc } \begin{cases} x = 0 \\ \ln^2 y + 2 \ln y = 0 \end{cases}$$

$$\text{donc } \begin{cases} x = 0 \\ \ln y (\ln y + 2) = 0 \end{cases}$$

$$\text{donc } \begin{cases} x = 0 \\ y = 1 \text{ ou } y = e^{-2} \end{cases}$$

On a donc deux points critiques : $(0, 1)$ et $(0, e^{-2})$.

f est \mathcal{C}^2 sur $\mathbb{R} \times \mathbb{R}_+^*$ donc

$$H_f(x, y) = \begin{pmatrix} 2y & 2x \\ 2x & 2\frac{\ln y}{y} + \frac{2}{y} \end{pmatrix}.$$

Donc

$$H_f(0, 1) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \in \mathcal{S}_2^{++}(\mathbb{R})$$

donc f admet un minimum local en $(0, 1)$.

De même on a

$$H_f(0, e^{-2}) = \begin{pmatrix} 2e^{-2} & 0 \\ 0 & -\frac{2}{e^{-2}} \end{pmatrix} \in \mathcal{S}_2(\mathbb{R})$$

or les valeurs propres non-nulles sont opposées donc f n'admet pas d'extremum en $(0, e^{-2})$.

Exercice 16.123

Soient $n \in \mathbb{N}^*$ et $a_1, \dots, a_n \in \mathbb{R}$.

On pose $f : (x, y) \mapsto \frac{1}{2x} \sum_{i=1}^n (y - a_i)^2 + \frac{n}{2} \ln x$.

Déterminez, s'ils existent, les extrema de f .

Correction 16.124

f est \mathcal{C}^2 sur $U = \mathbb{R}_+^* \times \mathbb{R}$.

On a

$$\partial_1 f(x, y) = -\frac{1}{2x^2} \sum_{i=1}^n (y - a_i)^2 + \frac{n}{2x} = \frac{1}{2x} \left(-\frac{1}{n} \sum_{i=1}^n (y - a_i)^2 + n \right)$$

et

$$\partial_2 f(x, y) = \frac{1}{x} \sum_{i=1}^n (y - a_i).$$

On a

$$\begin{aligned} \begin{cases} \frac{1}{2x} \left(-\frac{1}{n} \sum_{i=1}^n (y - a_i)^2 + n \right) = 0 \\ \frac{1}{x} \sum_{i=1}^n (y - a_i) = 0 \end{cases} &\iff \begin{cases} nx = \sum_{i=1}^n (y - a_i)^2 \\ \sum_{i=1}^n (y - a_i) = 0 \end{cases} \\ &\iff \begin{cases} y = m \\ x = \frac{1}{n} \sum_{i=1}^n (m - a_i)^2 = v \end{cases} \end{aligned} \quad \left. \begin{array}{l} \\ \end{array} \right\} m = \frac{1}{n} \sum_{i=1}^n a_i$$

Si $a_1 = \dots = a_n = m$, alors $x = 0$ donc f n'admet pas de point critique.

On suppose les a_i non tous égaux.

Réciproquement, avec $x = v$ et $y = m$, on a $df(x, y) = 0$.

Dans ce cas, il y a un point critique : (v, m) .

On a

$$\frac{\partial^2 f}{\partial x^2}(x, y) = \frac{1}{x^3} \sum_{i=1}^n (y - a_i)^2 - \frac{n}{2x^2} \quad \frac{\partial^2 f}{\partial y^2}(x, y) = \frac{n}{x} \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = -\frac{1}{x^2} \sum_{i=1}^n (y - a_i).$$

Donc

$$H_f(v, m) = \begin{pmatrix} \frac{1}{v^3} \sum_{i=1}^n (m - a_i)^2 - \frac{n}{2v^2} & -\frac{1}{v^2} \sum_{i=1}^n (m - a_i) \\ -\frac{1}{v^2} \sum_{i=1}^n (m - a_i) & \frac{n}{v} \end{pmatrix}.$$

Or $\sum_{i=1}^n (m - a_i) = nm - \sum_{i=1}^n a_i = 0$ donc

$$\frac{1}{v^3} \sum_{i=1}^n (m - a_i)^2 - \frac{n}{2v^2} = \frac{nv}{v^3} - \frac{n}{2v^2} = \frac{n}{2v^2}.$$

Donc

$$H_f(v, m) = \begin{pmatrix} \frac{n}{2v^2} & 0 \\ 0 & \frac{n}{v} \end{pmatrix} \in \mathcal{S}_2^{++}(\mathbb{R})$$

et donc f admet un minimum local en (v, m) .

De plus, on a

$$\begin{aligned} f(v, m) &= \frac{1}{2v} \sum_{i=1}^n (m - a_i)^2 + \frac{n}{2} \ln v \\ &= \frac{n}{2} (1 + \ln v). \end{aligned}$$

Soit $(a, b) \neq (0, 0)$ avec $a > 0$.

On pose $\varphi : t \mapsto f((v, m) + t(a, b))$ définie sur $\left] -\frac{v}{a} ; +\infty \right[$.

On a

$$\varphi(t) = \frac{1}{2(v+at)} \sum_{i=1}^n (m+tb-a_i)^2 + \frac{n}{2} \ln(v+at).$$

Quand $t \rightarrow -\frac{v}{a}$ par valeurs supérieures, on a $\ln(v+at) = o\left(\frac{1}{v+at}\right)$.

Donc

$$\varphi(t) \underset{t \rightarrow -v/a}{\sim} \frac{1}{2(v+at)} \sum_{i=1}^n (m+bt-a_i)^2 \xrightarrow[t \rightarrow -v/a]{} +\infty.$$

On a donc

t	$-\frac{v}{a}$	0	$+\infty$
$\varphi'(t)$	\parallel	$- \quad 0 \quad +$	
φ	$+\infty$	$f(v, m)$	$+\infty$

On pose $\psi : \mathbb{R} \longrightarrow \mathbb{R}$
 $t \longmapsto f((v, m) + t(0, 1))$

On a $\psi(t) = \frac{1}{2v} \sum_{i=1}^n (m - t + a_i)^2 + \frac{n}{2} \ln v$.

Donc, idem, on a

t	$-\infty$	0	$+\infty$
$\psi'(t)$	$-$	0	$+$
ψ	$+\infty$	$f(v, m)$	$+\infty$

Donc f admet un minimum global en (v, m) .

Exercice 16.125

Déterminez les extrema de la fonction $(x, y) \longmapsto x^3 + y^2$.

Correction 16.126

On pose $f : (x, y) \longmapsto x^3 + y^2$.

On a $\partial_1 f(x, y) = 3x^2$ et $\partial_2 f(x, y) = 2y$.

f possède donc un unique point critique : $(0, 0)$.

On a donc

$$H_f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 2 \end{pmatrix}.$$

Donc

$$H_f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \in \mathcal{S}_2^+(\mathbb{R}) \setminus \mathcal{S}_2^{++}(\mathbb{R})$$

donc on ne peut pas conclure.

Cependant, $f(x, 0) - f(0, 0) = x^3$ n'est pas de signe constant au voisinage de 0 donc f n'admet pas d'extremum.

Chapitre 17

Structures algébriques

Sommaire

17.1	Groupes.	531
17.1.1	Rappels	531
17.1.1.1	Généralités	531
17.1.1.2	Sous-groupes	532
17.1.1.3	Morphismes de groupes	533
17.1.2	Ordre d'un élément dans un groupe	534
17.1.3	Sous-groupe engendré par une partie	537
17.1.4	Théorème de Lagrange	541
17.2	Anneaux : rappels	542
17.2.1	Généralités	542
17.2.2	Sous-anneaux	543
17.2.3	Pièges calculatoires classiques	543
17.3	Anneau $\mathbb{Z}/n\mathbb{Z}$.	545
17.3.1	Définition	545
17.3.2	Propriétés	545
17.3.3	Théorème chinois des restes	547
17.3.4	Indicatrice d'Euler	550
17.4	Idéaux d'un anneau commutatif.	551
17.4.1	Généralités	551
17.4.2	Arithmétique dans un anneau principal	556
17.4.2.1	Pgcd et ppcm	556
17.4.2.2	Décomposition en facteurs irréductibles	558
17.5	\mathbb{K}-algèbre	559
17.5.1	Généralités	559
17.5.2	Sous-algèbres	560
17.5.3	Morphismes d'algèbres	560
17.5.4	Polynôme minimal d'un élément	560

Dans ce chapitre, comme toujours, \mathbb{K} désigne un sous-corps de \mathbb{C} , le plus souvent \mathbb{R} ou \mathbb{C} .

17.1 Groupes

17.1.1 Rappels

17.1.1.1 Généralités

Définition 17.1

Soit G un ensemble.

On dit que $(G, *)$ est un groupe quand

- ▷ $*$ est une loi de composition interne sur G
- ▷ $*$ est associative
- ▷ $*$ admet un neutre
- ▷ tout élément de G est symétrisable pour $*$ dans G .

Si, de plus, $*$ est commutative, on dit que $(G, *)$ est un groupe commutatif ou abélien.

Proposition 17.2 (Simplification)

Dans un groupe muni d'une loi $*$:

- ▷ si $x * y = x * z$, alors $y = z$
- ▷ si $y * x = z * x$, alors $y = z$.

Attention à l'ordre des éléments si le groupe n'est pas commutatif!

On ne note souvent la loi $+$ (on dit que la loi est notée additivement) que dans le cas où le groupe est abélien, le neutre est noté 0 et $-x$ le symétrique d'un élément x , appelé bien sûr l'opposé.

Sinon, la loi est notée \times ou $*$ ou $.$ en général, le neutre est noté 1 ou e ou I et le symétrique d'un élément x est noté x^{-1} (on dit que la loi du groupe est notée multiplicativement).

Exemple 17.3

- ▷ $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ et $(\mathbb{C}, +)$ sont des groupes abéliens, ainsi que $(\{-1, 1\}, \times)$, (\mathbb{K}^*, \times) , (\mathbb{R}_+^*, \times) et (\mathbb{U}, \times) .
- ▷ Soit A un ensemble. En notant \mathfrak{S}_A l'ensemble des bijections de A dans A (aussì appelées permutations de A), on a alors une structure de groupe (\mathfrak{S}_A, \circ) , non-abélien en général (dès que A a plus de trois éléments), appelé groupe symétrique de A . Un cas particulier important : $\mathfrak{S}_n = \mathfrak{S}_{[1;n]}$.

- Si E est un \mathbb{K} -espace vectoriel, on a le groupe $(\text{GL}(E), \circ)$, non-abélien en général (dès que $\dim E \geq 2$), appelé le groupe linéaire de E . De même, on a le groupe $(\text{GL}_n(\mathbb{K}), \times)$, non-abélien en général, appelé le groupe linéaire d'indice n .
- Attention ! $\mathcal{L}(E)$ n'est pas un groupe pour la loi \circ et $\mathcal{M}_n(\mathbb{K})$ n'est pas un groupe pour la loi \times . Ce sont des groupes pour la loi $+$.

Soient $(G, .)$ un groupe, $a \in G$ et $n \in \mathbb{N}^*$.

On pose $a^0 = e$ et $a^{n+1} = a^n . a = \underbrace{a . a \dots a}_{n \text{ facteurs}} . a$ et $a^{-n} = (a^{-1})^n$.

Dans le cas d'une loi additive, on note plutôt $n.a$ au lieu de a^n .

On vérifie que les règles de calcul habituelles restent valables.

Proposition 17.4

Pour tout $(m, n) \in \mathbb{Z}^2$, on a

$$a^{m+n} = a^m . a^n \quad \text{et} \quad a^{mn} = (a^m)^n .$$

En revanche, la règle habituelle de calcul $(a.b)^n = a^n . b^n$ n'est pas toujours vraie !

Définition 17.5

On dit que deux éléments a et b commutent quand $a.b = b.a$.

Proposition 17.6

Dans le groupe $(G, .)$, si a et b sont deux éléments qui commutent, alors

$$\forall n \in \mathbb{Z}, (a.b)^n = a^n . b^n .$$

Remarque 17.7

Dans le cas d'une notation additive, on a $(m+n).a = m.a + n.a$ et $(mn).a = m.(n.a)$, et comme $+$ est commutative, $m.(a+b) = m.a + m.b$.

17.1.1.2 Sous-groupes

Définition 17.8

Soient $(G, *)$ un groupe et H une partie de G .

On dit que H est un sous-groupe de G (sous-entendu pour la loi $*$) quand

- ▷ H est non-vide
- ▷ H est stable par $*$: $\forall (x, y) \in H^2, x * y \in H$
- ▷ H est stable par inversion pour $*$: $\forall x \in H, x^{-1} \in H$.

Une conséquence simple des trois propriétés est que H contient forcément le neutre. En pratique, pour montrer que H est non-vide, on montre donc souvent que H contient le neutre.

Il est facile de voir que $*$ induit une application de $H \times H$ dans H (qui est donc une loi de composition interne), qu'on note encore abusivement $*$, et que cette loi de composition interne $*$ dans H définit une structure de groupe, d'où le nom de sous-groupe.

Un sous-groupe d'un groupe est donc lui-même un groupe. En pratique, on démontre très rarement directement qu'un ensemble est un groupe. Dans presque tous les cas, pour montrer qu'un ensemble muni d'une loi de composition interne est un groupe, on montre que c'est un sous-groupe d'un groupe connu.

Un exemple fondamental : les sous-groupes de $(\mathbb{Z}, +)$.

Proposition 17.9

Les sous-groupes de $(\mathbb{Z}, +)$ sont exactement les ensembles $n\mathbb{Z}$ pour $n \in \mathbb{N}$.

Le résultat suivant a déjà été prouvé dans le cadre des sous-espaces vectoriels.

Proposition 17.10

Une intersection de sous-groupes est encore un sous-groupe.

17.1.1.3 Morphismes de groupes

Définition 17.11

Soient $(E, *)$ et $(F, .)$ deux groupes et $f : E \longrightarrow F$.

On dit que f est un morphisme de groupes de $(E, *)$ dans $(F, .)$ quand

$$\forall (a, b) \in E^2, f(a * b) = f(a) . f(b).$$

De plus, si f est bijective, alors on dit que f est un isomorphisme de groupes.

Exemple 17.12

- ▷ $x \mapsto 3x$ est un morphisme de $(\mathbb{Z}, +)$ dans $(\mathbb{Z}, +)$.
- ▷ $t \mapsto e^{it}$ est un morphisme de $(\mathbb{R}, +)$ dans (\mathbb{C}^*, \times) .
- ▷ \ln est un isomorphisme de (\mathbb{R}_+^*, \times) dans $(\mathbb{R}, +)$.
- ▷ \det est un morphisme de $(\mathrm{GL}_n(\mathbb{K}), \times)$ dans (\mathbb{K}^*, \times) .

On rappelle quelques résultats concernant les morphismes de groupes.

Proposition 17.13

*Soient $(E, *)$ et $(F, .)$ deux groupes, dont on note e et ε les neutres respectifs et f un morphisme de groupes de $(E, *)$ dans $(F, .)$.*

On a

$$f(e) = \varepsilon \quad \text{et} \quad \forall x \in E, \quad f(x^{-1}) = f(x)^{-1}.$$

Proposition 17.14

*Soient $(E, *)$ et $(F, .)$ deux groupes et f un morphisme de groupes de $(E, *)$ dans $(F, .)$.*

Si G est un sous-groupe de E , alors $f(G)$ est un sous-groupe de F .

Si H est un sous-groupe de F , alors $f^{-1}(H)$ est un sous-groupe de E .

Deux cas particuliers importants :

- ▷ $\mathrm{Im} f = f(E)$ est un sous-groupe de F .
 - ▷ $\ker f = f^{-1}(\{\varepsilon\}) = \{x \in E \mid f(x) = \varepsilon\}$ est un sous-groupe de E , appelé noyau de f .
-

Proposition 17.15

*Soient $(E, *)$ et $(F, .)$ deux groupes, dont on note e et ε les neutres respectifs et f un morphisme de groupes de $(E, *)$ dans $(F, .)$.*

f est injectif ssi $\ker f = \{e\}$.

17.1.2 Ordre d'un élément dans un groupe

Proposition 17.16

Soient $(G, .)$ un groupe et $x \in G$.

L'application $\varphi_x : \mathbb{Z} \longrightarrow G$ est un morphisme de groupes.
 $k \longmapsto x^k$

Il existe alors un unique entier naturel n tel que $\ker \varphi_x = n\mathbb{Z}$.

Démonstration 17.17

Pour $k, \ell \in \mathbb{Z}$, on a

$$\varphi_x(k + \ell) = x^{k+\ell} = x^k . x^\ell = \varphi_x(k) . \varphi_x(\ell)$$

donc φ_x est un morphisme de groupes.

D'après la Proposition 17.9, comme $\ker \varphi_x$ est un sous-groupe de $(\mathbb{Z}, +)$, il existe $n \in \mathbb{N}$ tel que

$$\ker \varphi_x = n\mathbb{Z} = \{k \in \mathbb{Z} \mid x^k = e\}.$$

■

Si $n > 0$, on dit que x est un élément d'ordre n , sinon on dit que x est d'ordre infini.

Avec les mêmes notations, l'image de φ_x est un sous-groupe de G , noté $\langle x \rangle$.

Proposition 17.18

On suppose que la loi du groupe est notée multiplicativement.

- Si x est d'ordre n , alors n est le plus petit entier strictement positif tel que $x^n = e$ et $\langle x \rangle$ est un sous-groupe fini d'ordre n : $\langle x \rangle = \{e, x, x^2, \dots, x^{n-1}\}$.

Autrement dit, l'équation $x^k = e$ d'inconnue $k \in \mathbb{Z}$ a une infinité de solutions : les multiples de n .

- Si x est d'ordre infini, alors l'équation $x^k = e$ d'inconnue $k \in \mathbb{Z}$ a pour unique solution 0 et le sous-groupe $\langle x \rangle$ est un ensemble infini.

Proposition 17.19

On suppose que la loi du groupe est notée additivement.

- Si x est d'ordre n , alors n est le plus petit entier strictement positif tel que $n.x = 0$ et $\langle x \rangle$ est un sous-groupe fini d'ordre n : $\langle x \rangle = \{0, x, 2.x, \dots, (n-1).x\}$.

Autrement dit, l'équation $k.x = 0$ d'inconnue $k \in \mathbb{Z}$ a une infinité de solutions : les multiples de n .

- Si x est d'ordre infini, alors l'équation $k.x = 0$ d'inconnue $k \in \mathbb{Z}$ a pour unique solution 0 et le sous-groupe $\langle x \rangle$ est un ensemble infini : $\mathbb{Z}.x = \{m.x \mid m \in \mathbb{Z}\}$.

Exercice 17.20

Quel est l'ordre de 3 dans le groupe \mathbb{Z} ?

Correction 17.21

On a

$$3x = 0 \iff x = 0$$

donc l'ordre de 3 est $+\infty$ dans \mathbb{Z} .

Remarque 17.22

Dans un groupe G , le neutre est toujours d'ordre 1 car

$$\forall k \in \mathbb{Z}, e^k = e$$

et c'est le seul élément d'ordre 1 car si $x \in G$ est d'ordre 1 alors

$$x^1 = x = e.$$

Exercice 17.23

Quel est l'ordre d'une symétrie dans le groupe $\text{GL}(E)$?

Correction 17.24

L'ordre d'une symétrie différente de id est 2.

Exercice 17.25

Soient $p, q \in \mathbb{N}^*$. Quel est l'ordre de $\exp\left(\frac{2ip\pi}{q}\right)$ dans le groupe \mathbb{C}^* ?

Correction 17.26
On pose $z = \exp\left(\frac{2ip\pi}{q}\right)$.

On remarque $z^q = 1$ donc l'équation $z^k = 1$ d'inconnue $k \in \mathbb{Z}$ possède une solution non-nulle.

Donc z est d'ordre n fini et $n \mid q$.

On pose $d = p \wedge q$. On veut montrer $n = \frac{q}{d}$.

On a $z = \exp\left(\frac{2i\pi}{q}\right)^p$ donc

$$z^k = 1 \iff \exp\left(\frac{2i\pi}{q}\right)^{pk} = 1$$

$$\iff q \mid pk$$

car q est l'ordre de $\exp\left(\frac{2i\pi}{q}\right)$.

$\left\{\frac{2k\pi}{q} \mid k \in \llbracket 0 ; q-1 \rrbracket\right\}$ est un ensemble à q éléments non-congrus modulo 2π .

On pose $p = dp'$ et $q = dq'$ avec $p' \wedge q' = 1$.

On a alors

$$\begin{aligned} z^k = 1 &\iff q \mid pk \\ &\iff q' \mid p'k \\ &\iff q' \mid k. \end{aligned} \quad \left. \vphantom{\begin{aligned} z^k = 1 \\ &\iff q' \mid p'k \\ &\iff q' \mid k. \end{aligned}} \right\} \text{théorème de Gauss}$$

Donc l'ordre de z est $q' = \frac{q}{d}$.

17.1.3 Sous-groupe engendré par une partie

Définition 17.27

Soient $(G, .)$ un groupe et A une partie de G .

On appelle sous-groupe engendré par A le plus petit sous-groupe de G contenant A , noté en général $\langle A \rangle$: c'est l'intersection de tous les sous-groupes de G qui contiennent A .

Si H est un sous-groupe de G tel que $H = \langle A \rangle$, on dit que A est une partie génératrice de H (ou que A engendre H).

Si $A = \{x_1, \dots, x_n\}$, on note plus simplement $\langle A \rangle = \langle x_1, \dots, x_n \rangle$.

Exercice 17.28

Que valent $\langle \emptyset \rangle$ et $\langle e \rangle$?

Correction 17.29

On a

$$\langle \emptyset \rangle = \langle e \rangle = \{e\}.$$

Exercice 17.30

Dans $(\mathbb{Z}, +)$, a et b étant deux entiers, que valent $\langle a \rangle$ et $\langle a, b \rangle$?

Correction 17.31

On a

$$\langle a \rangle = a\mathbb{Z}$$

et

$$\langle a, b \rangle = a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$$

où $d = a \wedge b$.

\supseteq

D'après le théorème de Bézout, il existe $(u, v) \in \mathbb{Z}^2$ tel que $d = au + bv$ donc $d \in a\mathbb{Z} + b\mathbb{Z}$.

Par stabilité par addition, on a donc $d\mathbb{Z} \subseteq a\mathbb{Z} + b\mathbb{Z}$.

\subseteq

On a $a = da'$ et $b = db'$ avec $a', b' \in \mathbb{Z}$.

Donc $a, b \in d\mathbb{Z}$.

Par stabilité par addition, on a donc $a\mathbb{Z} \subseteq d\mathbb{Z}$ et $b\mathbb{Z} \subseteq d\mathbb{Z}$.

Donc $a\mathbb{Z} + b\mathbb{Z} \subseteq d\mathbb{Z}$.

Finalement, $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$.

Exercice 17.32

Dans \mathfrak{S}_n , quel est le sous-groupe engendré par les transpositions ? par les produits de deux transpositions ?

Correction 17.33

On note \mathcal{T}_n l'ensemble des transpositions de \mathfrak{S}_n et \mathcal{T}_n' l'ensemble des produits de deux transpositions.

On a $\langle \mathcal{T}_n \rangle = \mathfrak{S}_n$ car

$$\forall \sigma \in \mathfrak{S}_n, \exists (\tau_1, \dots, \tau_r) \in \mathcal{T}_n^r, \sigma = \prod_{i=1}^r \tau_i.$$

On a $\langle \mathcal{T}_n' \rangle = \mathcal{A}_n$ car

$$\forall \tau \in \mathcal{T}_n', \varepsilon(\tau) = -1.$$

Un cas particulier important : les sous-groupes engendrés par un élément.

Définition 17.34

Un groupe est dit monogène quand il est engendré par un unique élément.

Un groupe est dit cyclique quand il est monogène et fini.

On peut reprendre l'énoncé de la Proposition 17.18.

Proposition 17.35

Soient $(G, .)$ un groupe et $x \in G$.

Le sous-groupe engendré par x est $\langle x \rangle = \{x^k \mid k \in \mathbb{Z}\}$.

Alors x est d'ordre fini ssi $\langle x \rangle$ est fini.

Dans ce cas, l'ordre n de x est le cardinal de $\langle x \rangle$ et ce sous-groupe est cyclique et égal à $\{e, x, x^2, \dots, x^{n-1}\}$.

Il n'y a que deux types de groupes monogènes.

Proposition 17.36

Si $(G, .)$ est un groupe monogène infini, il est isomorphe à $(\mathbb{Z}, +)$ et il possède exactement deux générateurs, inverses l'un de l'autre.

Si $(G, .)$ est un groupe cyclique d'ordre n , il est engendré par un élément x d'ordre n et les générateurs de G sont les x^p où p est un entier naturel premier avec n : ce sont les éléments d'ordre n dans G .

Démonstration 17.37 (Cas monogène infini)

On suppose que $(G, .)$ est monogène infini.

Il existe $x \in G$ tel que

$$G = \langle x \rangle = \{x^k \mid k \in \mathbb{Z}\}.$$

La fonction $\varphi_x : \begin{array}{ccc} \mathbb{Z} & \longrightarrow & G \\ k & \longmapsto & x^k \end{array}$ est un morphisme de groupes de $(\mathbb{Z}, +)$ dans $(G, .)$.

φ_x est surjectif car x engendre G .

$\ker \varphi_x$ est un sous-groupe de $(\mathbb{Z}, +)$ donc il est de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$.

Si $n > 0$, alors x est d'ordre n et donc $G = \{e, x, \dots, x^{n-1}\}$ est fini : contradiction.

Donc $n = 0$: x est d'ordre infini et φ_x est injectif.

Donc φ_x est un isomorphisme de groupes.

\mathbb{Z} n'a que deux générateurs : 1 et -1 , donc par isomorphisme, G a exactement deux générateurs : x et x^{-1} . ■

Démonstration 17.38 (Cas cyclique)

On suppose que $(G, .)$ est cyclique.

Il existe $x \in G$ tel que

$$G = \langle x \rangle = \{x^k \mid k \in \mathbb{Z}\}.$$

G est fini donc x est d'ordre fini et

$$G = \langle x \rangle = \{e, x, \dots, x^{n-1}\}$$

donc $|G| = n$.

Soit $y \in G$ un générateur de G :

$$\forall z \in G, \exists k \in \mathbb{Z}, z = y^k.$$

En particulier, il existe $q \in \mathbb{Z}$ tel que $x = y^q$.

De même, il existe $p \in \mathbb{Z}$ tel que $y = x^p$.

Donc $x = x^{pq}$ et donc $e = x^{pq-1}$.

Or n est l'ordre de x donc $n \mid pq - 1$ i.e. il existe $r \in \mathbb{Z}$ tel que $pq - 1 = nr$.

Donc $pq - nr = 1$ et d'après le théorème de Bézout, $n \wedge p = 1$.

Réciproquement, si $p \in \mathbb{Z}$ est premier avec n , alors il existe $(u, v) \in \mathbb{Z}^2$ tel que $up + vn = 1$.

Donc $x = x^1 = x^{up+vn} = (x^p)^u \cdot (x^n)^v$.

Or n est l'ordre de x donc $x^n = e$.

Donc $x = (x^p)^u$.

Donc pour tout $z \in G$, sachant qu'il existe $k \in \mathbb{Z}$ tel que $z = x^k$, on a

$$z = (x^p)^{ku}.$$

D'où

$$\forall z \in G, \exists \ell \in \mathbb{Z}, z = (x^p)^\ell.$$

Ceci prouve que x^p est générateur de G .

Conclusion : les générateurs de G sont les éléments x^p où $p \in \mathbb{Z}$ avec $p \wedge n = 1$. ■

Dans le dernier cas, on obtient exactement les générateurs de G en choisissant $p \in \llbracket 1 ; n \rrbracket$ et $p \wedge n = 1$.

Un exemple fondamental du groupe cyclique d'ordre n : l'ensemble des racines n -èmes de l'unité

$$\mathbb{U}_n = \left\{ e^{\frac{2ik\pi}{n}} \mid k \in \llbracket 0 ; n-1 \rrbracket \right\} = \left\{ \left(e^{\frac{2i\pi}{n}} \right)^k \mid k \in \llbracket 0 ; n-1 \rrbracket \right\}.$$

Il est engendré par $\omega = e^{\frac{2i\pi}{n}}$, mais aussi par toute puissance ω^k où k est premier avec n . Les générateurs de \mathbb{U}_n sont appelés les racines primitives n -èmes de l'unité.

Exemple 17.39

- Les générateurs de \mathbb{U}_3 sont j et j^2 .
- Les générateurs de \mathbb{U}_4 sont i et $-i$.
- Les générateurs de \mathbb{U}_6 sont $-j$ et $-j^2$.

- Les générateurs de \mathbb{U}_8 sont les quatre racines quatrièmes de -1 :

$$\exp\left(i\pi\left(\frac{\pi}{4} + \frac{k\pi}{2}\right)\right)$$

pour $k \in \llbracket 0 ; 3 \rrbracket$.

17.1.4 Théorème de Lagrange

Définition 17.40

Le cardinal d'un groupe est appelé l'ordre d'un groupe.

Théorème 17.41

Soit $(G, .)$ un groupe fini.

Pour tout sous-groupe H de G , $|H|$ divise $|G|$.

Démonstration 17.42

Supposons $|G| = n$.

Soit H un sous-groupe de G de cardinal p .

On veut montrer $p \mid n$.

Pour $(x, y) \in G^2$, on pose $x\mathcal{R}y$ quand $y^{-1}x \in H$.

On montre d'abord que \mathcal{R} est une relation d'équivalence sur G .

- \mathcal{R} est réflexive car $\forall x \in G, x^{-1}x = e \in H$ donc $x\mathcal{R}x$.

- \mathcal{R} est symétrique car H est stable par inversion :

$$\begin{aligned} \forall (x, y) \in G^2, x\mathcal{R}y &\iff y^{-1}x \in H \\ &\iff (y^{-1}x)^{-1} \in H \\ &\iff x^{-1}y \in H \\ &\iff y\mathcal{R}x. \end{aligned}$$

- \mathcal{R} est transitive car H est stable par produit :

$$\begin{aligned} \forall (x, y, z) \in G^3, [x\mathcal{R}y \text{ et } y\mathcal{R}z] &\implies [y^{-1}x \in H \text{ et } z^{-1}y \in H] \\ &\implies z^{-1}yy^{-1}x \in H \\ &\implies z^{-1}x \in H \\ &\implies x\mathcal{R}z. \end{aligned}$$

Donc les classes d'équivalence forment une partition de G .

Donc $n = |G|$ est la somme des cardinaux de toutes les classes d'équivalence.

On a

$$\begin{aligned} x \mathcal{R} y &\iff y^{-1}x \in H \\ &\iff x \in yH \end{aligned}$$

donc la classe de y est yH .

Or $\begin{array}{ccc} H & \xrightarrow{\sim} & H \\ h & \mapsto & yh \end{array}$ est une bijection de réciproque $h \mapsto y^{-1}h$.

Donc $|yH| = |H|$: toutes les classes d'équivalence ont pour cardinal p .

Donc en notant N le nombre de classes d'équivalence, on a

$$n = pN$$

i.e. $p \mid n$. ■

Corollaire 17.43

Dans un groupe fini G d'ordre n , tous les éléments sont d'ordres finis et leurs ordres divisent n .

Par conséquent, pour tout $x \in G$, $x^n = e$.

De plus, G est cyclique ssi il existe un élément d'ordre n .

17.2 Anneaux : rappels

17.2.1 Généralités

Définition 17.44

Un anneau est un ensemble A muni de deux lois de composition internes dont l'une est notée $+$ et l'autre est souvent notée \times ou \cdot tel que

- la loi $+$ est une loi de groupe commutatif
- \times est associative
- \times a un élément neutre, noté souvent 1_A ou 1 ou I ou ...
- \times est distributive pour $+$:

$$\forall (x, y, z) \in A^3, \quad x(y + z) = xy + xz \quad \text{et} \quad (x + y)z = xz + yz.$$

Si, de plus, la loi \times est commutative, on dit que A est un anneau commutatif.

Exemple 17.45

- ▷ \mathbb{Z} , \mathbb{Q} , \mathbb{R} et \mathbb{C} sont des anneaux commutatifs pour les lois habituelles.
- ▷ Si X est un ensemble, $\mathcal{F}(X, \mathbb{R})$ est un anneau commutatif.
Plus généralement, si A est un anneau, $\mathcal{F}(X, A)$ est un anneau pour les lois classiques déduites de celles de A .
- ▷ $\mathbb{K}[X]$ est un anneau commutatif pour les lois habituelles.
- ▷ L'ensemble $\mathcal{M}_n(\mathbb{K})$ est un anneau non-commutatif (sauf si $n = 1$).
Si E est un \mathbb{K} -espace vectoriel, l'ensemble $\mathcal{L}(E)$ est un anneau non-commutatif (sauf si $\dim E = 1$) pour les lois $+$ et \circ .

17.2.2 Sous-anneaux

Définition 17.46

Soient $(A, +, \cdot)$ un anneau et B une partie de A .

On dit que B est un sous-anneau de A (sous-entendu pour les lois $+$ et \cdot) quand

- ▷ $1_A \in B$
- ▷ B est stable par $+$: $\forall (x, y) \in B^2, x + y \in B$
- ▷ B est stable par opposition : $\forall x \in B, -x \in B$
- ▷ B est stable par \cdot : $\forall (x, y) \in B^2, x \cdot y \in B$.

Un sous-anneau d'un anneau est donc lui-même un anneau. En général, pour montrer qu'un ensemble muni de deux lois de composition internes est un anneau, on montre que c'est un sous-anneau d'un anneau connu (entre autres, les anneaux cités en exemples précédemment).

Proposition 17.47

Une intersection de sous-anneaux est un sous-anneau.

17.2.3 Pièges calculatoires classiques

Définition 17.48

Soit A un anneau.

Tout couple $(a, b) \in A^2$ tel que $ab = 0$, $a \neq 0$ et $b \neq 0$ est appelé couple de diviseurs de zéro.

L'anneau A est dit intègre quand il est commutatif et ne possède aucun diviseur de zéro.

Exemple 17.49

- \mathbb{Z} et, plus généralement, les anneaux de nombres sont intègres. L'anneau $\mathbb{K}[X]$ est intègre.
- L'anneau $\mathcal{F}(X, \mathbb{R})$ est un anneau commutatif non-intègre.
- Si $n \geq 2$, l'anneau $\mathcal{M}_n(\mathbb{K})$ est un anneau non-commutatif, qui possède des diviseurs de zéro.

Remarque 17.50

En général, dans un anneau quelconque, on ne peut pas simplifier dans un produit : $ab = ac$ n'implique pas en général $b = c$, même si $a \neq 0$.

Si l'anneau est intègre et si $a \neq 0$, alors la simplification est correcte.

Proposition 17.51

Soient A un anneau et $a, b \in A$ tels que $ab = ba$.

On a

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

et

$$\forall n \in \mathbb{N}^*, \quad a^n - b^n = (a - b) (a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + ab^{n-2} + b^{n-1}) = (a - b) \sum_{i=0}^{n-1} a^i b^{n-1-i}.$$

Remarque 17.52

Si les éléments ne commutent pas, c'est faux.

17.3 Anneau $\mathbb{Z}/n\mathbb{Z}$

17.3.1 Définition

Soit $n \in \mathbb{N}^*$. On note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes de congruence modulo n dans \mathbb{Z} (lire « \mathbb{Z} sur $n\mathbb{Z}$ »).

On sait que $\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$.

On munit cet ensemble de deux lois de composition $+$ et \times : pour tout $(\bar{a}, \bar{b}) \in \mathbb{Z}/n\mathbb{Z}$, on pose

$$\bar{a} + \bar{b} = \overline{a + b} \quad \text{et} \quad \bar{a} \times \bar{b} = \overline{ab}.$$

Il est alors facile de vérifier que $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un anneau commutatif dont le neutre pour $+$ est $\bar{0}$ et le neutre pour \times est $\bar{1}$: les propriétés de $+$ et \times dans $\mathbb{Z}/n\mathbb{Z}$ découlent de celles de $+$ et \times dans \mathbb{Z} et de la compatibilité de la congruence modulo n avec ces deux opérations.

Voici par exemple les tables d'addition et de multiplication dans $\mathbb{Z}/8\mathbb{Z}$:

+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{5}$	$\bar{5}$	$\bar{6}$	$\bar{7}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{6}$	$\bar{6}$	$\bar{7}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{7}$	$\bar{7}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$

\times	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{6}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{6}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{6}$	$\bar{1}$	$\bar{4}$	$\bar{7}$	$\bar{2}$	$\bar{5}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{0}$	$\bar{4}$
$\bar{5}$	$\bar{0}$	$\bar{5}$	$\bar{2}$	$\bar{7}$	$\bar{4}$	$\bar{1}$	$\bar{6}$	$\bar{3}$
$\bar{6}$	$\bar{0}$	$\bar{6}$	$\bar{4}$	$\bar{2}$	$\bar{0}$	$\bar{6}$	$\bar{4}$	$\bar{2}$
$\bar{7}$	$\bar{0}$	$\bar{7}$	$\bar{6}$	$\bar{5}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

On voit sur cet exemple quelques-unes des bizarreries qu'on peut trouver dans un anneau :

- des couples de diviseurs de zéro : par exemple, $(2, 4)$
- $\bar{4} + \bar{4} = \bar{0}$: l'opposé d'un élément peut être lui-même
- $\bar{5} \times \bar{5} = \bar{1}$: l'inverse d'un élément autre que $\bar{1}$ et $\overline{-1} = \bar{7}$ peut être égal à lui-même
- $\bar{4}^2 = \bar{0}$: le carré d'un élément non-nul peut être nul.

17.3.2 Propriétés

D'abord le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$.

Proposition 17.53

Le groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ est cyclique, il est engendré par les éléments \bar{k} tels que $1 \leq k \leq n$ et $k \wedge n = 1$.

Tout groupe cyclique d'ordre n est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$.

Puis l'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \times)$.

Proposition 17.54

Les inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$ sont les classes \bar{k} telles que $k \wedge n = 1$. Les autres éléments sont des diviseurs de zéro.

L'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre ssi n est premier.

$\mathbb{Z}/n\mathbb{Z}$ est un corps ssi n est premier.

Démonstration 17.55 (Éléments inversibles)

On a

$$\begin{aligned} \bar{k} \text{ inversible dans } \mathbb{Z}/n\mathbb{Z} &\iff \exists \ell \in \mathbb{Z}, \bar{k} \times \bar{\ell} = \bar{1} \\ &\iff \exists \ell \in \mathbb{Z}, \exists u \in \mathbb{Z}, k\ell = 1 + nu \\ &\iff \exists (\ell, u) \in \mathbb{Z}^2, k\ell - nu = 1 \\ &\iff k \wedge n = 1. \end{aligned} \quad \left. \vphantom{\begin{aligned} \bar{k} \text{ inversible dans } \mathbb{Z}/n\mathbb{Z} &\iff \exists \ell \in \mathbb{Z}, \bar{k} \times \bar{\ell} = \bar{1} \\ &\iff \exists \ell \in \mathbb{Z}, \exists u \in \mathbb{Z}, k\ell = 1 + nu \\ &\iff \exists (\ell, u) \in \mathbb{Z}^2, k\ell - nu = 1 \end{aligned}} \right\} \text{théorème de Bézout} \quad \blacksquare$$

Démonstration 17.56 (Éléments diviseurs de zéro)

Si $\bar{k} \in \mathbb{Z}/n\mathbb{Z}$ tel que $d = k \wedge n > 1$ alors on pose $k = dk'$ et $n = dn'$ où $k' \wedge n' = 1$.

On a donc $n = \frac{k}{k'}n'$ i.e. $nk' = kn'$.

Donc $\bar{0} = \overline{nk'} = \overline{kn'} = \bar{k} \times \bar{n'}$.

Or $\bar{k'} \neq \bar{0}$ et $\bar{n'} \neq \bar{0}$ donc (k', n') diviseurs de zéro. ■

Démonstration 17.57 ($\mathbb{Z}/n\mathbb{Z}$ corps)

Si n est premier, on a

$$\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$$

avec $\bar{1}, \dots, \overline{n-1}$ inversibles donc $\mathbb{Z}/n\mathbb{Z}$ est un corps.

Et réciproquement. ■

Démonstration 17.58 ($\mathbb{Z}/n\mathbb{Z}$ intègre)

Si n n'est pas premier, alors $n = pq$ avec $p, q > 1$ donc

$$\bar{0} = \bar{p} \times \bar{q}$$

donc $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre. ■

Quand p est premier, le corps $\mathbb{Z}/p\mathbb{Z}$ est noté aussi \mathbb{F}_p .

17.3.3 Théorème chinois des restes

Théorème 17.59

Soit $(m, n) \in (\mathbb{N}^*)^2$.

Si $m \wedge n = 1$, alors les anneaux $\mathbb{Z}/mn\mathbb{Z}$ et $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ sont isomorphes.

Démonstration 17.60

Supposons $m \wedge n = 1$.

On pose

$$\begin{aligned} \varphi : \mathbb{Z}/mn\mathbb{Z} &\longrightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ [x]_{mn} &\longmapsto ([x]_m, [x]_n) \end{aligned}$$

où $[x]_k$ désigne la classe de congruence de x modulo k .

On montre d'abord que cette définition n'est pas ambiguë.

Autrement dit, pour $(x, y) \in \mathbb{Z}^2$, si $[x]_{mn} = [y]_{mn}$ alors $[x]_m = [y]_m$ et $[x]_n = [y]_n$.

Autrement dit, le calcul ne dépend que des classes de congruence et pas des représentants desdites classes.

Si $[x]_{mn} = [y]_{mn}$ alors $x \equiv y [mn]$.

Donc $mn \mid x - y$.

Donc $m \mid x - y$ et $n \mid x - y$.

Donc $x \equiv y [m]$ et $x \equiv y [n]$.

Donc $[x]_m = [y]_m$ et $[x]_n = [y]_n$.

Montrons maintenant que φ est un morphisme d'anneaux.

► On a $\varphi([1]_{mn}) = ([1]_m, [1]_n)$.

► Pour $x, y \in \mathbb{Z}$, on a

$$\begin{aligned} \varphi([x]_{mn} + [y]_{mn}) &= \varphi([x + y]_{mn}) \\ &= ([x + y]_m, [x + y]_n) \\ &= ([x]_m + [y]_m, [x]_n + [y]_n) \\ &= ([x]_m, [x]_n) + ([y]_m, [y]_n) \\ &= \varphi([x]_{mn}) + \varphi([y]_{mn}). \end{aligned}$$

► De même avec \times .

Donc φ est un morphisme d'anneaux.

On a

$$\ker \varphi = \{[x]_{mn} \in \mathbb{Z}/mn\mathbb{Z} \mid \varphi([x]_{mn}) = ([0]_m, [0]_n)\}$$

donc

$$\begin{aligned} [x]_{mn} \in \ker \varphi &\iff \begin{cases} x \equiv 0 [m] \\ x \equiv 0 [n] \end{cases} \\ &\iff \begin{cases} m \mid x \\ n \mid x \end{cases} \end{aligned}$$

Or $m \wedge n = 1$ donc

$$\begin{aligned} [x]_{mn} \in \ker \varphi &\iff mn \mid x \\ &\iff x \equiv 0 [mn] \\ &\iff [x]_{mn} = [0]_{mn}. \end{aligned}$$

Donc φ est un morphisme injectif entre deux ensembles de même cardinal donc φ est un isomorphisme d'anneaux.

On peut expliciter la bijection réciproque :

D'après le théorème de Bézout, il existe $(u, v) \in \mathbb{Z}^2$ tel que $mu + nv = 1$.

Soit $([a]_m, [b]_n) \in \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$.

On pose $x = bmu + anv$ pour avoir $\begin{cases} x \equiv a [m] \\ x \equiv b [n] \end{cases}$

En effet, on a

$$\begin{aligned} bmu + anv &\equiv anv [m] \\ &\equiv a - amu [m] \\ &\equiv a [m] \end{aligned}$$

et idem pour b .

Donc $\varphi^{-1}([a]_m, [b]_n) = [bmu + anv]_{mn}$. ■

Autrement dit, dans ce cas, tout système de congruence $\begin{cases} x \equiv a [m] \\ x \equiv b [n] \end{cases}$ d'inconnue x est équivalent à une unique congruence $x \equiv c [mn]$.

Par récurrence, ce résultat se généralise à un nombre quelconque d'anneaux $\mathbb{Z}/a_i\mathbb{Z}$ à condition que a_1, \dots, a_k soient premiers entre eux deux à deux.

Exemple 17.61

Pour $x \in \mathbb{Z}$, $\begin{cases} x \equiv 5 [13] \\ x \equiv 8 [11] \end{cases} \iff x \equiv 96 [143]$.

Démonstration 17.62

On a $11 \wedge 13 = 1$ et la relation de Bézout suivante : $6 \times 11 - 5 \times 13 = 1$.

On a donc

$$\begin{aligned} \begin{cases} x \equiv 8 [11] \\ x \equiv 5 [13] \end{cases} &\iff x \equiv 5 \times 6 \times 11 - 8 \times 5 \times 13 [143] \\ &\iff x \equiv -190 [143] \\ &\iff x \equiv 96 [143]. \end{aligned}$$

Autre méthode :

Si $\begin{cases} x \equiv 8 [11] \\ x \equiv 5 [13] \end{cases}$ alors il existe $k, \ell \in \mathbb{Z}$ tels que

$$\begin{cases} x - 8 = 11k \\ x - 5 = 13\ell \end{cases}$$

donc $-3 = 11k - 13\ell$.

Or $6 \times 11 - 5 \times 13 = 1$ donc $-18 \times 11 + 15 \times 13 = -3$.

Donc $k = -18$ convient et donc $x = -190 \equiv 96 [143]$. ■

Exercice 17.63

Déterminez l'ensemble des entiers $x \in \mathbb{Z}$ tels que $\begin{cases} 3x \equiv 5 [13] \\ 5x \equiv 1 [11] \end{cases}$

Correction 17.64

On a

$$\begin{aligned} 3x \equiv 5 [13] &\iff \bar{3} \times \bar{x} = \bar{5} \text{ dans } \mathbb{Z}/13\mathbb{Z} \\ &\iff \bar{x} = (\bar{3})^{-1} \times \bar{5} \\ &\iff \bar{x} = \overline{45} = \bar{6}. \end{aligned} \quad \left. \vphantom{\begin{aligned} 3x \equiv 5 [13] \\ \bar{x} = (\bar{3})^{-1} \times \bar{5} \end{aligned}} \right) (\bar{3})^{-1} = \bar{9}$$

Idem, on a

$$5x \equiv 1 [11] \iff \bar{x} = \overline{-2} \text{ dans } \mathbb{Z}/11\mathbb{Z}.$$

Donc

$$\begin{cases} 3x \equiv 5 [13] \\ 5x \equiv 1 [11] \end{cases} \iff \begin{cases} x \equiv 6 [13] \\ x \equiv -2 [11] \end{cases}$$

Or $6 \times 11 - 5 \times 13 = 1$ donc

$$\begin{aligned} \begin{cases} x \equiv 6 [13] \\ x \equiv -2 [11] \end{cases} &\iff x \equiv -2 \times 13 \times (-5) + 6 \times 6 \times 11 [143] \\ &\iff x \equiv 526 [143] \\ &\iff x \equiv 97 [143]. \end{aligned}$$

17.3.4 Indicatrice d'Euler

Définition 17.65

Soit $n \in \mathbb{N}^*$.

On pose $\phi(n) = \text{Card} \{x \in \llbracket 1 ; n \rrbracket \mid x \wedge n = 1\}$.

L'application ϕ est appelée indicatrice d'Euler.

D'après la Proposition 17.35, $\phi(n)$ est le nombre de générateurs d'un groupe cyclique d'ordre n et le nombre de racines primitives n -èmes de l'unité.

Proposition 17.66

Soit $(m, n) \in (\mathbb{N}^*)^2$.

Si $m \wedge n = 1$, alors $\phi(mn) = \phi(m) \phi(n)$.

Démonstration 17.67

On reprend l'isomorphisme d'anneaux φ de la Démonstration 17.60.

On a

$$\begin{aligned} [x]_n \text{ inversible dans } \mathbb{Z}/mn\mathbb{Z} &\iff \exists [y]_{mn} \in \mathbb{Z}/mn\mathbb{Z}, [x]_{mn} \times [y]_{mn} = [1]_{mn} \\ &\iff \exists [y]_{mn} \in \mathbb{Z}/mn\mathbb{Z}, \varphi([x]_{mn}) \times \varphi([y]_{mn}) = \varphi([1]_{mn}) = ([1]_m, [1]_n) \\ &\iff \exists ([y]_m, [y]_n) \in \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}, ([x]_m, [x]_n) \times ([y]_m, [y]_n) = ([1]_m, [1]_n) \\ &\iff [x]_m \text{ inversible dans } \mathbb{Z}/m\mathbb{Z} \text{ et } [x]_n \text{ inversible dans } \mathbb{Z}/n\mathbb{Z}. \end{aligned}$$

Il y a donc autant d'inversibles des deux côtés :

$$\phi(mn) = \phi(m) \phi(n).$$

■

À titre culturel, on dit que ϕ est une fonction arithmétique.

Comme il est facile de calculer $\phi(p^\alpha)$ quand p est premier, il est alors possible de calculer $\phi(n)$ pour tout entier n à condition de savoir le décomposer en facteurs premiers.

Démonstration 17.68

Soient $p \in \mathbb{P}$, $\alpha \in \mathbb{N}^*$ et $x \in \llbracket 1 ; p^\alpha \rrbracket$.

On a

$$\begin{aligned} [x]_{p^\alpha} \text{ inversible dans } \mathbb{Z}/p^\alpha\mathbb{Z} &\iff x \wedge p^\alpha = 1 \\ &\iff x \wedge p = 1. \end{aligned}$$

Le nombre d'entiers $x \in \llbracket 1 ; p^\alpha - 1 \rrbracket$ divisibles par p est $p^{\alpha-1} - 1$.

En effet, en posant $x = kp$, on a

$$\begin{aligned} 1 \leq kp \leq p^\alpha - 1 &\iff 1 \leq k \leq \frac{p^\alpha - 1}{p} = p^{\alpha-1} - \frac{1}{p} \\ &\iff 1 \leq k \leq p^{\alpha-1} - 1. \end{aligned}$$

Donc

$$\begin{aligned} \phi(p^\alpha) &= (p^\alpha - 1) - (p^{\alpha-1} - 1) \\ &= p^\alpha - p^{\alpha-1} \\ &= p^\alpha \left(1 - \frac{1}{p}\right). \end{aligned}$$

■

On finit par le théorème d'Euler, qui étend le petit théorème de Fermat.

Proposition 17.69

Soient $x \in \mathbb{Z}$ et $n \in \mathbb{N}^*$.

Si $x \wedge n = 1$, alors $x^{\phi(n)} \equiv 1 [n]$.

Démonstration 17.70

Dans $\mathbb{Z}/n\mathbb{Z}$, on a

$$\bar{x} \text{ inversible} \iff x \wedge n = 1.$$

L'ensemble des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$ est un groupe pour la multiplication, qui a pour cardinal $\phi(n)$.

Donc $\bar{x}^{\phi(n)} = \bar{1}$ i.e.

$$x^{\phi(n)} \equiv 1 [n].$$

■

17.4 Idéaux d'un anneau commutatif

17.4.1 Généralités

Définition 17.71

Soient A un anneau commutatif et I une partie de A .

On dit que I est un idéal de A quand

- I est un sous-groupe de $(A, +)$
- I est stable par multiplication externe : $\forall (x, y) \in A \times I, x \cdot y \in I$.

Remarque 17.72

Bien que les définitions soient ressemblantes, les notions d'idéal et de sous-anneau sont distinctes : la seule partie de A qui est à la fois un idéal et un sous-anneau de A est A lui-même !

Proposition 17.73

Soient A, B deux anneaux commutatifs et $f : A \longrightarrow B$ un morphisme d'anneaux.

$\ker f = \{a \in A \mid f(a) = 0\}$ est un idéal de A (et donc n'est pas un sous-anneau en général).

Proposition 17.74

Soient A un anneau commutatif et $x \in A$.

L'ensemble des multiples de x , noté xA , est un idéal de A , dit idéal principal.

C'est le plus petit idéal contenant x , c'est-à-dire l'intersection des idéaux de A qui contiennent x : on l'appelle idéal engendré par x .

Définition 17.75

Soient A un anneau commutatif et $(x, y) \in A^2$.

On dit que x divise y et on note $x \mid y$ ssi il existe $z \in A$ tel que $y = xz$.

Proposition 17.76

Avec les mêmes notations, on a l'équivalence

$$x \mid y \iff yA \subseteq xA.$$

Un idéal engendré par un seul élément est appelé un idéal principal. Un anneau est dit principal quand il est intègre et que tous ses idéaux sont principaux.

Proposition 17.77

Les anneaux \mathbb{Z} et $\mathbb{K}[X]$ sont principaux.

Démonstration 17.78 (\mathbb{Z} anneau principal)

Si I est un idéal de \mathbb{Z} , alors, en particulier, I est un sous-groupe de $(\mathbb{Z}, +)$ donc il existe $n \in \mathbb{N}$ tel que $I = n\mathbb{Z}$.

Et réciproquement, tout ensemble $n\mathbb{Z}$ est un idéal principal de \mathbb{Z} . ■

Démonstration 17.79 ($\mathbb{K}[X]$ anneau principal)

Soit I un idéal de $\mathbb{K}[X]$.

Si $I = \{0\}$, alors $I = 0\mathbb{K}[X]$.

Si $I \neq \{0\}$, il existe $P \in I \setminus \{0\}$ de degré minimal.

Pour $Q \in I$, on effectue la division euclidienne de Q par P :

$$Q = PU + V$$

où $U, V \in \mathbb{K}[X]$ et $\deg V < \deg P$.

Comme $P \in I$ et I stable par multiplication extérieure, on a $PU \in I$ (*).

Comme $Q \in I$ et I sous-groupe de $(\mathbb{K}[X], +)$, on a $Q - PU \in I$.

Donc $V \in I$.

Or $\deg V < \deg P$ donc $V = 0$.

Donc $Q = PU \in P\mathbb{K}[X]$.

Ceci prouve $I \subseteq P\mathbb{K}[X]$.

On a montré l'inclusion réciproque au passage : (*).

Donc $I = P\mathbb{K}[X]$ i.e. I est principal. ■

Remarque 17.80

L'ensemble des entiers de Gauss

$$\mathbb{Z}[i] = \{a + ib \mid (a, b) \in \mathbb{Z}^2\}$$

est aussi un anneau principal.

Exercice 17.81

Soit A un anneau commutatif.

Montrez que A est un corps ssi A possède exactement deux idéaux (lesquels?).

Correction 17.82

Supposons que A est un corps. Soit I un idéal de A .

Soit $I = \{0\}$, soit $I \neq \{0\}$.

Dans ce dernier cas, on peut choisir $a \in I \setminus \{0\}$.

A est un corps donc a possède un inverse $a^{-1} \in A$.

Donc $1 = a^{-1}a \in I$.

Donc $\forall x \in A, 1x = x \in I$.

Donc $I = 1A = A$.

Réciproquement, soit A un anneau commutatif qui n'a pour idéaux que $\{0\}$ et A .

Soit $x \in A \setminus \{0\}$.

xA est alors un idéal de A .

Or $x \in xA$ et $x \neq 0$ donc $xA \neq \{0\}$.

Donc $xA = A$.

Or $1 \in A$ donc $1 \in xA$.

Donc il existe $y \in A$ tel que $xy = 1$ i.e. x est inversible dans A .

Donc A est un corps.

Exercice 17.83

Dans l'anneau $A = \mathcal{F}(\mathbb{R}, \mathbb{C})$, on pose $I = \{f \in A \mid f(0) = 0\}$.

Montrez que I est un idéal principal de A .

Correction 17.84

On cherche $\varphi \in A$ telle que

$$\forall f \in A, f \in I \iff f = \varphi g$$

où $g \in A$.

On pose $\varphi = 1_{\mathbb{R} \setminus \{0\}} : x \mapsto \begin{cases} 1 & \text{si } x \neq 0 \\ 0 & \text{sinon} \end{cases}$

On a bien $\varphi \in I$ et $\forall f \in I, f = \varphi f$ car

$$\forall x \in \mathbb{R}, f(x) = \begin{cases} f(x) \times 1 = f(x) \varphi(x) & \text{si } x \neq 0 \\ 0 = f(x) \times 0 = f(x) \varphi(x) & \text{sinon} \end{cases}$$

Donc $I = \varphi A$ i.e. I est principal.

Exercice 17.85

On reprend l'exercice précédent en posant cette fois-ci $A = \mathcal{C}^0(\mathbb{R}, \mathbb{C})$.

Montrez que I n'est pas principal.

Correction 17.86

I est clairement un idéal de A .

Si I est principal, il existe $\varphi \in I$ telle que

$$\forall f \in I, \exists g \in A, f = \varphi g$$

i.e. $\frac{f}{\varphi}$ est une fonction prolongeable par continuité en 0.

Or $\text{id} \in I$ donc il existe $g \in A$ telle que $\text{id} = \varphi g$ *i.e.*

$$\forall x \in \mathbb{R}, x = \varphi(x) g(x).$$

En particulier, pour $x \in \mathbb{R}^*$, $\varphi(x) g(x) \neq 0$ donc $\varphi(x) \neq 0$.

Donc φ ne s'annule qu'en 0.

De plus, $\sqrt{|\varphi|} \in I$ donc $\frac{\sqrt{|\varphi|}}{\varphi}$ a une limite finie en 0.

Or

$$\frac{\sqrt{|\varphi|}}{\varphi} = \frac{\sqrt{|\varphi|}}{|\varphi| \times \text{sg } \varphi} = \frac{1}{\sqrt{|\varphi|} \times \text{sg } \varphi} \xrightarrow{0} +\infty.$$

On a donc une contradiction.

Donc I n'est pas principal.

Exercice 17.87

On travaille dans $\mathbb{Z}[X]$, anneau des polynômes à coefficients dans \mathbb{Z} .

On pose $I = \{P \in \mathbb{Z}[X] \mid P(0) \text{ est pair}\}$.

Montrez que I est un idéal, que $I = 2\mathbb{Z}[X] + X\mathbb{Z}[X]$ (*i.e.* l'idéal engendré par 2 et X), et que I n'est pas principal.

Correction 17.88

Soit $(P, Q) \in I^2$.

$(P+Q)(0) = P(0) + Q(0)$ est la somme de deux nombres pairs donc est un nombre pair. Donc $P+Q \in I$.

Soient $P \in I$ et $Q \in \mathbb{Z}[X]$.

$(PQ)(0) = P(0)Q(0)$ est pair car $P(0)$ est pair et $Q(0) \in \mathbb{Z}$. Donc $PQ \in I$.

Donc I est un idéal de $\mathbb{Z}[X]$.

Soit $P \in 2\mathbb{Z}[X] + X\mathbb{Z}[X]$.

Il existe $(Q, R) \in \mathbb{Z}[X]^2$ tel que $P = 2Q + XR$.

Donc $P(0) = 2Q(0)$ est un entier pair.

Donc $P \in I$ i.e. $2\mathbb{Z}[X] + X\mathbb{Z}[X] \subseteq I$.

Réciproquement, soit $P \in I$.

On a $P = P(0) + a_1X + \dots + a_nX^n = 2k + a_1X + \dots + a_nX^n$ où $k \in \mathbb{Z}$.

On pose $Q = k$ et $R = a_1 + a_2X + \dots + a_nX^{n-1}$ de sorte que $P = 2Q + XR$.

Donc $I \subseteq 2\mathbb{Z}[X] + X\mathbb{Z}[X]$.

Donc $I = 2\mathbb{Z}[X] + X\mathbb{Z}[X]$.

Si I est un idéal principal, il existe $A \in \mathbb{Z}[X]$ tel que $I = A\mathbb{Z}[X]$.

Donc $A\mathbb{Z}[X] = 2\mathbb{Z}[X] + X\mathbb{Z}[X]$.

Or $2 = 2 \times 1 + 0 \times X$ donc $2 \in I$.

Donc il existe $B \in \mathbb{Z}[X]$ tel que $2 = AB$.

Donc $\deg 2 = 0 = \deg A + \deg B$.

Donc $\deg A = \deg B : A = \lambda \in \mathbb{Z}^*$ est constant non-nul.

De plus, $X = 2 \times 0 + 1 \times X$ donc il existe $C \in \mathbb{Z}[X]$ tel que $X = AC$.

Donc $\deg X = \deg A + \deg C$ i.e. $1 = \lambda \deg C$.

Donc $\lambda = \pm 1$.

Or $A \in I$ donc $\lambda = A(0)$ est pair : contradiction.

Donc I n'est pas principal.

17.4.2 Arithmétique dans un anneau principal

17.4.2.1 Pgcd et ppcm

Si I et J sont deux idéaux d'un anneau commutatif, on note $I + J = \{x + y \mid (x, y) \in I \times J\}$.

Il est facile de vérifier que $I + J$ est alors aussi un idéal, de même que $I \cap J$.

On généralise sans difficulté à une somme ou une intersection d'idéaux I_1, \dots, I_n .

Proposition 17.89

Soient A un anneau principal et $x, y \in A$.

Il existe un élément $z \in A$ tel que $zA = xA + yA$: on vérifie alors que les diviseurs communs à x et y sont exactement les diviseurs de z . On dit alors que z est un pgcd de x et y .

Il existe un élément $w \in A$ tel que $wA = xA \cap yA$: les multiples communs à x et y sont exactement les multiples de w . On dit alors que w est un ppcm de x et y .

Il n'y en général pas d'unicité de z et w dans les définitions précédentes : pour tout élément inversible u , $(uz)A = zA$.

Cependant, dans certains cas particuliers on peut ajouter une unicité en exploitant des propriétés spécifiques de l'anneau :

- ▷ si $A = \mathbb{Z}$, en choisissant $z > 0$
- ▷ si $A = \mathbb{K}[X]$, en choisissant z unitaire

dans les cas où $z \neq 0$.

Autrement dit :

- ▷ si $a_1, \dots, a_n \in \mathbb{Z}$, leur pgcd est l'unique entier positif d tel que $d\mathbb{Z} = a_1\mathbb{Z} + \dots + a_n\mathbb{Z}$
- ▷ si $P_1, \dots, P_n \in \mathbb{K}[X] \setminus \{0\}$, leur pgcd est l'unique polynôme unitaire D tel que $D\mathbb{K}[X] = P_1\mathbb{K}[X] + \dots + P_n\mathbb{K}[X]$.

La proposition précédente est donc une preuve non-effective de l'égalité de Bézout.

Définition 17.90

Deux éléments d'un anneau principal sont dits premiers entre eux quand un de leurs pgcd vaut 1.

On retrouve le théorème de Bézout.

Théorème 17.91

Soient A un anneau principal et $(a, b) \in A^2$.

a et b sont premiers entre eux ssi il existe $(u, v) \in A^2$ tel que $1 = au + bv$.

Dans un anneau principal, le théorème de Gauss est encore valable, de même que tous les résultats faisant intervenir la notion d'éléments premiers entre eux (voir le chapitre d'arithmétique de première année).

Théorème 17.92

Soient A un anneau principal et $(a, b, c) \in A^3$.

Si a divise bc et a et b sont premiers entre eux, alors a divise c .

17.4.2.2 Décomposition en facteurs irréductibles

Définition 17.93

Soit A un anneau.

On dit qu'un élément $a \in A$ est irréductible quand il n'existe pas $(b, c) \in A^2$ tel que $a = bc$ et b, c ne sont pas inversibles.

Autrement dit, a est irréductible quand toute écriture $a = bc$ implique que b ou c est un inversible, ou, ce qui revient au même, les seuls diviseurs de a sont les inversibles et ses associés ua , où u est inversible.

Exemple 17.94

- Les nombres premiers et leurs opposés sont les éléments irréductibles de l'anneau \mathbb{Z} .
- Dans $\mathbb{C}[X]$, les éléments irréductibles sont les polynômes de degré 1 (conséquence du théorème de D'Alembert-Gauss). Dans $\mathbb{R}[X]$, les éléments irréductibles sont les polynômes de degré 1 ou 2 sans racine réelle. Dans $\mathbb{Q}[X]$, il existe des polynômes irréductibles de tout degré.

Dans certains anneaux, l'irréductibilité est une notion piègeuse. Heureusement, dans les anneaux principaux, tout va bien : l'arithmétique est alors similaire à celle de \mathbb{Z} . On retrouve les résultats classiques, grâce surtout au théorème de Gauss.

Le théorème qui suit est souvent appelé théorème fondamental de l'arithmétique dans un anneau principal.

Théorème 17.95

Soit A un anneau principal.

Tout élément de A non-nul et non-inversible s'écrit comme un produit d'irréductibles. Cette écriture est unique à l'ordre près des irréductibles et à association près des irréductibles.

Autrement dit, pour tout $a \in A$, si $a \neq 0$ et a non-inversible, il existe un unique $k \in \mathbb{N}^$, un k -uplet $(p_1, \dots, p_k) \in A^k$ et $(\alpha_1, \dots, \alpha_k) \in (\mathbb{N}^*)^k$ tels que p_1, p_2, \dots, p_k sont irréductibles distincts et $a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$.*

Cette écriture s'appelle la décomposition en facteurs irréductibles de a .

De « l'unicité » de la décomposition en facteurs irréductibles, on tire le résultat suivant concernant les diviseurs d'un élément.

Proposition 17.96

Soit A un anneau principal.

Soit $a \in A$ de décomposition irréductible $a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ (les p_i sont irréductibles distincts).

Alors si b est un diviseur de a , alors il existe $(\beta_1, \dots, \beta_k) \in \mathbb{N}^k$ et u inversible tel que $b = up_1^{\beta_1} p_2^{\beta_2} \dots p_k^{\beta_k}$ et pour tout $i \in \llbracket 1 ; k \rrbracket$, $0 \leq \beta_i \leq \alpha_i$.

Réciproquement, tout élément de cette forme est un diviseur de a .

Remarque 17.97

Dans des cas spécifiques d'anneaux principaux, on impose en général des conditions supplémentaires sur les irréductibles :

- dans le cas de \mathbb{Z} , les irréductibles (*i.e.* les nombres premiers) sont positifs et l'écriture générale est alors $a = \varepsilon p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ où ε est un signe
- dans le cas de $\mathbb{K}[X]$, les irréductibles sont des polynômes unitaires et l'écriture générale est alors $a = \lambda p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ où $\lambda \in \mathbb{K}^*$ est le coefficient dominant de a .

17.5 \mathbb{K} -algèbre

17.5.1 Généralités

Définition 17.98

Un ensemble A est appelé \mathbb{K} -algèbre quand A est à la fois un anneau et un \mathbb{K} -espace vectoriel, dont les multiplications sont compatibles.

Il y a donc trois lois dans une \mathbb{K} -algèbre :

- une addition classique $+$
- une multiplication externe \mathbb{K} .
- une multiplication interne, compatible avec la précédente :

$$\forall (\lambda, a, b) \in \mathbb{K} \times A^2, \quad \lambda.(ab) = (\lambda.a)b = a(\lambda.b).$$

On qualifie les \mathbb{K} -algèbres par du vocabulaire des anneaux (algèbres intègres, algèbres principales, etc) ou des espaces vectoriels (algèbres de dimension finie, etc).

Exemple 17.99

- \mathbb{K} lui-même est une \mathbb{K} -algèbre, où les deux multiplications sont confondues ; \mathbb{C} est aussi une \mathbb{R} -algèbre de dimension 2.
- $\mathbb{K}[X]$ est une \mathbb{K} -algèbre intègre, commutative et de dimension finie.
- Si I un intervalle, $\mathcal{F}(I, \mathbb{K})$ est une \mathbb{K} -algèbre commutative non-intègre et de dimension infinie.
- Si $n \in \mathbb{N}^*$, $\mathcal{M}_n(\mathbb{K})$ est une \mathbb{K} -algèbre de dimension n^2 qui n'est ni intègre, ni commutative.
- Si E est un \mathbb{K} -espace vectoriel, alors $\mathcal{L}(E)$ est une \mathbb{K} -algèbre qui n'est ni intègre, ni commutative, de dimension finie ssi E l'est aussi.

17.5.2 Sous-algèbres

Définition 17.100

Une sous-algèbre d'une \mathbb{K} -algèbre est à la fois un sous-anneau et un sous-espace vectoriel.

Comme d'habitude, pour montrer qu'un ensemble est une \mathbb{K} -algèbre, on montre que c'est une sous-algèbre d'une \mathbb{K} -algèbre connue.

17.5.3 Morphismes d'algèbres

Définition 17.101

Soient A et B deux \mathbb{K} -algèbres.

Un morphisme d'algèbres de A dans B est une application $f : A \longrightarrow B$ qui est à la fois un morphisme d'anneaux et une application linéaire :

- ▷ $\forall (a, b) \in A^2, f(a + b) = f(a) + f(b)$ et $f(ab) = f(a)f(b)$
- ▷ $f(1_A) = 1_B$
- ▷ $\forall (\lambda, a) \in \mathbb{K} \times A, f(\lambda a) = \lambda f(a)$.

Proposition 17.102

Le noyau d'un morphisme d'algèbres est un idéal et un sous-espace vectoriel de l'algèbre de départ.

17.5.4 Polynôme minimal d'un élément

Proposition 17.103

Soient A une \mathbb{K} -algèbre et $a \in A$.

Pour $P = \sum_{i=0}^n c_i X^i \in \mathbb{K}[X]$, on pose $P(a) = \sum_{i=0}^n c_i a^i$.

L'application $\begin{array}{ccc} \mathbb{K}[X] & \longrightarrow & A \\ P & \longmapsto & P(a) \end{array}$ est alors un morphisme d'algèbres.

Puisque \mathbb{K} est un anneau principal, le noyau d'un tel morphisme, i.e. $\{P \in \mathbb{K}[X] \mid P(a) = 0\}$ est un idéal principal.

Dans le cas où il est non-nul, il est engendré par un unique polynôme unitaire M , appelé polynôme minimal de a . On a donc

$$\{P \in \mathbb{K}[X] \mid P(a) = 0\} = M\mathbb{K}[X]$$

c'est-à-dire

$$\forall P \in \mathbb{K}[X], P(a) = 0 \iff M \mid P.$$

Un exemple fondamental : celui des matrices carrées.

Théorème 17.104

Soit $n \in \mathbb{N}^*$.

Toute matrice carrée de $\mathcal{M}_n(\mathbb{K})$ possède un polynôme minimal.

Tout endomorphisme d'un \mathbb{K} -espace vectoriel de dimension n possède un polynôme minimal.

A priori ce polynôme minimal est de degré au plus n^2 mais on a montré qu'il est en fait de degré au plus n (cf. Corollaire 5.129).

Exemple 17.105

- Le polynôme minimal d'une matrice scalaire λI_n est $X - \lambda$.
 - Le polynôme minimal d'un projecteur non-trivial est $X^2 - X$, celui d'une symétrie non-triviale est $X^2 - 1$.
-

Exercice 17.106

Montrez que si $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ n'est pas une matrice scalaire, alors son polynôme minimal est $X^2 - (a + d)X + (ad - bc)$.

Deuxième partie

Exercices

Chapitre 1

Espaces vectoriels normés

★★ À venir ★★

Chapitre 2

Séries numériques et vectorielles : révisions et compléments

★★ À venir ★★

Chapitre 3

Familles sommables

★★ À venir ★★

Chapitre 4

Rappels et compléments d'algèbre linéaire

★★ À venir ★★

Chapitre 5

Réduction des endomorphismes

★★ À venir ★★

Chapitre 6

Intégrales généralisées

★★ À venir ★★

Chapitre 7

Intégrales à paramètre

★★ À venir ★★

Chapitre 8

Espaces préhilbertiens réels

★★ À venir ★★

Chapitre 9

Endomorphismes dans un espace euclidien

★★ À venir ★★

Chapitre 10

Fonctions vectorielles

★★ À venir ★★

Chapitre 11

Suites et séries de fonctions

★★ À venir ★★

Chapitre 12

Séries entières

★★ À venir ★★

Chapitre 13

Probabilités

★★ À venir ★★

Chapitre 14

Variables aléatoires discrètes

★★ À venir ★★

Chapitre 15

Équations différentielles linéaires

★★ À venir ★★

Chapitre 16

Calcul différentiel

★★ À venir ★★

Chapitre 17

Structures algébriques

★★ À venir ★★

