

Pierre COLMEZ

ÉLÉMENTS D'ANALYSE ET
D'ALGÈBRE

Pierre COLMEZ

C.M.L.S., École Polytechnique, 91128 Palaiseau Cedex, France.

ÉLÉMENTS D'ANALYSE ET D'ALGÈBRE

Pierre COLMEZ

TABLE DES MATIÈRES

Introduction	1
Notations standard	3
Bibliographie sommaire	3
Vocabulaire Mathématique	5
1. Grammaire élémentaire	6
1.1. L'anneau \mathbf{Z} des entiers relatifs	7
1.2. Parallélisme entre logique élémentaire et langage ensembliste	9
1.3. Ensembles dénombrables	9
2. Produits, sommes et quotients	11
2.1. Produits et sommes	11
2.1.1. Produits et sommes directes de groupes commutatifs	11
2.1.2. Le cas des espaces vectoriels	12
2.1.3. Produit et somme dans une catégorie	12
2.2. Relations d'équivalence	13
2.2.1. Relations d'équivalence et partitions	13
2.2.2. Passage au quotient par une relation d'équivalence	14
2.3. L'anneau $\mathbf{Z}/D\mathbf{Z}$ des entiers relatifs modulo D	15
2.4. Quotients d'espaces vectoriels	18
2.5. Anneaux quotients	18
2.6. Groupes quotients	20
2.6.1. Groupe opérant sur un ensemble	20
2.6.2. Classes de conjugaison	21
2.6.3. Quotients de groupes	22
3. Construction de nombres	23
3.1. Entiers naturels	23

3.2. Entiers relatifs, nombres rationnels	24
3.3. Nombres réels, nombres complexes	25
3.4. Nombres p -adiques	26
3.4.1. Le corps \mathbf{Q}_p	26
3.4.2. L'anneau \mathbf{Z}_p des entiers p -adiques	27
3.4.3. L'anneau des nombres complexes p -adiques	29
4. Groupes finis	29
4.1. Généralités sur les groupes	29
4.2. Groupes cycliques	30
4.2.1. Structure des groupes cycliques, ordre d'un élément	30
4.2.2. Sous-groupes des groupes cycliques	31
4.3. Groupes abéliens finis	31
4.4. Le théorème de Lagrange et ses variantes	32
4.5. Le groupe symétrique S_n	33
4.5.1. Permutations	33
4.5.2. Signature d'une permutation	35
4.5.3. Groupe alterné	36
4.6. Les théorèmes de Sylow	37
5. Algèbre linéaire	38
5.1. Généralités	38
5.1.1. Endomorphismes	38
5.1.2. Le théorème de Cayley-Hamilton	38
5.1.3. Automorphismes	39
5.1.4. Matrices	39
5.1.5. Espaces propres, espaces caractéristiques	39
5.1.6. Mise sous forme de Jordan	40
5.2. Modules de torsion sur $K[T]$ et réduction des endomorphismes	40
5.2.1. Anneaux et modules	40
5.2.2. Structure des modules de torsion sur $K[T]$	41
5.2.3. Exemples	42
5.2.4. Application à la réduction des endomorphismes	43
5.3. Modules de torsion sur les anneaux principaux	44
5.3.1. Généralités sur les idéaux	44
5.3.2. Anneaux principaux	44
5.3.3. Structure des modules de torsion sur un anneau principal	46
6. Topologie	48

6.1. Espaces topologiques	48
6.1.1. Ouverts, fermés, voisinages	48
6.1.2. Exemples	49
6.1.3. Comparaison de topologies	49
6.2. Espaces métriques	50
6.3. Continuité	51
6.4. Sous-espaces, produits, quotients	53
6.4.1. Topologie induite	53
6.4.2. Topologie produit	53
6.4.3. Topologie quotient	53
6.5. Espaces séparés	54
6.6. Intérieur, adhérence, densité	55
6.7. Suites dans un espace topologique	56
6.7.1. Suites, suites extraites	56
6.7.2. Suites et continuité	57
7. Compacité	57
7.1. Espaces compacts	57
7.2. Compacité et suites	58
7.3. Propriétés de base des compacts	59
7.3.1. Compacts d'un espace topologique	59
7.3.2. Compacts d'un espace métrique	61
7.4. La droite réelle achevée	63
7.4.1. Les espaces topologiques ordonnés $\overline{\mathbf{R}}$ et $\overline{\mathbf{R}}_+$	63
7.4.2. Limite supérieure, limite inférieure	63
7.5. L'espace topologique \mathbf{R}/\mathbf{Z}	64
8. Connexité	65
8.1. Ensembles connexes	65
8.2. Connexité par arcs	67
9. Complétude	68
9.1. Suites de Cauchy	68
9.2. Principales propriétés des espaces complets	70
9.3. Complétion d'un espace métrique	71
10. Convergence de fonctions	73
10.1. Convergence simple	73
10.2. Convergence uniforme	73
11. Espaces vectoriels normés	75

11.1. Normes et applications linéaires continues	75
11.2. La norme d'un opérateur	75
11.3. Normes équivalentes	76
11.4. La boule unité d'un espace vectoriel normé	77
11.5. Applications bilinéaires continues	77
12. Tératologie	78
12.1. Fonctions continues dérivables nulle part	78
12.2. L'escalier du diable	79
12.3. L'ensemble triadique de Cantor	81
12.4. La courbe de Peano	82
12.5. Ensembles connexes non connexes par arcs	83
12.5.1. Le graphe de $\sin \frac{1}{x}$	83
12.5.2. Le tipi de Cantor	84
13. Corrigé des exercices	86
Index du chapitre	97
I. Représentations des groupes finis	99
I.1. Représentations et caractères	101
1. Représentations de groupes, exemples	101
2. Caractère d'une représentation, exemples	103
2.1. Caractères linéaires	103
2.2. Sommes directes	104
2.3. Représentations de permutation, représentation régulière	105
3. Morphismes de représentations	105
3.1. La représentation $\text{Hom}(V_1, V_2)$	105
3.2. Opérateurs d'entrelacement, représentations isomorphes	106
I.2. Décomposition des représentations	107
1. Décomposition en somme directe de représentations irréductibles	107
2. Le lemme de Schur et ses conséquences immédiates	109
3. Orthogonalité des caractères	111
4. Applications du théorème principal	112
4.1. Nombre des représentations irréductibles	112
4.2. La décomposition canonique d'une représentation	112
4.3. Un critère d'irréductibilité	114
4.4. La décomposition de la représentation régulière	114
5. Le cas des groupes commutatifs	114
5.1. La transformée de Fourier	114

5.2. Le groupe dual	116
5.3. Le théorème de structure des groupes finis commutatifs	117
6. Table des caractères d'un groupe fini	117
I.3. Construction de représentations	121
1. Constructions tensorielles de représentations	121
1.1. Produit tensoriel d'espaces vectoriels de dimension finie	121
1.2. Produit tensoriel de représentations	123
1.3. Carré symétrique et carré extérieur d'une représentation	123
2. Représentations induites	124
2.1. Caractère d'une représentation induite	124
2.2. La formule de réciprocity de Frobenius	126
2.3. Transitivité des inductions	127
2.4. Les théorèmes d'Artin et de Brauer	127
3. Exercices	128
II. Espaces de Banach	131
II.1. Espaces de Banach	131
1. Convergence normale, séries sommables	131
2. Espaces de suites	133
3. Espaces de fonctions continues	134
4. Complétion d'espaces vectoriels normés	136
5. Applications linéaires continues entre espaces de Banach	137
6. Le dual d'un espace de Banach	139
II.2. Espaces de Hilbert	140
1. Espaces préhilbertiens	140
2. Espaces de Hilbert	141
2.1. Bases hilbertiennes	142
2.2. Projection orthogonale sur un sous-espace fermé	143
3. Le dual d'un espace de Hilbert	145
4. Le théorème de projection sur un convexe	146
II.3. Exercices	147
1. Espaces de Banach	147
2. Espaces de Hilbert	148
3. Séries de Fourier	149
III. Intégration	151
III.1. Intégrale de Lebesgue	151

1. Dallages et fonctions en escalier	151
2. Ensembles de mesure nulle	153
3. Fonctions mesurables, ensembles mesurables	155
4. Définition de l'intégrale de Lebesgue	156
5. Les théorèmes de convergence monotone et de convergence dominée	159
6. Premières applications	160
III.2. Quelques espaces fonctionnels	161
1. L'espace $L^1(X)$	161
2. L'espace $L^2(X)$	163
3. Convergence dans L^1 et L^2	164
4. Espaces L^p	166
III.3. Intégrales multiples	167
1. Le théorème de Fubini	167
2. La formule du changement de variable	169
3. L'intégrale de la gaussienne	172
4. Exercices	172
III.4. Construction de l'intégrale de Lebesgue	173
1. Le théorème de convergence dominée pour les fonctions en escalier bornées	174
2. Mesure et mesure extérieure des ensembles mesurables	176
3. Le théorème de convergence monotone pour les fonctions bornées à support compact	178
4. Limites simples p.p. de fonctions mesurables	179
5. Le théorème de convergence monotone et ses conséquences	180
IV. Transformée de Fourier	181
IV.1. Intégrales dépendant d'un paramètre	181
IV.2. Transformée de Fourier dans L^1	184
1. Caractères linéaires de \mathbf{R} et \mathbf{R}^m	184
2. Définition et premières propriétés	184
3. Le théorème de Riemann-Lebesgue	185
4. Transformée de Fourier et dérivation	186
IV.3. Formules d'inversion	187
1. Séries de Fourier	187
2. Séries de Fourier multidimensionnelles	191
2.1. Le cas du réseau \mathbf{Z}^m	191
2.2. Le cas d'un réseau quelconque	193
3. La formule de Poisson	194
4. La formule d'inversion de Fourier dans \mathcal{S}	195

5. Formules d'inversion dans L^1	196
6. Exercices	197
IV.4. Transformée de Fourier dans L^2	198
1. Transformée de Fourier des fonctions en escalier	198
2. Définition de la transformée de Fourier dans L^2	200
3. Comparaison des transformées de Fourier dans L^1 et L^2	201
4. Dérivation	201
V. Fonctions holomorphes	203
V.1. Fonctions holomorphes et fonctions analytiques complexes	203
1. Séries entières	203
2. Rayon de convergence d'une série entière	205
3. Premières propriétés des fonctions holomorphes	207
3.1. Définition	207
3.2. Théorème des zéros isolés et unicité du prolongement analytique	208
3.3. Principe du maximum	210
V.2. La formule intégrale de Cauchy et ses conséquences	211
1. Généralités sur les chemins	211
2. Intégration le long d'un chemin	212
3. Holomorphicité des fonctions dérivables au sens complexe	213
4. Construction de fonctions holomorphes	216
4.1. Séries de fonctions holomorphes	216
4.2. Produits infinis de fonctions holomorphes	217
4.3. Fonctions holomorphes définies par une intégrale	218
V.3. Structure locale des fonctions holomorphes	220
1. Le théorème d'inversion locale holomorphe	220
2. Logarithme et fonctions puissances	222
VI. La formule de Cauchy et celle des résidus (de Cauchy)	225
VI.1. Homotopie de lacets et formule de Cauchy	225
1. Vocabulaire de topologie algébrique	225
2. Un cas particulier de la formule de Stokes	226
3. Démonstration de la formule de Cauchy	229
4. Primitives	231
VI.2. La formule des résidus de Cauchy	232
1. Fonctions holomorphes sur une couronne	232
2. Fonctions holomorphes sur un disque épointé; résidus	235

3. Indice d'un lacet par rapport à un point	237
3.1. Définition	238
3.2. Détermination visuelle de l'indice d'un lacet par rapport à un point	238
4. La formule des résidus	240
5. Exercices	241
VI.3. Loi d'addition sur une courbe elliptique, problème corrigé	243
VII. Séries de Dirichlet	249
VII.1. Séries de Dirichlet	249
1. Abscisse de convergence absolue	249
2. Demi-plan de convergence d'une série de Dirichlet	251
VII.2. Séries de Dirichlet et transformée de Mellin	253
1. La fonction Γ dans le plan complexe	253
2. Une formule intégrale pour les séries de Dirichlet	255
VII.3. La fonction zêta de Riemann	258
1. Séries de Dirichlet attachées à des fonctions multiplicatives	258
2. Prolongement analytique de la fonction ζ	259
3. Équation fonctionnelle de la fonction zêta	260
4. Les zéros de la fonction ζ	264
VII.4. Fonctions L de Dirichlet	264
1. Caractères de Dirichlet et Fonctions L de Dirichlet	264
2. Conducteur et sommes de Gauss	265
3. Équation fonctionnelle des fonctions L de Dirichlet	266
VII.5. Autres exemples	268
1. La fonction de Moebius	268
2. La fonction τ de Ramanujan	269
VII.6. Exercices	270
A. Le théorème des nombres premiers	275
A.1. Introduction	275
A.2. Les fonctions ψ et ψ_1	278
1. Théorème des nombres premiers et comportement de ψ_1 en $+\infty$	278
2. Une formule intégrale pour ψ_1	279
A.3. Formules explicites	280
1. Énoncé du résultat	281
2. Les fonctions L et $\frac{L'}{L}$ en dehors de la bande critique	283
3. La fonction L dans la bande critique	284

4. La fonction $\frac{L'}{L}$ dans la bande critique	285
5. Conclusion	286
A.4. Démonstration du théorème des nombres premiers	288
1. Non annulation sur la droite $\text{Re}(s) = 1$	288
2. Conclusion	290
A.5. Compléments	290
1. L'hypothèse de Riemann et ses conséquences	290
2. L'hypothèse de Riemann et la fonction M de Mertens	291
3. L'hypothèse de Lindelöf	292
B. Groupes finis et représentations : exemples	293
B.1. p -Groupes	293
1. Généralités sur les p -groupes	293
2. Représentations des p -groupes	294
B.2. Représentations du groupe symétrique S_n	295
1. Partitions de n et représentations de S_n	295
2. Diagrammes de Young et représentations de S_n	297
3. Caractères de S_n	298
B.3. Représentations de $\mathbf{GL}_2(\mathbf{F})$	299
1. Le groupe $\mathbf{GL}_2(\mathbf{F})$	299
2. Construction de représentations de $\mathbf{GL}_2(\mathbf{F})$	299
3. Les classes de conjugaison de $\mathbf{GL}_2(\mathbf{F})$	300
4. La table des caractères de $\mathbf{GL}_2(\mathbf{F})$	302
5. Démonstrations	303
C. Le problème des nombres congruents	309
C.1. Introduction	309
C.2. Arithmétique des courbes elliptiques	311
C.3. L'heuristique de Birch et Swinnerton-Dyer	312
C.4. Fonction L d'une courbe elliptique	313
C.5. La stratégie de Tunnell	315
C.6. Formes modulaires	316
C.7. Courbes elliptiques et formes modulaires	317
D. Introduction au programme de Langlands	319
D.1. La conjecture d'Artin	321
1. Le groupe $\mathcal{G}_{\mathbf{Q}}$	321
2. Représentations de $\mathcal{G}_{\mathbf{Q}}$	322

3. Fonctions L d'Artin	324
4. Fonctions L de degré 2	325
4.1. Représentations impaires et formes modulaires	325
4.2. Représentations paires et formes de Maass	327
5. La théorie du corps de classes	328
D.2. Le théorème de Kronecker-Weber revisité	330
1. Adèles	331
1.1. Le théorème d'Ostrowski	331
1.2. L'anneau des adèles de \mathbf{Q}	332
1.3. Le groupe des idèles de \mathbf{Q}	333
2. La formule de Poisson adélique	334
2.1. Transformée de Fourier sur \mathbf{Q}_p	334
2.2. Transformée de Fourier adélique	336
3. Transformée de Mellin adélique et fonctions L	337
3.1. Intégration sur \mathbf{Q}_p^*	337
3.2. Intégration sur le groupe des idèles	337
3.3. Transformée de Mellin sur \mathbf{Q}_p	338
3.4. La transformée de Mellin adélique	339
3.5. Le théorème de Tate	340
4. Application aux fonctions L de Dirichlet	341
4.1. La fonction zêta de Riemann	342
4.2. Fonctions L de caractères de \mathbf{A}^*	342
4.3. Caractères de Dirichlet et caractères linéaires continus des idèles	344
D.3. Le programme de Langlands	344
1. Représentations automorphes	344
2. Des formes modulaires aux représentations automorphes	347
2.1. Préliminaires	347
2.2. La forme automorphe associée à une forme modulaire	347
2.3. La décomposition de $\mathbf{GL}_2(\mathbf{A})$	348
3. Quelques autres aspects du programme de Langlands	349
Index terminologique	351
Conjectures et théorèmes	355
Index des noms propres	357

INTRODUCTION

Les mathématiques sont à la fois un outil d'une puissance surprenante, utilisé à des degrés divers par les autres sciences, et une des plus incroyables constructions collectives de l'humanité, s'appuyant sur des bases consolidées génération après génération pour permettre à l'édifice de monter toujours plus haut.

Ce cours est une introduction à trois des théories qui servent de socle aux mathématiques. La première (chap. I) est la théorie des représentations des groupes finis et de leurs caractères, développée dans les années 1895-1905 par F. Frobenius, W. Burnside et I. Schur. Cette théorie est une extension de l'algèbre linéaire (il s'agit de comprendre l'action simultanée de plusieurs isomorphismes sur un espace vectoriel de dimension fini, et donc l'action du groupe qu'ils engendrent), mais la théorie des caractères est aussi une première approche de la transformée de Fourier dans un cadre fini où les difficultés analytiques sont absentes. La théorie des représentations des groupes joue un rôle central en mathématiques, dans certaines branches de la physique (par exemple en physique des particules) ou encore dans une petite partie de la chimie classique (cristallographie) ; le cas des groupes finis sert souvent de guide pour deviner ce que l'on est en droit d'espérer dans des cas plus compliqués.

La seconde (chap. II, III et IV) est l'analyse fonctionnelle des années 1900-1930 (espaces de Banach, intégration de Lebesgue, transformée de Fourier), dans laquelle se sont illustrés R. Baire, S. Banach, M. Fréchet, H. Hahn, D. Hilbert, H. Lebesgue, M. Plancherel, F. Riesz, H. Steinhaus... Cette théorie, née des préoccupations du siècle précédent concernant les équations différentielles, les équations aux dérivées partielles..., forme la base de l'analyse réelle moderne. Ses applications à l'étude des équations aux dérivées partielles provenant de la physique (équations de la chaleur, des ondes, de Schrödinger...) sont innombrables.

La dernière partie du cours (chap V, VI et VII) est consacrée à la théorie des fonctions analytiques d'une variable complexe, qui s'est développée entre les mains de A. Cauchy dans les années 1820-1840, mais a été revisitée régulièrement depuis ; la présentation suivie dans ce cours doit beaucoup aux apports de K. Weierstrass et de H. Poincaré datant de

la seconde moitié du XIX^e siècle. Cette théorie est probablement, avec la théorie générale des groupes, celle qui est utilisée dans le plus grand nombre des autres branches des mathématiques ou de la physique théorique. Par exemple, la représentation conforme des ouverts du plan, à laquelle nous ne ferons qu'une brève allusion (note 1 du chap. VI), a des applications à l'étude de l'équation de la chaleur avec conditions au bord dans un domaine plan, à l'aérodynamique (transformation de Joukovski), à l'étude du mouvement brownien ou celle des polymères, etc.

Le problème majeur d'un cours de ce type est que l'on est conduit à privilégier les résultats qui ont le plus d'applications futures et à reléguer en exercice tout ce qui fait le sel des mathématiques, ce qui revient un peu à visiter une cathédrale en ne s'intéressant qu'aux consolidations successives de la base de ses piliers. Pour essayer de lutter contre cette tendance, nous avons privilégié des objets analytiques, issus de la théorie des nombres, ayant la faculté étonnante d'interagir avec quasiment tous les domaines des mathématiques (voire de la physique théorique) et, ce faisant, de contribuer fortement au développement de ces domaines. Il s'agit des fonctions L, dont la fonction zêta de Riemann (définie par $\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$ pour $\text{Re}(s) > 1$) est le prototype. L'un des premiers résultats remarquables concernant ces objets est probablement la célèbre formule $\zeta(2) = \frac{\pi^2}{6}$ de L. Euler (1734), répondant à une question posée en 1644 et connue sous le nom de « problème de Bâle ». Le même Euler a mis au jour un lien heuristique entre la fonction ζ et la répartition des nombres premiers qui ne fut rigoureusement établi qu'en 1896 par J. Hadamard et C. de la Vallée Poussin en suivant une stratégie suggérée par B. Riemann en 1858. Entre-temps, G. Dirichlet avait introduit en 1837 les premières fonctions L pour démontrer l'existence d'une infinité de nombres premiers dans les progressions arithmétiques. L'annexe A, consacrée à ces résultats, fournit une illustration frappante de l'utilité des fonctions holomorphes pour attaquer des problèmes qui en semblent fort éloignés. Depuis, le monde des fonctions L s'est enrichi au point de former un édifice imposant dont l'annexe D essaie de donner une idée en partant de la constatation que, pour apprécier l'élégance et la majesté de la voûte de Notre-Dame, il n'est nul besoin de comprendre pourquoi elle ne s'écroule pas ni, a fortiori, comment on a fait pour la construire sans que tout tombe au fur et à mesure. Nous nous sommes restreint à l'aspect analytique des fonctions L ; celui-ci fait intervenir d'autres objets mathématiques ayant un don d'ubiquité assez époustoufflant, à savoir les formes modulaires que nous avons reléguées dans une série d'exercices en vertu du principe énoncé plus haut. Nous avons résisté à la tentation d'explorer les propriétés arithmétiques de ces fonctions L : leurs valeurs aux entiers cachent des trésors qui font l'objet de conjectures générales de P. Deligne (1977, dont la conjecture met en perspective le π^2 de la formule d'Euler, et la non apparition de π^3 pour $\zeta(3)$), de A. Beilinson (1985, qui vise, en particulier, à expliquer quels objets interviennent dans $\zeta(3)$) et de S. Bloch et K. Kato (1989, dont la conjecture donne une formule totalement générale fournissant, par exemple, une signification à l'apparition de 691 dans la formule $\zeta(12) = \frac{691 \pi^{12}}{3^6 \cdot 5^3 \cdot 7^2 \cdot 11 \cdot 13}$). Un exemple de ces trésors cachés est la conjecture de

Birch et Swinnerton-Dyer, qui date du début des années 1960, et à laquelle l'annexe C est consacrée.

Notations standard

On note \mathbf{N} l'ensemble des entiers naturels, \mathbf{Z} l'anneau des entiers relatifs, \mathbf{Q} le corps des nombres rationnels, \mathbf{R} le corps des nombres réels et \mathbf{C} le corps des nombres complexes. On note \mathbf{Q}^* , \mathbf{R}^* et \mathbf{C}^* les groupes multiplicatifs de \mathbf{Q} , \mathbf{R} et \mathbf{C} .

On note \mathbf{R}_+ (resp. \mathbf{R}_+^*) l'ensemble des nombres réels positifs (resp. strictement positifs) et \mathbf{R}_- (resp. \mathbf{R}_-^*) l'ensemble des nombres réels négatifs (resp. strictement négatifs).

On note $\overline{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$ la droite réelle achevée, et $\overline{\mathbf{R}}_+ = \mathbf{R}_+ \cup \{+\infty\}$ la demi-droite réelle achevée.

Si $t \in \mathbf{R}$, on note $[t]$ sa partie réelle, et $\{t\} = t - [t]$, sa partie fractionnaire.

Si X est un ensemble, on note $|X|$ son cardinal.

Si A est un anneau, et si $n \in \mathbf{N} - \{0\}$, on note $\mathbf{M}_n(A)$ l'anneau des matrices $n \times n$ à coefficients dans A , $\mathbf{GL}_n(A) \subset \mathbf{M}_n(A)$ le groupe des matrices inversibles (celles dont le déterminant est inversible dans A), et $\mathbf{SL}_n(A)$ le sous-groupe de $\mathbf{GL}_n(A)$ des matrices de déterminant 1.

Bibliographie sommaire

Le lecteur désirant approfondir⁽¹⁾ certains des thèmes développés dans ce cours est invité à consulter les ouvrages ci-dessous. Ces ouvrages partent à peu près au même niveau que le présent cours, mais sont plus spécialisés, ce qui leur permet d'aller plus loin.

P. Biane, J-B. Bost et P. Colmez, *La fonction zêta*, Presses de l'École Polytechnique.

Le lecteur y trouvera divers aspects de la fonction zêta en lien avec l'arithmétique ou les probabilités⁽²⁾

J-B. Bost, *Fonctions analytiques d'une variable complexe*, École Polytechnique.

Couvre les chapitres V à VII, et une partie de l'annexe A du cours, en suivant un rythme moins effréné, et propose des prolongements en direction de l'analyse.

D. Bump, *Automorphic forms and representations*, Cambridge University Press.

Version développée de l'annexe D ; sa lecture demande un investissement non négligeable.

H. Cartan, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Hermann.

Couvre les chapitres V et VI du cours, et poursuit en direction de la géométrie (surfaces de Riemann, et fonctions de plusieurs variables).

⁽¹⁾La manière standard pour fabriquer des exercices est de prendre des résultats démontrés dans des ouvrages plus spécialisés et de les découper en questions. Le lecteur trouvera donc dans ces ouvrages la solution de la plupart des exercices de ce cours...

⁽²⁾Le second texte de ce volume est loin d'être exempt de fautes de frappes ; admirer le pouvoir de la double négation...

W. Ellison, *Les nombres premiers*, Hermann.

Couvre l'annexe A, et bien plus.

W. Fulton et J. Harris, *Representation theory. A first course*, GTM 129, Springer-Verlag.

Débute par le chapitre I et l'annexe B du cours, et poursuit en direction des représentations des groupes de Lie.

R. Godement, *Analyse mathématique II, III et IV*, Springer-Verlag.

Couvre l'essentiel du cours, et de ce que j'aurais voulu y mettre, en prenant son temps, ce que son nombre de pages permet. Les formes modulaires y sont traitées avec le respect qu'elles méritent.

N. Koblitz, *Introduction to elliptic curves and modular forms*, GTM 97, Springer-Verlag.

Offre un voyage à travers la théorie des nombres en prenant comme fil conducteur le problème des nombres congruents (annexe C).

S. Patterson, *An introduction to the theory of the Riemann Zeta-function*, Cambridge University Press.

Couvre l'annexe A, et poursuit en direction des hypothèses de Riemann et Lindelöf.

W. Rudin, *Real and complex Analysis*, Mc Graw-Hill

Un cours d'analyse qui couvre en particulier la partie analyse du cours (chapitres II à VI), mais ne s'arrête pas là, loin s'en faut.

J-P. Serre, *Cours d'arithmétique*, Presses Universitaires de France.

Un fort joli livre pour en apprendre plus sur les formes quadratiques à coefficients rationnels et les formes modulaires.

J-P. Serre, *Représentations linéaires des groupes finis*, Hermann

Couvre le chapitre I et une partie de l'annexe B, et continue sur des sujets plus pointus concernant les représentations des groupes finis.

A. Weil, *Elliptic functions according to Eisenstein and Kronecker*, Springer-Verlag.

Un livre semi-historique très agréable à lire, illustrant à un niveau élémentaire les liens entre les fonctions holomorphes et la théorie des nombres.

VOCABULAIRE MATHÉMATIQUE

La nécessité de définir précisément les objets avec lesquels ils travaillent s'est imposée graduellement aux mathématiciens confrontés à des contradictions d'ordre presque métaphysique. L'avènement de la théorie des ensembles (à partir des travaux fondateurs de G. Cantor dont le début date des années 1870) et l'axiomatisation croissante des mathématiques ont d'une part fait disparaître un certain nombre d'obstacles psychologiques à la création d'objets nouveaux⁽³⁾, et d'autre part débouché sur la création d'un vocabulaire extrêmement précis, qui a rendu possible l'explosion des mathématiques au cours du XX^e siècle.

Ce mouvement a fini par atteindre l'enseignement avec l'introduction des « maths modernes » au collège (et même en grande section de maternelle). Dans les années 70, le programme enseigné dans le secondaire et dans les classes préparatoires reposait sur le slogan : « Dieu créa l'ensemble vide et l'homme fit le reste ». C'était un peu radical, mais avait le mérite de présenter les mathématiques de manière cohérente et de montrer que l'on pouvait créer de nouveaux objets à partir d'objets déjà existants. La présentation en était malheureusement extrêmement dogmatique, et l'impression qu'on en retirait était plutôt que Dieu avait créé l'ensemble vide et la théorie des ensembles, et sur sa lancée, les entiers, les entiers relatifs, les nombres rationnels, puis les groupes, les anneaux, les corps et les espaces vectoriels, puis les nombres réels, ensuite il avait introduit des ε et des δ ,

⁽³⁾Les nombres complexes ont mis près de deux siècles à être acceptés (et même les nombres négatifs ont eu leurs détracteurs ; un cas extrême est Augustus de Morgan qui continuait à les considérer, au milieu du XIX^e-siècle, comme dénués de tout fondement, et a passé une bonne partie de sa vie à essayer de prouver qu'on pouvait fort bien s'en passer), alors que, de nos jours, des objets nettement plus compliqués le sont dès qu'ils ont fait la preuve de leur utilité pour résoudre, ou même formuler proprement, certains problèmes ; c'est par exemple le cas de l'anneau des « nombres complexes p -adiques » construit par J.-M. Fontaine (1982). Les obstacles psychologiques n'ont toutefois pas complètement disparu ; l'apparition d'un objet nouveau ne se fait pas sans heurt, et provoque des conflits parfois brutaux entre les anciens, dont le point de vue « On a fait de très bonnes maths pendant 2000 ans sans avoir besoin de ces horreurs » reflète l'appréhension devant la perspective de devoir étudier un nouveau sujet « incompréhensible », et les modernes qui voient dans le nouvel objet la solution à tous les problèmes...

puis créé la topologie..., et quand il avait enfin été content du résultat, il avait fait don aux hommes d'une théorie immuable et parfaite, à la beauté froide et lisse.

Le dogme a changé vers le milieu des années 90, et on est reparti sur le mode : « Dieu a créé les nombres réels, puis les nombres complexes, et envoyé Gauss sur terre pour expliquer qu'il n'y avait pas besoin de chercher plus loin. ». Tout procédé de construction a été soigneusement banni du programme officiel, et une grande partie du vocabulaire mathématique de base a disparu ou a été vidé de sa substance⁽⁴⁾. C'est fort regrettable car la maîtrise du vocabulaire mathématique demande du temps : il décrit des concepts qui reposent souvent sur d'autres concepts, et il faut voir fonctionner ces concepts pour saisir véritablement le sens des mots. Or ce temps fait cruellement défaut une fois passée la période des classes préparatoires.

Ce chapitre essaie de pallier à ces disparitions ; la plus grande partie de son contenu n'est pas utilisée dans le texte principal⁽⁵⁾, mais est incluse car elle est susceptible de faire son apparition dans n'importe quel domaine utilisant des mathématiques. Il ne prend pas les mathématiques à leur début⁽⁶⁾, et le lecteur est supposé avoir déjà des notions même vagues de la plupart des sujets qui suivent. Plutôt qu'un cours organisé, il s'agit d'une espèce de dictionnaire, et comme dans un dictionnaire, il n'est pas rare que certains passages fassent appel à des notions définies ultérieurement.

1. Grammaire élémentaire

Si X est un ensemble, on note $|X|$ son cardinal.

L'expression $A \cong B$ signifie qu'il existe un isomorphisme entre A et B (la notion dépend donc de la structure mise sur A et B), ce qui est nettement moins précis (et donc plus

⁽⁴⁾Le programme de la filière PC est à cet égard assez catastrophique, puisque sa dernière mouture a vu l'introduction de faux concepts, et même de définitions fausses. Il est un peu difficile de comprendre l'idéologie qui a abouti à ce résultat ; peut-être peut-on mettre cela sur le compte de l'effet d'horizon (qui conduisait les premiers programmes jouant aux échecs à faire des sacrifices incompréhensibles visant à repousser hors de leur champ de vision un mat improbable) et de l'utilitarisme à court terme faisant des ravages dans les mentalités de l'époque.

⁽⁵⁾Les résultats exposés dans le cours sont en grande partie antérieurs à la mise en valeur des concepts présentés dans ce chapitre, ce qui fait que l'on peut les présenter, en se contorsionnant un peu, sans recourir à ces concepts. D'un autre côté, lire « Les misérables » ou les « Disquisitiones arithmeticae » à la lumière d'une lampe électrique est nettement plus confortable qu'à la lueur d'une chandelle, même si ces œuvres datent d'avant l'invention de l'ampoule électrique et si la chandelle a un charme certain...

⁽⁶⁾Il a été écrit de la manière suivante. J'ai d'abord, pour chaque concept de base, fait une liste des énoncés que j'utilise régulièrement sans me poser de question. C'est plus ou moins ce qui se trouve en gros caractères. J'ai ensuite rajouté les démonstrations (en général en petits caractères). Une exception à ce principe est le traitement de l'algèbre linéaire, où j'ai remplacé les démonstrations vues en classes préparatoires par d'autres, donnant des résultats plus puissants. J'ai aussi rajouté, pour les amateurs, une collection de monstres mathématiques, et quelques résultats plus culturels comme la construction des nombres p -adiques, les théorèmes de Sylow ou la simplicité de A_n .

souple) qu'une phrase du genre « u réalise un isomorphisme de A sur B » où un isomorphisme explicite est requis. Par exemple, dire que deux espaces vectoriels de dimension finie sur un corps K sont isomorphes revient juste à dire qu'ils ont la même dimension.

1.1. L'anneau \mathbf{Z} des entiers relatifs

- Si A est un sous groupe de \mathbf{Z} (muni de $+$), il existe $D \geq 0$ unique, tel que $A = D\mathbf{Z}$.

Si $A = \{0\}$, alors $D = 0$. Si $A \neq \{0\}$, alors A contient des éléments > 0 puisque A est stable par $x \mapsto -x$; soit D le plus petit de ces éléments. Une récurrence immédiate montre que A contient nD , pour tout $n \in \mathbf{N}$, et donc aussi pour tout $n \in \mathbf{Z}$ puisque A est stable par $x \mapsto -x$. Autrement dit, $A \supset D\mathbf{Z}$.

Maintenant, soit $a \in A$, et soit $r \in \{0, \dots, D-1\}$ le reste de la division euclidienne de a par D . Alors $a - r \in D\mathbf{Z} \subset A$, et donc $r = a - (a - r) \in A$. Comme D est par hypothèse le plus petit élément strictement positif de A , cela implique $r = 0$, et donc $a \in D\mathbf{Z}$. On en déduit l'inclusion $A \subset D\mathbf{Z}$ et l'égalité $A = D\mathbf{Z}$ que l'on cherchait à démontrer.

On écrit $a \mid b$ (pour a divise b) pour signifier que b est un multiple de a , et $a \nmid b$ pour signifier le contraire. Si $a, b \in \mathbf{Z}$, on définit le *plus grand diviseur commun* $\text{pgcd}(a, b)$ de a et b comme étant 0 si $a = b = 0$, et comme étant le plus grand entier $d > 0$ divisant à la fois a et b , si $a \neq 0$ ou $b \neq 0$. On dit que a et b sont *premiers entre eux*, si $\text{pgcd}(a, b) = 1$.

Un élément p de \mathbf{N} est *premier*, si $p \neq 1$ et si les seuls diviseurs de p sont 1 et p . On note $\mathcal{P} = \{2, 3, 5, \dots\}$ l'ensemble des nombres premiers. Il est clair que si $p \in \mathcal{P}$, et si $a \in \mathbf{N}$, alors soit $p \mid a$ auquel cas $\text{pgcd}(p, a) = p$, soit $p \nmid a$ auquel cas p est premier à a .

Remarquons que $a\mathbf{Z} + b\mathbf{Z} = \{ax + by, x, y \in \mathbf{Z}\}$ est un sous-groupe de \mathbf{Z} ; et c'est le plus petit sous-groupe de \mathbf{Z} contenant a et b (en effet, un sous-groupe de \mathbf{Z} contenant a et b contient ax et by et donc aussi $ax + by$, pour tous $x, y \in \mathbf{Z}$). On note (a, b) , l'élément de \mathbf{N} tel que $a\mathbf{Z} + b\mathbf{Z} = (a, b)\mathbf{Z}$; cet élément existe et est unique d'après le point ci-dessus.

- Si $a, b \in \mathbf{Z}$, alors $(a, b) = \text{pgcd}(a, b)$; en particulier, a et b sont premiers entre eux si et seulement si il existe $u, v \in \mathbf{Z}$ tels que $1 = au + bv$ (théorème de Bézout).

Si $a = b = 0$, le résultat est immédiat. Supposons donc $a \neq 0$ ou $b \neq 0$. Par définition de (a, b) , a et b sont des multiples de (a, b) , et donc $(a, b) \leq \text{pgcd}(a, b)$. Réciproquement, si $d \geq 1$ divise a et b , alors d divise $ax + by$, quels que soient $x, y \in \mathbf{Z}$; en particulier, d divise (a, b) et donc $d \leq (a, b)$. On en déduit l'inégalité $(a, b) \geq \text{pgcd}(a, b)$ qui permet de conclure.

- Si a est premier avec b et c , alors a est premier avec bc ; si a divise bc et si a est premier avec b , alors a divise c (lemme de Gauss).

Si $(a, b) = (a, c) = 1$, il existe u_1, v_1 tels que $au_1 + bv_1 = 1$ et u_2, v_2 tels que $au_2 + cv_2 = 1$. On a donc $1 = (au_1 + bv_1)(au_2 + cv_2) = au + bcv$, avec $u = au_1u_2 + bv_1u_2 + cu_1v_2$ et $v = v_1v_2$, ce qui prouve que $(a, bc) = 1$. On en déduit le premier énoncé.

Si $bc = ad$ et $au + bv = 1$, alors $acu + adv = c$, et donc $a(cu + dv) = c$, ce qui prouve que a divise c ; d'où le second énoncé.

- Si $n \in \mathbf{Z} - \{0\}$, il existe des nombres premiers p_1, \dots, p_r tels que $n = \text{sign}(n) p_1 \cdots p_r$; de plus, les p_i , pour $1 \leq i \leq r$, sont uniquement déterminés à l'ordre près. En d'autres

termes, n peut se factoriser de manière unique comme produit de facteurs premiers⁽⁷⁾ (théorème fondamental de l'arithmétique).

Le cas $n < 0$ se déduit du cas $n > 0$; on peut donc supposer $n > 0$.

L'existence se démontre par récurrence. C'est évident pour $n = 1$, auquel cas, on a $r = 0$ (un produit vide vaut 1 par définition). Maintenant, si $n \geq 2$ est premier, alors $n = n$ est une factorisation de n sous la forme voulue. Si $n \geq 2$ n'est pas premier, alors $n = ab$, avec $2 \leq a \leq n - 1$ et $2 \leq b \leq n - 1$. On peut donc appliquer l'hypothèse de récurrence à a et b , ce qui permet d'écrire a sous la forme $a = p_1 \cdots p_s$, et b sous la forme $b = p_{s+1} \cdots p_r$, où p_1, \dots, p_r sont des nombres premiers. On a alors $n = p_1 \cdots p_r$, ce qui prouve que n admet une factorisation sous la forme voulue.

L'unicité se démontre en utilisant le lemme de Gauss. Si $p_1 \cdots p_r = q_1 \cdots q_s$ où les p_i et les q_j sont des nombres premiers, le lemme de Gauss montre que p_r divise l'un des q_j et donc lui est égal. Quitte à permuter les q_j , on peut supposer que $p_r = q_s$, et en divisant les deux membres par $p_r = q_s$, on se ramène à $r - 1$ et $s - 1$, ce qui permet de conclure par récurrence.

• Si $n \in \mathbf{Z} - \{0\}$, et si p est un nombre premier, on note $v_p(n)$ le nombre de fois que p apparaît dans la décomposition en facteurs premiers de n ; alors $p^{v_p(n)}$ est aussi la plus grande puissance de p divisant n , et $v_p(n)$ est la *valuation p -adique* de n .

On étend cette définition à $n \in \mathbf{Z}$ en posant $v_p(0) = +\infty$. On dispose alors d'un critère de divisibilité assez utile : *a divise b si et seulement si $v_p(a) \leq v_p(b)$ pour tout nombre premier p* . En revenant à la définition de $\text{pgcd}(a, b)$, on en déduit la formule $\text{pgcd}(a, b) = \prod_p p^{\inf(v_p(a), v_p(b))}$.

Exercice 1.1. — Si $a, b \in \mathbf{Z}$, on définit le plus petit commun multiple $\text{ppcm}(a, b)$ de a et b comme le plus petit entier ≥ 0 , multiple à la fois de a et b .

(i) Montrer que $a\mathbf{Z} \cap b\mathbf{Z}$ est un sous-groupe de \mathbf{Z} , et que $a\mathbf{Z} \cap b\mathbf{Z} = \text{ppcm}(a, b)\mathbf{Z}$.

(ii) Montrer que $\text{ppcm}(a, b) = \prod_p p^{\sup(v_p(a), v_p(b))}$, si $a \neq 0$ et $b \neq 0$.

Exercice 1.2. — (i) Montrer que, si $a, b \in \mathbf{Z}$, alors $v_p(ab) = v_p(a) + v_p(b)$ et $v_p(a + b) \geq \inf(v_p(a), v_p(b))$.

(ii) Montrer que v_p a un unique prolongement à \mathbf{Q} tel que $v_p(xy) = v_p(x) + v_p(y)$, quels que soient $x, y \in \mathbf{Q}$, et que l'on a alors $v_p(x + y) \geq \inf(v_p(x), v_p(y))$, quels que soient $x, y \in \mathbf{Q}$.

Exercice 1.3. — Soient $n \geq 1$ et p un nombre premier. Montrer que $v_p(n!) = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots$. En déduire que $v_p(n!) = \frac{n - S_p(n)}{p - 1}$, où $S_p(n)$ est la somme des chiffres de n dans son développement en base p .

⁽⁷⁾Si n est le produit de deux nombres premiers ayant chacun un millier de chiffres, on peut prouver, avec l'aide d'un ordinateur, que n n'est pas premier, mais il est impossible, à l'heure actuelle, de retrouver les deux nombres premiers qui divisent n . Ceci est à la base de la sécurité du système RSA en vigueur pour les transactions sur Internet. C'est aussi une bonne illustration de la différence entre la théorie et la pratique, qui en théorie sont la même chose, mais en pratique...

1.2. Parallélisme entre logique élémentaire et langage ensembliste

- La négation $p \mapsto \bar{p}$ correspond au passage au complémentaire : $\{x, \bar{p}(x)\}$ est le complémentaire de $\{x, p(x)\}$.
- \wedge (“et”) correspond à l’intersection : $\{x, p(x) \wedge q(x)\} = \{x, p(x)\} \cap \{x, q(x)\}$.
- \vee (“ou”) correspond à la réunion : $\{x, p(x) \vee q(x)\} = \{x, p(x)\} \cup \{x, q(x)\}$.
- La formule $\overline{p \vee q} = \bar{p} \wedge \bar{q}$ (resp. $\overline{p \wedge q} = \bar{p} \vee \bar{q}$) devient : le complémentaire de la réunion (resp. l’intersection) est l’intersection (resp. la réunion) des complémentaires.
- \Rightarrow correspond à l’inclusion : $p \Rightarrow q$ si et seulement si $\{x, p(x)\} \subset \{x, q(x)\}$.
- \forall correspond à une intersection : $\{x, \forall i \in I, p_i(x)\} = \bigcap_{i \in I} \{x, p_i(x)\}$.
- \exists correspond à une réunion : $\{x, \exists i \in I, p_i(x)\} = \bigcup_{i \in I} \{x, p_i(x)\}$.

Considérons, par exemple, deux espaces métriques X et Y , et une suite de fonctions $(f_n)_{n \in \mathbf{N}}$ de X dans Y . Soit A l’ensemble des $x \in X$ tels que $f_n(x)$ converge. Alors A peut s’écrire sous la forme :

$$\begin{aligned} A &= \{x \in X, \exists y \in Y, \forall j \in \mathbf{N}, \exists N \in \mathbf{N}, \forall n \geq N, d(f_n(x), y) < 2^{-j}\} \\ &= \bigcup_{y \in Y} \bigcap_{j \in \mathbf{N}} \bigcup_{N \in \mathbf{N}} \bigcap_{n \geq N} f_n^{-1}(\{y' \in Y, d(y, y') < 2^{-j}\}). \end{aligned}$$

Si Y est complet, on peut utiliser le critère de Cauchy au lieu de donner un nom à la limite, et on obtient [en notant $f_{n,p} : X \rightarrow Y \times Y$ la fonction $x \mapsto (f_n(x), f_p(x))$] :

$$\begin{aligned} A &= \{x \in X, \forall j \in \mathbf{N}, \exists N \in \mathbf{N}, \forall n, p \geq N, d(f_n(x), f_p(x)) < 2^{-j}\} \\ &= \bigcap_{j \in \mathbf{N}} \bigcup_{N \in \mathbf{N}} \bigcap_{n, p \geq N} f_{n,p}^{-1}(\{(y, y') \in Y \times Y, d(y, y') < 2^{-j}\}). \end{aligned}$$

La seconde formulation a l’avantage de ne faire intervenir que des intersections et réunions indexées par des ensembles dénombrables.

1.3. Ensembles dénombrables

Un ensemble est *dénombrable* s’il est fini ou s’il peut être mis en bijection avec \mathbf{N} .

- Un sous-ensemble d’un ensemble dénombrable est dénombrable.

Il suffit de démontrer qu’un sous-ensemble X de \mathbf{N} , qui n’est pas fini, peut être mis en bijection avec \mathbf{N} . Si $x \in X$, soit $\varphi(x) = |\{y \in X, y < x\}|$. Si x_0 est le plus petit élément de X , on a $\varphi(x_0) = 0$, ce qui montre que $\varphi(X)$ contient 0. Si $\varphi(x) = n$, et x' est le plus petit élément de X strictement supérieur à x , on a $\varphi(x') = n + 1$, ce qui prouve que φ est surjective. Par ailleurs, φ est injective car strictement croissante (si $x_1 < x_2$, alors $\{y \in X, y < x_2\}$ contient $\{y \in X, y < x_1\}$ et x_1). Ceci permet de conclure.

- Si $\varphi : X \rightarrow Y$ est injective et si Y est dénombrable, alors X est dénombrable ; si $\varphi : X \rightarrow Y$ est surjective et si X est dénombrable, alors Y est dénombrable.

Si $\varphi : X \rightarrow Y$ est injective, alors φ réalise une bijection de X sur $\varphi(X)$ qui est dénombrable comme sous-ensemble d’un ensemble dénombrable, et donc X est dénombrable. Si $\varphi : X \rightarrow Y$ est surjective, on peut choisir, pour tout $y \in Y$, un antécédent $s(y) \in X$ de y par φ . Alors $s : Y \rightarrow X$ est injective car $s(y_1) = s(y_2)$ implique $y_1 = \varphi(s(y_1)) = \varphi(s(y_2)) = y_2$, et donc Y est dénombrable si X l’est, d’après ce qui précède.

- Un produit fini d’ensembles dénombrables est dénombrable.

Soient X_1, \dots, X_k des ensembles dénombrables, $X = X_1 \times \dots \times X_k$, et p_1, \dots, p_k des nombres premiers distincts. Soit $\varphi_i : X_i \rightarrow \mathbf{N}$ injective, pour tout $i \in \{1, \dots, k\}$. Alors $\varphi : X \rightarrow \mathbf{N}$, définie par $\varphi(x_1, \dots, x_k) = p_1^{\varphi_1(x_1)} \dots p_k^{\varphi_k(x_k)}$ est injective d'après le « théorème fondamental de l'arithmétique » (unicité de la factorisation d'un entier naturel non nul en produit de nombres premiers).

- Une réunion dénombrable d'ensembles dénombrables est dénombrable.

Soit $(X_i)_{i \in I}$, avec I dénombrable et chacun des X_i aussi. Soient $\varphi_i : X_i \rightarrow \mathbf{N}$, pour $i \in I$, des applications injectives, et soit $Y \subset I \times \mathbf{N}$ l'ensemble des couples $(i, \varphi_i(x))$, pour $i \in I$ et $x \in X_i$. Alors Y est dénombrable comme sous-ensemble de l'ensemble dénombrable $I \times \mathbf{N}$, et l'application $(i, y) \mapsto \varphi_i^{-1}(y)$ de Y dans $\cup_{i \in I} X_i$ est surjective, ce qui prouve que $\cup_{i \in I} X_i$ est dénombrable.

- \mathbf{Z} , \mathbf{N}^d , \mathbf{Z}^d , si $d \in \mathbf{N}$, et \mathbf{Q} sont dénombrables.

L'application $(a, b) \mapsto a - b$ est une surjection de $\mathbf{N} \times \mathbf{N}$ sur \mathbf{Z} , et comme $\mathbf{N} \times \mathbf{N}$ est dénombrable, en tant que produit fini d'ensembles dénombrables, il en est de même de \mathbf{Z} . Les ensembles \mathbf{N}^d , \mathbf{Z}^d sont dénombrables puisque ce sont des produits finis d'ensembles dénombrables. Enfin, $(a, b) \mapsto \frac{a}{b}$ induit une surjection de $\mathbf{Z} \times (\mathbf{Z} - \{0\})$ sur \mathbf{Q} qui, de ce fait est dénombrable, \mathbf{Z} et $\mathbf{Z} - \{0\}$ l'étant.

- \mathbf{R} et l'ensemble $\{0, 1\}^{\mathbf{N}}$ des suites à valeurs dans $\{0, 1\}$ ne sont pas dénombrables.

Supposons que $\{0, 1\}^{\mathbf{N}}$ est dénombrable. Il existe donc une bijection $n \mapsto x_n$ de \mathbf{N} sur $\{0, 1\}^{\mathbf{N}}$. Chaque x_n est une suite $x_n = (x_{n,k})_{k \in \mathbf{N}}$, où $x_{n,k} \in \{0, 1\}$, ce qui permet de considérer la suite $y = (y_k)_{k \in \mathbf{N}}$, où $y_k = 1 - x_{k,k}$. Par construction, la suite y a sa n -ième valeur distincte de celle de x_n , pour tout n , et on a donc $y \neq x_n$, quel que soit $n \in \mathbf{N}$, ce qui est en contradiction avec l'hypothèse selon laquelle $n \mapsto x_n$ est surjective; c'est donc que $\{0, 1\}^{\mathbf{N}}$ n'est pas dénombrable. Cet argument est l'*argument diagonal* de Cantor (1891).

Pour démontrer que \mathbf{R} n'est pas dénombrable, il suffit de constater que si X désigne le sous-ensemble de $[0, 1]$ des nombres dont le développement décimal ne comportent que des 0 et des 1, alors X est en bijection avec $\{0, 1\}^{\mathbf{N}}$, et donc n'est pas dénombrable. Il en est a fortiori de même de \mathbf{R} , qui contient X .

Exercice 1.4. — Montrer que l'ensemble $\mathcal{P}(\mathbf{N})$ des parties de \mathbf{N} n'est pas dénombrable, mais que l'ensemble des parties finies de \mathbf{N} est dénombrable.

Exercice 1.5. — On rappelle que $x \in \mathbf{C}$ est *algébrique* s'il existe $P \in \mathbf{Q}[X]$ non nul tel que $P(x) = 0$, et que $x \in \mathbf{C}$ est *transcendant* s'il n'est pas algébrique. Montrer que l'ensemble $\overline{\mathbf{Q}}$ des nombres algébriques est dénombrable. En déduire qu'il existe des nombres transcendants.

Exercice 1.6. — Soit $(B_j)_{j \in I}$ une famille de disques ouverts non vides de \mathbf{C} . Montrer que si les B_j sont deux à deux disjoints, alors I est dénombrable.

Exercice 1.7. — Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ une fonction croissante.

(i) Montrer que f admet une limite à droite et une limite à gauche en tout point et que, si $x_0 \in \mathbf{R}$, alors $f(x_0^+) = \inf_{x > x_0} f(x)$ et $f(x_0^-) = \sup_{x < x_0} f(x)$; en déduire que $f(x_0^-) \leq f(x_0) \leq f(x_0^+)$. A quelle condition f est-elle continue en x_0 ?

(ii) Montrer que, si $x_0 < x_1$, alors $f(x_0^+) \leq f(x_1^-)$.

(iii) Montrer que l'ensemble D des points où f est discontinue est dénombrable.

Exercice 1.8. — (difficile) Un « huit » est la réunion de deux cercles dans le plan, de même rayon (non nul), tangents en un point. Montrer que l'on peut mettre dans le plan au plus un nombre dénombrable de huit deux à deux disjoints.

Exercice 1.9. — (difficile) Un « tripode » est la figure formée de trois segments $[G, A]$, $[G, B]$ et $[G, C]$, où A, B, C sont les sommets d'un triangle équilatéral (non réduit à un point) et G est le centre de gravité du triangle. Montrer que l'on peut mettre dans le plan au plus un nombre dénombrable de tripodes deux à deux disjoints.

2. Produits, sommes et quotients

2.1. Produits et sommes

2.1.1. Produits et sommes directes de groupes commutatifs

Si $(A_i)_{i \in I}$ est une famille de groupes (de lois notées multiplicativement), on munit leur produit $\prod_{i \in I} A_i$ d'une structure de groupe, en faisant le produit composante par composante [i.e en posant $(x_i)_{i \in I} (y_i)_{i \in I} = (x_i y_i)_{i \in I}$]. L'élément neutre est alors $(e_i)_{i \in I}$, si e_i désigne l'élément neutre de A_i , et l'inverse de $(x_i)_{i \in I}$ est $(x_i^{-1})_{i \in I}$. On dispose, pour tout i , d'une surjection naturelle $p_i : \prod_{i \in I} A_i \rightarrow A_i$ envoyant $(x_i)_{i \in I}$ sur x_i , qui est un morphisme de groupes de manière évidente.

• Le produit vérifie la *propriété universelle* suivante : si B est un groupe, et si $f_i : B \rightarrow A_i$ est un morphisme de groupes pour tout $i \in I$, il existe un unique morphisme de groupes $f : B \rightarrow \prod_{i \in I} A_i$ tel que $p_i \circ f = f_i$, quel que soit $i \in I$.

On doit poser $f(x) = (f_i(x))_{i \in I}$. Il est alors évident que f est un morphisme de groupes, et que l'on a $p_i \circ f = f_i$, quel que soit $i \in I$.

Si $(A_i)_{i \in I}$ est une famille de groupes commutatifs (de loi notée additivement), on définit leur *somme directe* $\oplus_{i \in I} A_i$ comme le sous-ensemble du produit $\prod_{i \in I} A_i$ des $(x_i)_{i \in I}$ vérifiant $x_i = 0$ pour presque tout i (i.e. à l'exception d'un nombre fini de i). On dispose alors, pour tout i , d'une injection naturelle $\iota_i : A_i \rightarrow \oplus_{i \in I} A_i$, envoyant $a \in A_i$ sur $(x_i)_{i \in I}$, avec $x_i = a$ et $x_j = 0$, si $j \neq i$.

• Si I est fini, la somme directe est égale au produit, mais pas si I est infini⁽⁸⁾.

• La somme directe vérifie la *propriété universelle* suivante : si B est un groupe commutatif, et si $f_i : A_i \rightarrow B$ est un morphisme de groupes pour tout $i \in I$, il existe un unique morphisme de groupes $f : \oplus_{i \in I} A_i \rightarrow B$ tel que $f \circ \iota_i = f_i$, quel que soit $i \in I$.

On doit poser $f((x_i)_{i \in I}) = \sum_{i \in I} f_i(x_i)$, ce qui a un sens car la somme est en fait finie. Il est alors évident que f est un morphisme de groupes, et que l'on a $f \circ \iota_i = f_i$, quel que soit $i \in I$.

⁽⁸⁾Le lecteur désireux de comprendre plus en profondeur la différence entre les notions de produit et de somme est invité à se munir d'une loupe et à consulter l'alinéa 2.1.3.

- Si A est un groupe commutatif, et si $(A_i)_{i \in I}$ est une famille de sous-groupe de A , on dispose d'un morphisme de groupes naturel de $\bigoplus_{i \in I} A_i$ dans A induite par l'identité sur A_i , pour tout i . On note $\sum_{i \in I} A_i$ l'image de ce morphisme; c'est le sous-groupe de A engendré par les A_i . On dit que les A_i sont *en somme directe*, si l'application naturelle de $\bigoplus_{i \in I} A_i$ dans A est un isomorphisme. De manière plus concrète, les A_i sont en somme directe, si tout élément x de A peut s'écrire de manière unique sous la forme $x = \sum_{i \in I} x_i$, avec $x_i \in A_i$ pour tout i , et $x_i = 0$ pour presque tout i .
- Si B et C sont deux sous-groupes d'un groupe commutatif A , alors B et C sont en somme directe, si et seulement si $B \cap C = \{0\}$ et tout élément de A est somme d'un élément de B et d'un élément de C .

2.1.2. Le cas des espaces vectoriels

Soit K un corps (sous-entendu commutatif). Un K -espace vectoriel est en particulier un groupe commutatif, et tout ce que l'on a dit à l'alinéa précédent s'applique. On dispose en plus d'une action de K , définie par $\lambda(x_i)_{i \in I} = (\lambda x_i)_{i \in I}$, si $\lambda \in K$, sur le produit et la somme directe, ce qui en fait des K -espaces vectoriels. Si $(E_i)_{i \in I}$ est une famille de K -espaces vectoriels, ces objets vérifient alors les propriétés universelles suivantes.

- Si F est un K -espace vectoriel, et si $f_i : F \rightarrow E_i$ est une application linéaire pour tout $i \in I$, il existe une unique application linéaire $f : F \rightarrow \prod_{i \in I} E_i$ telle que $p_i \circ f = f_i$, quel que soit $i \in I$.
- si F est un K -espace vectoriel commutatif, et si $f_i : E_i \rightarrow F$ est une application linéaire pour tout $i \in I$, il existe une unique application linéaire $f : \bigoplus_{i \in I} E_i \rightarrow F$ telle que $f \circ \iota_i = f_i$, quel que soit $i \in I$.
- Si I est fini, les espaces vectoriels $\bigoplus_{i \in I} E_i$ et $\prod_{i \in I} E_i$ sont isomorphes.

2.1.3. Produit et somme dans une catégorie

On définit la notion de *catégorie* pour mettre sous un même chapeau les objets ayant les mêmes propriétés. Le lecteur connaît déjà, sans en avoir forcément conscience, un certain nombre de ces catégories (celle des ensembles, celle des groupes ou celle des espaces vectoriels sur \mathbf{R} ou \mathbf{C} par exemple; il y en a beaucoup d'autres comme celle des espaces topologiques, des espaces de Banach...).

Une catégorie C est une collection d'objets (les objets de la catégorie), et de flèches entre ces objets (les morphismes de la catégorie) : si X et Y sont deux objets de C , on note $\text{Hom}_C(X, Y)$ les morphismes de X vers Y dans la catégorie C . On impose que l'identité id_X soit un morphisme de X dans X , et que l'on puisse composer les morphismes : si X, Y et Z sont trois objets de C , on dispose d'une application $(f, g) \mapsto f \circ g$ de $\text{Hom}_C(X, Y) \times \text{Hom}_C(Y, Z) \rightarrow \text{Hom}_C(X, Z)$ vérifiant les propriétés évidentes :

$$f \circ \text{id}_X = f, \text{id}_Y \circ f = f \text{ et } (f \circ g) \circ h = f \circ (g \circ h).$$

Les exemples les plus simples de catégories sont les suivants :

- La catégorie des ensembles; les morphismes de X dans Y sont les applications Y^X de X dans Y .
- La catégorie des groupes; les morphismes sont les morphismes de groupes.
- La catégorie des groupes commutatifs; les morphismes sont les morphismes de groupes.
- La catégorie des anneaux commutatifs; les morphismes sont les morphismes d'anneaux.
- La catégorie des K -espaces vectoriels, K un corps; les morphismes sont les applications K -linéaires.
- La catégorie des espaces topologiques; les morphismes sont les applications continues.
- La catégorie des espaces métriques; les morphismes sont les applications continues.
- La catégorie des \mathbf{K} -espaces de Banach, $\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$; les morphismes sont les applications \mathbf{K} -linéaires continues.

Dans une catégorie, on définit les notions de *produit* et *somme* par les propriétés universelles suivantes (la propriété universelle implique l'unicité d'un tel objet, mais pas son existence qui doit se démontrer cas par cas).

Si C est une catégorie et les $(X_i)_{i \in I}$ sont des objets de C , le produit $X = \prod_{i \in I} X_i$ des X_i est un objet de C muni de morphismes $p_i \in \text{Hom}_C(X, X_i)$, pour $i \in I$, tel que, si Y est n'importe quel objet de C , et si $f_i \in \text{Hom}_C(Y, X_i)$, pour tout $i \in I$, alors il existe $f \in \text{Hom}_C(Y, X)$ unique, tel que $p_i \circ f = f_i$ pour tout $i \in I$.

La somme $X = \coprod_{i \in I} X_i$ des X_i est un objet de C muni de morphismes $\iota_i \in \text{Hom}_C(X_i, X)$, pour $i \in I$, tel que, si Y est n'importe quel objet de C , et si $f_i \in \text{Hom}_C(X_i, Y)$, pour tout $i \in I$, alors il existe $f \in \text{Hom}_C(X, Y)$ unique, tel que $f \circ \iota_i = f_i$ pour tout $i \in I$.

Montrons par exemple l'unicité du produit. Si X (resp. X') muni des $p_i : X \rightarrow X_i$ (resp. des $p'_i : X' \rightarrow X_i$) est un produit des X_i , alors en particulier, il existe $f : X' \rightarrow X$ unique tel que $p_i \circ f = p'_i$ pour tout i , et il existe $g : X \rightarrow X'$ unique tel que $p'_i \circ g = p_i$ pour tout i . Alors $f \circ g : X \rightarrow X$ vérifie $p_i \circ (f \circ g) = p_i$ pour tout i , ce qui implique que $f \circ g = \text{id}_X$ puisque id_X vérifie la même propriété, et que par hypothèse, il n'y a qu'un seul morphisme de X dans X ayant cette propriété. Pour la même raison, on a $g \circ f = \text{id}_{X'}$, ce qui prouve que X et X' sont isomorphes (à isomorphisme unique près puisque f et g étaient uniques). Cette démonstration s'étend à tout objet solution d'un problème universel.

On dit qu'une catégorie admet des produits (resp. des sommes), si tout couple (et donc toute famille finie) d'objets de la catégorie admet un produit (resp. une somme). Toutes les catégories ci-dessus admettent des produits, car on peut munir le produit ensembliste de deux objets des structures additionnelles demandées. Elles admettent aussi toutes une somme, mais celle-ci peut prendre des formes assez variées.

- Dans la catégorie des ensembles, la somme d'une famille $(X_i)_{i \in I}$ d'ensemble est leur réunion disjointe $\cup_{i \in I} (\{i\} \times X_i)$.
- Dans la catégorie des K -espaces vectoriels, ou dans celle des groupes commutatifs, la somme est la somme directe, et la somme d'un nombre fini d'objets est isomorphe à leur produit comme on l'a vu ci-dessus.
- Dans la catégorie des groupes, la somme de deux groupes A et B est leur produit libre $A \star B$: les éléments de $A \star B$ sont les mots finis composés d'éléments de A et B modulo la relation d'équivalence selon laquelle on peut remplacer toute lettre x dans un mot par deux lettres x_1, x_2 appartenant au même groupe si $x_1 x_2 = x$, et réciproquement, on peut remplacer deux lettres consécutives appartenant au même groupe par leur produit. La somme de deux groupes commutatifs n'est donc pas la même dans la catégorie des groupes que dans celle des groupes commutatif. Par exemple, on a $(\mathbf{Z}/2\mathbf{Z}) \oplus (\mathbf{Z}/3\mathbf{Z}) = (\mathbf{Z}/6\mathbf{Z})$, alors que $(\mathbf{Z}/2\mathbf{Z}) \star (\mathbf{Z}/3\mathbf{Z})$ est un groupe infini, isomorphe au groupe $\mathbf{PSL}_2(\mathbf{Z})$, quotient de $\mathbf{SL}_2(\mathbf{Z})$ par son centre $\{\pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\}$.

2.2. Relations d'équivalence

2.2.1. Relations d'équivalence et partitions

Si E est un ensemble, une *partition* de E est une famille de sous-ensembles non vides de E , deux à deux disjoints, dont la réunion est E .

Par exemple, $\{\mathbf{R}_+^*, \mathbf{R}_-^*, \{0\}\}$ est une partition de \mathbf{R} . Si $D \in \mathbf{N} - \{0\}$, les $r + D\mathbf{Z}$, pour $r \in \{0, \dots, D-1\}$ forment une partition de \mathbf{Z} .

Une *relation* R sur E est un sous-ensemble de $E \times E$. Si $(x, y) \in E \times E$, on écrit souvent xRy pour signifier $(x, y) \in R$.

Une relation R sur E est une *relation d'équivalence* si elle est *réflexive* (xRx quel que soit $x \in E$), *symétrique* (xRy implique yRx) et *transitive* (xRy et yRz impliquent xRz).

Si R est une relation d'équivalence sur E , et si $x \in E$, la *classe d'équivalence* de x est l'ensemble $C_x = \{y \in E, yRx\}$. Un sous-ensemble C de E est une *classe d'équivalence* (pour R), s'il existe $x \in E$ tel que $C = C_x$. Si $x, y \in E$, alors $C_x \cap C_y \neq \emptyset$ si et seulement si xRy , et on a alors $C_x = C_y$. Les classes d'équivalence forment donc une partition de E .

Réciproquement, si les $(C_i)_{i \in I}$ forment une partition de E , alors la relation R définie par xRy si et seulement si il existe $i \in I$ tel que $\{x, y\} \subset C_i$ est une relation d'équivalence dont les classes d'équivalence sont les C_i . En d'autres termes, il revient au même de munir un ensemble d'une relation d'équivalence ou de faire une partition de cet ensemble.

Par exemple, la partition de \mathbf{R} ci-dessus correspond à la relation d'équivalence « $x \sim y$ si et seulement si x et y ont même signe » ; celle de \mathbf{Z} correspond à la relation d'équivalence « $a \sim b$ si et seulement si a et b ont même reste dans la division euclidienne par D ».

2.2.2. Passage au quotient par une relation d'équivalence. — Si R est une relation d'équivalence sur E , on définit *le quotient* E/R de E par la relation d'équivalence R comme l'ensemble des classes d'équivalence. On dispose d'une application naturelle de E dans E/R , à savoir l'application qui à x associe sa classe d'équivalence (souvent notée \bar{x}) ; cette application est surjective par construction de E/R . Un sous-ensemble S de E est *un système de représentants* de E/R , s'il contient un et un seul élément de chaque classe d'équivalence. Autrement dit, $S \subset E$ est un système de représentants de E/R si et seulement si l'application naturelle de E dans E/R induit une bijection de S sur E/R .

Cette manière de définir de nouveaux objets en passant au quotient par une relation d'équivalence est une des plus universelles qui soit⁽⁹⁾. Pour le petit enfant, le nombre 5 est la classe d'équivalence des ensembles pouvant être mis en bijection avec l'ensemble {un, deux, trois, quatre, cinq} (ce n'est pas une raison pour le lui définir de cette manière...). Pour le commun des mortels, un nombre réel est un nombre avec une infinité de chiffres derrière la virgule, et comme certains nombres ont deux écritures, il faut passer au quotient...

En général, on aime bien que E/R hérite des propriétés que pouvait avoir E (i.e. on aime bien que les propriétés de E *passent au quotient*), ce qui impose des contraintes aux relations d'équivalence que l'on peut considérer. Par exemple, une fonction $f : E \rightarrow X$ passe au quotient si et seulement si on a $f(x) = f(y)$ pour tout couple d'éléments de E vérifiant xRy (si c'est le cas, on définit $\bar{f} : E/R \rightarrow X$ par $\bar{f}(z) = f(x)$ pour n'importe quel élément x de E ayant pour image z dans E/R). Si $\pi : E \rightarrow E/R$ est l'application naturelle, on a $f = \bar{f} \circ \pi$; on dit que f *se factorise à travers* E/R ou que f se factorise à travers π , ce qui est une terminologie assez parlante puisqu'elle signifie que l'équation $f = g \circ \pi$ a une solution $g = \bar{f}$.

⁽⁹⁾L'expérience montre que les premiers passages au quotient que l'on rencontre sont un peu traumatisants, mais on finit par s'y faire... Il fut un temps pas si lointain, où l'on définissait \mathbf{Z} comme le quotient de $\mathbf{N} \times \mathbf{N}$ par la relation d'équivalence $(a, b) \sim (a', b')$ si et seulement si $a + b' = a' + b$, l'idée étant que (a, b) représente l'entier relatif $a - b$. Au bout de 3 semaines, on avait enfin le droit d'écrire $2 - 3 + 5 - 7 = -3$, ce que n'importe qui ayant regardé un thermomètre comprend très bien. Pour en arriver là, il avait fallu passer par $(2, 0) + (0, 3) + (5, 0) + (0, 7) = (7, 10) = (0, 3)$, puis par $(+2) + (-3) + (+5) + (-7) = (-3)$. On achevait de traumatiser les élèves (et leur parents) en définissant, en classe de 4^{ème}, un vecteur comme une classe d'équipollence de bipoints (un bipoint (i.e. un couple de points) (A, B) est *équipollent* à (C, D) , si (A, B, D, C) est un parallélogramme). Dans une période plus récente, les aléas de la conjoncture ayant provoqué un tarissement des vocations de professeurs de mathématiques, on s'est retrouvé avec une pénurie que l'on a traitée en diminuant l'horaire de mathématiques dans l'enseignement, et on en a profité pour jeter allègrement à la poubelle toutes ces horribles mathématiques modernes...

2.3. L'anneau $\mathbf{Z}/D\mathbf{Z}$ des entiers relatifs modulo D

Dans tout ce qui suit, D est un entier ≥ 1 . On note $D\mathbf{Z}$ l'ensemble des multiples de D . On définit une relation *de congruence modulo D* sur \mathbf{Z} , en disant que a est congru à b modulo D (ou modulo $D\mathbf{Z}$), ce qui se note $a \equiv b [D]$ ou $a \equiv b \pmod{D}$, si $b - a \in D\mathbf{Z}$.

- La relation de congruence modulo D est une relation d'équivalence sur \mathbf{Z} . On note $\mathbf{Z}/D\mathbf{Z}$ l'ensemble des classes d'équivalence. L'image d'un entier dans $\mathbf{Z}/D\mathbf{Z}$ est sa *réduction modulo D* .

Cette relation est réflexive car 0 est un multiple de D , symétrique car si $b - a$ est un multiple de D , il en est de même de $a - b$, et transitive car si $b - a$ et $c - b$ sont des multiples de D , il en est de même $c - a = (c - b) + (b - a)$.

- Un système naturel de représentants de $\mathbf{Z}/D\mathbf{Z}$ dans \mathbf{Z} est l'ensemble $\{0, 1, \dots, D - 1\}$; en particulier, $\mathbf{Z}/D\mathbf{Z}$ est de cardinal D .

Si $a, b \in \{0, 1, \dots, D - 1\}$ sont distincts, et si $b > a$, alors $1 \leq b - a \leq D - 1$. En particulier $b - a$ n'est pas un multiple de D , ce qui prouve que b et a sont dans des classes distinctes modulo D , et donc que l'application naturelle de $\{0, 1, \dots, D - 1\}$ dans $\mathbf{Z}/D\mathbf{Z}$ est injective. Par ailleurs, si $a \in \mathbf{Z}$ est quelconque, et si $r \in \{0, 1, \dots, D - 1\}$ est le reste de la division de a par D , alors $a - r$ est un multiple de D et a est dans la même classe que r modulo D ; l'application naturelle de $\{0, 1, \dots, D - 1\}$ dans $\mathbf{Z}/D\mathbf{Z}$ est donc surjective.

- L'addition et la multiplication sur \mathbf{Z} passent au quotient, et $\mathbf{Z}/D\mathbf{Z}$ muni des lois d'addition et multiplication ainsi définies est un anneau commutatif⁽¹⁰⁾.

Si $x - x'$ et $y - y'$ sont divisibles par D , alors $(x + y) - (x' + y') = (x - x') + (y - y')$ et $xy - x'y' = x(y - y') + y'(x - x')$ sont divisibles par D , ce qui prouve que le résultat modulo D de l'addition et la multiplication de deux entiers ne dépend que de leurs réductions modulo D ; en d'autres termes, l'addition et la multiplication passent au quotient. Par ailleurs, les identités à vérifier pour prouver que $\mathbf{Z}/D\mathbf{Z}$ est un anneau sont déjà vraies dans l'anneau \mathbf{Z} ; elles le sont donc, a fortiori, dans $\mathbf{Z}/D\mathbf{Z}$.

- $a \in \mathbf{Z}$ est inversible (pour la multiplication) dans $\mathbf{Z}/D\mathbf{Z}$ si et seulement si a est premier à D . On note $(\mathbf{Z}/D\mathbf{Z})^*$ l'ensemble des éléments inversibles; c'est un groupe dont le cardinal est traditionnellement noté $\varphi(D)$, et la fonction φ est la *fonction indicatrice d'Euler*.

Si a est premier à D , il existe, d'après le théorème de Bézout, $u, v \in \mathbf{Z}$ tels que $au + Dv = 1$, ce qui prouve que a est inversible dans $\mathbf{Z}/D\mathbf{Z}$, d'inverse u . Réciproquement, si $ab = 1$ dans $\mathbf{Z}/D\mathbf{Z}$, cela signifie que $ab - 1$ est divisible par D , et donc qu'il existe $v \in \mathbf{Z}$ tel que $ab + Dv = 1$; d'après le théorème de Bézout, cela implique que a et D sont premiers entre eux, ce qui permet de conclure.

- D est premier si et seulement si $\mathbf{Z}/D\mathbf{Z}$ est un corps.

⁽¹⁰⁾La manière qui est probablement la plus efficace pour penser à l'anneau $\mathbf{Z}/D\mathbf{Z}$ est de le voir comme étant l'anneau \mathbf{Z} auquel on a rajouté la relation $D = 0$; on fait donc les additions et les multiplications comme si on était dans \mathbf{Z} , mais on se permet d'enlever le multiple de D que l'on veut au résultat. Par exemple, dans $\mathbf{Z}/21\mathbf{Z}$, on a $6 \times 14 = 4 \times 21 = 0$ et $4 \times 16 = 1 + 3 \times 21 = 1$, ce qui montre que 6 et 14 sont des diviseurs de 0 , alors que 4 est inversible d'inverse 16 .

L'anneau $\{0\}$ n'est pas un corps (si K est corps, $K - \{0\}$ forme un groupe pour la multiplication et en particulier est non vide) et 1 n'est pas un nombre premier ; on peut donc supposer $D \geq 2$.

Si $D \geq 2$ n'est pas premier, on peut le factoriser sous la forme $D = ab$, avec $a \in \{2, \dots, D-1\}$ et $b \in \{2, \dots, D-1\}$. Donc a et b ne sont pas nuls dans $\mathbf{Z}/D\mathbf{Z}$ alors que $ab = D = 0$ dans $\mathbf{Z}/D\mathbf{Z}$; l'anneau $\mathbf{Z}/D\mathbf{Z}$ admet donc des diviseurs de 0 et n'est pas un corps.

Si D est premier, et si a n'est pas divisible par D , alors a est premier à D et donc inversible dans $\mathbf{Z}/D\mathbf{Z}$ d'après le point précédent. Ceci permet de conclure.

Un nombre premier a tendance à être noté p , et si on veut insister sur le fait que $\mathbf{Z}/p\mathbf{Z}$ est un corps, on le note \mathbf{F}_p . Par exemple, on parlera d'espaces vectoriels sur \mathbf{F}_2 au lieu d'espaces vectoriels sur $\mathbf{Z}/2\mathbf{Z}$ pour parler des objets qui peuplent Internet⁽¹¹⁾ et dans lesquels vivent les codes correcteurs d'erreurs.

- Tout corps de cardinal p est isomorphe à \mathbf{F}_p ; autrement dit \mathbf{F}_p est le corps à p éléments⁽¹²⁾.

Soit K un corps à p élément. On dispose d'un morphisme d'anneaux $f : \mathbf{Z} \rightarrow K$ envoyant 1 sur 1. Ce morphisme d'anneaux est en particulier un morphisme de groupes additifs. Son image est donc un sous-groupe du groupe $(K, +)$, et son cardinal est un diviseur de $|K| = p$, d'après le théorème de Lagrange, et comme cette image a au moins deux éléments, à savoir 0 et 1, c'est K tout entier. On en déduit que f induit un isomorphisme de $\mathbf{Z}/\text{Ker } f$ sur K , et comme $|K| = p$, on a $\text{Ker } f = p\mathbf{Z}$, et donc $K \cong \mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$. Ceci permet de conclure.

- Si D' est un diviseur de D , alors l'application naturelle $\mathbf{Z} \rightarrow (\mathbf{Z}/D'\mathbf{Z})$ se factorise à travers une application naturelle $(\mathbf{Z}/D\mathbf{Z}) \rightarrow (\mathbf{Z}/D'\mathbf{Z})$ qui est un morphisme d'anneaux.

Si D' est un diviseur de D , alors un multiple de D est aussi un multiple de D' . On en déduit que, si $a \equiv b \pmod{D}$, alors $a \equiv b \pmod{D'}$; autrement dit l'application naturelle $\mathbf{Z} \rightarrow (\mathbf{Z}/D'\mathbf{Z})$ se factorise à travers une application naturelle $(\mathbf{Z}/D\mathbf{Z}) \rightarrow (\mathbf{Z}/D'\mathbf{Z})$. On obtient un morphisme d'anneaux car les identités à vérifier sont déjà valables en remontant à \mathbf{Z} .

- Si a et b sont premiers entre eux, l'application naturelle $\mathbf{Z}/ab\mathbf{Z} \rightarrow (\mathbf{Z}/a\mathbf{Z}) \times (\mathbf{Z}/b\mathbf{Z})$ est un isomorphisme d'anneaux qui induit un isomorphisme de groupes de $(\mathbf{Z}/ab\mathbf{Z})^*$ sur $(\mathbf{Z}/a\mathbf{Z})^* \times (\mathbf{Z}/b\mathbf{Z})^*$ (théorème des restes chinois).

L'application naturelle $\mathbf{Z}/ab\mathbf{Z} \rightarrow (\mathbf{Z}/a\mathbf{Z}) \times (\mathbf{Z}/b\mathbf{Z})$ est un morphisme d'anneaux d'après le point précédent. Il est injectif car, si $x \in \mathbf{Z}$ a une réduction modulo ab qui est dans le noyau,

⁽¹¹⁾Internet aime beaucoup $\mathbf{Z}/D\mathbf{Z}$. Non content de faire voyager des milliards de \mathbf{F}_2 -espaces vectoriels, Internet est très gourmand de grands nombres premiers, par exemple pour le système RSA de sécurité à clé publique. Ce système et la fabrication de grands nombres premiers reposent sur l'arithmétique dans $\mathbf{Z}/D\mathbf{Z}$ qui s'avère être nettement plus subtile que ce que l'on pourrait attendre d'un objet aussi petit. On peut saluer la clairvoyance de la commission ayant accouché des programmes actuels, qui a fait disparaître $\mathbf{Z}/D\mathbf{Z}$ des programmes de la filière PC au moment même où Internet prenait son envol...

⁽¹²⁾Plus généralement, si q est une puissance d'un nombre premier, il y a, à isomorphisme près, un unique corps à q élément, et ce corps est noté \mathbf{F}_q . On a beaucoup fantasmé ces dernières années autour du corps \mathbf{F}_1 « à 1 élément » dont on voit la trace dans plusieurs phénomènes sans comprendre quel genre d'objet cela pourrait bien être (pas l'anneau $\{0\}$ en tout cas). Certains y voient la clef d'une démonstration de l'hypothèse de Riemann.

c'est que x est divisible par a et par b , et donc par ab puisqu'on a supposé a et b premiers entre eux; autrement dit, le noyau est réduit à 0. Comme les deux ensembles considérés ont même cardinal ab , une application injective est aussi bijective, ce qui montre que $\mathbf{Z}/ab\mathbf{Z} \rightarrow (\mathbf{Z}/a\mathbf{Z}) \times (\mathbf{Z}/b\mathbf{Z})$ est un isomorphisme. On conclut en remarquant que si A et B sont deux anneaux, alors $(A \times B)^* = A^* \times B^*$.

En fait, on peut décrire explicitement l'isomorphisme inverse. Comme a et b sont premiers entre eux, il existe $u, v \in \mathbf{Z}$ tels que $1 = au + bv$. Si $(x, y) \in (\mathbf{Z}/a\mathbf{Z}) \times (\mathbf{Z}/b\mathbf{Z})$, et si $\tilde{x}, \tilde{y} \in \mathbf{Z}$ ont pour image x et y dans $\mathbf{Z}/a\mathbf{Z}$ et $\mathbf{Z}/b\mathbf{Z}$ respectivement, alors l'image de $bv\tilde{x} + au\tilde{y}$ dans $\mathbf{Z}/ab\mathbf{Z}$ ne dépend pas des choix de \tilde{x} et \tilde{y} et s'envoie sur (x, y) dans $(\mathbf{Z}/a\mathbf{Z}) \times (\mathbf{Z}/b\mathbf{Z})$, comme le montre un petit calcul immédiat. On remarque que $x \mapsto bv\tilde{x}$ induit un isomorphisme de $\mathbf{Z}/a\mathbf{Z}$ sur le sous-groupe $b\mathbf{Z}/ab\mathbf{Z}$ de $\mathbf{Z}/ab\mathbf{Z}$ et que $y \mapsto au\tilde{y}$ induit un isomorphisme de $\mathbf{Z}/b\mathbf{Z}$ sur le sous-groupe $a\mathbf{Z}/ab\mathbf{Z}$ de $\mathbf{Z}/ab\mathbf{Z}$. On en déduit le résultat suivant :

- Si a et b sont premiers entre eux, alors $\mathbf{Z}/ab\mathbf{Z}$ est la somme directe de ses sous-groupes $b\mathbf{Z}/ab\mathbf{Z}$ et $a\mathbf{Z}/ab\mathbf{Z}$; de plus, on a des isomorphismes de groupes additifs $b\mathbf{Z}/ab\mathbf{Z} \cong \mathbf{Z}/a\mathbf{Z}$ et $a\mathbf{Z}/ab\mathbf{Z} \cong \mathbf{Z}/b\mathbf{Z}$, et donc $\mathbf{Z}/ab\mathbf{Z}$ est isomorphe à $(\mathbf{Z}/a\mathbf{Z}) \oplus (\mathbf{Z}/b\mathbf{Z})$, comme groupe additif.

Exercice 2.1. — Montrer que si $a \neq 0$ et $b \neq 0$ ne sont pas premiers entre eux, les groupes additifs $\mathbf{Z}/ab\mathbf{Z}$ et $(\mathbf{Z}/a\mathbf{Z}) \oplus (\mathbf{Z}/b\mathbf{Z})$ ne sont pas isomorphes.

Exercice 2.2. — Résoudre les équations $4x + 3 = 0$, $14x + 2 = 0$ et $14x + 7 = 0$ dans $\mathbf{Z}/21\mathbf{Z}$.

Exercice 2.3. — Résoudre l'équation $x^2 + x + 1 = 0$ dans $\mathbf{Z}/91\mathbf{Z}$. (Comme 91 est relativement petit⁽¹³⁾, on peut tester chaque élément de $\mathbf{Z}/91\mathbf{Z}$ et voir lesquels conviennent, mais c'est un peu fastidieux...)

Exercice 2.4. — (i) Soit $p \in \mathcal{P}$. Montrer que si $p \neq 3$, et si l'équation $x^2 + x + 1 = 0$ a une solution dans \mathbf{F}_p , alors elle en a deux.

(ii) (difficile) Montrer qu'il existe une infinité de nombres premiers p tels que l'équation $x^2 + x + 1 = 0$ ait deux solutions dans \mathbf{F}_p .

(iii) En déduire que quel que soit $M > 0$, il existe $D \in \mathbf{N}$ tel que $x^2 + x + 1 = 0$ ait plus de M solutions dans $\mathbf{Z}/D\mathbf{Z}$.

⁽¹³⁾Essayer de résoudre de la même manière l'équation $x^2 = 5$ dans $\mathbf{Z}/D\mathbf{Z}$, avec $D = 2^{2802} - 2^{521} - 2^{2281} + 1$ est voué à l'échec, même avec l'aide d'un ordinateur. Par contre, en partant de $D = (2^{521} - 1)(2^{2281} - 1)$, si on sait que $p_1 = 2^{521} - 1$ et $p_2 = 2^{2281} - 1$ sont premiers (ce sont des *nombres premiers de Mersenne* découverts par Robinson en 1952), alors on peut sans trop d'effort calculer le nombre de solutions de l'équation et, avec l'aide d'un ordinateur et d'algorithmes astucieux, calculer explicitement ces solutions. Le calcul du nombre de solutions repose sur la *loi de réciprocité quadratique* conjecturée par Euler en 1783 et démontrée par Gauss en 1801. Si p est un nombre premier, et si $a \in \mathbf{Z}$ n'est pas divisible par p , on pose $(\frac{a}{p}) = 1$, si a est un carré modulo p (i.e. si l'équation $x^2 = a$ a des solutions dans \mathbf{F}_p) et $(\frac{a}{p}) = -1$ si a n'est pas un carré modulo p (si l'équation $x^2 = a$ n'a pas de solutions dans \mathbf{F}_p). La loi de réciprocité quadratique s'énonce alors ainsi : *si p et q sont deux nombres premiers impairs distincts, alors $(\frac{q}{p}) = (-1)^{(p-1)(q-1)/4} (\frac{p}{q})$* . On applique ce qui précède à $p = p_1$ et $q = 5$. Comme $p_1 = 2^{521} - 1 = 2^{4 \cdot 130 + 1} - 1 = 2 \cdot (2^4)^{130} - 1 = 2 - 1 = 1$ dans \mathbf{F}_5 , on a $(\frac{p_1}{5}) = 1$ et donc $(\frac{5}{p_1}) = 1$, d'après la loi de réciprocité quadratique. On en déduit que l'équation $x^2 = 5$ a deux solutions dans \mathbf{F}_{p_1} . Pour la même raison, elle en a aussi 2 dans \mathbf{F}_{p_2} et donc 4 dans $\mathbf{Z}/D\mathbf{Z}$.

Exercice 2.5. — Montrer que si $p \in \mathcal{P}$, alors $\mathbf{Z}/p^n\mathbf{Z}$ a $p^n - p^{n-1}$ éléments inversibles. En déduire que, si $D \geq 2$, alors $\varphi(D) = D \cdot \prod_{p|D} (1 - \frac{1}{p})$, où φ est la fonction indicatrice d'Euler.

2.4. Quotients d'espaces vectoriels

Soit E un espace vectoriel sur un corps K , soit R une relation d'équivalence sur E , et soit $F \subset E$ la classe d'équivalence de 0 . Pour que la structure d'espace vectoriel de E passe au quotient, on doit en particulier avoir $\lambda x \in F$ si $\lambda \in K$ et $x \in F$ (puisque $\lambda 0 = 0$ dans E/R) et $x + y \in F$ si $x, y \in F$ (puisque $0 + 0 = 0$ dans E/R); en d'autres termes, F doit être un sous-espace vectoriel de E . De plus, comme $a + 0 = a$ dans E/R , les classes d'équivalence doivent être de la forme $a + F$.

Réciproquement, si F est un sous-espace vectoriel de E , la relation \sim_F , définie sur E par $x \sim_F y$ si et seulement si $x - y \in F$, est une relation d'équivalence. Le quotient E/\sim_F est traditionnellement noté E/F . Comme « $x - y \in F$ » \Rightarrow « $\lambda x - \lambda y \in F$ », et comme « $x - y \in F$ et $x' - y' \in F$ » \Rightarrow « $(x + x') - (y + y') \in F$ », la structure d'espace vectoriel sur E passe au quotient.

Si $F' \subset E$ est un sous-espace vectoriel supplémentaire de F , les classes d'équivalence pour \sim_F sont les $a + F$, pour $a \in F'$, et l'application naturelle $F' \rightarrow E/F$ est un isomorphisme d'espaces vectoriels. En d'autres termes, *dans le cas des espaces vectoriels*, un quotient est toujours isomorphe à un sous-objet, mais il est très souvent nocif de remplacer, dans les raisonnements, un quotient par un sous-objet qui lui est isomorphe. Par exemple, si F est un sous-espace vectoriel d'un espace vectoriel E , le dual F^* de F (i.e. l'ensemble des formes linéaires de F dans K) est naturellement un quotient du dual E^* de E (on peut restreindre à F une forme linéaire sur E , et F^* est le quotient de E^* par le sous-espace des formes linéaires identiquement nulles sur F), et n'est pas, en général, un sous-espace de E^* , de manière naturelle.

Dans le cas général, si $u : E \rightarrow E'$ est une application linéaire, alors u se factorise à travers $E/\text{Ker } u$ (ce qui signifie qu'il existe $\bar{u} : E/\text{Ker } u \rightarrow E'$ telle que $u = \bar{u} \circ \pi$, où $\pi : E \rightarrow E/\text{Ker } u$ est l'application associant à $x \in E$ sa classe d'équivalence \bar{x} modulo $\text{Ker } u$), et l'application induite $\bar{u} : E/\text{Ker } u \rightarrow \text{Im } u$ est un isomorphisme d'espaces vectoriels.

2.5. Anneaux quotients

Soit A un anneau commutatif (un anneau a toujours un élément unité 1), soit R une relation d'équivalence sur A , et soit $I \subset A$ la classe d'équivalence de 0 . Pour que la structure d'anneau de A passe au quotient, on doit en particulier avoir $\lambda x \in I$ si $\lambda \in A$ et $x \in I$ (puisque $\lambda 0 = 0$ dans A/R) et $x + y \in I$ si $x, y \in I$ (puisque $0 + 0 = 0$ dans A/R); un sous-ensemble de A vérifiant ces deux propriétés est un *idéal* de A . De plus, comme $a + 0 = a$ dans A/R , les classes d'équivalence doivent être de la forme $a + I$.

Réciproquement, si I est un idéal de A , la relation \sim_I , définie sur A par $x \sim_I y$ si et seulement si $x - y \in I$, est une relation d'équivalence. Le quotient A/\sim_I est traditionnellement noté A/I . Comme « $x - y \in I$ et $x' - y' \in I$ » \Rightarrow « $(x + x') - (y + y') \in I$ », et

comme « $x - y \in I$ et $x' - y' \in I$ » \Rightarrow « $xx' - yy' = x(y - y') + y'(x - x') \in I$ » la structure d'anneau sur A passe au quotient.

Contrairement à ce qui se passe dans le cas des espaces vectoriels, l'anneau A/I n'est pas, en général, isomorphe à un sous-anneau de A . Par exemple $\mathbf{Z}/D\mathbf{Z}$ n'est pas isomorphe à un sous-anneau de \mathbf{Z} , ni même à un sous-groupe additif.

Dans le cas général, si $f : A \rightarrow A'$ est un morphisme d'anneaux, alors $\text{Ker } f$ est un idéal de A , f se factorise à travers $A/\text{Ker } f$ et l'application induite $\bar{f} : A/\text{Ker } f \rightarrow \text{Im } f$ est un isomorphisme d'anneaux.

Le lecteur connaît déjà beaucoup d'anneaux définis de cette manière. Par exemple,

- le corps des nombres complexes⁽¹⁴⁾ $\mathbf{C} = \mathbf{R}[X]/(X^2 + 1)$ (prendre le quotient de $\mathbf{R}[X]$ par l'idéal⁽¹⁵⁾ $(X^2 + 1)$ revient à rajouter à \mathbf{R} un élément X vérifiant $X^2 + 1 = 0$, et donc X devient une racine carrée de -1 dans le quotient) ;
- l'anneau $\mathbf{Z}[X]/(10X - 1)$ des nombres décimaux (prendre le quotient de $\mathbf{Z}[X]$ par $(10X - 1)$ revient à rajouter à \mathbf{Z} un élément X vérifiant $10X = 1$, et $a_n X^n + \dots + a_0 \in \mathbf{Z}[X]$ devient le nombre décimal $\frac{a_n}{10^n} + \dots + \frac{a_1}{10} + a_0$) ;
- le corps \mathbf{F}_p , quotient de \mathbf{Z} par l'idéal $p\mathbf{Z}$ (p étant un nombre premier) ;
- l'anneau $\mathbf{Z}/D\mathbf{Z}$ quotient de \mathbf{Z} par l'idéal $D\mathbf{Z}$ (D entier quelconque).

On en rencontre beaucoup d'autres, par exemple les anneaux suivants.

- $\mathbf{Z}[X]/(X^2 + 1)$, anneau des *entiers de Gauss* ; en envoyant X sur i ou $-i$, cet anneau s'identifie au sous-anneau des $\{a + ib, a, b \in \mathbf{Z}\}$ de \mathbf{C} .
- $\mathbf{Z}[X]/(X^3 - 2)$. Il s'identifie à un sous-anneau de \mathbf{C} de trois manières différentes : on peut envoyer X sur $\sqrt[3]{2}$, ou sur $e^{2i\pi/3}\sqrt[3]{2}$ ou sur $e^{4i\pi/3}\sqrt[3]{2}$. Dans le premier cas, l'image est un sous-anneau de \mathbf{R} , dans les autres cas, elle n'est pas incluse dans \mathbf{R} .
- L'anneau $K[\varepsilon]/(\varepsilon^2)$ des *nombres duaux*, où K est un corps ; ε est alors l'analogue algébrique d'un infiniment petit⁽¹⁶⁾.
- L'anneau $\mathbf{C}[X, Y]/(DY^2 - (X^3 - X))$ des fonctions rationnelles sur la courbe algébrique C_D d'équation $DY^2 = X^3 - X$ dans \mathbf{C}^2 (si $f \in \mathbf{C}[X, Y]$, la restriction de f à C_D ne dépend que de l'image de f modulo l'idéal engendré par $P(X, Y) = DY^2 - (X^3 - X)$, puisque P est identiquement nul sur C_D).

Exercice 2.6. — Montrer que, si D' est un diviseur de D , alors $\mathbf{Z}/D'\mathbf{Z}$ est le quotient de $\mathbf{Z}/D\mathbf{Z}$ par l'idéal engendré par D' .

⁽¹⁴⁾Cette définition de \mathbf{C} est due à Cauchy.

⁽¹⁵⁾De manière générale, si A est un anneau, et si a est un élément de A , on note souvent (a) l'idéal de A engendré par a ; on a donc $(a) = aA$.

⁽¹⁶⁾On peut difficilement faire plus petit puisque $\varepsilon \neq 0$, alors que $\varepsilon^2 = 0$; si $P \in K[X]$ est un polynôme, on a $P(X + \varepsilon) = P(X) + P'(X)\varepsilon$ dans $K[\varepsilon]/(\varepsilon^2)$, comme le montre la formule de Taylor pour les polynômes. Peut-on rêver de développements limités plus sympathiques ?

2.6. Groupes quotients

2.6.1. Groupe opérant sur un ensemble. — Soit G un groupe d'élément neutre 1 , et soit X un ensemble. On dit que G opère à gauche sur X ou que l'on a une action à gauche de G sur X si on dispose d'une application $(g, x) \mapsto g \cdot x$ de $G \times X$ dans X telle que $1 \cdot x = x$, quel que soit $x \in X$, et $g \cdot (g' \cdot x) = gg' \cdot x$, quels que soient $g, g' \in G$ et $x \in X$. On remarquera que si $g \in G$, alors $x \mapsto \sigma_g(x) = g \cdot x$ est une bijection de X dans X , la bijection réciproque étant $x \mapsto \sigma_{g^{-1}}(x) = g^{-1} \cdot x$, et que l'on a $\sigma_{gg'} = \sigma_g \circ \sigma_{g'}$, quels que soient $g, g' \in G$. Définir une action de G sur X revient donc à se donner un morphisme de G dans le groupe des permutations de X (i.e. les bijections de X dans X) muni de la composition.

On dit que G opère à droite sur X si on a une application $(g, x) \mapsto x \star g$ de $G \times X$ dans X telle que $x \star 1 = x$, quel que soit $x \in X$, et $(x \star g) \star g' = x \star gg'$, quels que soient $g, g' \in G$ et $x \in X$. On peut toujours transformer une action à gauche en action à droite (et vice-versa), en posant $x \star g = g^{-1} \cdot x$.

Par exemple, si K est un corps commutatif, le groupe $\mathbf{GL}_n(K)$ opère naturellement (à gauche) sur beaucoup d'objets :

- par définition, il opère sur l'espace vectoriel K^n ;
- comme l'action est linéaire, elle transforme une droite vectorielle en droite vectorielle et donc $\mathbf{GL}_n(K)$ opère sur l'ensemble $\mathbf{P}^{n-1}(K)$ des droites vectorielles de K^n (espace projectif de dimension $n - 1$ sur K) ;
- il opère sur l'ensemble $\mathbf{M}_n(K)$ des matrices $n \times n$ à coefficients dans K , par multiplication à gauche (i.e. $A \cdot M = AM$), par multiplication à droite (i.e. $A \cdot M = MA^{-1}$) et par similitude (i.e. $A \cdot M = AMA^{-1}$) ;
- Il opère sur les ensembles des matrices symétriques et antisymétriques par $A \cdot M = AM^t A$;
- Le groupe $\mathbf{GL}_n(\mathbf{C})$ opère sur l'ensemble des matrices auto-adjointes (i.e. vérifiant ${}^t M = \overline{M}$) par $A \cdot M = AMA^*$, avec $A^* = {}^t \overline{A}$.

Exercice 2.7. — Soit K un corps commutatif. On rajoute à K un élément ∞ , et on étend l'arithmétique de K en posant $\frac{a}{0} = \infty$, si $a \neq 0$ (on ne donne pas de sens à $\frac{0}{0}$), et $\frac{a\infty+b}{c\infty+d} = \frac{a}{c}$, si $a \neq 0$ ou $c \neq 0$.

(i) Montrer que l'application qui à $v = (x, y) \in K^2 - \{(0, 0)\}$ associe $\lambda(v) = \frac{x}{y} \in K \cup \{\infty\}$ induit une bijection de la droite projective $\mathbf{P}^1(K)$, ensemble des droites vectorielles de K^2 , sur $K \cup \{\infty\}$.

(ii) Montrer que l'action de $\mathbf{GL}_2(K)$ sur $K \cup \{\infty\}$ qui s'en déduit est donnée par $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}$.

Si G opère (à gauche ou à droite) sur X , et si $x \in X$, un *translaté* de x est un point de X dans l'image de $G \times \{x\}$, et l'*orbite* O_x de x est l'ensemble des translatés de x (i.e., l'image de $G \times \{x\}$ dans X). Une *orbite* pour l'action de G est un sous-ensemble O de X de la forme O_x pour un certain $x \in X$.

• La relation \sim_G définie sur X par « $x \sim_G y$ si et seulement si il existe $g \in G$ tel que $y = g \cdot x$ (si l'action est à gauche) ou $y = x \star g$ (si l'action est à droite) » est une relation d'équivalence sur X dont les classes d'équivalence sont les orbites.

On peut se contenter de traiter le cas d'une action à gauche. On a $x = 1 \cdot x$, et donc \sim_G est reflexive. Si $y = g \cdot x$, alors $x = g^{-1} \cdot y$, et donc \sim_G est symétrique. Enfin, si $y = g \cdot x$ et $z = h \cdot y$, alors $z = hg \cdot x$, et donc \sim_G est transitive. Cela prouve que \sim_G est une relation d'équivalence sur X . La classe d'équivalence de x est O_x par définition de O_x , ce qui prouve que les classes d'équivalence sont les orbites.

L'espace quotient X/\sim_G , ensemble des orbites, est traditionnellement noté $G\backslash X$ si l'action est à gauche, et X/G si l'action est à droite. Un système de représentants de $G\backslash X$ ou X/G dans X est parfois appelé⁽¹⁷⁾ un *domaine fondamental*.

- Si $x \in X$, l'ensemble G_x des $g \in G$ fixant x est un sous-groupe de G , appelé *stabilisateur* de x .

Comme $1 \cdot x = x$, on a $1 \in G_x$. Si $g \cdot x = x$, alors $x = g^{-1} \cdot (g \cdot x) = g^{-1} \cdot x$, et donc G_x est stable par passage à l'inverse. Enfin, si $g \cdot x = x$ et $h \cdot x = x$, alors $gh \cdot x = g \cdot (h \cdot x) = g \cdot x = x$, ce qui prouve que G_x est stable par la loi de groupe de G , et donc est un sous-groupe de G .

On fabrique des tas de groupes intéressants en considérant les stabilisateurs d'éléments d'ensembles munis d'actions de groupes.

— Si M est une matrice symétrique, le stabilisateur de M dans $\mathbf{GL}_n(\mathbf{K})$ pour l'action $A \cdot M = AM^tA$ est le *groupe orthogonal* associé à M ; si $M = I_n$, ce groupe est noté $\mathbf{O}_n(\mathbf{K})$. Si $\mathbf{K} = \mathbf{R}$, si $p + q = n$, et si M est la matrice diagonale avec p fois 1 et q fois -1 sur la diagonale, le groupe obtenu est noté $\mathbf{O}(p, q)$; en particulier $\mathbf{O}(n) = \mathbf{O}_n(\mathbf{R})$.

— Si M est une matrice antisymétrique, le stabilisateur de M dans $\mathbf{GL}_n(\mathbf{K})$ pour l'action $A \cdot M = AM^tA$ est le *groupe symplectique* associé à M ; si $n = 2m$ est pair, et si M est la matrice par bloc $\begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}$, ce groupe est noté $\mathbf{Sp}_n(\mathbf{K})$.

— Le stabilisateur de I_n pour l'action $A \cdot M = AMA^*$ de $\mathbf{GL}_n(\mathbf{C})$ est le *groupe unitaire* $\mathbf{U}(n)$.

Exercice 2.8. — Montrer que, si $y = g \cdot x$, alors $G_y = gG_xg^{-1} = \{gxg^{-1}, x \in G_x\}$. En déduire que, si G est fini, le cardinal du stabilisateur est constant dans chaque orbite.

Exercice 2.9. — (i) Montrer que le groupe D_4 des isométries du carré de sommets $A = (1, 1)$, $B = (-1, 1)$, $C = (-1, -1)$ et $D = (1, -1)$ est un groupe d'ordre 8 dont on explicitera les éléments.

(ii) Soit $O = (0, 0)$, et soit $S = \{O, A, B, C, D\}$. Montrer que S est stable sous l'action de D_4 , et déterminer les orbites sous l'action de D_4 , ainsi que le stabilisateur d'un des éléments de chaque orbite.

(iii) Soit T l'ensemble des paires d'éléments distincts de S . Déterminer les orbites de T sous l'action de D_4 , ainsi que le stabilisateur d'un des éléments de chaque orbite.

(iv) Quel lien y-a-t'il entre le cardinal d'une orbite et celui du stabilisateur dans tous les cas ci-dessus ?

2.6.2. Classes de conjugaison

- Si G est un groupe, alors $(g, x) \mapsto g \cdot x = gxg^{-1}$ est une action (à gauche) de G sur lui-même.

Si $g, h, x \in G$, alors $gh \cdot x = ghx(gh)^{-1} = ghxh^{-1}g^{-1} = g \cdot (h x h^{-1}) = g \cdot (h \cdot x)$.

L'action de G sur lui-même ainsi définie est l'action *par conjugaison*. L'orbite de $x \in G$ est alors la *classe de conjugaison* de x , et l'ensemble $\text{Conj}(G)$ des orbites est *l'ensemble des classes de conjugaison* de G . Le stabilisateur Z_x de x pour cette action est appelé le *centralisateur* de x ; c'est l'ensemble des $g \in G$ qui commutent à x .

- G est commutatif si et seulement si les classes de conjugaison sont réduites à un élément.

⁽¹⁷⁾Du moins dans le cas où X est un espace topologique, l'action de G est continue, et le système de représentants n'est pas trop moche...

La classe de conjugaison de $x \in G$ est l'ensemble des gxg^{-1} , pour $g \in G$. Comme elle contient x , elle est réduite à un élément si et seulement si $gxg^{-1} = x$, quel que soit $g \in G$, et donc si et seulement si x commute à tous les éléments de G . Ceci permet de conclure.

- Le *centre* Z de G est l'ensemble des $x \in G$ commutant à tout élément de G ; c'est aussi l'ensemble des $x \in G$ dont la classe de conjugaison est réduite à un point, et c'est un sous-groupe de G .

Si $xg = gx$ et $yg = gy$ quel que soit $g \in G$, alors $xyg = xgy = gxy$, ce qui montre que xy commute à tous les éléments de G et donc que Z est stable par la loi de groupe. De même, si $xg = gx$ quel que soit $g \in G$, alors $gx^{-1} = x^{-1}xgx^{-1} = x^{-1}gxx^{-1} = x^{-1}g$, ce qui prouve que Z est stable par passage à l'inverse. Comme il contient l'élément neutre; c'est un sous-groupe de G . Le reste ayant été démontré ci-dessus, cela permet de conclure.

2.6.3. Quotients de groupes. — Si G est un groupe, et si H est un sous-groupe de G , on peut utiliser la multiplication dans G pour faire agir H sur G à gauche ($h \cdot x = hx$) et à droite ($x \star h = xh$). Une classe à gauche est alors de la forme $Hx = \{hx, h \in H\}$, pour $x \in G$, et une classe à droite, de la forme $xH = \{xh, h \in H\}$, pour $x \in G$. Les quotients $H \backslash G$ (à gauche) et G/H (à droite) de G par H ne sont, en général, pas des groupes, mais la multiplication dans G les munit d'actions de G (à droite pour $H \backslash G$ et à gauche pour G/H). Réciproquement, si R est une relation d'équivalence sur G telle que la multiplication dans G induise une action à gauche (resp. à droite) de G sur G/R , et si H est la classe d'équivalence de e , alors H est un sous-groupe de G et $G/R = G/H$ (resp. $G/R = H \backslash G$).

- Si G opère (à gauche) sur un ensemble X , si $x \in X$, et si G_x est le stabilisateur de x dans G , alors $g \mapsto g \cdot x$ induit un isomorphisme de G/G_x sur l'orbite O_x de x (c'est un isomorphisme d'ensembles munis d'une action de G).

Commençons par remarquer que, si g_1, g_2 ont même image dans G/G_x , alors il existe $h \in G_x$ tel que $g_2 = g_1h$, ce qui implique que $g_2 \cdot x = (g_1h) \cdot x = g_1 \cdot (h \cdot x) = g_1 \cdot x$; l'application $g \mapsto g \cdot x$ passe donc au quotient et nous définit une application $\iota : G/G_x \rightarrow O_x$ qui est surjective par définition de O_x . Maintenant, si $g_1 \cdot x = g_2 \cdot x$, alors $g_2^{-1}g_1 \cdot x = x$ et donc $g_2^{-1}g_1 \in G_x$; on en déduit que $g_1 \in g_2G_x$ et donc que g_1 et g_2 ont même image dans G/G_x , ce qui prouve que ι est injective et donc bijective. Enfin, si $h \in G$ et $g \in G/G_x$, alors $h \cdot \iota(g) = h \cdot (g \cdot x) = hg \cdot x = \iota(hg)$, ce qui prouve que ι commute à l'action de G et donc est un morphisme de G -ensembles.

- La classe de conjugaison de x est isomorphe à G/Z_x , où Z_x est le centralisateur de x .
C'est un cas particulier du point précédent.

Pour que la structure de groupe de G passe au quotient G/H , il faut et il suffit que, quels que soient $x, x' \in G$ et $h, h' \in H$, on puisse trouver $h'' \in H$ tel que $xhx'h' = xx'h''$. Comme $h''(h')^{-1} = (x')^{-1}hx'$, on voit que la condition précédente est équivalente à ce que H soit laissé stable par la conjugaison $h \mapsto ghg^{-1}$, quel que soit $g \in G$. Si tel est le cas on dit que H est *distingué* (*normal* en "français") dans G .

Un *groupe simple* est un groupe dont les seuls sous-groupes distingués sont $\{1\}$ et le groupe lui-même.

Si $u : G \rightarrow G'$ est un morphisme de groupes, alors $\text{Ker } u$ est distingué dans G et u se factorise à travers $G/\text{Ker } u$ et induit un isomorphisme de groupes de $G/\text{Ker } u$ sur $\text{Im } u$. Si G est simple, alors u est soit injectif soit trivial ($u(g) = 1$, quel que soit $g \in G$).

3. Construction de nombres

Dans ce §, purement culturel, on explique rapidement (sans démonstration) comment construire toutes les quantités usuelles à partir d'un système minimal d'axiomes. Cette problématique n'est apparue que relativement récemment dans l'histoire des mathématiques puisqu'il a fallu attendre 1872 pour que Weierstrass s'aperçoive que les nombres réels n'avaient pas été définis, ce qui aurait pu avoir des conséquences fâcheuses... Une des raisons qui ont poussé les mathématiciens à s'intéresser à ces questions de fondements a été l'apparition de monstres montrant que l'intuition pouvait se révéler fort trompeuse, et de paradoxes menaçant de faire s'écrouler tout l'édifice.

3.1. Entiers naturels

La première présentation axiomatique des entiers remonte à 1888 (Dedekind), simplifiée l'année suivante par Peano. L'ensemble des nombres entiers qui semble être constitué des objets les plus évidents (tout le monde comprend ce que sont $0, 1, 2, \dots$; le problème est dans le « ... »), ne peut pas être construit; on est plus ou moins forcé de postuler son existence et d'espérer que le ciel ne va pas nous tomber sur la tête.

- La présentation la plus efficace postule l'existence d'un ensemble \mathbf{N} , l'ensemble des entiers naturels, muni d'un élément 0 et d'une application « successeur » $s : \mathbf{N} \rightarrow \mathbf{N}$, vérifiant les axiomes (de Peano) suivants :

(A1) l'application s est injective;

(A2) 0 n'est le successeur d'aucun entier naturel;

(A3) Si $X \subset \mathbf{N}$ est tel que $0 \in X$ et $s(n) \in X$ pour tout $n \in X$, alors $X = \mathbf{N}$ (axiome de récurrence).

On définit alors l'addition et la multiplication par récurrence par $a + 0 = a$ et $a + s(b) = s(a + b)$ (pour l'addition); $a \cdot 0 = 0$ et $a \cdot s(b) = ab + a$ (pour la multiplication). On pose $1 = s(0)$, et on a $s(a) = s(a + 0) = a + s(0) = a + 1$, ce qui permet de supprimer l'application successeur et de la remplacer par $a \mapsto a + 1$. Vérifier, à partir des axiomes de Peano, que l'addition et la multiplication sont commutatives et que la multiplication est distributive par rapport à l'addition, est un exercice un peu répétitif mais très satisfaisant pour l'esprit

- On obtient une présentation plus intuitive en partant de l'idée que se fait le petit enfant du nombre 5. On dit qu'un ensemble X est fini s'il ne peut pas être mis en bijection avec $X \cup \{x\}$, si $x \notin X$. On postule l'existence d'un ensemble infini Ω (axiome de l'infini), et on munit l'ensemble des parties de Ω de la relation d'équivalence

\sim définie par $X_1 \sim X_2$ s'il existe une bijection de X_1 sur X_2 . On définit alors l'ensemble \mathbf{N} des entiers naturels comme l'ensemble des classes d'équivalence de parties finies de Ω pour cette relation d'équivalence. Si X est une partie finie de Ω , on note $|X| \in \mathbf{N}$ sa classe d'équivalence ; c'est un entier naturel que l'on appelle le cardinal de X , et une analyse *a posteriori* de la construction précédente, montre que l'on a défini l'entier n comme la classe d'équivalence de tous les ensembles (inclus dans notre Ω) de cardinal n .

On note 0 le cardinal de l'ensemble vide, 1 celui d'un singleton (i.e. un ensemble X tel que $x, y \in X \Rightarrow x = y$)... Si $a, b \in \mathbf{N}$, on choisit $X, Y \subset \Omega$ disjoints, de cardinaux respectifs a et b , et on définit $a + b$ comme le cardinal de $X \cup Y$, et ab comme le cardinal de (tout sous-ensemble de Ω pouvant être mis en bijection avec) $X \times Y$. Il est alors quasi-immédiat que $a + b = b + a$ (car $X \cup Y = Y \cup X$), que $ab = ba$ (car $(x, y) \mapsto (y, x)$ induit une bijection de $X \times Y$ sur $Y \times X$), que $0 \cdot a = a$ pour tout $a \in \mathbf{N}$ (l'ensemble $\emptyset \times X$ est vide quel que soit X), et que $a(b + c) = ab + ac$ (car $X \times (Y \cup Z) = (X \times Y) \cup (X \times Z)$).

Les choses se compliquent quand on essaie de montrer que \mathbf{N} vérifie l'axiome de récurrence pour l'application successeur $x \mapsto x + 1$.

3.2. Entiers relatifs, nombres rationnels

En partant des entiers naturels, on fait les constructions suivantes.

- On construit \mathbf{Z} comme quotient de $\mathbf{N} \times \mathbf{N}$ par la relation d'équivalence $(a, b) \sim (a', b')$ si et seulement si $a + b' = a' + b$, l'idée étant que (a, b) représente l'entier relatif $a - b$. L'application $n \mapsto (n, 0)$ induit une injection de \mathbf{N} dans \mathbf{Z} , ce qui permet de voir \mathbf{N} comme un sous-ensemble de \mathbf{Z} .

L'addition $(a, b) + (a', b') = (a + a', b + b')$ passe au quotient, et définit une loi qui fait de \mathbf{Z} un groupe commutatif, l'élément neutre étant (la classe de) $(0, 0)$ (ou de (a, a) , pour tout $a \in \mathbf{N}$), et l'opposé de (a, b) étant (b, a) . L'opposé $-n$ de n est donc représenté par $(n, 0)$, et on peut maintenant définir $a - b$, si a et $b \in \mathbf{Z}$, et si $a, b \in \mathbf{N}$, on a $a - b = (a, 0) + (0, b) = (a, b)$, ce que l'on cherchait à obtenir.

La multiplication⁽¹⁸⁾ $(a, b)(a', b') = (aa' + bb', ab' + ba')$ passe au quotient, et \mathbf{Z} muni de l'addition et de la multiplication est un anneau commutatif.

Enfin, on dit que $a \geq b$, si $a - b \in \mathbf{N}$, et on obtient de la sorte une relation d'ordre totale sur \mathbf{Z} .

- On construit \mathbf{Q} comme quotient de $\mathbf{Z} \times (\mathbf{Z} - \{0\})$ par la relation d'équivalence $(a, b) \sim (a', b')$ si et seulement si $ab' = a'b$, l'idée étant que (a, b) représente le nombre rationnel $\frac{a}{b}$. L'application $n \mapsto (n, 1)$ induit une injection de \mathbf{Z} dans \mathbf{Q} , ce qui permet de voir \mathbf{Z} comme un sous-ensemble de \mathbf{Q} .

L'addition et la multiplication sur \mathbf{Q} sont définies par les formules $(a, b) + (a', b') = (ab' + ba', bb')$ et $(a, b)(a', b') = (aa', bb')$ qui passent au quotient, et \mathbf{Q} muni de l'addition et de la multiplication est un corps commutatif : l'élément neutre pour $+$ est 0, la classe de $(0, b)$, pour tout $b \in \mathbf{N}$, l'opposé de (a, b) est $(-a, b)$, l'élément neutre pour \times est 1, classe de (b, b) , pour tout $b \in \mathbf{Z} - \{0\}$, et l'inverse de (a, b) , si $(a, b) \neq 0$ (ce qui équivaut

⁽¹⁸⁾ On rappelle que (a, b) représente $b - a$, et donc que $(aa' + bb', ab' + ba')$ représente $aa' + bb' - ab' - ba' = (a - b)(a' - b')$, ce qui explique comment la formule pour la multiplication a été obtenue.

à $a \neq 0$) est (b, a) . Si $a \in \mathbf{Z}$ et $b \in \mathbf{Z} - \{0\}$, on peut maintenant diviser a par b dans \mathbf{Q} , et $b^{-1}a$ est la classe de $(1, b)(a, 1) = (a, b)$, ce que l'on cherchait à obtenir.

Enfin, on dit que q est positif, si q a un représentant (a, b) avec $b \geq 0$ et $a \geq 0$, et que $q_1 \geq q_2$, si $q_1 - q_2$ est positif. On obtient de la sorte une relation d'ordre total sur \mathbf{Q} .

3.3. Nombres réels, nombres complexes

Pour construire \mathbf{R} à partir de \mathbf{Q} , on dispose essentiellement de trois possibilités.

- On peut utiliser les *coupures de Dedekind*, c'est-à-dire l'ensemble des couples (A, B) de parties non vides de \mathbf{Q} tels que $A \cup B = \mathbf{Q}$, et tout élément de A est \leq à tout élément de B . L'idée étant que si $r \in \mathbf{R}$, alors r correspond à la coupure (A_r, B_r) donnée par $A_r = \{x \in \mathbf{Q}, x \leq r\}$ et $B_r = \{x \in \mathbf{Q}, x \geq r\}$. Les rationnels s'identifient aux coupures (A, B) telles que $A \cap B$ est non vide. Il est alors facile de montrer que l'ensemble \mathbf{R} ainsi construit vérifie la propriété de la borne supérieure (toute partie majorée non vide admet une borne supérieure), puis qu'il est complet.

- On peut aussi, comme G. Cantor, compléter \mathbf{Q} pour la valeur absolue, en rajoutant de force les limites des suites de Cauchy d'éléments de \mathbf{Q} . Cela peut se faire, par exemple, de la manière suivante. On considère l'ensemble $\text{Cauchy}(\mathbf{Q})$ des suites de Cauchy d'éléments de \mathbf{Q} (i.e. l'ensemble des suites $(a_n)_{n \in \mathbf{N}}$, avec $a_n \in \mathbf{Q}$, telles que, pour tout $j \in \mathbf{N}$, il existe $N_j \in \mathbf{N}$ tel que $|a_p - a_n| < 2^{-j}$, quels que soient $n, p \geq N_j$). Alors $\text{Cauchy}(\mathbf{Q})$ est un anneau pour l'addition et la multiplication terme à terme dans lequel l'ensemble I des suites tendant vers 0 est un idéal. On définit alors \mathbf{R} comme le quotient de $\text{Cauchy}(\mathbf{Q})$ par I , ce qui revient (moralement) à identifier deux suites de Cauchy ayant même limite (i.e. dont la différence tend vers 0), et donc « à identifier une suite de Cauchy avec sa limite ». Le résultat \mathbf{R} est un corps⁽¹⁹⁾, muni d'une relation d'ordre stricte $<$ totale⁽²⁰⁾, dans lequel \mathbf{Q} (identifié à l'image des suites constantes) est dense⁽²¹⁾.

- On peut aussi utiliser la construction de « l'homme de la rue » qui part du fait qu'un réel a un développement décimal. Cela conduit à définir \mathbf{R} comme l'ensemble des développements décimaux $a_n \dots a_0, a_{-1}a_{-2} \dots$ avec un nombre fini de chiffres avant la virgule et un nombre infini après, modulo la relation d'équivalence \sim , identifiant $a_n \dots a_m 999999 \dots$ à $a_n \dots a_{m+1}(a_m + 1)00000 \dots$, si $a_m \neq 9$. Nous laissons au lecteur le soin de définir l'addition et la multiplication de deux réels de « l'homme de la rue »...

⁽¹⁹⁾Si $(a_n)_{n \in \mathbf{N}}$ est une suite de Cauchy qui ne tend pas vers 0, alors $a_n \neq 0$ si $n \geq n_0$, et la suite $(1, \dots, 1, a_{n_0}^{-1}, \dots, a_n^{-1}, \dots)$ est une suite de Cauchy dont l'image dans \mathbf{R} est l'inverse de celle de la suite $(a_n)_{n \in \mathbf{N}}$.

⁽²⁰⁾Si $a, b \in \mathbf{R}$, on dit que $a < b$, si $a \neq b$ et si, pour tous représentants $(a_n)_{n \in \mathbf{N}}$ et $(b_n)_{n \in \mathbf{N}}$ de a et b , on a $a_n < b_p$ si n et p sont assez grands; on constate sans problème que si c'est vrai pour un choix de représentants, alors c'est vrai pour tous.

⁽²¹⁾Cela signifie qu'entre deux éléments de \mathbf{R} on peut toujours trouver un élément de \mathbf{Q} . Si $a < b$ sont deux éléments de \mathbf{R} , et si $(a_n)_{n \in \mathbf{N}}$ et $(b_n)_{n \in \mathbf{N}}$ sont des représentants de a et b , alors $a_n < b_p$, si $n, p \geq n_0$, et $r = \frac{a_{n_0} + b_{n_0}}{2}$ est un élément de \mathbf{Q} vérifiant $a < r < b$.

• Une fois les nombres réels construits, on obtient le corps des nombres complexes \mathbf{C} en rajoutant à \mathbf{R} une racine carrée i de -1 , ce qui revient à poser $\mathbf{C} = \mathbf{R}[X]/(X^2 + 1)$. Le résultat est un corps complet pour la norme $|z| = \sqrt{x^2 + y^2}$, si $z = x + iy$, et qui est algébriquement clos (résultat connu sous le nom de « théorème fondamental de l'algèbre » bien qu'il n'existe aucune démonstration de ce résultat qui n'utilise pas de technique d'analyse).

3.4. Nombres p -adiques

3.4.1. *Le corps \mathbf{Q}_p .* — La construction de \mathbf{R} de G. Cantor, bien que plus compliquée, est nettement plus souple que celle de R. Dedekind, et se généralise facilement. Il n'a fallu que 25 ans après la construction des nombres réels (qui avait pris quelque deux millénaires...), pour que K. Hensel envisage la construction des nombres p -adiques, et une petite dizaine d'années pour qu'il leur donne une forme maniable. De nos jours, on procède de la manière suivante.

Soit p un nombre premier. Si $a \in \mathbf{Z} - \{0\}$, on définit la valuation p -adique $v_p(a)$ comme le plus grand entier v tel que p^v divise a . On a $v_p(ab) = v_p(a) + v_p(b)$ si $a, b \in \mathbf{Z} - \{0\}$, ce qui permet d'étendre v_p à \mathbf{Q} en posant $v_p(\frac{a}{b}) = v_p(a) - v_p(b)$, si $a, b \in \mathbf{Z} - \{0\}$, et $v_p(0) = +\infty$. On a alors $v_p(x + y) \geq \min(v_p(x), v_p(y))$, si $x, y \in \mathbf{Q}$ car, si x et y sont divisibles par p^v , il en est de même de $x + y$. On en déduit le fait que, si on pose $|x|_p = p^{-v_p(x)}$, alors $|x + y|_p \leq \sup(|x|_p, |y|_p)$ et donc que $d_p(x, y) = |x - y|_p$ est une distance sur \mathbf{Q} (la *distance p -adique*), l'inégalité ci-dessus, dite *ultramétrique*, étant plus forte que l'inégalité triangulaire.

On définit \mathbf{Q}_p , corps des nombres p -adiques, comme le complété de \mathbf{Q} pour la norme p -adique $| \cdot |_p$, c'est-à-dire que l'on prend, comme pour définir \mathbf{R} , l'anneau des suites de Cauchy (pour la norme $| \cdot |_p$) d'éléments de \mathbf{Q} , et on quotiente par l'idéal des suites tendant vers 0. Si $x \in \mathbf{Q}_p$, et si $(a_n)_{n \in \mathbf{N}}$ est un représentant de x , alors $|a_n|_p$ tend vers une limite dans \mathbf{R} (et même dans $p^{\mathbf{Z}} \cup \{0\}$, car tous ses termes sont dans $p^{\mathbf{Z}} \cup \{0\}$ qui est fermé dans \mathbf{R}_+) qui ne dépend que de x , et qu'on note $|x|_p$. Par construction, $| \cdot |_p$ est une norme ultramétrique sur \mathbf{Q}_p , ce qui signifie que $|x|_p = 0$ si et seulement si $x = 0$, que $|xy|_p = |x|_p |y|_p$, quels que soient $x, y \in \mathbf{Q}_p$, et que $|x + y|_p \leq \sup(|x|_p, |y|_p)$, et donc $d_p(x, y) = |x - y|_p$ est une distance ultramétrique sur \mathbf{Q}_p pour laquelle \mathbf{Q}_p est complet. On étend v_p à \mathbf{Q}_p par continuité, et on a encore $|x|_p = p^{-v_p(x)}$, si $x \in \mathbf{Q}_p$.

• Dans \mathbf{Q}_p , une suite $(x_n)_{n \in \mathbf{N}}$ converge si et seulement si $x_{n+1} - x_n$ tend vers 0 et une série $\sum_{n \in \mathbf{N}} u_n$ converge si et seulement si u_n tend vers 0.

D'après l'inégalité ultramétrique, on a $|x_{n+k} - x_n|_p \leq \sup_{0 \leq i \leq k-1} |x_{n+i+1} - x_{n+i}|_p$, ce qui montre que si $|x_{n+1} - x_n|_p$ tend vers 0, alors la suite est de Cauchy. La complétude de \mathbf{Q}_p permet de conclure (l'argument est le même pour une série).

Exercice 3.1. — Montrer que la série $1 + 2 + 4 + 8 + \dots$ converge vers -1 dans \mathbf{Q}_2 .

Exercice 3.2. — (i) Montrer que, si $|x|_p > |y|_p$, alors $|x + y|_p = |x|_p$.

(ii) Montrer que si $u_n \rightarrow 0$, et si $|u_0|_p > |u_n|_p$, pour tout $n \geq 1$, alors $\sum_{n \in \mathbf{N}} u_n \neq 0$, et $|\sum_{n \in \mathbf{N}} u_n|_p = |u_0|_p$.

- La topologie de \mathbf{Q}_p possède des propriétés un peu déroutantes au premier abord.
- (i) Tout point d'une boule de \mathbf{Q}_p en est "le" centre.
- (ii) Deux boules de \mathbf{Q}_p sont soit disjointes soit l'une est contenue dans l'autre (comme des billes de mercure).
- (iii) Les boules de \mathbf{Q}_p sont à la fois ouvertes et fermées.
- (iv) La topologie de \mathbf{Q}_p est totalement discontinue.

Si $x_1 \in B(x_0, r)$ (ouverte ou fermée) et $y \in B(x_1, r)$, alors $d_p(x_0, y) \leq \sup(d_p(x_0, x_1), d_p(x_1, y)) \leq r$ (ou $< r$ si on parle de boules ouvertes), et donc $B(x_1, r) \subset B(x_0, r)$. L'inclusion dans l'autre sens s'obtient en échangeant les rôles de x_0 et x_1 , ce qui permet de démontrer le (i).

D'après le (i), si deux boules ont une intersection non vide, tout élément de l'intersection est le centre des deux boules, ce qui démontre le (ii).

Si B est une boule ouverte de rayon r , le complémentaire de B contient la boule ouverte de rayon r autour de chacun de ses points d'après le (ii), ce qui montre que ce complémentaire est ouvert et donc que B est fermée. Si B est une boule fermée de rayon non nul, alors B est un voisinage de chacun de ses points puisque ceux-ci en sont "le" centre. On en déduit le (iii).

Enfin, si $x \in \mathbf{Q}_p$, si C_x est la composante connexe de x , et si $r > 0$, alors $C_x \cap B(x, r)$ est à la fois ouvert et fermé dans C_x , et non vide puisque contenant x . Comme C_x est connexe, cela implique $C_x \cap B(x, r) = C_x$, quel que soit $r > 0$, et donc $C_x = \{x\}$. On en déduit le (iv).

3.4.2. L'anneau \mathbf{Z}_p des entiers p -adiques

- L'ensemble $\mathbf{Z}_p = \{x \in \mathbf{Q}_p, |x|_p \leq 1\}$ est un sous-anneau fermé de \mathbf{Q}_p qui contient \mathbf{Z} .

La multiplicativité de $|\cdot|_p$ montre que \mathbf{Z}_p est stable par multiplication et l'inégalité ultramétrique montre que \mathbf{Z}_p est stable par addition. C'est donc un sous-anneau de \mathbf{Q}_p qui contient \mathbf{Z} de manière évidente et qui est fermé puisque c'est l'image inverse de $[0, 1]$ par l'application $x \mapsto |x|_p$.

- Le sous-groupe \mathbf{Z}_p^* des unités de \mathbf{Z}_p est l'ensemble des $x \in \mathbf{Z}_p$ vérifiant $|x|_p = 1$; c'est aussi $\mathbf{Z}_p - p\mathbf{Z}_p$.

Si $x \in \mathbf{Z}_p - \{0\}$, l'inverse x^{-1} de x dans \mathbf{Q}_p vérifie $|x^{-1}|_p |x|_p = 1$. Comme $|x|_p \leq 1$, cet inverse appartient à \mathbf{Z}_p si et seulement si $|x|_p = 1$. Maintenant, pour les mêmes raisons que ci-dessus, l'ensemble des $x \in \mathbf{Z}_p$ vérifiant $|x|_p < 1$ est un idéal de \mathbf{Z}_p , et comme $|x|_p < 1$ implique $|x|_p \leq p^{-1}$, c'est l'idéal $p\mathbf{Z}_p$. On a donc $\mathbf{Z}_p^* = \mathbf{Z}_p - p\mathbf{Z}_p$.

- L'application naturelle de $\mathbf{Z}/p^n\mathbf{Z}$ dans $\mathbf{Z}_p/p^n\mathbf{Z}_p$ est un isomorphisme.

Si x est un élément de $\mathbf{Z} \cap p^n\mathbf{Z}_p$, on a $v_p(x) \geq n$, ce qui signifie que x est divisible par p^n dans \mathbf{Z} . On en déduit l'injectivité. Prouvons la surjectivité. Soit $\bar{x} \in \mathbf{Z}_p/p^n\mathbf{Z}_p$ et $x \in \mathbf{Z}_p$ ayant pour image \bar{x} modulo p^n . Comme \mathbf{Q} est dense dans \mathbf{Q}_p , il existe $r \in \mathbf{Q}$ vérifiant $v_p(x - r) \geq n$; en particulier $v_p(r) \geq 0$. Écrivons r sous la forme $\frac{a}{b}$ avec $a, b \in \mathbf{Z}$. Comme $v_p(r) \geq 0$, on a $v_p(b) \leq v_p(a)$ et quitte à tout diviser par $p^{v_p(b)}$, on peut supposer $v_p(b) = 0$, et donc $(b, p) = 1$, ce qui implique que b est premier à p^n et donc est inversible dans $\mathbf{Z}/p^n\mathbf{Z}$. Soit \bar{c} l'inverse de b dans $\mathbf{Z}/p^n\mathbf{Z}$ et $c \in \mathbf{Z}$ dont la réduction modulo p^n est \bar{c} . On a alors $v_p(r - ac) =$

$v_p(a) - v_p(b) + v_p(1 - bc) \geq n$ et donc $v_p(x - ac) \geq n$, ce qui prouve que ac a pour image \bar{x} dans $\mathbf{Z}_p/p^n\mathbf{Z}_p$, et permet de conclure.

- Tout élément de \mathbf{Q}_p peut s'écrire de manière unique sous la forme $x = \sum_{i=-k}^{+\infty} a_i p^i$, avec $a_i \in \{0, \dots, p-1\}$ pour tout i . Il admet donc une unique *écriture en base p*

$$x = \dots a_{n-1} \dots a_0, a_{-1} \dots a_{-k},$$

et on a $|x|_p = p^k$, si $a_{-k} \neq 0$. Une différence avec les nombres réels est qu'il y a une infinité de chiffres avant la virgule et un nombre fini après. Les éléments de \mathbf{Z}_p sont ceux dont l'écriture en base p n'a pas de chiffre après la virgule (du point de vue de l'écriture en base p , ils correspondent au segment $[0, 1]$ de \mathbf{R}).

Si $n \in \mathbf{N}$, alors $\{0, \dots, p^n - 1\}$ est un système de représentants de $\mathbf{Z}/p^n\mathbf{Z}$. Soit alors $x \in \mathbf{Q}_p^*$, et soit $k = -v_p(x)$ de telle sorte que $y = p^k x \in \mathbf{Z}_p^*$. Si $n \geq -k$, soit $y_n \in \{0, \dots, p^{n+k} - 1\}$ le représentant de l'image de y dans $\mathbf{Z}_p/p^{n+k}\mathbf{Z}_p \cong \mathbf{Z}/p^{n+k}\mathbf{Z}$ (en particulier, $y_{-k} = 0$ et $y_{1-k} \neq 0$, car $y \notin p\mathbf{Z}_p$). Alors $y_{n+1} - y_n$ est divisible par p^{n+k} , ce qui permet de définir $a_n \in \{0, \dots, p-1\}$ par $a_n = p^{-n-k}(y_{n+1} - y_n)$. On a alors $y_n = \sum_{i=0}^{n+k-1} a_{i-k} p^i$; autrement dit, $a_{n-1} a_{n-2} \dots a_{-k}$ est l'écriture de y_n en base p . Par suite, $a_{n-1} \dots a_0, a_{-1} \dots a_{-k}$ est l'écriture de $x_n = p^{-k} y_n$ en base p . Or $y_n - p^k x \in p^{n+k}\mathbf{Z}_p$ par construction, ce qui se traduit par $|y_n - p^k x|_p \leq p^{-(n+k)}$, ou encore par $|x_n - x|_p \leq p^{-n}$, et montre que $x_n \rightarrow x$ dans \mathbf{Q}_p . On a donc $x = \sum_{i=-k}^{+\infty} a_i p^i$ (la somme converge puisque son terme général tend vers 0). On en déduit l'existence d'une écriture sous la forme voulue.

Pour démontrer l'unicité, il suffit de constater que si $\sum_{i=-k}^{+\infty} a_i p^i = \sum_{i=-k}^{+\infty} b_i p^i$, alors en multipliant les deux membres par p^k , et en regardant modulo $p\mathbf{Z}_p$, on obtient $a_{-k} - b_{-k} \in p\mathbf{Z}_p$. Comme les a_i et les b_i sont dans un système de représentants modulo $p\mathbf{Z}_p$, cela prouve que $a_{-k} = b_{-k}$. Une récurrence immédiate permet d'en déduire que $a_i = b_i$ pour tout i . Le reste découle de la manière dont les a_i ont été construits ci-dessus.

- \mathbf{N} et \mathbf{Z} sont denses dans \mathbf{Z}_p et $\mathbf{Z}[\frac{1}{p}]$ est dense dans \mathbf{Q}_p .

Cela suit de l'existence de l'écriture en base p d'un nombre p -adique (si on coupe cette écriture au n -ième chiffre avant la virgule, on obtient un élément x de \mathbf{N} (resp. $\mathbf{Z}[\frac{1}{p}]$), si on est parti d'un élément de \mathbf{Z}_p (resp. \mathbf{Q}_p), et la suite de nombres ainsi obtenue converge vers x).

- \mathbf{Z}_p est compact.

Comme \mathbf{Z}_p est un espace métrique, il suffit de vérifier que toute suite d'éléments de \mathbf{Z}_p admet une sous-suite convergente dans \mathbf{Z}_p . Soit donc $(x_n)_{n \in \mathbf{N}}$ une telle suite, et soit $\sum_{i=0}^{+\infty} a_{n,i} p^i$ l'écriture de x_n en base p . Il existe alors $a_0 \in \{0, \dots, p-1\}$ tel que $a_{n,0} = a_0$ pour une infinité de n . Ceci permet d'extraire de la suite $(x_n)_{n \in \mathbf{N}}$ une sous-suite $(x_{\varphi_0(n)})_{n \in \mathbf{N}}$ telle que $a_{\varphi_0(n),0} = a_0$, quel que soit $n \in \mathbf{N}$. Pour la même raison, il existe alors $a_1 \in \{0, \dots, p-1\}$, et une sous-suite $(x_{\varphi_1(n)})_{n \in \mathbf{N}}$, extraite de $(x_{\varphi_0(n)})_{n \in \mathbf{N}}$, telle que $a_{\varphi_1(n),0} = a_1$, quel que soit $n \in \mathbf{N}$. Par récurrence, cela permet de définir $a_k \in \{0, \dots, p-1\}$ et une sous-suite $(x_{\varphi_k(n)})_{n \in \mathbf{N}}$ extraite de $(x_{\varphi_{k-1}(n)})_{n \in \mathbf{N}}$, tels que $a_{\varphi_k(n),i} = a_i$, quels que soient $n \in \mathbf{N}$ et $i \leq k$. La suite $(x_{\varphi_n(n)})_{n \in \mathbf{N}}$ est alors extraite (procédé d'*extraction diagonale*) de $(x_n)_{n \in \mathbf{N}}$ et, par construction, on a $a_{\varphi_n(n),i} = a_i$, si $i \leq n$, ce qui se traduit par $|x_{\varphi_n(n)} - \sum_{i=0}^{+\infty} a_i p^i|_p \leq p^{-n-1}$, et montre que $x_{\varphi_n(n)} \rightarrow \sum_{i=0}^{+\infty} a_i p^i$ dans \mathbf{Z}_p . Ceci permet de conclure.

- \mathbf{Q}_p est localement compact.

Une boule ouverte $B(a, r^-)$ de \mathbf{Q}_p est aussi de la forme $a + p^n \mathbf{Z}_p$, où n est le plus grand élément de \mathbf{Z} tel que $p^{-n} < r$. Elle est donc homéomorphe à \mathbf{Z}_p , et la compacité de \mathbf{Z}_p permet de conclure.

3.4.3. L'anneau des nombres complexes p -adiques

On peut essayer d'imiter la construction de \mathbf{C} à partir de \mathbf{R} pour obtenir des nombres complexes p -adiques. On procède donc de la manière suivante. On commence par rajouter à \mathbf{Q}_p toutes les racines des polynômes à coefficients dans \mathbf{Q}_p , et on obtient ainsi un corps $\overline{\mathbf{Q}_p}$ algébriquement clos auquel on étend la norme p -adique $|\cdot|_p$ (cette étape ne se fait pas toute seule). Une différence avec le cas réel est qu'on est forcé de rajouter une infinité d'éléments et que le résultat n'est pas complet. On complète donc $\overline{\mathbf{Q}_p}$ pour la norme $|\cdot|_p$, et on obtient un corps \mathbf{C}_p qui est complet et algébriquement clos, et qui est abstraitement isomorphe à \mathbf{C} . Le seul problème est que J. Tate (1968) a démontré que \mathbf{C}_p ne contient pas d'analogue raisonnable de $2i\pi$, ce qui est un peu ennuyeux vu le rôle joué par $2i\pi$ dans le monde usuel (cf. la formule de Cauchy par exemple). Le problème a été résolu par J.-M. Fontaine (1982) qui a construit un anneau \mathbf{B}_{dR}^+ (sa construction est assez compliquée...), l'anneau des nombres complexes p -adiques, qui contient un $2i\pi$ naturel, et qui est muni d'un morphisme d'anneaux surjectif $\theta : \mathbf{B}_{\text{dR}}^+ \rightarrow \mathbf{C}_p$ dont le noyau est engendré par le $2i\pi$ de Fontaine (ce qui explique qu'on ne le voit pas dans \mathbf{C}_p).

4. Groupes finis

4.1. Généralités sur les groupes

Un *groupe* G est un ensemble non vide, muni d'une loi $(g, h) \rightarrow gh$ de $G \times G \rightarrow G$ qui est *associative* (i.e. $(g_1 g_2) g_3 = g_1 (g_2 g_3)$, quels que soient $g_1, g_2, g_3 \in G$), possède un *élément neutre* (i.e. il existe $e \in G$ tel que $eg = ge = g$, pour tout $g \in G$), et telle que tout élément g admet un *inverse* (i.e. il existe $g^{-1} \in G$ tel que $gg^{-1} = g^{-1}g = e$).

Un groupe G est *commutatif* ou *abélien* si $gh = hg$ quels que soient $g, h \in G$. La loi de groupe d'un groupe commutatif est souvent notée $+$, auquel cas l'élément neutre est noté 0 et l'inverse de $x \in G$ est noté $-x$ et appelé *l'opposé* de x . Une loi notée $+$ ou \oplus ou \boxplus est implicitement commutative, à moins que l'auteur n'ait vraiment décidé de rendre son texte illisible. Si la loi de groupe est notée multiplicativement, l'élément neutre de G est en général noté 1 au lieu de e ; s'il s'agit d'un groupe de bijections d'un ensemble X , l'élément neutre est l'identité de X , et est souvent noté id .

Si G est un groupe d'élément neutre 1 , si $x \in G$, on définit x^n , pour $n \in \mathbf{Z}$, en posant $x^0 = 1$, et $x^{n+1} = x^n x$, si $n \in \mathbf{N}$, et $x^n = (x^{-1})^{-n}$, si $n \leq 0$. Si G est commutatif et si la loi est notée $+$, l'élément x^n est noté nx , et on a $0x = 0$ et $(-1)x = -x$. On vérifie facilement que si $n \in \mathbf{Z}$, alors $x^{n+1} = x^n x$ et $x^{n-1} = x^n x^{-1}$, ce qui permet de montrer, par récurrence sur m , que $x^{m+n} = x^m x^n$ quels que soient $m, n \in \mathbf{N}$. Autrement dit, $n \mapsto x^n$ est un *morphisme de groupes de \mathbf{Z} dans G* .

Si x et y commutent, on a $(xy)^n = x^n y^n$, mais s'ils ne commutent pas, c'est en général faux (et si $n = 2$ ou si $n = -1$, cela n'est vrai que si x et y commutent).

Un *sous-groupe* H d'un groupe G est une partie de H qui contient l'élément neutre, est stable par la loi de groupe ($h_1, h_2 \in H \Rightarrow h_1 h_2 \in H$) et par passage à l'inverse ($h \in H \Rightarrow h^{-1} \in H$); c'est donc un groupe pour la loi induite par celle de G .

Si les $(H_i)_{i \in I}$ sont des sous-groupes d'un groupe G , alors $\cap_{i \in I} H_i$ est un sous-groupe de G . Ceci permet de définir le *sous-groupe* $\langle X \rangle$ de G engendré par une partie X de G comme étant l'intersection de tous les sous-groupes de G qui la contiennent. Par exemple, si $x \in G$, le sous-groupe $\langle x \rangle$ engendré par x est l'ensemble des x^n , pour $n \in \mathbf{Z}$.

En effet, d'une part un sous-groupe qui contient x contient x^n , pour $n \in \mathbf{N}$, comme le montre une récurrence immédiate, et comme il contient x^{-1} , il contient aussi x^n , pour $n \leq 0$; d'autre part, l'ensemble des x^n , pour $n \in \mathbf{Z}$, est un groupe qui contient x , puisque c'est l'image de \mathbf{Z} par le morphisme $x \mapsto x^n$.

4.2. Groupes cycliques

4.2.1. Structure des groupes cycliques, ordre d'un élément

Un groupe est *cyclique* s'il peut être engendré par un seul élément. Autrement dit, G est cyclique, s'il existe $x \in G$ tel que le morphisme $x \mapsto x^n$ de \mathbf{Z} dans G soit surjectif. Si G est cyclique, un *générateur* de G est un élément x de G tel que le morphisme $x \mapsto x^n$ de \mathbf{Z} dans G soit surjectif.

- Le groupe \mathbf{Z} est cyclique, et il admet deux générateurs 1 et -1 . Si $D \geq 1$, le groupe $\mathbf{Z}/D\mathbf{Z}$ est cyclique et les générateurs de $\mathbf{Z}/D\mathbf{Z}$ sont les éléments de $(\mathbf{Z}/D\mathbf{Z})^*$, c'est-à-dire les (réductions modulo D des) entiers premiers à D .

L'énoncé concernant \mathbf{Z} est immédiat. Il est aussi immédiat que $\mathbf{Z}/D\mathbf{Z}$ est cyclique et que 1 en est un générateur. Maintenant, si $a \in \mathbf{Z}/D\mathbf{Z}$ en est un générateur, alors il existe en particulier $b \in \mathbf{Z}$ tel que $ba = 1$, ce qui fait que la réduction modulo D de b est un inverse de a , et donc que a est inversible. Réciproquement, si a est inversible, alors $n \mapsto na$ est bijectif de $\mathbf{Z}/D\mathbf{Z}$ dans $\mathbf{Z}/D\mathbf{Z}$ et donc $n \mapsto na$ est surjectif de \mathbf{Z} dans $\mathbf{Z}/D\mathbf{Z}$, ce qui prouve que a est un générateur de $\mathbf{Z}/D\mathbf{Z}$.

- Le groupe μ_D des racines D -ièmes de l'unité dans \mathbf{C} est cyclique, engendré par $e^{2i\pi/D}$, et $n \mapsto e^{2i\pi n/D}$ induit un isomorphisme de groupes $\mathbf{Z}/D\mathbf{Z} \cong \mu_D$. Un générateur de μ_D est une *racine primitive* D -ième de l'unité, et les racines primitives D -ièmes de l'unité sont, d'après le point précédent, les racines de la forme $e^{2i\pi a/D}$, pour a premier à D .
- Un groupe cyclique infini est isomorphe à \mathbf{Z} ; un groupe cyclique de cardinal D est isomorphe à $\mathbf{Z}/D\mathbf{Z}$. En particulier, un groupe cyclique est commutatif.

Soit G un groupe cyclique, et soit x un générateur de G . Alors $f : \mathbf{Z} \rightarrow G$ défini par $f(n) = x^n$ est un morphisme surjectif, et il y a deux cas :

- f est injectif et alors G est isomorphe à \mathbf{Z} ;
- le noyau de f est non nul et donc de la forme $D\mathbf{Z}$, avec $D \geq 1$, puisque c'est un sous-groupe de \mathbf{Z} ; alors f se factorise à travers $\bar{f} : \mathbf{Z}/D\mathbf{Z} \rightarrow G$, et \bar{f} est surjectif puisque f l'est et injectif puisqu'on a factorisé modulo $\text{Ker } f$; autrement dit \bar{f} est un isomorphisme de $\mathbf{Z}/D\mathbf{Z}$ sur G et, en particulier, G et $\mathbf{Z}/D\mathbf{Z}$ ont même cardinal.

Ceci permet de conclure.

- Si G est un groupe quelconque et $x \in G$, le sous-groupe $\langle x \rangle$ de G engendré par x est cyclique par définition. On définit *l'ordre de x* comme le cardinal du groupe $\langle x \rangle$. Si x est d'ordre D , le noyau du morphisme $n \rightarrow x^n$ de \mathbf{Z} dans G est $D\mathbf{Z}$ d'après ce qui précède, ce qui fait que *l'ordre de x est aussi le plus petit entier $n > 0$ tel que x^n soit égal à l'élément neutre*.

4.2.2. Sous-groupes des groupes cycliques

- Si $D \geq 1$, l'application $d \mapsto d\mathbf{Z}/D\mathbf{Z}$ est une bijection de l'ensemble des diviseurs de D sur celui des sous-groupes de $\mathbf{Z}/D\mathbf{Z}$.

Si G est un sous-groupe de $\mathbf{Z}/D\mathbf{Z}$, on peut considérer son image inverse dans \mathbf{Z} , qui est un sous-groupe de \mathbf{Z} contenant $D\mathbf{Z}$; on obtient ainsi une bijection de l'ensemble des sous-groupes de $\mathbf{Z}/D\mathbf{Z}$ dans celui des sous-groupes de \mathbf{Z} contenant $D\mathbf{Z}$, la bijection inverse étant $\tilde{G} \rightarrow \tilde{G}/D\mathbf{Z}$. Comme un sous-groupe de \mathbf{Z} contenant $D\mathbf{Z}$ est de la forme $d\mathbf{Z}$, avec d diviseur de D , cela permet de conclure.

- Si G est un groupe cyclique, tous les sous-groupes de G sont cycliques, et si G est de cardinal D , alors G admet exactement un sous-groupe de cardinal D' , pour tout diviseur D' de D .

Si G est infini, alors G est isomorphe à \mathbf{Z} , et tous les sous-groupes non nuls de G sont isomorphes à \mathbf{Z} , et donc cycliques.

Si G est fini de cardinal D , alors G est isomorphe à $\mathbf{Z}/D\mathbf{Z}$, et on sait que les sous-groupes de $\mathbf{Z}/D\mathbf{Z}$ sont de la forme $d\mathbf{Z}/D\mathbf{Z}$, pour d diviseur de D . Or $n \mapsto dn$ induit une surjection de \mathbf{Z} sur $d\mathbf{Z}/D\mathbf{Z}$ dont le noyau est $D'\mathbf{Z}$, où $D' = D/d$, ce qui montre que $d\mathbf{Z}/D\mathbf{Z} \cong \mathbf{Z}/D'\mathbf{Z}$. Comme $d \mapsto D' = D/d$ est une permutation de l'ensemble des diviseurs de D , cela permet de conclure.

4.3. Groupes abéliens finis

Soit \mathcal{P} l'ensemble des nombres premiers. D'après le théorème des restes chinois, si $D \in \mathbf{N} - \{0\}$, alors $\mathbf{Z}/D\mathbf{Z} \cong \bigoplus_{p \in \mathcal{P}} (\mathbf{Z}/p^{v_p(D)}\mathbf{Z})$. La somme ci-dessus est en fait une somme finie car $v_p(D) = 0$, sauf pour un nombre fini de nombres premiers. Ce résultat se généralise à tous les groupes abéliens finis sous la forme (cf. n° 5.3 du § 5 pour la démonstration).

Théorème 4.1. — (Kronecker, 1867) *Soit G un groupe abélien fini et, si $p \in \mathcal{P}$, soit G_p l'ensemble des éléments de G d'ordre une puissance de p .*

- (i) G_p est un sous-groupe de G , nul pour presque tout p , et $G = \bigoplus_{p \in \mathcal{P}} G_p$.
- (ii) Si $p \in \mathcal{P}$, il existe une suite finie d'entiers $a_{p,i} \geq 1$, décroissante et uniquement déterminée, telle que l'on ait $G_p \cong \bigoplus_i (\mathbf{Z}/p^{a_{p,i}}\mathbf{Z})$.

Remarque 4.2. — Avec les notations du théorème, on a $|G| = \prod_p \prod_i p^{a_i}$, et donc $|G|$ est un multiple de p^{a_i} , pour tous p et i , ce qui prouve que la multiplication par $|G|$ annule tout élément de G , puisqu'elle annule tous les $\mathbf{Z}/p^{a_{p,i}}\mathbf{Z}$. Autrement dit, dans un groupe commutatif, l'ordre d'un élément divise l'ordre du groupe (cas particulier du théorème de Lagrange).

Exercice 4.3. — Décomposer $(\mathbf{Z}/108\mathbf{Z})^*$ et $(\mathbf{Z}/200\mathbf{Z})^*$ sous la forme ci-dessus.

4.4. Le théorème de Lagrange et ses variantes

Si G est un groupe fini, et si H est un sous-groupe de G , alors $h \mapsto xh$ induit une bijection de H sur xH , ce qui fait que les classes à droite xH ont toutes le même cardinal $|H|$. Comme G est la réunion disjointe des xH , pour $x \in G/H$, on en déduit la formule

$$|G| = |G/H| \cdot |H|.$$

En particulier, $|H|$ divise $|G|$, ce qui se traduit par :

- Si G est un groupe fini, alors le cardinal de tout sous-groupe de G divise celui de G (théorème de Lagrange).

On peut spécialiser cela au sous-groupe engendré par un élément x de G : le cardinal de ce sous-groupe est, par définition, l'ordre de x , ce qui nous donne :

- Dans un groupe fini l'ordre d'un élément divise le cardinal du groupe.

Finalement, on remarque que $|G/H|$ aussi divise $|G|$. Si X est un ensemble sur lequel G agit, si O est une orbite, si $x \in O$, et si H est le stabilisateur de x , on sait que $O \cong G/H$. On en déduit que :

- Dans un ensemble sur lequel agit un groupe fini, le cardinal d'une orbite divise le cardinal du groupe et, plus précisément, le produit du cardinal de l'orbite par celui du stabilisateur d'un de ses éléments est égal au cardinal du groupe.

En particulier, en appliquant ceci à l'action de G sur lui-même par conjugaison intérieure, on obtient :

- Dans un groupe fini, le cardinal d'une classe de conjugaison divise le cardinal du groupe.
- Si X est un ensemble fini sur lequel agit un groupe fini G , on peut découper X en orbites pour cette action. Si on choisit un élément par orbite, et si on utilise l'isomorphisme $O_x \cong G/G_x$, où G_x est le stabilisateur de x , on obtient la *formule des classes* :

$$|X| = \sum_{x \in G \backslash X} |O_x| = |G| \cdot \sum_{x \in G \backslash X} \frac{1}{|G_x|}.$$

Exercice 4.4. — Montrer que tout élément $x \in \mathbf{F}_p^*$ vérifie $x^{p-1} = 1$. En déduire le petit théorème de Fermat (si $n \in \mathbf{Z}$, alors $n^p - n$ est divisible par p).

Exercice 4.5. — Soit K un corps fini commutatif⁽²²⁾. Montrer que le groupe K^* est cyclique (on pourra considérer le nombre de solutions de l'équation $x^p = 1$, pour p nombre premier divisant $|K^*|$ et utiliser le th. 4.1).

⁽²²⁾Cette hypothèse est en fait superflue car tout corps fini est commutatif (théorème de Wedderburn).

Exercice 4.6. — (i) Soit p un nombre premier. Montrer que, si $x \equiv 1 + p^k a \pmod{p^{k+1}}$, et si $k \geq 1$ ($k \geq 2$, si $p = 2$), alors $x^p \equiv 1 + p^{k+1} a \pmod{p^{k+2}}$. En déduire que $(1+p)^{p^{n-2}} \neq 1$ dans $(\mathbf{Z}/p^n \mathbf{Z})^*$, si $p \neq 2$ et $n \geq 2$, et que $(1+4)^{p^{n-3}} \neq 1$ dans $(\mathbf{Z}/2^n \mathbf{Z})^*$, si $n \geq 3$.

(ii) Soit N le noyau de la réduction modulo p de $(\mathbf{Z}/p^n \mathbf{Z})^*$ dans \mathbf{F}_p^* (dans $(\mathbf{Z}/4\mathbf{Z})^*$, si $p = 2$). Montrer que N est isomorphe à $\mathbf{Z}/p^{n-1} \mathbf{Z}$ (à $\mathbf{Z}/2^{n-2} \mathbf{Z}$, si $p = 2$).

(iii) En utilisant le résultat de l'ex. 4.5, montrer que $(\mathbf{Z}/p^n \mathbf{Z})^* \cong (\mathbf{Z}/(p-1)\mathbf{Z}) \oplus (\mathbf{Z}/p^{n-1} \mathbf{Z})$ en tant que groupe commutatif, si $p \neq 3$ et $n \geq 1$.

(iv) Montrer que $(\mathbf{Z}/2^n \mathbf{Z})^* \cong (\mathbf{Z}/2\mathbf{Z}) \oplus (\mathbf{Z}/2^{n-2} \mathbf{Z})$, si $n \geq 2$.

4.5. Le groupe symétrique S_n

4.5.1. Permutations

Si $n \in \mathbf{N} - \{0\}$, on note S_n le groupe des bijections de $\{1, \dots, n\}$. Un élément de S_n est une *permutation*. Si $\sigma \in S_n$, on définit le *support* de σ comme l'ensemble des $i \in \{1, \dots, n\}$ tels que $\sigma(i) \neq i$. Il est plus ou moins évident que *deux permutations de supports disjoints permutent entre elles*.

On peut représenter une permutation σ de S_n sous la forme d'une matrice à 2 lignes et n colonnes en mettant les nombres de 1 à n sur la première ligne et leurs images par σ juste en-dessous. Cette représentation est très commode pour faire le produit de deux permutations (en n'oubliant pas que c'est la matrice de droite qui agit en premier). Par exemple, si σ et τ sont les permutations de S_6 définies par $\sigma(1) = 2, \sigma(2) = 4, \sigma(3) = 5, \sigma(4) = 6, \sigma(5) = 1$ et $\sigma(6) = 3$, et $\tau(1) = 4, \tau(2) = 2, \tau(3) = 1, \tau(4) = 6, \tau(5) = 5$ et $\tau(6) = 3$, alors

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 5 & 6 & 1 & 3 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 1 & 6 & 5 & 3 \end{pmatrix} \quad \text{et} \quad \sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 5 & 6 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 1 & 6 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 4 & 2 & 3 & 1 & 5 \end{pmatrix}.$$

Une permutation $\sigma \in S_n$ est un *k-cycle* s'il existe i_1, \dots, i_k distincts, tels que

$$\sigma(i_1) = i_2, \sigma(i_2) = i_3, \dots, \sigma(i_k) = i_1, \quad \text{et} \quad \sigma(j) = j, \text{ si } j \notin \{i_1, \dots, i_k\}.$$

On note (i_1, i_2, \dots, i_k) le *k-cycle* défini ci-dessus ; son support est l'ensemble $\{i_1, \dots, i_k\}$; il est d'ordre k . On remarquera que le *k-cycle* (i_1, i_2, \dots, i_k) est aussi égal au *k-cycle* $(i_a, i_{a+1}, \dots, i_{a+k-1})$, si on écrit les indices modulo k , et a est n'importe quel élément de $\mathbf{Z}/k\mathbf{Z}$. Pour rétablir une unicité de l'écriture, il suffit d'imposer que i_1 soit le plus petit élément de $\{i_1, \dots, i_k\}$. Il est commode d'étendre la notation ci-dessus aux « cycles de longueur 1 » (qui sont tous égaux à l'identité...).

- Une permutation peut s'écrire comme un produit de cycles de supports disjoints.

Si σ est une permutation, on fabrique une partition de $\{1, \dots, n\}$ en prenant les orbites O_1, \dots, O_s sous l'action de σ (i.e. sous l'action du sous-groupe cyclique de S_n engendré par σ). Si O_i est une de ces orbites, de cardinal k_i , on peut considérer le cycle $c_i = (a, \sigma(a), \dots, \sigma^{k_i-1}(a))$, où a est le plus petit élément de O_i ; c'est un cycle de longueur k_i et de support O_i , et σ est le produit des c_i , pour $i \in \{1, \dots, s\}$.

Comme des cycles ayant des supports deux à deux disjoints commutent entre eux, on peut faire le produit dans n'importe quel ordre dans la décomposition d'une permutation en cycles de supports disjoints. Par exemple, soit $\sigma \in S_6$ la permutation $\sigma(1) = 3, \sigma(2) = 2, \sigma(3) = 5, \sigma(4) = 6, \sigma(5) = 1$ et $\sigma(6) = 4$. Alors on a $\sigma = (1, 3, 5)(4, 6)(2) = (4, 6)(2)(1, 3, 5) \dots$ Très souvent on omet les cycles de longueur 1 de la décomposition ; on écrit plutôt la permutation précédente sous la forme $\sigma = (1, 3, 5)(4, 6)$ ou $\sigma = (4, 6)(1, 3, 5)$.

- Tout élément de S_n est conjugué à un unique élément de la forme

$$(1, \dots, \ell_1)(\ell_1 + 1, \dots, \ell_1 + \ell_2) \cdots (\ell_1 + \cdots + \ell_{s-1} + 1, \dots, \ell_1 + \cdots + \ell_{s-1} + \ell_s),$$

où (ℓ_1, \dots, ℓ_s) est une *partition* de n , c'est-à-dire une suite décroissante d'entiers ≥ 1 dont la somme est n . Les classes de conjugaison de S_n sont donc en bijection naturelle avec les partitions de n .

Soit $\sigma \in S_n$. La conjugaison $\sigma \mapsto \alpha\sigma\alpha^{-1}$ par un élément α de S_n transforme un k -cycle $i_1 \mapsto i_2 \mapsto \cdots \mapsto i_k \mapsto i_1$, en le k -cycle $\alpha(i_1) \mapsto \alpha(i_2) \mapsto \cdots \mapsto \alpha(i_k) \mapsto \alpha(i_1)$. On en déduit en particulier que les longueurs des cycles apparaissant dans les décompositions de deux permutations conjuguées sont les mêmes, ce qui implique l'unicité d'un conjugué de la forme voulue, les ℓ_j étant les longueurs des cycles apparaissant dans la décomposition de σ rangées dans l'ordre décroissant. Écrivons donc σ comme un produit de cycles $\tau_1 \dots \tau_s$ à supports disjoints. Soit ℓ_j la longueur de τ_j . On a $\ell_1 + \cdots + \ell_s = n$, et quitte à permuter les τ_j , on peut supposer que $\ell_1 \geq \ell_2 \geq \cdots \geq \ell_s$. On peut alors écrire τ_j sous la forme $\tau_j = (i_{\ell_1 + \cdots + \ell_{j-1} + 1}, \dots, i_{\ell_1 + \cdots + \ell_{j-1} + \ell_j})$, et $k \mapsto i_k$ nous définit une permutation α de $\{1, \dots, n\}$, car les supports des τ_j forment une partition de $\{1, \dots, n\}$. Alors $\alpha^{-1}\sigma\alpha$ est un conjugué de σ de la forme voulue, ce qui permet de conclure.

- Un 2-cycle est appelé une *transposition*, et S_n est engendré par les transpositions ; plus précisément, tout élément de S_n est produit de moins de $n - 1$ transpositions.

La démonstration se fait par récurrence sur n . Pour $n = 1$ (et $n = 2$), le résultat est immédiat. Si $n \geq 2$, et si $\sigma \in S_n$ vérifie $\sigma(n) \neq n$, alors $\tau = (\sigma(n), n)$ est une transposition et $\tau\sigma$ fixe n , et donc peut être vu comme un élément de S_{n-1} . D'après l'hypothèse de récurrence, $\tau\sigma$ est un produit de moins de $n - 2$ transpositions à support dans $\{1, \dots, n - 1\}$, et donc $\sigma = \tau(\tau\sigma)$ est un produit de moins de $n - 1$ transpositions. On en déduit le résultat, le cas $\sigma(n) = n$ se traitant directement.

Exercice 4.7. — Calculer $(1, 2)(2, 3)(3, 4)(4, 5)$ dans S_5 .

Exercice 4.8. — Montrer que S_n est engendré par les transpositions $(1, 2), (2, 3), \dots, (n - 1, n)$.

Exercice 4.9. — Soit $\sigma \in S_n$ dont la décomposition en cycles disjoints est $\tau_1 \cdots \tau_s$, chaque τ_i étant de longueur ℓ_i . Quel est l'ordre de σ ?

Exercice 4.10. — (i) Combien y-a-t'il de cycles de longueur k dans S_n .

(ii) Montrer que le nombre moyen de cycles dans la décomposition d'un éléments de S_n tend vers l'infini avec n . (On pourra commencer par se demander dans combien de permutations un cycle donné apparaît.)

Exercice 4.11. — (difficile mais très surprenant) Le DGAE voulant tester le niveau de compréhension des X a décidé de les mettre à l'épreuve. Pour ce faire, il réunit les 500 membres de la promotion dans l'amphi Poincaré et leur tient ce langage : « J'ai disposé dans l'amphi Arago vos 500 noms dans des casiers numérotés de 1 à 500, à raison d'un par casier. Je vais vous appeler un par un, et demander à chacun d'entre vous d'ouvrir des casiers un par un à la recherche de son nom puis de les refermer sans changer le contenu et de regagner sa chambre sans possibilité de communiquer quoi que ce soit à ses camarades restés dans l'amphi Poincaré. Si tout le monde trouve son nom dans les 250 premiers casiers qu'il a ouverts, vous pouvez partir en vacances. Si l'un d'entre vous ne trouve pas son nom, on recommence le jour suivant (et je change le contenu des casiers bien évidemment). Voilà, vous avez deux heures pour concevoir une stratégie. » Désespoir des X qui se rendent compte que chacun a une chance sur deux de tomber sur son nom, et qu'au total ils ont une chance sur 2^{500} de partir en vacances au bout d'un jour, et donc qu'ils ne partiront pas en vacances. Pourtant au bout d'un certain temps, l'un de nos X déclare : « pas de panique, avec un peu de discipline, on a 9 chances sur 10 de partir en vacances avant la fin de la semaine ». Saurez-vous retrouver son raisonnement ?

4.5.2. Signature d'une permutation

Si $\sigma \in S_n$, on définit la *signature* $\text{sign}(\sigma)$ de σ par la formule

$$\text{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j}.$$

- $\text{sign} : S_n \rightarrow \{\pm 1\}$ est un morphisme de groupes.

Si $\sigma, \tau \in S_n$, on a

$$\text{sign}(\sigma\tau) = \prod_{1 \leq i < j \leq n} \frac{\sigma\tau(i) - \sigma\tau(j)}{i - j} = \left(\prod_{1 \leq i < j \leq n} \frac{\sigma\tau(i) - \sigma\tau(j)}{\tau(i) - \tau(j)} \right) \left(\prod_{1 \leq i < j \leq n} \frac{\tau(i) - \tau(j)}{i - j} \right).$$

Le second terme est égal à $\text{sign}(\tau)$, et le premier à $\text{sign}(\sigma)$ car on a $\frac{\sigma\tau(i) - \sigma\tau(j)}{\tau(i) - \tau(j)} = \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)}$, ce qui permet d'écrire le produit sous la forme $\prod_{1 \leq \tau(i) < \tau(j) \leq n} \frac{\sigma(\tau(i)) - \sigma(\tau(j))}{\tau(i) - \tau(j)} = \text{sign}(\sigma)$.

- Si τ est un k -cycle, alors $\text{sign}(\tau) = (-1)^{k-1}$.

On a $\text{sign}(\alpha\sigma\alpha^{-1}) = \text{sign}(\alpha)\text{sign}(\sigma)\text{sign}(\alpha)^{-1} = \text{sign}(\sigma)$, ce qui prouve que la signature est invariante par conjugaison et donc que tous les k -cycles ont même signature. Cela permet de prendre $\tau = (n-1, n)$ pour calculer la signature d'une transposition. On a alors

$$\begin{aligned} \text{sign}(\tau) &= \left(\prod_{1 \leq i < j \leq n-2} \frac{\tau(i) - \tau(j)}{i - j} \right) \left(\prod_{i \leq n-2} \frac{\tau(i) - \tau(n-1)}{i - (n-1)} \right) \left(\prod_{i \leq n-2} \frac{\tau(i) - \tau(n)}{i - n} \right) \cdot \frac{\tau(n-1) - \tau(n)}{(n-1) - n} \\ &= \left(\prod_{i \leq n-2} \frac{i - n}{i - (n-1)} \right) \left(\prod_{i \leq n-2} \frac{i - (n-1)}{i - n} \right) \cdot (-1) = -1, \end{aligned}$$

ce qui prouve le résultat pour une transposition. Maintenant, le k -cycle $\sigma_k = (i_1, \dots, i_k)$ est le produit des transpositions $(i_1, i_2) \cdots (i_{k-1}, i_k)$, et comme il y en a $k-1$, on a $\text{sign}(\sigma_k) = (-1)^{k-1}$, ce qu'on cherchait à démontrer.

Exercice 4.12. — Montrer que $\text{sign}(\sigma) = (-1)^{n-\omega(\sigma)}$, où $\omega(\sigma)$ est le nombre d'orbites de σ .

Exercice 4.13. — Si $\sigma \in S_n$, on note u_σ l'endomorphisme de \mathbf{C}^n envoyant un élément e_i de la base canonique de \mathbf{C}^n sur $e_{\sigma(i)}$.

- (i) Montrer que $\sigma \mapsto u_\sigma$ est un morphisme de groupes de S_n dans $GL(\mathbb{C}^n)$.
- (ii) Montrer que si τ est une transposition, alors u_τ est une symétrie par rapport à un hyperplan que l'on déterminera. Que vaut $\det u_\tau$?
- (iii) En déduire que $\det u_\sigma = \text{sign}(\sigma)$ pour tout $\sigma \in S_n$.

4.5.3. Groupe alterné

On définit le *groupe alterné* A_n comme le noyau de la signature. Un k -cycle est donc dans A_n , si et seulement si k est impair.

- A_n est engendré par les 3-cycles.

Cela se démontre par récurrence sur n . Le résultat est évident (et vide) si $n \leq 2$. Soit $n \geq 3$, et soit $\sigma \in A_n$. Si $\sigma(n) \neq n$, on peut choisir un 3-cycle $(n, \sigma(n), c)$, où $c \notin \{n, \sigma(n)\}$, et alors $\tau^{-1}\sigma$ fixe n et peut être écrit comme un produit de 3-cycles à support dans $\{1, \dots, n-1\}$ d'après l'hypothèse de récurrence; donc $\sigma = \tau(\tau^{-1}\sigma)$ peut être écrit comme un produit de 3-cycles. On en déduit le résultat, le cas $\sigma(n) = n$ se traitant directement.

- Si $n \geq 5$, tous les 3-cycles sont conjugués dans A_n .

Il suffit de prouver qu'ils sont tous conjugués à $\sigma_0 = (1, 2, 3)$. Soit σ un 3-cycle. Comme les 3-cycles sont tous conjugués dans S_n , il existe $\alpha \in S_n$ tel que $\sigma = \alpha\sigma_0\alpha^{-1}$. Si $\alpha \in A_n$, on a gagné. Sinon, $\tau = (4, 5)$ commute avec σ_0 puisque leurs supports sont disjoints, et $\beta = \alpha\tau \in A_n$ vérifie $\beta\sigma_0\beta^{-1} = \alpha\tau\sigma_0\tau^{-1}\alpha^{-1} = \alpha\sigma_0\alpha^{-1} = \sigma$, ce qui montre que σ est conjugué à σ_0 dans A_n .

- Le groupe A_5 est un groupe simple.

Soit H un sous-groupe distingué de A_5 non réduit à l'identité. On veut prouver que $H = A_5$ et il suffit de prouver que H contient un 3-cycle, car ceci implique qu'il contient tous les 3-cycles puisque ceux-ci sont conjugués dans A_5 , et donc $H = A_5$ puisque les 3-cycles engendrent A_5 .

Soit donc $\sigma \in H - \{1\}$. Il y a trois possibilités : σ est un 3-cycle et il n'y a rien à faire, ou σ est un 5-cycle ou c'est le produit de 2 transpositions de supports disjoints.

- Si σ est le 5-cycle (a, b, c, d, e) , soit $\tau = (a, b, c)$. Alors H contient $\tau^{-1}\sigma^{-1}\tau$ puisqu'il est distingué et donc aussi $h = \sigma\tau^{-1}\sigma^{-1}\tau$. Or τ^{-1} est le 3-cycle (c, b, a) et $\sigma\tau^{-1}\sigma^{-1}$ est le 3-cycle $(\sigma(c), \sigma(b), \sigma(a)) = (d, c, b)$. Donc $h = (d, c, b)(a, b, c)$ laisse fixe e et $a \mapsto b \mapsto d$, $b \mapsto c \mapsto b$, $c \mapsto a \mapsto a$, et $d \mapsto d \mapsto c$; c'est donc le 3-cycle (a, d, c) .

- Si $\sigma = \sigma_1\sigma_2$, avec $\sigma_1 = (a, b)$, $\sigma_2 = (c, d)$, et a, b, c, d distincts deux à deux, et si $\tau = (c, d, e)$, où $e \notin \{a, b, c, d\}$, alors H contient $h = \sigma\tau^{-1}\sigma^{-1}\tau$. Or σ_1 commute à σ_2 et τ , donc $h = \sigma_2\tau^{-1}\sigma_2^{-1}\tau$. Maintenant, $\tau^{-1}\sigma_2^{-1}\tau$ est la transposition $(\tau^{-1}(c), \tau^{-1}(d)) = (e, c)$ et donc $h = (c, d)(e, c) = (c, e, d)$ est un 3-cycle.

- Si $n \geq 5$, le groupe A_n est un groupe simple⁽²³⁾.

Soit $n \geq 5$, soit H un sous-groupe distingué de A_n , et soient $\sigma \neq \text{id}$ un élément de H et $\tau = (a, b, c)$ un 3-cycle. Alors H contient $h = \tau\sigma\tau^{-1}\sigma^{-1}$ qui est le produit des 3-cycles τ et $\sigma\tau^{-1}\sigma^{-1} = (\sigma(c), \sigma(b), \sigma(a))$. Soit alors $b = \sigma(a)$, et soit $c \notin \{a, \sigma(a), \sigma^2(a)\}$, non fixé par σ si jamais σ échange a et $\sigma(a)$ (un tel c existe toujours, sinon σ serait une transposition, ce qui est impossible puisque $\sigma \in A_n$). La condition mise sur c fait que $h \neq \text{id}$, et celle mise sur

⁽²³⁾Ce résultat, conjugué avec la théorie de Galois, explique que l'on ne puisse pas trouver de formule générale donnant les racines d'une équation de degré n , si $n \geq 5$.

b implique que le support de h est inclus dans $\{a, \sigma(a), \sigma^2(a), c, \sigma(c)\}$ et donc comporte au plus 5 éléments. Soit X de cardinal 5 contenant le support de h , et soit $\text{Perm}(X)$ le groupe des permutations de X . Alors $H \cap \text{Perm}(X)$ est un sous-groupe distingué de $\text{Perm}(X)$, et donc contient un 3-cycle d'après l'étude du cas $n = 5$. On en déduit que $H = A_n$ comme-ci-dessus puisque A_n est engendré par les 3-cycles qui sont tous conjugués dans A_n .

Exercice 4.14. — (i) Montrer que si G est un groupe fini abélien, et si d divise $|G|$, alors G a un sous-groupe de cardinal d . (On pourra utiliser le théorème de structure.)

(ii) Montrer que si $f : S_5 \rightarrow S_3$ est un morphisme de groupes, alors $\text{Im}(f)$ a 1 ou 2 éléments.

(iii) Montrer que S_5 n'a pas de sous-groupe d'ordre 40.

4.6. Les théorèmes de Sylow

Cauchy a démontré que, si G est un groupe fini d'ordre (*l'ordre d'un groupe* est, par définition, son cardinal) divisible par p , alors G contient un élément d'ordre p (et donc un sous-groupe (cyclique) d'ordre p). D'un autre côté, l'ordre d'un sous-groupe divisant l'ordre du groupe (théorème de Lagrange), tout sous- p -groupe (un p -groupe est un groupe dont l'ordre est une puissance de p) de G d'ordre p^a vérifie $a \leq v_p(|G|)$. Un p -Sylow de G est un sous-groupe d'ordre $p^{v_p(|G|)}$. (Dans le cas $v_p(|G|) = 0$, un tel sous-groupe est donc réduit à l'élément neutre.)

• Si G est un groupe commutatif d'ordre divisible par p , alors G contient un sous-groupe cyclique d'ordre p .

Si $x \in G$, soit n_x l'ordre de x . Par définition cela signifie que le morphisme de groupes de \mathbf{Z} dans G envoyant $a \in \mathbf{Z}$ sur x^a admet pour noyau $n_x \mathbf{Z}$, et donc induit un isomorphisme de $\mathbf{Z}/n_x \mathbf{Z}$ sur le sous-groupe de G engendré par x . Soit alors $X \subset G$ engendrant G (on peut prendre $X = G$ par exemple). Comme G est commutatif, l'application $\bigoplus_{x \in X} (\mathbf{Z}/n_x \mathbf{Z}) \rightarrow G$, envoyant $(a_x)_{x \in X}$ sur $\prod_{x \in X} x^{a_x}$ est un morphisme de groupes, et comme X engendre G , ce morphisme est surjectif. L'ordre de G est donc un diviseur de $\prod_{x \in X} n_x$. Comme p divise $|G|$, cela implique que p divise un des n_x , et alors $y = x^{n_x/p}$ est d'ordre p et le sous-groupe de G engendré par y est d'ordre p . Ceci permet de conclure.

Théorème 4.15. — (Sylow, 1872) *Si G est un groupe fini, l'ensemble des p -Sylow de G est non vide. De plus :*

- (i) *Tous les p -Sylow de G sont conjugués.*
- (ii) *Si Q est un sous- p -groupe de G , alors il existe un p -Sylow de G contenant Q ; en particulier, tout élément d'ordre p est contenu dans un p -Sylow de G .*

La démonstration se fait par récurrence sur $|G|$, le cas $|G| = 1$ étant évident (et vide). Soit Z le centre de G , et soit $k = v_p(|G|)$.

• Si p divise l'ordre de Z , alors Z contient, d'après le point précédent, un sous-groupe cyclique C d'ordre p . On peut appliquer l'hypothèse de récurrence à $H = G/C$ qui est d'ordre mp^{k-1} . Si P_H est un p -Sylow de H , l'image inverse de P_H dans G est un sous-groupe d'ordre $|P_H| \cdot |C| = p^{k-1}p = p^k$; c'est donc un p -Sylow de G .

• Si p ne divise pas $|Z|$, on fait agir G par conjugaison intérieure ($g \cdot x = gxg^{-1}$) sur G . Par définition du centre, les orbites (qui ne sont autres que les classes de conjugaison de G)

ne comportant qu'un seul élément pour cette action, sont exactement les $\{c\}$, pour $c \in Z$. Comme $|Z|$ est premier à p et comme $|G|$ est divisible par p , il y a une orbite O , non réduite à un élément, de cardinal premier à p . Si $x \in O$, et si H est l'ensemble des éléments de G commutant à x , on a $O = G/H$. On en déduit que $|H| = \frac{|G|}{|O|}$. Comme $v_p(|O|) = 0$, on a $v_p(|H|) = v_p(|G|) = k$, et comme $|O| > 1$, on a $|H| < |G|$. L'hypothèse de récurrence montre que H contient un sous-groupe d'ordre p^k , et donc que G aussi.

Maintenant, si P est un p -Sylow de G , et si Q est un sous- p -groupe de G , on peut faire agir Q sur G/P par translation à gauche. Comme G/P est de cardinal premier à p , puisque P est un p -Sylow, au moins une des orbites O a un cardinal premier à p . Mais O est de la forme Q/H , où H est un sous-groupe de Q , et comme Q est un p -groupe, on a $|Q/H|$ premier à p si et seulement si $H = Q$. Il existe donc $x \in G/P$ fixe par Q tout entier. En prenant un représentant \tilde{x} de x dans G , cela se traduit par $Q\tilde{x}P \subset \tilde{x}P$, ou encore par $Q \subset \tilde{x}P\tilde{x}^{-1}$.

Si Q est un p -Sylow, on en déduit que $Q = \tilde{x}P\tilde{x}^{-1}$ pour des raisons de cardinal, ce qui démontre le (i). Si Q est un sous- p -groupe quelconque, cela montre que Q est contenu dans un sous-groupe d'ordre p^k , c'est-à-dire dans un p -Sylow. Ceci démontre le (ii) et permet de conclure.

5. Algèbre linéaire

Dans le n° 5.1, on rappelle (et complète) sans démonstration les résultats vus en classe préparatoire concernant la réduction des endomorphismes (diagonalisation, mise sous forme de Jordan...). Au n° 5.2, on explique comment on peut retrouver ces résultats en utilisant le théorème de structure des modules de torsion sur les anneaux principaux démontré au n° 5.3. L'intérêt de cette nouvelle approche est de ne rien supposer sur le corps K , alors que l'approche vue en classe préparatoire impose plus ou moins à K d'être algébriquement clos (ce qui, il faut le reconnaître, est le cas de \mathbf{C} , mais est loin d'être celui de \mathbf{F}_2).

5.1. Généralités

Soit K un corps commutatif, et soit V un K -espace vectoriel de dimension finie.

5.1.1. Endomorphismes

On note $\text{End}(V)$ l'ensemble des *endomorphismes* de V , c'est-à-dire, l'ensemble des applications $u : V \rightarrow V$ qui sont linéaires. Muni de l'addition définie par $(u_1 + u_2)(v) = u_1(v) + u_2(v)$, et de la composition des endomorphismes, $\text{End}(V)$ est un anneau non commutatif (sauf en dimension 1), possédant un élément unité (que nous noterons 1) en la personne de l'application identité $\text{id} : V \rightarrow V$. L'*homothétie de rapport* λ est l'application $v \mapsto \lambda v$. On la note simplement λ , ce qui est compatible avec le fait que l'identité (que l'on a notée 1) peut aussi être vue comme l'homothétie de rapport 1.

5.1.2. Le théorème de Cayley-Hamilton

Si $u \in \text{End}(V)$, on note $\det(u) \in K$ son déterminant. On a $\det(u_1 u_2) = \det(u_1) \det(u_2)$, si $u_1, u_2 \in \text{End}(V)$. On note $\text{Car}_u(X)$ le *polynôme caractéristique* $\text{Car}_u(X) = \det(X - u)$

de u . Si V est de dimension d , c'est un polynôme de degré d , dont le développement est donné par

$$\text{Car}_u(X) = X^d - \text{Tr}(u)X^{d-1} + \cdots + (-1)^d \det(u),$$

où $\text{Tr}(u)$ est, par définition, la *trace* de u . On a $\text{Tr}(u_1 u_2) = \text{Tr}(u_2 u_1)$, si $u_1, u_2 \in \text{End}(V)$.

L'ensemble des $P \in K[X]$ tels que $P(u) = 0$ est un idéal de $K[X]$, non nul car $\text{End}(V)$ est de dimension $(\dim V)^2$ et donc $1, u, \dots, u^{(\dim V)^2}$ forment une famille liée. On note Min_u le générateur unitaire de cet idéal. C'est le *polynôme minimal* de u et, d'après le théorème de Cayley-Hamilton, Car_u annule u ; autrement dit, Car_u est un multiple de Min_u .

5.1.3. Automorphismes

Si $u \in \text{End}(V)$, le noyau $\text{Ker}(u) = \{v \in V, u(v) = 0\}$ et l'image $\text{Im}(u) = \{v \in V, \exists v' \in V, u(v') = v\}$ de u sont des sous-espaces vectoriels de V , et

$$\det u \neq 0 \Leftrightarrow \text{Ker}(u) = 0 \Leftrightarrow u \text{ est injectif} \Leftrightarrow u \text{ est bijectif} \Leftrightarrow u \text{ est surjectif} \Leftrightarrow \text{Im}(u) = V.$$

Un *automorphisme* de V est un élément de $\text{End}(V)$ vérifiant les conditions ci-dessus. On note $\text{GL}(V) \subset \text{End}(V)$ l'ensemble des automorphismes de V ; c'est le groupe des éléments inversibles de l'anneau $\text{End}(V)$.

5.1.4. Matrices

Si V est de dimension d , et si on choisit une base e_1, \dots, e_d de V , on peut associer à tout élément u de $\text{End}(V)$ sa matrice dans la base e_1, \dots, e_d . C'est l'élément $(a_{i,j})_{1 \leq i,j \leq d}$ de $\mathbf{M}_d(K)$ défini par $u(e_j) = \sum_{i=1}^d a_{i,j} e_i$. La trace de u est alors la somme $\sum_{i=1}^d a_{i,i}$ des coefficients diagonaux de la matrice de u ; cette somme ne dépend donc pas du choix de la base. Le groupe $\text{GL}(V)$ s'identifie au groupe $\mathbf{GL}_d(K)$ des matrices $d \times d$ inversibles (ce qui équivaut à ce que le déterminant soit non nul) à coefficients dans K .

5.1.5. Espaces propres, espaces caractéristiques

Soit $u \in \text{End}(V)$. On dit que $\lambda \in K$ est une *valeur propre* de u , si $u - \lambda$ n'est pas inversible, ce qui équivaut à $\text{Ker}(u - \lambda) \neq 0$, et donc à l'existence de $v \in V$, non nul, tel que $u(v) = \lambda v$; un tel v est un *vecteur propre* de u pour la valeur propre λ . Le *spectre* $\text{Spec}(u)$ de u est l'ensemble des valeurs propres de u . C'est aussi l'ensemble des racines du polynôme caractéristique $\text{Car}_u(X) = \det(X - u)$ de u .

Si $\lambda \in \text{Spec}(u)$, le noyau de $\text{Ker}(u - \lambda)$ est l'*espace propre* associé à la valeur propre λ . On dit que u est *diagonalisable*, si $V = \bigoplus_{\lambda \in \text{Spec}(u)} \text{Ker}(u - \lambda)$. Ceci équivaut à l'existence d'une base $(e_i)_{i \in I}$ de V (constituée de vecteurs propres) dans laquelle la matrice de u est une *matrice diagonale* (i.e. $a_{i,j} = 0$ si $i \neq j$). Le polynôme minimal de u est alors le produit des $(X - \lambda)$, pour $\lambda \in \text{Spec}(u)$; en particulier, tous ses zéros sont dans K et ces zéros sont simples. Réciproquement, *s'il existe $P \in K[X]$, dont tous les zéros sont simples et qui se scinde dans K , avec $P(u) = 0$, alors u est diagonalisable.*

Si $\lambda \in \text{Spec}(u)$, la suite des $\text{Ker}(u - \lambda)^k$ est croissante, et donc stationnaire (i.e. constante à partir d'un certain rang). On note e_λ le plus petit k tel que $\text{Ker}(u - \lambda)^{k'} =$

$\text{Ker}(u - \lambda)^k$, quel que soit $k' \geq k$. Alors $\text{Ker}(u - \lambda)^{e_\lambda}$ est le *sous-espace caractéristique* associé à λ . Si K est algébriquement clos, V est la somme directe $\bigoplus_{\lambda \in \text{Spec}(u)} V_\lambda$ de ses sous-espaces caractéristiques. On note d_λ la dimension de V_λ ; c'est la *multiplicité* de la valeur propre λ , et c'est aussi la multiplicité de λ en tant que racine de Car_u .

5.1.6. Mise sous forme de Jordan

Un *bloc de Jordan* $J_{\lambda,r}$ d'ordre r pour λ est une matrice $r \times r$ avec des λ sur la diagonale, des 1 juste au-dessus de la diagonale et des 0 partout ailleurs. Les polynômes minimal et caractéristique de $J_{\lambda,r}$ sont tous deux égaux à $(X - \lambda)^r$. Une matrice est *sous forme de Jordan* si elle est diagonale par blocs, et si chacun des blocs est un bloc de Jordan (on ne demande pas aux blocs d'être de la même taille, ni d'être associés au même λ).

On peut trouver une base de V_λ dans laquelle la matrice de u est sous forme de Jordan. La taille des blocs $r_{\lambda,1} \geq r_{\lambda,2} \geq \dots \geq r_{\lambda,k_\lambda}$ est alors indépendante du choix de la base, et on a $r_{\lambda,1} = e_\lambda$ et $\sum_{j=1}^{k_\lambda} r_{\lambda,j} = d_\lambda \geq e_\lambda$. En juxtaposant les bases des V_λ , pour $\lambda \in \text{Spec}(u)$, cela permet, si Car_u a toutes ses racines dans K (ce qui est automatique si K est algébriquement clos), de mettre la matrice de u sur V sous forme de Jordan. On en déduit que les polynômes minimal Min_u et caractéristique Car_u de u sont donnés par

$$\text{Min}_u(X) = \prod_{\lambda \in \text{Spec}(u)} (X - \lambda)^{e_\lambda} \quad \text{et} \quad \text{Car}_u(X) = \prod_{\lambda \in \text{Spec}(u)} (X - \lambda)^{d_\lambda}.$$

On déduit aussi de l'existence de la forme de Jordan que $\text{Tr}(u)$ (resp. $\det(u)$) est la somme (resp. le produit) des valeurs propres de u , comptées avec multiplicité.

5.2. Modules de torsion sur $K[T]$ et réduction des endomorphismes

5.2.1. Anneaux et modules. — Si A est un anneau (avec élément unité 1), un A -module M est un groupe commutatif A -module pour une loi $+$, muni d'une action $(a, x) \mapsto ax$ de A , vérifiant :

$$0x = 0, \quad 1x = x, \quad a(x + y) = ax + ay, \quad (a + b)x = ax + bx, \quad (ab)x = a(bx),$$

quels que soient $x, y \in M$ et $a, b \in A$.

- Si A est un corps commutatif, on retombe sur la définition d'un espace vectoriel, et il y a de grandes similarités entre la théorie des modules sur un anneau commutatif et celle des espaces vectoriels sur un corps commutatif.
- Tout groupe commutatif est naturellement un \mathbf{Z} -module, en définissant nx par récurrence sur n par $0x = 0$, $(n+1)x = nx + x$ si $n \in \mathbf{N}$, et $nx = -((-n)x)$, si $n \leq 0$. (Montrer que ceci définit bien une action de \mathbf{Z} est un exercice fastidieux qui n'est pas sans rappeler la démonstration du fait que \mathbf{Z} est un anneau en partant des axiomes de Peano.)
- Si A est commutatif, un sous- A -module de A n'est autre qu'un idéal de A .
- Si K est un corps commutatif, et si V est un K -espace vectoriel, alors V est un module sur l'anneau $\text{End}(V)$ (non commutatif si $\dim V \geq 2$).

- Si $(M_i)_{i \in I}$ est une famille de A -modules, alors les groupes commutatifs $\bigoplus_{i \in I} M_i$ et $\prod_{i \in I} M_i$ sont naturellement munis d'une action de A (comme dans le cas des espaces vectoriels), et sont donc des A -modules.
- Si $M' \subset M$ sont deux A -modules, le groupe commutatif quotient M/M' est muni d'une action de A et donc est un A -module.
- Un *morphisme* $u : M_1 \rightarrow M_2$ de A -modules est un morphisme de groupes additifs commutant à l'action de A (i.e. $u(ax) = au(x)$, si $x \in M_1$ et $a \in A$); si A est un corps commutatif, on retombe sur la définition d'une application linéaire entre espaces vectoriels.
- Si $u : M_1 \rightarrow M_2$ est un morphisme de A -modules, alors $\text{Ker } u$ et $\text{Im } u$ sont des A -modules, et u induit un isomorphisme de A -modules de $M_1/\text{Ker } u$ sur $\text{Im } u$.

Si M est un A -module et si les M_i , pour $i \in I$, sont des sous- A -modules de M , alors l'intersection des M_i est un A -module. Ceci permet de définir le *sous- A -module engendré* par une famille $(e_j)_{j \in J}$ d'éléments de M , comme l'intersection de tous les sous- A -modules de M contenant les e_j . Comme dans le cas des espaces vectoriels, ce module est l'ensemble des combinaisons linéaires finies, à coefficients dans A , en les e_j .

Un A -module M est *de type fini* si on peut trouver un ensemble fini e_1, \dots, e_d d'éléments de M tels que l'application $(a_1, \dots, a_d) \mapsto a_1 e_1 + \dots + a_d e_d$ soit une surjection de A^d sur M ; autrement dit, M est de type fini s'il admet une famille génératrice finie. Une différence essentielle avec le cas des espaces vectoriels sur un corps est qu'un A -module ne possède pas, en général, de base sur A .

Un A -module M est *de torsion* si, pour tout $x \in M$, on peut trouver $a \in A - \{0\}$, tel que $ax = 0$. Un A -module de torsion non nul est un exemple de module ne possédant pas de base puisque toute famille ayant plus d'un élément est liée. Si A est commutatif, un exemple typique de A -module de torsion est un module de la forme A/I , où I est un idéal de A distinct de $\{0\}$ et A ; par exemple, $\mathbf{Z}/D\mathbf{Z}$ est un \mathbf{Z} -module de torsion, si $D \geq 2$.

5.2.2. Structure des modules de torsion sur $K[T]$. — Soit K un corps commutatif. Comme le montre la discussion suivant le th 5.2 ci-dessous, un K -espace vectoriel de dimension finie muni d'un endomorphisme K -linéaire u est la même chose qu'un $K[T]$ -module de torsion et de type fini. Ce changement de point de vue est particulièrement intéressant à cause du théorème de structure (th. 5.2) ci-dessous, que le lecteur pourra comparer avec le théorème de structure (th. 4.1) pour les groupes finis abéliens (nous démontrons les deux simultanément au n° 5.3).

Un polynôme $P \in K[T]$ est dit *irréductible* s'il est de degré ≥ 1 et si on ne peut pas le factoriser sous la forme $P = Q_1 Q_2$, avec $Q_1, Q_2 \in K[T]$ et $\deg Q_1 \geq 1$, $\deg Q_2 \geq 1$. Un corps K est algébriquement clos si et seulement si les polynômes irréductibles de $K[T]$ sont de degré 1; les polynômes irréductibles de $\mathbf{R}[T]$ sont de degré 1 ou 2, ceux de $\mathbf{Q}[T]$

ou de $\mathbf{F}_p[T]$ ont des degrés arbitraires. On note $\mathcal{P}_{K[T]}$ l'ensemble des polynômes unitaires irréductibles de degré ≥ 1 .

Si $Q \in K[T]$, on note $K[T]/Q$ (au lieu de $K[T]/QK[T]$ ou $K[T]/(Q)$) le quotient de $K[T]$ par l'idéal engendré par Q .

Exercice 5.1. — Montrer que si $Q \in \mathcal{P}_{K[T]}$, alors $K[T]/Q$ est un corps.

Théorème 5.2. — Soit M un $K[T]$ -module de torsion et de type fini. Si $P \in \mathcal{P}_{K[T]}$, soit M_P l'ensemble des $x \in M$ tués par une puissance de P .

(i) M_P est un sous- $K[T]$ -module de M , nul sauf pour un nombre fini de $P \in \mathcal{P}_{K[T]}$, et $M = \bigoplus_{P \in \mathcal{P}_{K[T]}} M_P$.

(ii) Il existe $r_P \in \mathbf{N}$ et une unique famille décroissante d'entiers $a_{P,i} \geq 1$, tels que $M_P = \bigoplus_{1 \leq i \leq r_P} K[T]/P^{a_{P,i}}$.

5.2.3. Exemples. — Soit M un $K[T]$ -module de torsion et de type fini, et soient e_1, \dots, e_d engendrant M . Par définition, cela veut dire que l'application $(x_1, \dots, x_d) \mapsto x_1 e_1 + \dots + x_d e_d$ de $(K[T])^d$ dans M est surjective. Par ailleurs, si $P_i \in K[T] - \{0\}$, pour $i \in \{1, \dots, d\}$, vérifie $P_i e_i = 0$ (de tels P_i existent puisque M est de torsion), alors le noyau de l'application précédente contient $(P_1) \times \dots \times (P_d)$, et donc M est un quotient de $K[T]/P_1 \times \dots \times K[T]/P_d$, qui est un K -espace vectoriel de dimension finie $\deg P_1 \dots \deg P_d$. On en déduit que M est un K -espace vectoriel de dimension finie. De plus, la multiplication par T sur M est K -linéaire, ce qui munit M d'un élément privilégié u_M de $\text{End}(M)$.

Réciproquement, si V est un K -espace vectoriel de dimension finie, et si u est un endomorphisme de V , alors $P \mapsto P(u)$ induit un morphisme d'anneaux de $K[T]$ dans $\text{End}(V)$. Comme V est un $\text{End}(V)$ -module, cela muni V d'une action de $K[T]$ (où $P \in K[T]$ agit par $P(u) \in \text{End}(V)$), ce qui permet de voir V comme un $K[T]$ -module; par construction, on a $u_V = u$. De plus, le $K[T]$ -module V est de torsion car $\text{Min}_u \in K[T]$ tue tous les éléments de V puisque, par définition, Min_u agit par $\text{Min}_u(u)$ sur V , et que $\text{Min}_u(u) = 0$.

Exemple 5.3. — (Modules cycliques) Soit $Q = T^d + a_{d-1}T^{d-1} + \dots + a_0 \in K[T]$, avec $d \geq 1$, et soit $M = K[T]/Q$. Alors la matrice de u_M dans la base $1, T, \dots, T^{d-1}$ est

$$A_Q = \begin{pmatrix} 0 & \dots & 0 & -a_0 \\ 1 & \ddots & \vdots & -a_1 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \dots & 1 & -a_{d-1} \end{pmatrix}$$

et les polynômes minimal et caractéristique de u_M sont tous deux égaux à Q .

En effet, par construction $Q(T)$ est la multiplication par 0 sur M , et donc $Q(u_M) = 0$, ce qui implique que le polynôme minimal de u_M divise Q . Par ailleurs, si $P(u_M) = 0$, alors en particulier, $P(u_M) \cdot 1 = P(T)$ est nul dans $M = K[T]/Q$, et donc P est un multiple de Q . Ceci prouve que le polynôme minimal de u_M est bien Q .

Le polynôme caractéristique de u_M , qui n'est autre que le déterminant de $X - u_M$, peut se calculer en développant par rapport à la dernière colonne. Le coefficient de $X + a_{d-1}$ est le déterminant d'une matrice $(d-1) \times (d-1)$, triangulaire inférieure, avec des X sur la diagonale, et donc est égal à X^{d-1} . Si $i \geq 2$, le coefficient de a_{d-i} est $(-1)^{i-1} \times$ le déterminant d'une matrice diagonale par blocs, un des blocs de dimension $(d-i) \times (d-i)$ étant triangulaire inférieur avec des X sur la diagonale, et l'autre, de dimension $(i-1) \times (i-1)$, étant triangulaire supérieur avec des -1 sur la diagonale ; il est donc égal à $(-1)^{i-1} X^{d-i} (-1)^{i-1} = X^{d-i}$. On en déduit que

$$\det(X - u_M) = (X + a_{d-1})X^{d-1} + a_{d-2}X^{d-2} + \cdots + a_0 = Q(X).$$

Exemple 5.4. — (Modules nilpotents) Soit $\lambda \in K$, et soit $M = K[T]/(X - \lambda)^d$. Alors la matrice de u_M dans la base $f_1 = (X - \lambda)^{d-1}, f_2 = (X - \lambda)^{d-2}, \dots, f_d = 1$ est un bloc de Jordan $J_{\lambda,d}$. En effet, on a $X(X - \lambda)^{d-i} = (X - \lambda)^{d-(i-1)} + \lambda(X - \lambda)^{d-i}$, ce qui se traduit par $u_M(f_i) = f_{i-1} + \lambda f_i$, si $i \neq 1$, et par $u_M(f_1) = \lambda f_1$ car $(X - \lambda)^d = 0$ dans M .

5.2.4. Application à la réduction des endomorphismes

Corollaire 5.5. — Si K est algébriquement clos, si V est un K -espace vectoriel de dimension finie, et si $u \in \text{End}(V)$, alors il existe une base de V dans laquelle la matrice de u est sous forme de Jordan.

Démonstration. — On peut supposer que V est un $K[T]$ -module de torsion et de type fini, et que u est la multiplication par T . Maintenant, comme K est algébriquement clos, $\mathcal{P}_{K[T]}$ est l'ensemble des $T - \lambda$, avec $\lambda \in K$. On déduit du th. 5.2 une décomposition de V sous la forme $V = \bigoplus_{i \in I} K[T]/(T - \lambda_i)^{a_i}$ (dans laquelle plusieurs λ_i peuvent être égaux). On conclut en utilisant le résultat de l'exemple 5.4, selon lequel la matrice de la multiplication par T sur $K[T]/(T - \lambda_i)^{a_i}$ peut se mettre sous forme de Jordan.

Remarque 5.6. — La décomposition de V sous la forme $\bigoplus_{\lambda} V_{T-\lambda}$ fournie par le (i) du th. 5.2 n'est autre que la décomposition de V comme somme directe de sous-espaces caractéristiques.

Lemme 5.7. — Soit $(Q_i)_{i \in I}$ une famille finie d'éléments de $K[T]$ de degrés ≥ 1 . Si $M = \bigoplus_{i \in I} K[T]/Q_i$, alors le polynôme minimal de u_M est le ppcm des Q_i , pour $i \in I$, et le polynôme caractéristique de u_M est le produit des Q_i , pour $i \in I$.

Démonstration. — Le polynôme minimal de u_M doit en particulier annuler $K[T]/Q_i$ pour tout i ; il doit donc être divisible par Q_i d'après les résultats de l'exemple 5.3, et donc aussi par le ppcm des Q_i . Réciproquement, le ppcm des Q_i est divisible par Q_i ; il annule donc $K[T]/Q_i$ pour tout i et est un multiple du polynôme minimal de u_M ; d'où le résultat en ce qui concerne le polynôme minimal de u_M .

Pour calculer le polynôme caractéristique de u_M , on remarque que chaque $K[T]/Q_i$ est stable par u_M , et donc que la matrice de u_M est diagonale par blocs, avec un bloc pour

chaque i correspondant à l'action de u_M sur $K[T]/Q_i$. Comme le polynôme caractéristique d'une matrice diagonale par blocs est le produit des polynômes caractéristiques des blocs, les résultats de l'exemple 5.3 permettent de conclure.

Corollaire 5.8. — (Cayley-Hamilton) *Si V est un K -espace vectoriel de dimension finie, et si $u \in \text{End}(V)$, le polynôme minimal de u divise le polynôme caractéristique de u .*

Démonstration. — On peut supposer que V est un $K[T]$ -module de torsion et de type fini, et que u est la multiplication par T . Reprenons les notations du théorème 5.2. Si on note $\text{Spec}(u)$ l'ensemble des $P \in \mathcal{P}_{K[T]}$ tels que $V_P \neq 0$, on déduit du lemme 5.7 que

$$\text{Min}_u(X) = \prod_{P \in \text{Spec}(u)} P^{a_{P,1}}, \quad \text{et} \quad \text{Car}_u(X) = \prod_{P \in \text{Spec}(u)} P^{a_{P,1} + \dots + a_{P,r_P}},$$

ce qui permet de conclure.

5.3. Modules de torsion sur les anneaux principaux

5.3.1. Généralités sur les idéaux

Dans tout ce qui suit, les anneaux considérés sont supposés commutatifs. Un anneau A est *intègre* s'il n'est pas réduit à 0 (i.e. si $0 \neq 1$ dans A), et s'il ne possède pas de *diviseur de 0* (i.e. $xy = 0 \Rightarrow x = 0$ ou $y = 0$). Un idéal I de A est dit *premier* si l'anneau A/I est intègre, ce qui équivaut, en remontant dans A , à « $I \neq A$ et $xy \in I \Rightarrow x \in I$ ou $y \in I$ ». En particulier, l'*idéal nul* $\{0\}$ est premier si et seulement si A est intègre.

Lemme 5.9. — *Les conditions suivantes sont équivalentes pour un idéal I d'un anneau A .*

- (i) A/I est un corps.
- (ii) Si $x \in A - I$, alors l'idéal engendré par I et x contient 1.
- (iii) Les seuls idéaux de A contenant I , sont A et I .

Démonstration. — Si I vérifie (iii) et si $x \notin A$, alors l'idéal engendré par A et x contient strictement A et donc est égal à A ; en particulier, il contient 1, ce qui démontre l'implication (iii) \Rightarrow (ii).

Si I vérifie (ii), et si $x \notin I$, alors il existe $b \in I$ et $u \in A$ tels que $b + ux = 1$. On en déduit que x est inversible dans A/I d'inverse u , et donc que tout élément non nul de A/I est inversible; autrement dit, A/I est un corps. D'où l'implication (ii) \Rightarrow (i).

Finalement, si A/I est un corps, et si J est un idéal de A contenant I , alors J/I est un idéal de A/I , et donc est soit réduit à 0 (ce qui implique $J = I$), soit égal à A/I (ce qui implique $J = A$). On en déduit l'implication (i) \Rightarrow (iii), ce qui permet de conclure.

Un idéal satisfaisant les propriétés du lemme est dit *maximal*. Un corps étant intègre, un idéal maximal est premier, mais la réciproque est fausse. Par exemple, l'idéal (X) de $\mathbf{Z}[X]$ est premier puisque $\mathbf{Z}[X]/(X) = \mathbf{Z}$ est intègre, mais il n'est pas maximal puisque \mathbf{Z} n'est pas un corps.

5.3.2. Anneaux principaux. — Si A est un anneau, un idéal de A est *principal* s'il est engendré par un élément. Un *anneau principal* est un anneau intègre dans lequel tout idéal est principal.

Par exemple, \mathbf{Z} est un anneau principal. En effet, un idéal est en particulier un sous-groupe pour l'addition, et on a vu que tout sous-groupe de \mathbf{Z} est de la forme $D\mathbf{Z}$, avec $D \in \mathbf{N}$; c'est donc aussi un idéal principal, et tout idéal de \mathbf{Z} est principal.

De même, $K[T]$ est un anneau principal. En effet, soit I un idéal de $K[T]$ non réduit à 0, et soit $B \in I - \{0\}$ de degré minimal. Soit $P \in I$, et soit R le reste de la division euclidienne de P par B . Alors $R = P - BQ \in I$ puisque $P \in I$ et $B \in I$, et $\deg R < \deg B$ par définition du reste. Ceci implique que $R = 0$, par construction de B , et donc P est un multiple de B et $I = (B)$ est principal.

Proposition 5.10. — *Si A est un anneau principal, et si I est un idéal premier non nul de A , alors A/I est un corps. Autrement dit, tout idéal premier non nul d'un anneau principal est maximal.*

Démonstration. — Soit J un idéal de A contenant strictement I . Soit a un générateur de J et p un générateur de I . Comme $I \subset J$, il existe $b \in A$ tel que $p = ab$. Comme $J \neq I$, on a $a \notin I$, et comme I est premier, l'égalité $p = ab$ implique que $b \in I$, et donc qu'il existe $c \in A$ tel que $b = pc$. On a alors $p(1 - ac) = 0$, et comme A est intègre et $p \neq 0$, cela implique que a est inversible dans A d'inverse c , et donc que $J = A$. On en déduit que I est maximal, ce qui permet de conclure.

Lemme 5.11. — *Toute suite croissante d'idéaux de A est stationnaire. (Un anneau vérifiant cette propriété est dit noethérien, et donc un anneau principal est noethérien.)*

Démonstration. — Soit $(I_n)_{n \in \mathbf{N}}$ une suite croissante d'idéaux de A , et soit $I = \bigcup_{n \in \mathbf{N}} I_n$. Si $a, b \in I$, il existe $n, m \in \mathbf{N}$ tels que $a \in I_n$ et $b \in I_m$, et comme la suite est croissante, a et b appartiennent à $I_{\sup(n,m)}$, et donc $a + b \in I_{\sup(n,m)} \subset I$. Comme I est aussi stable par multiplication par $\lambda \in A$, cela montre que I est un idéal. Maintenant, I est principal puisqu'on a supposé A principal; il est donc de la forme (λ) , pour un certain $\lambda \in I$, et il existe $n \in \mathbf{N}$ tel que $\lambda \in I_n$. On a alors $(\lambda) \subset I_n \subset I = (\lambda)$, ce qui montre que $I_n = I$, quel que soit $m \geq n$. Ceci permet de conclure.

Lemme 5.12. — *Tout idéal propre de A est contenu dans un idéal maximal.*

Démonstration. — Supposons le contraire. Soit $I \neq A$ un idéal de A contenu dans aucun idéal maximal. En particulier, I n'est pas maximal et il existe $I_1 \neq A$ contenant strictement I . Alors I_1 n'est contenu dans aucun idéal maximal, sinon un idéal maximal qui contiendrait I_1 contiendrait aussi I , ce qui permet de réitérer le processus, et donc de construire une suite strictement croissante $(I_n)_{n \in \mathbf{N}}$ d'idéaux de A . Comme ceci est contraire au lemme 5.11, cela permet de conclure.

Lemme 5.13. — *Si $b \in A - \{0\}$, et si p est premier et divise b , alors l'idéal (b/p) contient strictement (b) .*

Démonstration. — Supposons le contraire. Il existe alors $a \in A$ tel que $b/p = ba$, et donc $b(1 - ap) = 0$. Comme A est intègre, cela implique que p est inversible dans A d'inverse a , ce qui est contraire à l'hypothèse selon laquelle p est premier. Ceci permet de conclure.

On dit que a et b sont *premiers entre eux*, si l'idéal (a, b) de A engendré par a et b est égal à A , ce qui équivaut à l'existence de $u, v \in A$ tels que $au + bv = 1$ puisque $(a, b) = \{au + bv, u, v \in A\}$, et qu'un idéal de A contenant 1 est égal à A . On écrit souvent $(a, b) = 1$, pour dire que a et b sont premiers entre eux.

Lemme 5.14. — (lemme de Gauss)

- (i) *Si a est premier avec b et c , alors a est premier avec bc .*
- (ii) *Si a divise bc et si a est premier avec b , alors a divise c .*

Démonstration. — Si $(a, b) = (a, c) = 1$, il existe u_1, v_1 tels que $au_1 + bv_1 = 1$ et u_2, v_2 tels que $au_2 + cv_2 = 1$. On a donc $1 = (au_1 + bv_1)(au_2 + cv_2) = au + bcv$, avec $u = au_1u_2 + bv_1u_2 + cu_1v_2$ et $v = v_1v_2$, ce qui prouve que $(a, bc) = 1$. On en déduit le premier énoncé.

Si $bc = ad$ et $au + bv = 1$, alors $acu + adv = c$, et donc $a(cu + dv) = c$, ce qui prouve que a divise c ; d'où le second énoncé.

Théorème 5.15. — Si $a \in A - \{0\}$, il existe une unité u de A et $p_1, \dots, p_r \in \mathcal{P}_A$ tels que $a = u p_1 \cdots p_r$; de plus, les p_i , pour $1 \leq i \leq r$, sont uniquement déterminés à l'ordre près. En d'autres termes, a peut se factoriser de manière unique comme produit de facteurs premiers.

Démonstration. — Commençons par montrer l'existence d'une telle factorisation. Si a est une unité, il n'y a rien à faire puisque $a = a$ est une factorisation sous la forme souhaitée. Si a n'est pas une unité, alors il existe $p_1 \in \mathcal{P}_A$ divisant a , et on pose $a_1 = a/p_1$, ce qui fait que, d'après le lemme 5.13, l'idéal (a_1) contient strictement (a) . En réitérant le processus, on construit une suite d'éléments p_i de \mathcal{P}_A et une suite d'éléments a_i de A , avec $a_{i+1}p_{i+1} = a_i$. La suite d'idéaux (a_i) est alors strictement croissante, ce qui implique, d'après le lemme 5.11, que le procédé s'arrête. Autrement dit, il existe s tel que a_s soit une unité de A , et $a = a_s p_1 \cdots p_s$ est une factorisation de a sous la forme voulue.

L'unicité se démontre en utilisant le lemme de Gauss. Si $u p_1 \cdots p_r = v q_1 \cdots q_s$ où les p_i et les q_j sont des nombres premiers et u, v des unités de A , le lemme de Gauss montre que p_r divise l'un des q_j et donc lui est égal. Quitte à permuter les q_j , on peut supposer que $p_r = q_s$, et en divisant les deux membres par $p_r = q_s$ (ce qui est licite car A est intègre), on se ramène à $r - 1$ et $s - 1$, ce qui permet de conclure par récurrence.

On note $v_p(a)$ le nombre de fois que p apparaît dans la factorisation de a en facteurs premiers. Alors $p^{v_p(a)}$ est la plus grande puissance de p divisant a ; on a donc $v_p(ab) = v_p(a) + v_p(b)$ et $v_p(a + b) \geq \inf(v_p(a), v_p(b))$.

Si $a_1, \dots, a_n \in A - \{0\}$, on définit $\text{pgcd}(a_1, \dots, a_n)$ par $\text{pgcd}(a_1, \dots, a_n) = \prod_p p^{\inf_i v_p(a_i)}$, ce qui fait de $\text{pgcd}(a_1, \dots, a_n)$ le plus grand diviseur commun des a_i (à multiplication près par une unité de A)

Lemme 5.16. — (théorème de Bézout) $\text{pgcd}(a_1, \dots, a_n)$ est un générateur de l'idéal (a_1, \dots, a_n) engendré par les a_i .

Démonstration. — Commençons par démontrer le résultat pour $n = 2$, et posons $a_1 = a$ et $a_2 = b$. Il est clair que tout élément de (a, b) est un multiple de $\text{pgcd}(a, b)$; il suffit donc de prouver que $d = \text{pgcd}(a, b) \in (a, b)$. Pour cela, écrivons a et b sous la forme $a = u p_1 \cdots p_r$ et $b = v q_1 \cdots q_s$, où u, v sont des unités de A et $p_1, \dots, p_r, q_1, \dots, q_s$ des éléments de \mathcal{P}_A . Par définition de d , on a $p_i \neq q_j$ quels que soient i et j , ce qui prouve, d'après le lemme de Gauss, que a/d et b/d sont premiers entre eux. Il existe donc $x, y \in A$ tels que $(a/d)x + (b/d)y = 1$, et alors $d = ax + by \in (a, b)$, ce que l'on cherchait à démontrer.

Maintenant, comme $\inf_{i \leq n} v_p(a_i) = \inf(\inf_{i \leq n-1} v_p(a_i), v_p(a_n))$, on a

$$\text{pgcd}(a_1, \dots, a_n) = \text{pgcd}(\text{pgcd}(a_1, \dots, a_{n-1}), a_n).$$

De même, l'idéal (a_1, \dots, a_n) est l'idéal engendré par (a_1, \dots, a_{n-1}) et par a_n , ce qui permet de déduire, par récurrence, le cas général du cas $n = 2$.

On dit que les a_i sont *premiers entre eux dans leur ensemble* si $\text{pgcd}(a_1, \dots, a_n) = 1$, ce qui équivaut, d'après le lemme 5.16, à l'existence de $\alpha_1, \dots, \alpha_n \in A$ tels que $\alpha_1 a_1 + \cdots + \alpha_n a_n = 1$.

5.3.3. Structure des modules de torsion sur un anneau principal. — Les anneaux \mathbf{Z} et $\mathbf{K}[T]$ sont principaux, ce qui fait que le théorème 5.17 ci-dessous a pour conséquences les th. 4.1 et 5.2.

Soit A un anneau principal, et soit \mathcal{P}_A l'ensemble des idéaux premiers non nuls de A . Choisissons pour tout élément de \mathcal{P}_A un générateur, et identifions \mathcal{P}_A à l'ensemble de ces générateurs. Alors tout élément non nul x de A se factorise, de manière unique, sous la forme $x = u \prod_{p \in \mathcal{P}_A} p^{v_p(x)}$, où u est inversible dans A . De plus, x et y sont premiers entre eux (ce qui signifie que l'idéal de A engendré par

x et y contient 1), si et seulement si $\inf(v_p(x), v_p(y)) = 0$ quel que soit $p \in \mathcal{P}_A$, et A/p est un corps quel que soit $p \in \mathcal{P}_A$.

Si M est un A -module, et si $a \in A$, on note $aM \subset M$ l'image du morphisme $x \mapsto ax$ de A -modules ; c'est un sous- A -module de M . De plus, M/aM est, par construction, tué par a , et donc l'action de A sur M/aM se factorise à travers A/a , ce qui fait de M/aM un A/a -module. En particulier, si $p \in \mathcal{P}_A$, alors M/pM est un espace vectoriel sur le corps A/p .

Théorème 5.17. — *Soit M un A -module de torsion et de type fini. Si $p \in \mathcal{P}_A$, soit M_p l'ensemble des $x \in M$ tués par une puissance de p .*

- (i) M_p est un sous- A -module de M , nul sauf pour un nombre fini de p , et $M = \bigoplus_{p \in \mathcal{P}_A} M_p$.
- (ii) Si $r_p = \dim_{A/p}(M/pM)$, alors il existe une unique famille décroissante d'entiers $a_{p,i} \geq 1$, pour $1 \leq i \leq r_p$, telle que $M_p = \bigoplus_{1 \leq i \leq r_p} A/p^{a_{p,i}}$.

Démonstration. — Si $p^a x = 0$ et $p^b y = 0$, alors $p^{\sup(a,b)}(\lambda x + \mu y) = 0$ quels que soient $\lambda, \mu \in A$. On en déduit que M_p est un sous- A -module de M .

Soient x_1, \dots, x_d engendrant M . Si $i \in \{1, \dots, d\}$, soit $\lambda_i \in A$ tel que $\lambda_i x_i = 0$, et soit $\lambda = \lambda_1 \cdots \lambda_d$. On a $\lambda x = 0$ quel que soit $x \in M$. Si $p \in \mathcal{P}_A$ ne divise pas λ , et si $x \in M_p$ est tué par p^a , alors x est tué par tout élément de l'idéal (λ, p^a) de A engendré par λ et p^a , c'est-à-dire par A , puisque λ et p^a sont premiers entre eux. On a donc $x = 0$, et $M_p = 0$ si p ne divise pas λ .

Soit $\mathcal{P}_A(\lambda) \subset \mathcal{P}_A$ l'ensemble des diviseurs premiers de λ , et soit $\lambda = \prod_{p \in \mathcal{P}_A(\lambda)} p^{n_p}$ la factorisation de λ en facteurs premiers. Les $\frac{\lambda}{p^{n_p}}$, pour $p \in \mathcal{P}_A(\lambda)$ sont premiers entre eux dans leur ensemble. Il existe donc, d'après le théorème de Bézout, des éléments α_p de A tels que l'on ait $\sum_{p \in \mathcal{P}_A(\lambda)} \alpha_p \frac{\lambda}{p^{n_p}} = 1$. On en déduit que l'on peut décomposer tout élément x de M sous la forme $\sum_{p \in \mathcal{P}_A(\lambda)} x_p$, avec $x_p = \frac{\alpha_p \lambda}{p^{n_p}} x$, et $x_p \in M_p$ car x_p est tué par p^{n_p} . En résumé, $M = \sum_{p \in \mathcal{P}_A} M_p$.

Finalement, si $x_p \in M_p$, pour $p \in \mathcal{P}_A(\lambda)$, et si $\sum_{p \in \mathcal{P}_A(\lambda)} x_p = 0$, alors $x_p = -\sum_{\ell \neq p} x_\ell$ est à la fois tué par p^{n_p} et par $p^{-n_p} \lambda$, qui sont premiers entre eux par définition de n_p . On a donc $x_p = 0$ quel que soit p , ce qui termine de démontrer le (i).

Passons à la démonstration du (ii). Commençons par montrer que l'on peut calculer r_p en ne considérant que M_p . Si $\ell \in \mathcal{P}_A$ est distinct de p , la multiplication par p induit une surjection sur M_ℓ : en effet, il existe n tel que $\ell^n M_\ell = 0$, et comme p et ℓ^n sont premiers entre eux, il existe $a, b \in A$ tels que $ap + b\ell^n = 1$. Les multiplications par a et p sont inverses l'une de l'autre sur M_ℓ , et donc $M_\ell/pM_\ell = 0$. Il en résulte que r_p est aussi la dimension de M_p/pM_p sur A/p .

La démonstration du (ii) va se faire en deux étapes. On commence par démontrer, par récurrence sur $r = r_p$ (le cas $r = 0$ étant vide), l'existence d'une décomposition sous la forme voulue, puis on démontre, toujours par récurrence, l'unicité de la famille $a_{p,i}$.

Si $x \in M_p$, on note $n(x)$ le plus petit $n \in \mathbb{N}$ tel que $p^n x = 0$. Donc $p^{n(x)} x = 0$ et $p^{n(x)-1} x \neq 0$, si $n(x) \geq 1$. Soit $e_1 \in M_p$ réalisant le maximum de $n(x)$, pour $x \in M_p$ (comme $n(x) \leq n_p$, quel que soit $x \in M_p$, il existe un tel e_1), et soit $a_1 = n(e_1)$. Soit $N = M_p/(A/p^{a_1})e_1$. Alors N/pN est le quotient de M_p/pM_p par le sous- (A/p) -espace vectoriel engendré par l'image de e_1 , et comme cette image est non nulle (sinon, on aurait $e_1 = pf$ et $n(f) = n(e_1) + 1 > n(e_1)$), on en déduit que $\dim_{A/p}(N/pN) = r - 1$, ce qui permet d'appliquer l'hypothèse de récurrence à N . Il existe donc $\bar{e}_2, \dots, \bar{e}_r \in N$ et $a_2 \geq \dots \geq a_r$ tels que $N = \bigoplus_{2 \leq i \leq r} (A/p^{a_i})\bar{e}_i$.

Soit $e'_i \in M_p$ un relèvement quelconque de \bar{e}_i . On a alors $p^{a_i} e'_i = b_i e_1$, avec $b_i \in A$, bien défini modulo p^{a_1} . Comme $p^{a_1} e'_i = 0$, on en déduit que $p^{a_1-a_i} b_i \in p^{a_1} A$, et donc que $b_i \in p^{a_i} A$. Soit $c_i = p^{-a_i} b_i \in A$, et soit $e_i = e'_i - c_i e_1$. On a alors $p^{a_i} e_i = 0$. Maintenant, soit $x \in M_p$, et soit \bar{x} son image dans N . Il existe alors $\lambda_2 \in A/p^{a_2}, \dots, \lambda_r \in A/p^{a_r}$, uniques, tels que $\bar{x} = \lambda_2 \bar{e}_2 + \dots + \lambda_r \bar{e}_r$.

Comme $p^{a_i}e_i = 0$, l'élément $\lambda_i e_i$ de M_p est bien défini, et $x - \sum_{i=2}^r \lambda_i e_i \in (A/p^{a_1})e_1$, ce qui prouve que $M_p = (A/p^{a_1})e_1 \oplus (A/p^{a_2})e_2 \oplus \cdots \oplus (A/p^{a_r})e_r$. Comme $a_1 \geq a_2$, cela fournit une décomposition de M_p sous la forme voulue.

Il reste à prouver l'unicité de la suite des $a_{p,i}$. Supposons donc que l'on ait

$$M_p = \bigoplus_{1 \leq i \leq r} (A/p^{a_i})e_i = \bigoplus_{1 \leq j \leq s} (A/p^{b_j})f_j,$$

avec $a_1 \geq a_2 \geq \cdots a_r \geq 1$ et $b_1 \geq b_2 \geq \cdots b_s \geq 1$. Soit $n(M_p)$ le maximum des $n(x)$, pour $x \in M_p$. Alors $n(M_p) = a_1$ et $n(M_p) = b_1$, et donc $a_1 = b_1$. Maintenant, on peut écrire e_1 sous la forme $e_1 = \sum_{j=1}^s \lambda_j f_j$, et comme $p^{a_1-1}e_1 \neq 0$, cela implique qu'il existe j tel que $p^{a_1-1}\lambda_j f_j \neq 0$. En particulier, on a $p^{a_1-1}f_j \neq 0$, ce qui prouve que $b_j \geq a_1 = b_1$ et donc que $b_j = b_1$. Quitte à permuter les f_j , on peut donc supposer $j = 1$. La propriété $p^{a_1-1}\lambda_1 f_1 \neq 0$ implique alors (car $a_1 = b_1$) que $\lambda_1 \notin pA$, et donc que λ_1 est premier à p et p^{a_1} , et est inversible dans $A/p^{a_1}A$. En notant μ_1 son inverse, cela permet d'écrire f_1 sous la forme $\mu_1 e_1 - \sum_{j=2}^s \mu_1 \lambda_j f_j$, ce qui prouve que l'on a aussi $M_p = (A/p^{b_1})e_1 \oplus \bigoplus_{2 \leq j \leq s} (A/p^{b_j})f_j$. On en déduit que

$$M_p / (A/p^{b_1})e_1 = \bigoplus_{2 \leq i \leq r} (A/p^{a_i})e_i = \bigoplus_{2 \leq j \leq s} (A/p^{b_j})f_j,$$

et une récurrence immédiate permet d'en conclure que l'on a $a_i = b_i$ quel que soit i (et donc aussi que $r = s$). Ceci termine la démonstration.

6. Topologie

Les notions de topologie générale interviennent directement dans toutes les branches des mathématiques, comme on s'en est aperçu graduellement à partir des travaux de Hausdorff (1906). Parmi les espaces topologiques, les espaces métriques (dont les espaces vectoriels normés sont un cas particulier fondamental) forment une catégorie d'objets aux propriétés particulièrement agréables. Les suites y jouent un rôle privilégié permettant souvent de simplifier les démonstrations qui, pour un espace topologique général, utilisent le langage de la théorie des ensembles. Chaque fois que c'est le cas, nous avons doublé la démonstration dans le cas général d'une démonstration propre aux espaces métriques afin de diversifier les approches.

6.1. Espaces topologiques

6.1.1. Ouverts, fermés, voisinages

Si X est un ensemble, une *topologie* \mathcal{T} sur X est un sous-ensemble de l'ensemble des parties de X , contenant X et \emptyset , stable par intersection finie et par réunion quelconque. Avec des quantificateurs, cela se traduit par :

- $\emptyset \in \mathcal{T}$ et $X \in \mathcal{T}$;
- si I est un ensemble fini, et si $U_i \in \mathcal{T}$, pour $i \in I$, alors $\bigcap_{i \in I} U_i \in \mathcal{T}$;
- si I est un ensemble quelconque, et si $U_i \in \mathcal{T}$, pour $i \in I$, alors $\bigcup_{i \in I} U_i \in \mathcal{T}$.

Si (X, \mathcal{T}) est un *espace topologique* (i.e. un ensemble X muni d'une topologie \mathcal{T}), les éléments de \mathcal{T} sont les *ouverts*. On dit que $F \subset X$ est *fermé*, si son complémentaire est ouvert. Donc X et \emptyset sont des fermés, et les fermés sont stables par réunion finie et intersection quelconque.

Une *base d'ouverts* pour une topologie \mathcal{T} est un sous-ensemble \mathcal{B} de \mathcal{T} tel que tout élément de \mathcal{T} soit réunion d'éléments de \mathcal{B} . Par exemple, dans un espace métrique (voir plus loin), les boules ouvertes forment une base d'ouverts.

Si (X, \mathcal{T}) est un espace topologique, et si $x \in X$, un *voisinage* V de x est un sous-ensemble de X contenant un ouvert contenant x . Un ensemble est donc ouvert si et seulement si il est voisinage de chacun de ses points.

Une *base de voisinages* de x est une famille de voisinages de x telle que tout ouvert contenant x contienne un élément de la famille. Par exemple, dans un espace métrique, les boules ouvertes de centre x ou les boules fermées de centre x et de rayon non nul forment une base de voisinages de x .

6.1.2. Exemples

- La *topologie discrète* sur un ensemble X est celle pour laquelle $\mathcal{T} = \mathcal{P}(X)$ ensemble des parties de X . De manière équivalente, X est muni de la topologie discrète si les singletons sont des ouverts (en effet toute partie de X est la réunion des singletons qu'elle contient).
- La *topologie grossière* sur X est la topologie dont les seuls ouverts sont X et \emptyset .
- La topologie naturelle sur \mathbf{R} est celle pour laquelle les segments ouverts forment une base d'ouverts.
- Si E est un espace vectoriel sur \mathbf{R} ou \mathbf{C} muni d'une norme $\| \cdot \|$, la topologie sur E associée à $\| \cdot \|$ est celle pour laquelle les boules ouvertes forment une base d'ouverts.
- La *topologie de Zariski* sur \mathbf{C}^n est définie de la manière suivante : $F \subset \mathbf{C}^n$ est un *fermé de Zariski* si et seulement si il existe une famille de polynômes $P_i \in \mathbf{C}[X_1, \dots, X_n]$, pour $i \in I$, telle que F soit l'ensemble des zéros communs des P_i (i.e. $F = \cap_{i \in I} \{z \in \mathbf{C}^n, P_i(z) = 0\}$). Alors \mathbf{C}^n est un fermé de Zariski (en prenant une famille vide), \emptyset est un fermé de Zariski (en prenant $P_1 = X_1$ et $P_2 = X_1 - 1$), et une intersection quelconque des fermés de Zariski est un fermé de Zariski (si F_j , pour $j \in J$, est l'ensemble des zéros communs de la famille $(P_{i,j})_{i \in I_j}$, alors $\cap_{j \in J} F_j$ est l'ensemble des zéros communs de la famille $(P_{i,j})_{j \in J, i \in I_j}$), ce qui montre qu'en définissant un ouvert de \mathbf{C}^n pour la topologie de Zariski comme le complémentaire d'un fermé de Zariski, on obtient bien une topologie dont les fermés sont les fermés de Zariski.
- On peut munir un ensemble quelconque, de la *topologie du filtre des complémentaires des parties finies* pour laquelle une partie non vide est un ouvert si et seulement si elle a un complémentaire fini.

6.1.3. Comparaison de topologies

Si \mathcal{T}_1 et \mathcal{T}_2 sont deux topologies sur X , on dit que \mathcal{T}_1 est *plus fine* que \mathcal{T}_2 si \mathcal{T}_1 contient \mathcal{T}_2 . Le summum de la finesse est donc la discrétion ; à l'opposé, la topologie la moins fine est la topologie grossière. On fera attention au fait que, si on prend deux

topologies quelconques, il n'y a aucune raison pour qu'il y en ait une qui soit plus fine que l'autre (cf. ex. 11.2).

6.2. Espaces métriques

Si X est un ensemble, une application $d : X \times X \rightarrow \mathbf{R}_+$ est une *distance* sur X si elle vérifie les propriétés suivantes :

- $d(x, y) = 0$ si et seulement si $x = y$ (séparation) ;
- $d(x, y) = d(y, x)$ quels que soient $x, y \in X$;
- $d(x, z) \leq d(x, y) + d(y, z)$ quels que soient $x, y, z \in X$ (inégalité triangulaire).
- Si la distance vérifie l'inégalité $d(x, z) \leq \sup(d(x, y), d(y, z))$ plus forte que l'inégalité triangulaire, on dit qu'elle est *ultramétrique* ou *non archimédienne*.

Si $x \in X$ et $r > 0$, on note $B(x, r) = \{y \in X, d(x, y) \leq r\}$ la *boule fermée de centre x et de rayon r* , et $B(x, r^-) = \{y \in X, d(x, y) < r\}$ la *boule ouverte de centre x et de rayon r* .

- Une boule ouverte contient une boule ouverte centrée en chacun de ses points.

L'inégalité triangulaire montre que, si $r > 0$, si $y \in B(x, r^-)$, et si $s = r - d(x, y)$, alors $B(y, s^-) \subset B(x, r^-)$.

- L'ensemble \mathcal{T}_d constitué de \emptyset et des réunions (quelconques) de boules ouvertes est une topologie sur X , et $U \in \mathcal{T}_d$ si et seulement si, quel que soit $x \in U$, il existe $r > 0$ tel que $B(x, r^-) \subset U$.

Par construction \mathcal{T}_d contient \emptyset et X et est stable par réunion quelconque. Il suffit donc de prouver que \mathcal{T}_d est stable par intersection finie. Soit $U \in \mathcal{T}_d$ non vide, et soit $x \in U$. Par définition de \mathcal{T}_d , il existe $y \in X$ et $r > 0$ tels que $B(y, r^-) \subset U$ et $x \in B(y, r^-)$; le point ci-dessus montre qu'il existe $s > 0$ tel que $B(x, s^-) \subset B(y, r^-) \subset U$, ce qui prouve le second énoncé. La stabilité par intersection finie s'en déduit puisque si $(U_i)_{i \in I}$ est une famille finie d'éléments de \mathcal{T}_d , et si $x \in \bigcap_{i \in I} U_i$, alors pour tout i , il existe $s_i > 0$ tel que $B(x, s_i^-) \subset U_i$, ce qui fait que $\bigcap_{i \in I} U_i$ contient $B(x, s^-)$, si $s = \inf_{i \in I} s_i$ (et $s \neq 0$ car I est fini).

On note en général (X, d) au lieu de (X, \mathcal{T}_d) l'espace topologique ainsi obtenu. Un espace topologique obtenu de cette manière est appelé un *espace métrique*. Par construction, les boules ouvertes forment une base d'ouverts de la topologie.

Deux distances sur X sont *équivalentes* si elles définissent la même topologie.

Un espace topologique (X, \mathcal{T}) est *métrisable* s'il existe une distance d sur X telle que l'on ait $\mathcal{T} = \mathcal{T}_d$.

- Dans un espace métrique, les boules fermées sont des fermés.

Si $x \notin B(x_0, r)$, et si $s = d(x, x_0) - r$, alors $s > 0$ et le complémentaire de $B(x_0, r)$ contient $B(x, s^-)$. On en déduit que ce complémentaire est ouvert et donc que $B(x_0, r)$ est fermée.

- Si (X, d) est un espace métrique, et si $x \in X$, les $B(x, r^-)$ et les $B(x, r)$, pour $r > 0$, forment une base de voisinages de x .

On a vu ci-dessus que si U est un ouvert non vide contenant x , alors U contient une boule ouverte $B(x, r^-)$, avec $r > 0$, ce qui prouve que les $B(x, r^-)$ forment une base de voisinages de x . De plus, $B(x, r^-)$ contient $B(x, r/2)$ qui contient $B(x, (r/2)^-)$, ce qui prouve que les $B(x, r)$ forment aussi une base de voisinages de x .

- Deux distances d_1 et d_2 sur un ensemble X sont équivalentes si et seulement si, pour tout $x \in X$, toute boule ouverte de centre x pour d_1 contient une boule ouverte de centre x pour d_2 et réciproquement.

La première condition équivaut à ce que l'identité $(X, d_2) \rightarrow (X, d_1)$ est continue, ainsi que son inverse, et la seconde condition est une traduction de cette bicontinuité utilisant le fait que les boules ouvertes de centre x forment une base de voisinages de x .

Exercice 6.1. — Montrer que, si (X, d) est un espace métrique, et si $x \in X$, les $B(x, 2^{-j})$, pour $j \in \mathbf{N}$, forment une base de voisinages de x .

Exercice 6.2. — Soit X un ensemble. Montrer que $d : X \times X \rightarrow \mathbf{R}_+$, définie par $d(x, y) = 0$ si $x = y$ et $d(x, y) = 1$, si $x \neq y$, est une distance (la distance triviale) sur X . Quelle est la topologie associée ?

Exercice 6.3. — Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ définie par $f(x) = \frac{x}{1+|x|}$. Montrer que $(x, y) \mapsto d'(x, y) = |f(x) - f(y)|$ est une distance sur \mathbf{R} , qui est équivalente à la distance usuelle $d(x, y) = |x - y|$.

6.3. Continuité

Si X et Y sont deux espaces topologiques, si $f : X \rightarrow Y$ est une application, et si $x \in X$, on dit que f est *continue en x* , si quel que soit l'ouvert V de Y contenant $f(x)$, il existe un ouvert U de X contenant x tel que $f(U) \subset V$. De manière équivalente, f est continue en x si, quel que soit le voisinage V de $f(x)$ dans Y , il existe un voisinage U de x tel que $f(U) \subset V$. Il suffit de vérifier ceci pour V dans une base de voisinages de $f(x)$.

On dit que $f : X \rightarrow Y$ est *continue*, si elle est continue en tout point $x \in X$.

On dit que $f : X \rightarrow Y$ est un *homéomorphisme* si f est continue bijective, et si sa réciproque $f^{-1} : Y \rightarrow X$ est continue. On dit que X et Y sont *homéomorphes*⁽²⁴⁾ s'il existe un homéomorphisme $f : X \rightarrow Y$.

Si (X, d) est un espace métrique, si (Y, \mathcal{T}) est un espace topologique, si f est une application de X dans Y et si $x \in X$, on voit, en revenant à la définition, que $f : X \rightarrow Y$ est continue en x_0 si et seulement si quel que soit U ouvert de Y contenant $f(x_0)$, il existe $\delta > 0$ tel que $d(x_0, x) < \delta$ implique $f(x) \in U$. Si Y est aussi métrique, cela se traduit par :

quel que soit $\varepsilon > 0$, il existe $\delta = \delta(x, \varepsilon) > 0$ tel que $d_X(x_0, x) < \delta \Rightarrow d_Y(f(x_0), f(x)) < \varepsilon$;

ou encore par :

pour tout $j \in \mathbf{N}$, il existe $\delta = \delta(x, j) > 0$ tel que $d_X(x_0, x) < \delta \Rightarrow d_Y(f(x_0), f(x)) \leq 2^{-j}$.

⁽²⁴⁾ Montrer que deux espaces topologiques ne sont pas homéomorphes est loin d'être évident en général (le lecteur est invité à essayer de prouver qu'un pneu et un ballon de football ne sont pas homéomorphes) ; la topologie algébrique (Analysis in situ de Poincaré) fournit des tas d'outils permettant de le faire.

On dit que $f : X \rightarrow Y$ est *uniformément continue* sur X , si pour tout $\varepsilon > 0$ il existe $\delta = \delta(\varepsilon) > 0$ tel que $d_X(x, x') < \delta$ implique $d_Y(f(x), f(x')) < \varepsilon$. La différence entre la continuité et la continuité uniforme est que δ ne dépend pas de x ; en particulier, une application uniformément continue est continue.

Si $\kappa \in \mathbf{R}_+$, on dit que $f : X \rightarrow Y$ est κ -lipschitzienne (ou *lipschitzienne de rapport κ*), si on a $d_Y(f(x), f(x')) \leq \kappa d_X(x, x')$, quels que soient $x, x' \in X$. Une application lipschitzienne est uniformément continue et donc aussi continue.

Exercice 6.4. — Soit (X, d) un espace métrique. Montrer que $d : X \rightarrow \mathbf{R}$ est continue.

• Les conditions suivantes sont équivalentes :

- (i) $f : X \rightarrow Y$ est continue ;
- (ii) il existe une base d'ouverts \mathcal{B} de Y telle que l'image réciproque par f de tout $U \in \mathcal{B}$ est un ouvert de X ;
- (iii) l'image réciproque par f de tout ouvert de Y est un ouvert de X ;
- (iv) l'image réciproque par f de tout fermé de Y est un fermé de X .

L'équivalence de (iii) et (iv) vient juste de ce que l'image réciproque du complémentaire est le complémentaire de l'image réciproque (si $A \subset Y$, alors $f^{-1}(Y - A) = X - f^{-1}(A)$).

Si f est continue, si V est un ouvert de Y , et si $y \in V \cap f(X)$, il existe, pour tout $x \in X$ vérifiant $f(x) = y$, un ouvert U_x de X qui contient x et vérifie $f(U_x) \subset V$. Alors $U = \bigcup_{y \in V \cap f(X)} \left(\bigcup_{x \in f^{-1}(y)} U_x \right)$ est un ouvert qui contient $\bigcup_{y \in V \cap f(X)} f^{-1}(y) = f^{-1}(V)$, et qui vérifie $f(U) \subset V$, ce qui prouve que $f^{-1}(V) = U$ et donc que $f^{-1}(V)$ est ouvert. On en déduit l'implication (i) \Rightarrow (iii), et comme l'implication (iii) \Rightarrow (i) est immédiate (si V est un ouvert contenant $f(x)$, alors $U = f^{-1}(V)$ est un ouvert de X qui contient x et qui vérifie $f(U) \subset V$), cela prouve que les propriétés (i) et (iii) sont équivalentes.

L'implication (iii) \Rightarrow (ii) est immédiate. Réciproquement, soit \mathcal{B} une base d'ouverts de Y , et soit V un ouvert de Y . Il existe alors une famille $(V_i)_{i \in I}$ d'éléments de \mathcal{B} telle que $V = \bigcup_{i \in I} V_i$. On a alors $f^{-1}(V) = \bigcup_{i \in I} f^{-1}(V_i)$, et si $f^{-1}(V_i)$ est ouvert pour tout i , il en est de même de $f^{-1}(V)$. On en déduit l'équivalence des propriétés (ii) et (iii).

Ceci permet de conclure.

• Soient X, Y, Z des espaces topologiques. Si $f : X \rightarrow Y$ est continue en x , et si $g : Y \rightarrow Z$ est continue en $f(x)$, alors $g \circ f : X \rightarrow Z$ est continue en x ; si $f : X \rightarrow Y$ et $g : Y \rightarrow Z$ sont continues, alors $g \circ f : X \rightarrow Z$ est continue.

Soit W un ouvert de Z contenant $g(f(x))$. Comme g est continue en $f(x)$, il existe un ouvert V de Y qui contient $f(x)$ et qui vérifie $g(V) \subset W$, et comme f est continue en x , il existe un ouvert U de X qui contient x et qui vérifie $f(U) \subset V$. Alors $g \circ f(U) \subset W$, ce qui permet de démontrer le premier énoncé; le second en est une conséquence immédiate.

6.4. Sous-espaces, produits, quotients

6.4.1. Topologie induite

Si (X, \mathcal{T}) est un espace topologique, et si $Y \subset X$, alors $\mathcal{T}_Y = \{U \cap Y, U \in \mathcal{T}\}$ est une topologie sur Y appelée la *topologie induite*. Autrement dit, tout sous-ensemble d'un espace topologique est naturellement un espace topologique.

6.4.2. Topologie produit

Si $(X_i, \mathcal{T}_i)_{i \in I}$ est une famille (éventuellement infinie) d'espaces topologiques, on appelle *topologie produit* sur $X = \prod_{i \in I} X_i$, la topologie la moins fine rendant continues les projections $p_i : X \rightarrow X_i$, pour $i \in I$. De manière explicite, une base d'ouverts pour cette topologie est constituée des $\prod_{i \in J} U_i \times \prod_{i \in I-J} X_i$, où J décrit les sous-ensembles *finis* de I et, U_i est, si $i \in J$, un ouvert de X_i .

- Si Y est un espace topologique, alors $f : Y \rightarrow \prod_{i \in I} X_i$ est continue si et seulement si $p_i \circ f : Y \rightarrow X_i$ est continue, quel que soit $i \in I$.

Comme la composée d'applications continues est continue, si $f : Y \rightarrow \prod_{i \in I} X_i$ est continue, alors $p_i \circ f : Y \rightarrow X_i$ est continue, quel que soit $i \in I$. Réciproquement, si les $p_i \circ f$, pour $i \in I$, sont continues, et si $U = \prod_{i \in J} U_i \times \prod_{i \in I-J} X_i$, où $J \subset I$ est fini, est un élément de la base d'ouverts ci-dessus, alors $f^{-1}(U) = \cap_{i \in J} (p_i \circ f)^{-1}(U_i)$ est un ouvert comme intersection finie d'ouverts. Ceci implique que f est continue, ce qui permet de conclure.

- Si (X, d_X) et (Y, d_Y) sont deux espaces métriques, alors l'espace topologique $X \times Y$ est métrisable, la topologie produit pouvant être définie par la distance $d_{X \times Y}((x, y), (x', y')) = \sup(d_X(x, x'), d_Y(y, y'))$, ou par toute autre distance équivalente, comme par exemple $\sqrt{d_X(x, x')^2 + d_Y(y, y')^2}$.

La distance $d_{X \times Y}$ fait qu'une boule de $X \times Y$ est le produit d'une boule de X et d'une boule de Y , ce qui prouve que la topologie qu'elle définit est bien la topologie produit.

6.4.3. Topologie quotient

Si X est un espace topologique et \sim est une relation d'équivalence sur X , on définit la *topologie quotient* sur X/\sim en disant que U est ouvert dans X/\sim si et seulement si son image inverse dans X est ouverte dans X . C'est la topologie la plus fine rendant continue la surjection canonique $\pi : X \rightarrow X/\sim$.

- Si Y est un espace topologique, alors $f : X/\sim \rightarrow Y$ est continue si et seulement si $f \circ \pi : X \rightarrow Y$ est continue.

$f : X/\sim \rightarrow Y$ est continue si et seulement si $f^{-1}(U)$ est ouvert pour tout ouvert U de Y , ce qui équivaut, par définition de la topologie quotient, à ce que $\pi^{-1}(f^{-1}(U))$ est ouvert dans X , pour tout ouvert U de Y , et donc à ce que $f \circ \pi : X \rightarrow Y$ soit continue.

Exercice 6.5. — Quelle est la topologie quotient sur \mathbf{R}/\mathbf{Q} ?

Voici quelques espaces que l'on peut construire par des passages au quotient. Le lecteur est invité à s'armer de ciseaux et de colle pour voir à quoi ressemblent les trois premiers espaces, et à chercher sur

Internet (par exemple à <http://www.mathcurve.com/surfaces/surfaces.shtml>) des images des deux derniers (on ne peut pas les plonger physiquement dans \mathbf{R}^3).

— Le *cylindre* : c'est le quotient de $[0, 1] \times [0, 1]$ par la relation d'équivalence $(x, 0) \sim (x, 1)$, si $x \in [0, 1]$.

— La *bande de Moebius* : c'est le quotient de $[0, 1] \times [0, 1]$ par la relation d'équivalence $(x, 0) \sim (1 - x, 1)$, si $x \in [0, 1]$.

— Le *tore* : c'est le quotient de $[0, 1] \times [0, 1]$ par la relation d'équivalence $(x, 0) \sim (x, 1)$, si $x \in [0, 1]$ et $(0, y) \sim (1, y)$, si $y \in [0, 1]$. C'est aussi le quotient de \mathbf{R}^2 par \mathbf{Z}^2 ou encore le produit $(\mathbf{R}/\mathbf{Z})^2$ de deux cercles.

— La *bouteille de Klein* : c'est le quotient de $[0, 1] \times [0, 1]$ par la relation d'équivalence $(x, 0) \sim (x, 1)$, si $x \in [0, 1]$ et $(0, y) \sim (1, 1 - y)$, si $y \in [0, 1]$.

— Le *plan projectif réel* : c'est le quotient de la sphère unité de \mathbf{R}^3 par la relation d'équivalence $x \sim -x$; il est homéomorphe au quotient de $[0, 1] \times [0, 1]$ par la relation d'équivalence $(x, 0) \sim (1 - x, 1)$, si $x \in [0, 1]$ et $(0, y) \sim (1, 1 - y)$, si $y \in [0, 1]$.

6.5. Espaces séparés

Une topologie est *séparée* si, quels que soient $x, y \in X$, avec $x \neq y$, on peut trouver des ouverts U, V de X , avec $x \in U$, $y \in V$, et $U \cap V = \emptyset$. Par exemple, la topologie discrète est séparée (prendre $U = \{x\}$ et $V = \{y\}$), et la topologie grossière est on ne peut moins séparée (sauf si X a 0 ou 1 élément). Dans un espace séparé, les points sont fermés, mais la réciproque n'est pas vraie⁽²⁵⁾.

- Un espace métrique est séparé.

Si $x \neq y$, on a $d(x, y) > 0$, et si $r = \frac{1}{2}d(x, y)$, alors $B(x, r^-) \cap B(y, r^-) = \emptyset$, d'après l'inégalité triangulaire.

- Si les X_i sont séparés, alors $X = \prod_{i \in I} X_i$ est séparé.

Si $x = (x_i)_{i \in I}$ et $y = (y_i)_{i \in I}$ sont deux éléments distincts de X , il existe $j \in I$ tel que $x_j \neq y_j$, et comme X_j est séparé, il existe des ouverts disjoints U_j et V_j de X_j contenant x_j et y_j respectivement. Alors $U = U_j \times \prod_{i \neq j} X_i$ et $V = V_j \times \prod_{i \neq j} X_i$ sont des ouverts disjoints de X contenant x et y respectivement. On en déduit la séparation de X .

- X est séparé si et seulement si la diagonale $\Delta = \{(x, x), x \in X\}$ est fermée dans $X \times X$.

Si X est séparé, alors quels que soient $x, y \in X$ distincts, il existe des ouverts $U_{x,y}$, $V_{x,y}$ disjoints, avec $x \in U_{x,y}$ et $y \in V_{x,y}$. La condition « $U_{x,y}$, $V_{x,y}$ disjoints » est équivalente à ce que l'ouvert $W_{x,y} = U_{x,y} \times V_{x,y}$ de $X \times X$ ne rencontre pas Δ . De plus, $W_{x,y}$ contient (x, y) , ce qui fait que la réunion des $W_{x,y}$, pour $x \neq y$, est égale à $(X \times X) - \Delta$ qui est donc ouvert en tant que réunion d'ouverts. On en déduit que Δ est fermée.

⁽²⁵⁾Par exemple, dans \mathbf{C}^n muni de la topologie de Zariski, les points sont fermés puisque $z = (z_1, \dots, z_n)$ est l'ensemble des zéros communs de la famille de polynômes $X_i - z_i$, pour $i \in I$, mais la topologie de Zariski est fort peu séparée puisque tout ouvert de Zariski non vide est dense (pour la topologie de Zariski et aussi pour la topologie usuelle de \mathbf{C}^n). Il a fallu attendre les travaux de A. Weil (1952) et J-P. Serre (1956) pour que l'on se rende compte que cette topologie, loin d'être une curiosité pathologique, permet de retrouver, de manière algébrique, la plupart des invariants que l'on peut définir en utilisant la topologie usuelle. Ceci servit de point de départ à la révolution grothendieckienne.

Réciproquement, si Δ est fermée, alors $(X \times X) - \Delta$ est ouvert. Par définition de la topologie produit, cela implique que si $(x, y) \in (X \times X) - \Delta$ (i.e. si $x \neq y$), alors il existe U, V ouverts de X tels que $U \times V \subset (X \times X) - \Delta$ et $(x, y) \in U \times V$. Alors $x \in U$, $y \in V$ et $U \cap V = \emptyset$. On en déduit la séparation de X .

Exercice 6.6. — Montrer que, si $f : X \rightarrow Y$ est injective et continue, et si Y est séparé, alors X est séparé.

Un espace métrique est séparé grâce à la *condition de séparation* « $d(x, y) = 0 \Rightarrow x = y$ ». Si on supprime la condition de séparation, on obtient une *semi-distance* qui permet encore de définir une topologie \mathcal{T}_d dans laquelle un ouvert non vide est une réunion (quelconque) de boules ouvertes. L'espace topologique (X, \mathcal{T}_d) n'est plus forcément séparé (si $x \neq y$, mais $d(x, y) = 0$, alors tout ouvert de X contenant x contient aussi y). C'est le cas des espaces $\mathcal{L}^1(\mathbf{R}^m)$ et $\mathcal{L}^2(\mathbf{R}^m)$ du § III.2, par exemple.

On peut fabriquer un espace séparé à partir de (X, d) , en identifiant deux points dont la distance est nulle. De manière précise, on définit une relation \sim sur X par $x \sim y$ si et seulement si $d(x, y) = 0$; la relation \sim est une relation d'équivalence grâce à la symétrie de d et à l'inégalité triangulaire. De plus, on a $d(x, y) = d(x', y')$ si $x \sim x'$ et $y \sim y'$, toujours grâce à l'inégalité triangulaire. On en déduit le fait que d définit une distance sur l'ensemble X/\sim des classes d'équivalence pour la relation \sim , et le séparé de (X, d) est l'ensemble X/\sim muni de la distance induite par d .

Un exemple de cette construction est le passage de $\mathcal{L}^1(\mathbf{R}^m)$ à $L^1(\mathbf{R}^m)$ ou de $\mathcal{L}^2(\mathbf{R}^m)$ à $L^2(\mathbf{R}^m)$ rencontré dans le cours (cf. § III.2).

6.6. Intérieur, adhérence, densité

Si X est un espace topologique, et $Y \subset X$, alors la réunion $\overset{\circ}{Y}$ de tous les ouverts de X contenus dans Y est un ouvert, et donc est le plus grand ouvert contenu dans Y ; c'est l'*intérieur* de Y . On dit que Y est d'*intérieur vide* si $\overset{\circ}{Y} = \emptyset$.

De même, l'intersection \overline{Y} de tous les fermés de X contenant Y est un fermé appelé l'*adhérence* de Y . On dit que Y est *dense dans* X si $\overline{Y} = X$. De manière équivalente, Y est dense dans X si et seulement si $Y \cap U \neq \emptyset$ quel que soit U ouvert non vide de X ou encore si et seulement si tout point de X admet au moins un point de Y dans chacun de ses voisinages.

- \mathbf{Q} est dense dans \mathbf{R} et \mathbf{Q}_p (par construction).
- Les polynômes sont denses dans l'espace des fonctions continues sur $[0, 1]$ muni de la norme $\|\phi\|_\infty = \sup_{x \in [0, 1]} |\phi(x)|$ de la convergence uniforme (théorème de Weierstrass).
- Si X est muni de la topologie grossière, tout point est dense dans X .
- Si Y est dense dans X , si Z est séparé, et si $f, g : X \rightarrow Z$ sont continues et coïncident sur Y , alors $f = g$.

Il suffit de prouver que l'ensemble A des $x \in X$ vérifiant $f(x) = g(x)$ est fermé dans X , puisque A contenant Y , et Y étant dense dans X , cela implique $A = X$. Or A est l'image inverse de la diagonale $\Delta = \{(x, x), x \in X\}$ dans $X \times X$ par l'application $x \mapsto (f(x), g(x))$, qui est continue, et l'hypothèse Z séparé est équivalente à ce que Δ soit fermé dans $X \times X$, ce qui fait que A est fermé comme image inverse d'un fermé par une application continue.

Exercice 6.7. — Soit X un espace topologique. Montrer que $Y \subset X$ est d'intérieur vide si et seulement si son complémentaire est dense dans X .

Exercice 6.8. — (i) Montrer que si Y_1 est dense dans X_1 et si Y_2 est dense dans X_2 , alors $Y_1 \times Y_2$ est dense dans $X_1 \times X_2$.

(ii) Soit $f : Y \rightarrow Z$ une application continue entre espaces métriques. Montrer que si X est dense dans Y , et si la restriction de f à X est une isométrie, alors f est une isométrie.

Exercice 6.9. — (i) Montrer que si U est ouvert, alors l'intérieur de l'adhérence de U contient U , et qu'on n'a pas toujours égalité, mais que l'adhérence de l'intérieur de l'adhérence de U est l'adhérence de U .

(ii) Montrer que, si F est fermé, alors l'adhérence de l'intérieur de F est contenu dans F , et qu'on n'a pas toujours égalité, mais que l'intérieur de l'adhérence de l'intérieur de F est l'intérieur de F .

Exercice 6.10. — Montrer que $A = \{(n, e^n), n \in \mathbf{N}\}$ est dense dans \mathbf{C}^2 muni de la topologie de Zariski. Est-t-il dense dans \mathbf{C}^2 pour la topologie usuelle?

6.7. Suites dans un espace topologique

6.7.1. Suites, suites extraites

Soit X un espace topologique. Si $(x_n)_{n \in \mathbf{N}}$ est une suite d'élément de X , et si $a \in X$, on dit que x_n *tend vers* a ou que x_n *a pour limite* a , si pour tout voisinage V de a , il existe $N \in \mathbf{N}$, tel que $x_n \in V$, si $n \geq N$. Il suffit bien évidemment de vérifier ceci pour V dans une base de voisinages de a .

Si X est séparé, une suite a au plus une limite comme on le constate aisément en revenant à la définition d'espace séparé. On prendra garde au fait que ce n'est plus forcément le cas, si l'espace n'est pas séparé. On dit qu'une suite est *convergente* si elle a au moins une limite. On réserve la notation $\lim_{n \rightarrow +\infty} x_n = a$ au cas où l'espace est séparé et donc la limite est unique.

On obtient une traduction agréable de la notion de suite convergente en introduisant l'espace topologique $\overline{\mathbf{N}} = \mathbf{N} \cup \{+\infty\}$, muni de la topologie pour laquelle les ouverts sont les parties de \mathbf{N} auxquelles on a rajouté les complémentaires dans $\overline{\mathbf{N}}$ des parties finies de \mathbf{N} . C'est alors un simple exercice de traduction de montrer que $\lim_{n \rightarrow +\infty} x_n = a$ si et seulement si la suite x_n se prolonge en une fonction continue de $\overline{\mathbf{N}}$ dans X prenant la valeur a en $+\infty$ (i.e. l'application de $\overline{\mathbf{N}}$ dans X obtenue en envoyant n sur x_n et $+\infty$ sur a est continue).

Une suite $(y_n)_{n \in \mathbf{N}}$ est dite *extraite* de $(x_n)_{n \in \mathbf{N}}$ s'il existe $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ tendant vers $+\infty$ quand n tend vers $+\infty$, telle que $y_n = x_{\varphi(n)}$, pour tout $n \in \mathbf{N}$.

- Si a est une limite de $x = (x_n)_{n \in \mathbf{N}}$, alors a est aussi limite de toute suite extraite.

Soit $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ tendant vers $+\infty$ quand n tend vers $+\infty$, ce qui se traduit par le fait que φ peut s'étendre par continuité à $\overline{\mathbf{N}}$, en posant $\varphi(+\infty) = +\infty$. Si a est une limite de x , alors x peut aussi s'étendre par continuité à $\overline{\mathbf{N}}$, en posant $x(+\infty) = a$ et donc $x \circ \varphi$ est continue sur $\overline{\mathbf{N}}$, ce qui se traduit par le fait que a est limite de la suite extraite $(x_{\varphi(n)})_{n \in \mathbf{N}}$.

On peut aussi se passer de $\overline{\mathbf{N}}$, et revenir à la définition. Si V est un voisinage de a , alors il existe $N \in \mathbf{N}$ tel que $x_n \in V$, pour tout $n \geq N$. Par ailleurs, si $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ tend vers $+\infty$

quand n tend vers $+\infty$, il existe $N' \in \mathbf{N}$ tel que $\varphi(n) \geq N$, si $n \geq N'$. On a donc $x_{\varphi(n)} \in V$, pour tout $n \geq N'$, ce qui permet de montrer que $(x_{\varphi(n)})_{n \in \mathbf{N}}$ tend vers a .

6.7.2. Suites et continuité

- Si $f : X \rightarrow Y$ est continue, et si $x = (x_n)_{n \in \mathbf{N}}$ est une suite d'éléments de X admettant a comme limite, alors $(f(x_n))_{n \in \mathbf{N}}$ admet $f(a)$ pour limite.

La suite x se prolonge en une fonction continue de $\overline{\mathbf{N}}$ dans X prenant la valeur a en $+\infty$, et comme f est continue, $f \circ x$ est continue sur $\overline{\mathbf{N}}$, ce qui se traduit par le fait que $f(a)$ est limite de la suite $(f(x_n))_{n \in \mathbf{N}}$.

On peut aussi se passer de $\overline{\mathbf{N}}$, et dire que si V est un voisinage de $f(a)$, alors $f^{-1}(V)$ contient un voisinage U de a puisque f est continue, et qu'il existe $N \in \mathbf{N}$ tel que $x_n \in U$, si $n \geq N$, ce qui implique $f(x_n) \in V$, si $n \geq N$.

- Si X est un espace métrique, alors $f : X \rightarrow Y$ est continue en x si et seulement si pour toute suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de X tendant vers x , la suite $(f(x_n))_{n \in \mathbf{N}}$ tend vers $f(x)$.

On a déjà démontré (dans le cas d'espaces topologiques généraux) que si $f : X \rightarrow Y$ est continue en x , alors pour toute suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de X tendant vers x , la suite $(f(x_n))_{n \in \mathbf{N}}$ tend vers $f(x)$. Maintenant, si f est non continue en x , il existe un voisinage V de $f(x)$, tel que, pour tout $n \in \mathbf{N}$, il existe $x_n \in B(x, 2^{-n})$ avec $f(x_n) \notin V$. Alors $x_n \rightarrow x$ dans X , tandis que $f(x_n) \not\rightarrow f(x)$. En prenant la contraposée, on en déduit que, si pour toute suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de X tendant vers x , la suite $(f(x_n))_{n \in \mathbf{N}}$ tend vers $f(x)$, alors f est continue en x . Ceci permet de conclure.

On prendra garde au fait que cette caractérisation de la continuité par les suites n'est pas valable pour un espace topologique général.

Exercice 6.11. — Soit X un espace métrique (ou métrisable).

(i) Soit $Z \subset X$. Montrer que $a \in X$ est dans l'adhérence \overline{Z} de Z si et seulement si il existe une suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de Z , ayant a pour limite.

(ii) Montrer que Z est dense dans X si et seulement si tout point de X est la limite d'une suite d'éléments de Z .

(iii) Montrer que, si Y est un espace métrique, si f, g sont deux applications continues de X dans Y , telles que l'on ait $f(x) = g(x)$, pour tout $x \in Z$, où Z est dense dans X , alors $f = g$.

7. Compacité

7.1. Espaces compacts

Un espace topologique X est dit *compact* s'il est non vide, séparé, et si de tout recouvrement de X par des ouverts, on peut extraire un sous-recouvrement fini. Autrement dit, X (non vide et séparé) est compact si, quelle que soit la famille $(U_i)_{i \in I}$ d'ouverts de X telle que $\cup_{i \in I} U_i = X$, il existe $J \subset I$ fini tel que $\cup_{i \in J} U_i = X$. En passant aux complémentaires,

on voit que la compacité de X (séparé) est équivalente à ce que de toute famille de fermés de X d'intersection vide, on puisse extraire une famille finie d'intersection vide.

- Un ensemble fini, muni de la topologie discrète, est compact.
- L'espace $\overline{\mathbf{N}} = \mathbf{N} \cup \{+\infty\}$, muni de la topologie pour laquelle les ouverts sont les parties de \mathbf{N} et les complémentaires dans $\overline{\mathbf{N}}$ des parties finies de \mathbf{N} , est un espace compact.

$\overline{\mathbf{N}}$ est séparé car, si $x \neq y$, alors $x \neq +\infty$ ou $y \neq +\infty$, ce qui fait que l'un des deux singletons $\{x\}$ ou $\{y\}$ est ouvert, ainsi que son complémentaire. Par ailleurs, si les $(U_i)_{i \in I}$ forment un recouvrement ouvert de $\overline{\mathbf{N}}$, alors un des U_i contient $+\infty$, et son complémentaire est fini ; on peut donc extraire du recouvrement par les U_i , un sous-recouvrement fini.

- Le segment $[0, 1]$ est compact.

Soit $(U_i)_{i \in I}$ une famille d'ouverts de $[0, 1]$ formant un recouvrement. Soit A l'ensemble des $a \in [0, 1]$ tels que $[0, a]$ puisse être recouvert par un nombre fini de U_i , et soit M la borne supérieure de A . Par hypothèse, il existe $i(M) \in I$ et $\varepsilon > 0$ tels que $]M - \varepsilon, M + \varepsilon[\cap [0, 1] \subset U_{i(M)}$, et par définition de M , il existe $a \in]M - \varepsilon, M[$ et $J \subset I$ fini, tels que $[0, a] \subset \cup_{i \in J} U_i$. Mais alors $[0, b] \subset \cup_{i \in J \cup \{i(M)\}} U_i$, quel que soit $b \in [M, M + \varepsilon[\cap [0, 1]$, et donc $[M, M + \varepsilon[\cap [0, 1] \subset A$. Par définition de M , ceci implique $M = 1$ et permet de conclure.

7.2. Compacité et suites

Si X est un espace topologique, et si $(x_n)_{n \in \mathbf{N}}$ est une suite d'éléments de X , on dit que $a \in X$ est une *valeur d'adhérence* de la suite $(x_n)_{n \in \mathbf{N}}$, si tout voisinage de a contient une infinité de termes de la suite. Ceci équivaut à ce que a soit dans l'adhérence F_k de $\{x_n, n \geq k\}$, pour tout $k \in \mathbf{N}$. En particulier, l'ensemble des valeurs d'adhérence d'une suite est un fermé, puisque c'est l'intersection des fermés F_k , pour $k \in \mathbf{N}$.

- Si X est un espace métrique, alors a est une valeur d'adhérence de la suite $(x_n)_{n \in \mathbf{N}}$, si et seulement si on peut extraire une sous-suite de la suite $(x_n)_{n \in \mathbf{N}}$ ayant pour limite a .

Si on peut extraire de $(x_n)_{n \in \mathbf{N}}$, une sous-suite $(x_{\varphi(n)})_{n \in \mathbf{N}}$ de limite a , et si V est un voisinage de a , alors $x_{\varphi(n)} \in V$, pour tout n assez grand, ce qui prouve que a est une valeur d'adhérence de la suite (noter que ce sens n'utilise pas le fait que X est métrique). Réciproquement, si X est métrique, et si a est une valeur d'adhérence de $(x_n)_{n \in \mathbf{N}}$, alors pour tout $n \in \mathbf{N}$, il existe une infinité de termes de la suite dans $B(a, 2^{-n})$, et donc on peut choisir $\varphi(n) \geq n$ tel que $x_{\varphi(n)} \in B(a, 2^{-n})$. La suite $(x_{\varphi(n)})_{n \in \mathbf{N}}$ est alors extraite de la suite $(x_n)_{n \in \mathbf{N}}$ et converge vers a . Ceci permet de conclure.

- Dans un compact, toute suite admet une valeur d'adhérence ; dans un compact métrique, on peut extraire de toute suite une sous-suite convergente.

Soit X un compact, et soit $(x_n)_{n \in \mathbf{N}}$ une suite d'éléments de X . Soit F_n , si $n \in \mathbf{N}$ l'adhérence de l'ensemble $\{x_{n+p}, p \in \mathbf{N}\}$; l'intersection des F_n est, par définition ou presque, l'ensemble des valeurs d'adhérence de la suite $(x_n)_{n \in \mathbf{N}}$. Comme l'intersection d'un nombre fini de F_n est toujours non vide puisqu'elle contient les x_n , pour n assez grand, la compacité de X assure que l'intersection des fermés F_n , pour $n \in \mathbf{N}$, est non vide, ce qui permet de conclure.

Exercice 7.1. — (i) Montrer que dans un compact, une suite ayant une seule valeur d'adhérence a une limite.

(ii) Donner un exemple de suite à valeurs dans \mathbf{R} ayant une seule valeur d'adhérence mais ne convergeant pas.

• Un espace métrique est compact si et seulement si toute suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de X admet une valeur d'adhérence⁽²⁶⁾ (théorème de Borel-Lebesgue).

On sait déjà que dans un compact (même non métrique), toute suite admet une valeur d'adhérence ; montrons la réciproque dans le cas d'un espace métrique. Soit $(U_i)_{i \in I}$ un recouvrement ouvert de X . Alors, quel que soit $x \in X$, il existe $k(x) \geq 0$ et $i \in I$, tels que $B(x, r(x)^-) \subset U_i$, où $r(x) = 2^{-k(x)}$. On cherche à prouver qu'on peut extraire du recouvrement par les U_i un recouvrement fini, et il suffit de prouver qu'on peut en faire autant du recouvrement par les $B(x, r(x)^-)$.

Pour cela, construisons par récurrence une suite x_n d'éléments de X vérifiant :

- $x_n \in Y_n$, où Y_n est le fermé complémentaire de $\cup_{j \leq n-1} B(x_j, r(x_j)^-)$,
- $k(x_n) \leq k(y)$, quel que soit $y \in Y_n$.

Si la construction s'arrête, c'est que les $B(x_j, r(x_j)^-)$, pour $j \leq n-1$ recouvrent X , ce que l'on veut. Sinon, la suite $(x_n)_{n \in \mathbf{N}}$ a une valeur d'adhérence y_0 , et on a $y_0 \in Y_n$, quel que soit $n \in \mathbf{N}$, car Y_n est fermé et $x_{n+p} \in Y_n$, quel que soit $p \in \mathbf{N}$. Par construction de la suite $(x_n)_{n \in \mathbf{N}}$, on a $d(x_n, x_{n+p}) \geq 2^{-k(x_n)}$, quels que soient $n, p \in \mathbf{N}$. Comme on peut extraire une sous-suite de Cauchy de la suite $(x_n)_{n \in \mathbf{N}}$, on en déduit que $k(x_n) \rightarrow +\infty$. En particulier, il existe n tel que $k(x_n) \geq k(y_0) + 1$, en contradiction avec la construction de x_n (puisque $y_0 \in Y_n$). Ceci permet de conclure.

Exercice 7.2. — Montrer que $[0, 1]$ est compact en passant par les suites.

Exercice 7.3. — (i) Montrer que pour tout $\varepsilon > 0$, il existe une suite de segments ouverts $]a_n, b_n[$ telle que $\sum_{n \in \mathbf{N}} (b_n - a_n) < \varepsilon$ et l'adhérence de $\cup_{n \in \mathbf{N}}]a_n, b_n[$ contienne $[0, 1]$.

(ii) Soit $]a_n, b_n[$, pour $n \in \mathbf{N}$, une suite de segments ouverts tels que $[0, 1] \subset \cup_{n \in \mathbf{N}}]a_n, b_n[$. Montrer que $\sum_{n \in \mathbf{N}} (b_n - a_n) > 1$. (On pourra admettre que le résultat est vrai pour une famille finie.)

7.3. Propriétés de base des compacts

Les énoncés qui suivent sont d'un usage constant.

7.3.1. Compacts d'un espace topologique

• Si X est compact, alors $Y \subset X$ est compact, si et seulement si Y est fermé.

⁽²⁶⁾Cette caractérisation est parfois prise comme définition des espaces compacts. Elle est effectivement d'un maniement plus facile que la caractérisation en termes de recouvrements ouverts si on cherche à vérifier qu'un espace (métrique) est compact. Par contre, si on veut utiliser la compacité d'un espace pour en tirer des conséquences, c'est en général la caractérisation par les recouvrements ouverts qui est la plus naturelle et la plus puissante.

Supposons Y fermé. Soit $(U_i)_{i \in I}$ un recouvrement⁽²⁷⁾ ouvert de Y . Par définition, il existe, pour tout $i \in I$, un ouvert V_i de X tel que $U_i = V_i \cap Y$, et comme $U = X - Y$ est ouvert, les V_i , pour $i \in I$ et U forment un recouvrement ouvert de X . Comme X est supposé compact, il existe $J \subset I$ fini, tel que $X \subset U \cup (\cup_{i \in J} V_i)$, et les U_i , pour $i \in J$ forment un recouvrement ouvert de Y extrait du recouvrement initial. On en déduit la compacité de Y .

Réciproquement, supposons $Y \subset X$ compact. Soit $a \notin Y$. Comme X est séparé, pour tout $y \in Y$, il existe des ouverts U_y, V_y tels que $y \in U_y, a \in V_y$ et $U_y \cap V_y = \emptyset$. Les U_y , pour $y \in Y$, forment un recouvrement ouvert de Y ; il existe donc $J \subset Y$ fini tel que $Y \subset \cup_{y \in J} U_y$. Mais alors $V = \cap_{y \in J} V_y$ est un ouvert de X contenant a et ne rencontrant pas Y , ce qui prouve que a n'appartient pas à l'adhérence \bar{Y} de Y . On a donc $\bar{Y} \subset Y$, ce qui prouve que Y est fermé.

- Si X_1 et X_2 sont compacts, alors $X_1 \times X_2$ est compact.

Soit $(U_i)_{i \in I}$ une famille d'ouverts de $X_1 \times X_2$ formant un recouvrement⁽²⁸⁾. Si $y \in X_2$, soit $I(y)$ l'ensemble des $i \in I$ tels que $U_i \cap (X_1 \times \{y\})$ soit non vide. Si $i \in I(y)$, et si $(a, y) \in U_i$, il existe $V_{i,y,a}$ ouvert de X_1 contenant a et $W_{i,y,a}$ ouvert de X_2 contenant y tels que $U_i \supset V_{i,y,a} \times W_{i,y,a}$. Les U_i , pour i dans I , forment un recouvrement de $X_1 \times X_2$, les $V_{i,y,a}$, pour $i \in I(y)$ et $(a, y) \in U_i$, forment un recouvrement de X_1 . Comme X_1 est compact, il existe un ensemble fini $J(y)$ de couples (i, a) , avec $i \in I(y)$ et $(a, y) \in U_i$ tels que $X_1 = \cup_{(i,a) \in J(y)} V_{i,y,a}$. Soit alors $W_y = \cap_{(i,a) \in J(y)} W_{i,y,a}$. C'est un ouvert de X_2 contenant y , et U_i contient $V_{i,y,a} \times W_y$, quel que soit $(i, a) \in J(y)$. Comme X_2 est compact, on peut trouver Y fini tel que $X_2 = \cup_{y \in Y} W_y$, et alors

$$\cup_{y \in Y} \cup_{(i,a) \in J(y)} U_i \supset \cup_{y \in Y} (\cup_{(i,a) \in J(y)} V_{i,y,a} \times W_y) = \cup_{y \in Y} (X_1 \times W_y) = X_1 \times X_2,$$

ce qui montre que l'on peut extraire du recouvrement par les U_i un sous-recouvrement fini.

- L'image d'un compact X par une application continue $f : X \rightarrow Y$, où Y est séparé, est un compact.

Soit $(U_i)_{i \in I}$ un recouvrement⁽²⁹⁾ ouvert de $f(X)$. Par définition, si $i \in I$, il existe U'_i ouvert de Y tel que $U_i = U'_i \cap f(X)$, et comme f est continue, $V_i = f^{-1}(U'_i)$ est ouvert dans X , et $(V_i)_{i \in I}$ est donc un recouvrement ouvert de X . Comme X est compact, il existe $J \subset I$ fini tels

⁽²⁷⁾ Si X est un espace métrique, on peut passer par les suites. Comme X est compact, une suite $(y_n)_{n \in \mathbf{N}}$ d'éléments de Y a une valeur d'adhérence dans X , et si Y est fermé, cette valeur d'adhérence est dans Y , ce qui prouve que Y est compact. Réciproquement, si Y est compact, si a est dans l'adhérence de Y , il existe une suite $(y_n)_{n \in \mathbf{N}}$ d'éléments de Y ayant pour limite a dans X , et sa seule valeur d'adhérence dans X est alors a . Comme Y est supposé compact, cette suite admet une valeur d'adhérence dans Y , et comme sa seule valeur d'adhérence dans X est a , cela implique $a \in Y$. On en déduit que Y est fermé.

⁽²⁸⁾ Si X_1 et X_2 sont des espaces métriques, on peut raisonner en termes de suites. Soit $(x_n, y_n)_{n \in \mathbf{N}}$ une suite d'éléments de $X_1 \times X_2$. Comme X_1 est compact, on peut extraire de la suite $(x_n)_{n \in \mathbf{N}}$ une sous-suite $(x_{\varphi(n)})_{n \in \mathbf{N}}$ ayant une limite a dans X_1 . Comme X_2 est compact, on peut extraire de la suite $(y_{\varphi(n)})_{n \in \mathbf{N}}$ une sous-suite $(y_{\psi(n)})_{n \in \mathbf{N}}$ ayant une limite b dans X_2 , et alors $(x_{\psi(n)}, y_{\psi(n)})_{n \in \mathbf{N}}$ admet (a, b) comme limite dans $X_1 \times X_2$ puisque $(x_{\psi(n)})_{n \in \mathbf{N}}$ est extraite de $(x_{\varphi(n)})_{n \in \mathbf{N}}$, et donc tend vers a dans X_1 . Autrement dit la suite $(x_n, y_n)_{n \in \mathbf{N}}$ admet une valeur d'adhérence.

⁽²⁹⁾ Si X et Y sont des espaces métriques, on peut raisonner en termes de suites. Soit $(y_n)_{n \in \mathbf{N}}$ une suite d'éléments de $f(Y)$, et, si $n \in \mathbf{N}$, soit $x_n \in X$ tel que $y_n = f(x_n)$. Comme X est compact, la suite $(x_n)_{n \in \mathbf{N}}$ admet une valeur d'adhérence $a \in X$, et comme f est continue, $f(a)$ est une valeur d'adhérence de la suite $(y_n)_{n \in \mathbf{N}}$. On en déduit la compacité de $f(X)$.

que les V_i , pour $i \in J$, recouvrent X , et les U_i , pour $i \in J$, forment alors un recouvrement ouvert fini de $f(X)$ extrait du recouvrement initial. On en déduit la compacité de $f(X)$.

- Si X est compact, et si $f : X \rightarrow Y$ est bijective continue avec Y séparé, alors f est un homéomorphisme.

Notons $g : Y \rightarrow X$ l'application réciproque de f . Il suffit de prouver que, si F est fermé dans X , alors $g^{-1}(F)$ est fermé dans Y . Or $g^{-1}(F) = \{y \in Y \mid \exists x \in F, g(y) = x\} = \{y \in Y \mid \exists x \in F, y = f(g(y)) = f(x)\} = f(F)$, et comme F est compact puisque fermé dans un compact, et que Y est séparé, $f(F)$ est compact et donc fermé. Ceci permet de conclure.

7.3.2. Compacts d'un espace métrique

- Si E est un espace métrique, un compact X de E est fermé dans E et borné, mais la réciproque est en générale fausse.

On a déjà vu qu'un compact est toujours fermé. Par ailleurs, si X est compact, et si $x_0 \in X$, alors $x \mapsto d(x_0, x)$ est continue sur X et donc est bornée puisque toute fonction continue à valeurs réelles sur un compact est bornée. Autrement dit, il existe $M \in \mathbf{R}_+$ tel que $X \subset B(x_0, M)$, et X est borné.

Soit E le segment $[-1, 1[$ de \mathbf{R} muni de la distance induite par la valeur absolue sur \mathbf{R} ; c'est un espace métrique parfaitement respectable. Alors $X = [0, 1[$ est fermé dans E puisque c'est l'intersection de E avec le fermé \mathbf{R}_+ de \mathbf{R} , et il est borné. Il n'est pas compact car on ne peut pas extraire de recouvrement fini du recouvrement de X par les ouverts $U_n = X \cap]\frac{1}{2}, 1 - \frac{1}{n}[$.

- Si X est compact, et $f : X \rightarrow \mathbf{R}$ est continue, alors f atteint son maximum et son minimum.

Comme X est compact et f continue, cela implique que $f(X)$ est compact, et donc admet des borne inférieure et supérieure finies car $f(X)$ est borné, et les contient car il est fermé.

- Si E est un espace vectoriel *de dimension finie* sur \mathbf{R} ou \mathbf{C} , alors les compacts de E sont les fermés bornés⁽³⁰⁾.

Par définition de la norme $\|\cdot\|_\infty$ sur \mathbf{R}^n , un borné de \mathbf{R}^n est inclus dans $[-M, M]^n$, si M est assez grand. Or $[-M, M]$ est compact, puisque c'est l'image de $[0, 1]$ par l'application continue $x \mapsto (2x - 1)M$, et donc $[-M, M]^n$ est compact comme produit de compacts. Comme un fermé d'un compact est compact, on en déduit qu'un fermé borné de $(\mathbf{R}^n, \|\cdot\|_\infty)$ est compact. Le résultat dans le cas d'un \mathbf{R} ou \mathbf{C} -espace vectoriel de dimension finie quelconque s'en déduit si on sait que deux normes sur un \mathbf{R} -espace vectoriel de dimension finie sont équivalentes (cf. n° 11.3), et donc que les fermés bornés sont les mêmes, quelle que soit la norme.

- Une fonction continue sur un compact d'un espace métrique est uniformément continue.

$f : X \rightarrow Y$, où X et Y sont des espaces métriques, est uniformément continue si

$$\forall \varepsilon > 0, \exists \delta > 0, \text{ tel que } d_X(x, x') < \delta \Rightarrow d_Y(y, y') < \varepsilon.$$

⁽³⁰⁾En filière PC, cette propriété est prise comme définition de compact; on peut difficilement imaginer un point de vue plus nocif : être fermé est une notion relative (un ensemble est toujours fermé dans lui-même), alors que la compacité est une notion intrinsèque. Qui plus est, cette propriété devient fausse en dimension infinie, et les espaces de dimension infinie ne sont pas qu'une lubie de mathématicien.

Supposons X compact. Soit $\varepsilon > 0$. Comme f est continue, pour tout $x \in X$, il existe $\delta_x > 0$ tel que $d_X(x, x') < 2\delta_x \Rightarrow d_Y(f(x), f(x')) < \frac{\varepsilon}{2}$. Les $B_X(x, \delta_x)$ forment un recouvrement⁽³¹⁾ ouvert de X ; on peut donc en extraire un recouvrement fini $X \subset \bigcup_{x \in J} B_X(x, \delta_x)$, où $J \subset X$ est fini. Alors, par construction, si $x' \in X$, il existe $x \in J$ tel que $d_X(x, x') < \delta_x$. Soit alors $\delta = \inf_{x \in J} \delta_x$. Si $x_1, x_2 \in X$ vérifient $d_X(x_1, x_2) < \delta$, et si $x \in J$ est tel que $d_X(x, x_1) < \delta_x$, alors $d_X(x, x_2) < 2\delta_x$, et donc $d_Y(f(x), f(x_1)) < \frac{\varepsilon}{2}$, $d_Y(f(x), f(x_2)) < \frac{\varepsilon}{2}$ et $d_Y(f(x_2), f(x_1)) < \varepsilon$. Ceci montre que f est uniformément continue.

Exercice 7.4. — Soit $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie. On dit que $f : E \rightarrow \mathbf{C}$ tend vers 0 à l'infini, si pour tout $\varepsilon > 0$, il existe $M > 0$, tel que $|f(x)| < \varepsilon$, si $\|x\| \geq M$. Montrer que, si $f : E \rightarrow \mathbf{C}$ est continue et tend vers 0 à l'infini, alors f est bornée et $|f|$ atteint son maximum.

Exercice 7.5. — Soit (X, d) un espace métrique. Si $F \subset X$, et si $x \in X$, on définit la distance $d(x, F)$ de x à F comme la borne inférieure des $d(x, y)$, pour $y \in F$.

- (i) Montrer que $x \mapsto d(x, F)$ est continue et même 1-lipschitzienne sur X .
- (ii) Montrer que $d(x, F) = 0$ si et seulement si x est dans l'adhérence \overline{F} de F .
- (iii) En déduire que si F_1 et F_2 sont des fermés disjoints de X , il existe des ouverts disjoints U_1, U_2 de X avec $F_1 \subset U_1$ et $F_2 \subset U_2$.
- (iv) On définit la distance $d(F_1, F_2)$ entre F_1 et F_2 comme la borne inférieure de $d(x, y)$, pour $x \in F_1$ et $y \in F_2$. Montrer que si F_1 et F_2 sont des compacts disjoints, alors $d(F_1, F_2) > 0$.
- (v) Montrer que si $F_1, F_2 \subset X$, si $F_1 \cap F_2 = \emptyset$, si F_1 est fermé et si F_2 est compact, alors $d(F_1, F_2) \neq 0$.
- (vi) Construire des fermés disjoints de \mathbf{R} ou \mathbf{R}^2 dont la distance est nulle.

Exercice 7.6. — Soient X un compact métrique et $f : X \rightarrow X$ une application contractante (i.e. vérifiant $d(f(x), f(y)) < d(x, y)$, quels que soient $x \neq y$).

- (i) Montrer que f a un unique point fixe x_0 .
- (ii) Montrer que si $x \in X$, et si $f^n = f \circ \dots \circ f$ (n fois), alors $f^n(x) \rightarrow x_0$.
- (iii) Montrer que $f^n \rightarrow f$ uniformément sur X .

Exercice 7.7. — (difficile) Soit X un espace métrique. Montrer que si toute fonction continue de X dans \mathbf{R} est bornée, alors X est compact.

⁽³¹⁾ Comme on travaille avec des espaces métriques, on peut aussi passer par les suites. Supposons donc que X est compact, que $f : X \rightarrow Y$ est continue mais pas uniformément continue. En niant la définition de la continuité uniforme rappelée ci-dessus, on voit qu'il existe $\varepsilon > 0$, tel que, quel que soit $n \in \mathbf{N}$, il existe $(x_n, x'_n) \in X \times X$ tels que $d_X(x_n, x'_n) \leq 2^{-n}$ et $d_Y(f(x_n), f(x'_n)) \geq \varepsilon$. Comme X est supposé compact, il en est de même de $X \times X$, et la suite $(x_n, x'_n)_{n \in \mathbf{N}}$ admet une valeur d'adhérence (a, b) dans $X \times X$. De plus, comme $d_X(x_n, x'_n) \rightarrow 0$, on a $a = b$, et comme f est continue, $(f(a), f(b))$ est une valeur d'adhérence de la suite $(f(x_n), f(x'_n))_{n \in \mathbf{N}}$ dans $Y \times Y$. Comme $f(a) = f(b)$, cela est en contradiction avec le fait que $d_Y(f(x_n), f(x'_n)) \geq \varepsilon$, quel que soit $n \in \mathbf{N}$ (en effet, $(y, y') \mapsto d_Y(y, y')$ est continue sur $Y \times Y$, et une valeur d'adhérence (c, c') de la suite $(f(x_n), f(x'_n))_{n \in \mathbf{N}}$ doit donc vérifier $d_Y(c, c') \geq \varepsilon > 0$). Ceci permet de conclure.

7.4. La droite réelle achevée

7.4.1. Les espaces topologiques ordonnés $\overline{\mathbf{R}}$ et $\overline{\mathbf{R}}_+$

On note $\overline{\mathbf{R}} = \mathbf{R} \cup \{\pm\infty\}$ la *droite réelle achevée*. On étend \leq de manière naturelle en une relation d'ordre totale sur $\overline{\mathbf{R}}$, en convenant que $-\infty \leq a \leq +\infty$, quel que soit $a \in \mathbf{R}$. On fait de $\overline{\mathbf{R}}$ un espace topologique, en prenant les $]a, b[$, pour $a < b \in \mathbf{R}$, et les $[-\infty, a[$ et $]a, +\infty]$, pour $a \in \mathbf{R}$, comme base d'ouverts. La topologie induite sur \mathbf{R} est donc la topologie usuelle.

- Une suite de nombres réels x_n tend vers $+\infty$ dans $\overline{\mathbf{R}}$ si et seulement si x_n tend vers $+\infty$ au sens classique. (Idem pour $-\infty$.)

Les $]a, +\infty]$ forment une base de voisinages de $+\infty$, et donc $x_n \rightarrow +\infty$ dans $\overline{\mathbf{R}}$ si et seulement si, quel que soit $a \in \mathbf{R}$, il existe $N \in \mathbf{N}$ tel que $x_n \in]a, +\infty]$, si $n \geq N$.

- L'espace topologique $\overline{\mathbf{R}}$ est isomorphe à $[-1, 1]$ en tant qu'espace ordonné et en tant qu'espace topologique ; en particulier, il est compact et métrisable, et tout sous-ensemble non vide de $\overline{\mathbf{R}}$ admet une borne inférieure et une borne supérieure.

L'application $x \mapsto f(x)$, avec $f(x) = \frac{x}{1+|x|}$, si $x \in \mathbf{R}$, $f(+\infty) = 1$ et $f(-\infty) = -1$, est un homéomorphisme strictement croissant de $\overline{\mathbf{R}}$ sur $[-1, 1]$, dont l'inverse est g défini par $g(x) = \frac{x}{1-|x|}$, si $x \in \mathbf{R}$, $g(1) = +\infty$, $g(-1) = -\infty$ (nous laissons au lecteur le soin de vérifier que f et g sont bien des applications continues inverses l'une de l'autre).

- Une suite $(x_n)_{n \in \mathbf{N}}$ croissante (resp. décroissante) d'éléments de $\overline{\mathbf{R}}$ converge vers la borne supérieure (resp. inférieure) de $\{x_n, n \in \mathbf{N}\}$.
- Si $X \subset \overline{\mathbf{R}}$ est non vide, alors $\sup X$ et $\inf X$ sont dans l'adhérence de X .

En utilisant l'homéomorphisme $f : \overline{\mathbf{R}} \rightarrow [-1, 1]$, qui est strictement croissant, on se ramène à démontrer le même énoncé pour $X \subset [-1, 1]$ ce qui permet de traiter tous les cas de la même manière. Maintenant, si la borne supérieure M de X appartient à X , elle appartient a fortiori à son adhérence. Si M n'appartient pas à X , alors pour tout $n > 0$, il existe $x_n \in X$ avec $M - 2^{-n} < x_n < M$, ce qui prouve que M est limite d'une suite d'éléments de X et donc est dans son adhérence. Ceci permet de conclure.

On note $\overline{\mathbf{R}}_+$ la *demi-droite achevée*. C'est l'ensemble des $x \in \overline{\mathbf{R}}$ vérifiant $x \geq 0$. On étend l'addition à $\overline{\mathbf{R}}_+$ de la manière évidente, en posant $x + (+\infty) = +\infty$, si $x \in \overline{\mathbf{R}}_+$.

Comme toute suite croissante d'éléments de $\overline{\mathbf{R}}_+$ admet une limite dans $\overline{\mathbf{R}}_+$, on en déduit que :

- Toute série $\sum_{n \in \mathbf{N}} u_n$ à termes dans $\overline{\mathbf{R}}_+$ converge dans $\overline{\mathbf{R}}_+$. Si les u_n sont dans \mathbf{R}_+ , alors $\sum_{n \in \mathbf{N}} u_n < +\infty$ si et seulement si la série $\sum_{n \in \mathbf{N}} u_n$ converge au sens usuel.

7.4.2. Limite supérieure, limite inférieure

- Toute suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de $\overline{\mathbf{R}}$ admet une plus grande valeur d'adhérence $\limsup x_n$, *limite supérieure* de la suite x_n et une plus petite valeur d'adhérence $\liminf x_n$, *limite inférieure* de la suite x_n . De plus, $(x_n)_{n \in \mathbf{N}}$ converge si et seulement si ses limites supérieure

et inférieure sont égales, et la limite de la suite est alors la valeur commune des limites supérieure et inférieure⁽³²⁾.

La compacité de $\overline{\mathbf{R}}$ implique que l'ensemble des valeurs d'adhérence d'une suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de $\overline{\mathbf{R}}$ est non vide. Comme cet ensemble est fermé, les bornes inférieure et supérieure de cet ensemble sont encore des valeurs d'adhérence; autrement dit toute suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de $\overline{\mathbf{R}}$ admet une plus grande valeur d'adhérence. De plus, comme $\overline{\mathbf{R}}$ est un espace compact métrisable, une suite converge si et seulement si elle a une seule valeur d'adhérence et donc si et seulement si ses limites supérieure et inférieure sont égales. On en déduit le résultat.

- On a aussi $\limsup x_n = \inf_{k \in \mathbf{N}} \left(\sup_{n \geq k} x_n \right)$ et $\liminf x_n = \sup_{k \in \mathbf{N}} \left(\inf_{n \geq k} x_n \right)$.

Pour éviter d'avoir à traiter séparément les cas où une des limites est infinie, on utilise l'homéomorphisme $f : \overline{\mathbf{R}} \rightarrow [-1, 1]$ ci-dessus pour se ramener au cas de suites à valeurs dans $[-1, 1]$. Soient $a = \limsup x_n$ et $b = \inf_{k \in \mathbf{N}} \left(\sup_{n \geq k} x_n \right)$, et soit $\varepsilon > 0$. Comme a est une valeur d'adhérence, il existe pour tout $k \in \mathbf{N}$, un entier $n \geq k$ tel que $|x_n - a| < \varepsilon$. On a donc $\sup_{n \geq k} x_n \geq a - \varepsilon$, pour tout k , et donc $b \geq a - \varepsilon$, pour tout $\varepsilon > 0$. On en déduit que $b \geq a$. Par ailleurs, comme a est la plus grande valeur d'adhérence, il n'y a qu'un nombre fini de n tels que $x_n \geq a + \varepsilon$, et donc $\sup_{n \geq k} x_n \leq a + \varepsilon$, si k est assez grand, et $b \leq a + \varepsilon$, pour tout $\varepsilon > 0$. On en déduit que $b \leq a$, ce qui permet de démontrer la première égalité. La seconde se démontre de même en renversant les inégalités.

7.5. L'espace topologique \mathbf{R}/\mathbf{Z}

\mathbf{Z} étant un sous-groupe de \mathbf{R} pour l'addition, on peut considérer le quotient \mathbf{R}/\mathbf{Z} qui est un groupe commutatif; on le munit de la topologie quotient, ce qui en fait un espace topologique.

- L'application $x \mapsto \exp(2i\pi x)$ induit des homéomorphismes de \mathbf{R}/\mathbf{Z} et $[0, 1]/(0 \sim 1)$, munis de la topologie quotient, sur le cercle⁽³³⁾ $S^1 = \{z \in \mathbf{C}, |z| = 1\}$ muni de la topologie induite par celle de \mathbf{C} . En particulier, \mathbf{R}/\mathbf{Z} est un espace compact métrisable.

Notons $\pi : \mathbf{R} \rightarrow \mathbf{R}/\mathbf{Z}$ l'application naturelle et $f : \mathbf{R} \rightarrow S^1$ l'application $x \mapsto \exp(2i\pi x)$. Comme f est périodique de période 1, elle induit une application \bar{f} de \mathbf{R}/\mathbf{Z} dans S^1 qui est bijective de manière évidente, et on a $f = \bar{f} \circ \pi$ par construction. De plus, f est continue de \mathbf{R} dans \mathbf{C} , et donc \bar{f} est continue de \mathbf{R}/\mathbf{Z} (muni de la topologie quotient) dans S^1 (muni de la topologie induite par celle de \mathbf{C}). Comme f est injective et comme S^1 est séparé car métrique, on en déduit que \mathbf{R}/\mathbf{Z} est séparé (cf. ex. 6.6).

Maintenant, l'application $x \mapsto x$ de $[0, 1]$ dans \mathbf{R} est continue, et donc la composée avec π est une application continue de $[0, 1]$ dans \mathbf{R}/\mathbf{Z} qui est surjective. Comme la seule relation modulo \mathbf{Z} entre les éléments de $[0, 1]$ est $0 \sim 1$, cette application continue induit, par passage au quotient, une injection continue $\iota : [0, 1]/(0 \sim 1) \rightarrow \mathbf{R}/\mathbf{Z}$, et comme elle est surjective, c'est une bijection continue de $[0, 1]/(0 \sim 1)$ sur \mathbf{R}/\mathbf{Z} . Comme \mathbf{R}/\mathbf{Z} est séparé, on en déduit, par le même argument que ci-dessus, que $[0, 1]/(0 \sim 1)$ est séparé. Comme $[0, 1]$ est compact

⁽³²⁾Ça a l'air un peu tautologique, mais il est très utile de disposer des quantités $\limsup x_n$ et $\liminf x_n$ sans aucune hypothèse sur la suite $(x_n)_{n \in \mathbf{N}}$.

⁽³³⁾Visuellement, si on prend un segment et qu'on attache ses deux extrémités, on obtient un cercle.

et comme l'application naturelle de $[0, 1]$ dans $[0, 1]/(0 \sim 1)$ est continue par définition de la topologie quotient, on en déduit que :

- $[0, 1]/(0 \sim 1)$ est compact ;
- $\iota : [0, 1]/(0 \sim 1) \rightarrow \mathbf{R}/\mathbf{Z}$ est un homéomorphisme et \mathbf{R}/\mathbf{Z} est compact ;
- $\bar{f} : \mathbf{R}/\mathbf{Z} \rightarrow S^1$ est un homéomorphisme.

Ceci permet de conclure.

Ces diverses identifications permettent de voir un lacet γ dans un espace topologique X comme, au choix :

- une application continue $\gamma : S^1 \rightarrow X$,
- une application continue $\gamma : \mathbf{R} \rightarrow X$, périodique de période 1,
- une application continue $\gamma : \mathbf{R}/\mathbf{Z} \rightarrow X$,
- une application continue $\gamma : [0, 1] \rightarrow X$ vérifiant $\gamma(1) = \gamma(0)$.

C'est cette dernière description qui est utilisée la plupart du temps dans le cours.

8. Connexité

8.1. Ensembles connexes

- Si X est un espace topologique, les propriétés suivantes sont équivalentes :
 - (i) toute application continue de X dans $\{0, 1\}$ (muni de la topologie discrète) est constante ;
 - (ii) toute application continue de X dans un espace topologique discret Y est constante ;
 - (iii) X ne peut pas s'écrire comme réunion de deux ouverts non vides disjoints ;
 - (iv) X ne peut pas s'écrire comme réunion de deux fermés non vides disjoints ;
 - (v) si $Y \subset X$ est à la fois ouvert et fermé, alors $Y = \emptyset$ ou $Y = X$.

L'implication (ii) \Rightarrow (i) suit juste de ce que $\{0, 1\}$ est un ensemble discret. Réciproquement, si Y est discret, toute application $g : Y \rightarrow \{0, 1\}$ est continue ; on en déduit que si X vérifie (i), et $f : X \rightarrow Y$ est continue, alors toute application composée $g \circ f : X \rightarrow \{0, 1\}$ est constante, ce qui implique que f est constante. Les conditions (i) et (ii) sont équivalentes.

Maintenant, si $f : X \rightarrow \{0, 1\}$ est continue, alors $U_1 = f^{-1}(\{0\})$ et $U_2 = f^{-1}(\{1\})$ sont ouverts puisque $\{0\}$ et $\{1\}$ sont ouverts dans $\{0, 1\}$, sont disjoints, et $X = U_1 \cup U_2$. Réciproquement, si U_1 et U_2 sont ouverts, disjoints et si $X = U_1 \cup U_2$, l'application $f : X \rightarrow \{0, 1\}$ définie par $f(x) = 0$, si $x \in U_1$ et $f(x) = 1$ si $x \in U_2$ est continue. On en déduit qu'il existe $f : X \rightarrow \{0, 1\}$ continue non constante si et seulement si on peut écrire X comme réunion de deux ouverts non vides disjoints. On en déduit l'équivalence de (i) et (iii). L'équivalence des autres propriétés avec (iii) est immédiate.

Un espace topologique X est *connexe* s'il est non vide et vérifie une des (et donc toutes les) propriétés équivalentes précédentes.

- Si X_1 et X_2 sont deux ensembles connexes avec $X_1 \cap X_2 \neq \emptyset$, alors $X_1 \cup X_2$ est connexe.

Soit $f : X_1 \cup X_2 \rightarrow \{0, 1\}$ continue. Les restrictions de f à X_1 et X_2 sont continues et donc constantes. Comme on a supposé $X_1 \cap X_2 \neq \emptyset$, on peut choisir $y \in X_1 \cap X_2$, et f vaut $f(y)$ sur X_1 et X_2 ; par suite elle est constante sur $X_1 \cup X_2$. On en déduit la connexité de $X_1 \cup X_2$.

Ceci permet, si X est un espace topologique quelconque, et $x \in X$, de définir la *composante connexe* C_x de x dans X comme le plus grand sous-ensemble connexe de X contenant x ; c'est la réunion de tous les connexes de X contenant x . On appelle *composante connexe de X* tout sous-ensemble de la forme C_x , pour $x \in X$. On a $y \in C_x$ si et seulement si $C_y = C_x$, ce qui fait que les composantes connexes de X forment une partition de X , la *partition en composantes connexes*. Un ensemble est *totalement discontinu* si les composantes connexes sont réduites à un point.

- Dans \mathbf{R} , les connexes sont les segments (tous les segments, i.e. les $[a, b]$, $[a, b[$, $]a, b]$, $]a, b[$, pour $a, b \in \mathbf{R}$, ainsi que les demi-droites ou \mathbf{R} tout entier obtenus en permettant à a ou b de prendre les valeurs $\pm\infty$).

Si $X \subset \mathbf{R}$ n'est pas un segment, c'est qu'il existe $a \notin X$ et $x_1, x_2 \in X$, avec $x_1 < a$ et $x_2 > a$. Alors $U_1 = X \cap]-\infty, a[$ et $U_2 = X \cap]a, +\infty[$ sont des ouverts de X , qui sont non vides, disjoints, et dont la réunion est X , ce qui prouve que X n'est pas connexe. Autrement dit, si X est connexe, alors X est un segment.

Maintenant, soient $a \leq b$, et soit $f : [a, b] \rightarrow \{0, 1\}$ continue. Quitte à remplacer f par $1 - f$, on peut supposer que $f(a) = 0$. Soit $X = \{x \in [a, b], f(x) = 1\}$, et soit c la borne inférieure de X , si X n'est pas vide. Par définition de c , il existe une suite d'éléments de X (qui peut être la suite constante c , si $c \in X$) ayant pour limite c , et comme f est continue, on a $f(c) = 1$. En particulier, on a $c \neq a$, et si $x \in [a, c[$, alors $f(x) = 0$, par définition de c . Comme f est continue et comme c est dans l'adhérence de $[a, c[$, cela implique que $f(c) = 0$. D'où une contradiction qui prouve que X est vide et donc que f est constante sur $[a, b]$. On en déduit la connexité du segment $[a, b]$.

Pour prouver la connexité de $[a, b[$, on prend une suite croissante b_n tendant vers b , et on écrit $[a, b[$ comme réunion croissante des segments $[a, b_n]$ qui sont connexes d'après ce qui précède. Comme une réunion de connexe dont l'intersection est non vide est connexe, cela prouve que $[a, b[$ est connexe. Les autres cas se traitant de la même manière, cela permet de conclure.

- L'image d'un ensemble connexe par une application continue est un ensemble connexe.

Si X est connexe, si $f : X \rightarrow Y$ est continue, et si $g : f(X) \rightarrow \{0, 1\}$ est continue, alors $g \circ f : X \rightarrow \{0, 1\}$ est continue, et donc constante puisque X est connexe. Comme $f : X \rightarrow f(X)$ est surjective, cela implique que g est constante. On en déduit la connexité de $f(X)$.

- Soit $f : [a, b] \rightarrow \mathbf{R}$ continue. Si $f(a)$ et $f(b)$ sont de signes opposés, alors il existe $x \in [a, b]$ tel que $f(x) = 0$ (théorème des valeurs intermédiaires).

Comme $[a, b]$ est connexe, son image par f l'est aussi et donc est un segment de \mathbf{R} , et comme cette image contient des réels négatifs et positifs par hypothèse, elle contient 0.

- Si X et Y sont connexes, alors $X \times Y$ est connexe.

Soit $f : X \times Y \rightarrow \{0, 1\}$ continue. Si $x \in X$, la restriction de f à $\{x\} \times Y$ est continue et donc constante, et si $y \in Y$, la restriction de f à $X \times \{y\}$ est continue et donc constante. Ceci implique que si $(x_1, y_1), (x_2, y_2) \in X \times Y$, alors $f(x_2, y_2) = f(x_2, y_1) = f(x_1, y_1)$, et donc que f est constante. On en déduit la connexité de $X \times Y$.

- Si X est un espace topologique, et si $Y \subset X$ est connexe, alors l'adhérence de Y dans X est connexe.

Soit $f : \bar{Y} \rightarrow \{0, 1\}$ continue. Comme Y est connexe, la restriction de f à Y est constante. Soit $a \in \{0, 1\}$ l'image de Y . Alors $f^{-1}(a)$ est un fermé de \bar{Y} contenant Y , et donc est égal à \bar{Y} par définition de l'adhérence. Autrement dit, f est constante. On en déduit la connexité de \bar{Y} .

- Les composantes connexes d'un espace topologique sont fermées.

8.2. Connexité par arcs

Un espace topologique X est dit *connexe par arcs* si, quels que soient $x, y \in X$, il existe $u : [0, 1] \rightarrow X$ continue, avec $u(0) = x$ et $u(1) = y$ (i.e. si on peut joindre n'importe quelle paire d'éléments de X par un chemin continu). Si X_1 et X_2 sont connexes par arcs, et si $X_1 \cap X_2$ est non vide, alors $X_1 \cup X_2$ est connexe par arcs puisqu'on peut joindre n'importe quel point de $X_1 \cup X_2$ à un point de l'intersection par un chemin continu, et donc n'importe quel couple de points de $X_1 \cup X_2$. Ceci permet, comme ci-dessus, de parler des *composantes connexes par arcs* de X .

- Un espace connexe par arcs est connexe⁽³⁴⁾, mais il existe des ensembles connexes qui ne sont pas connexes par arcs.

Soit X connexe par arc, et soit $x_0 \in X$. Par hypothèse, il existe, pour tout $x \in X$, une application continue $u : [0, 1] \rightarrow X$ avec $u(0) = x_0$ et $u(1) = x$. Comme $[0, 1]$ est connexe et comme l'image d'un connexe par une application continue est connexe, cela montre que x est dans la composante connexe de x_0 . Par suite la composante connexe de x_0 est X tout entier qui, de ce fait, est connexe.

Pour des exemples de connexes non connexes par arcs, voir la rubrique tératologie.

- Un ouvert connexe de \mathbf{R}^n est connexe par arcs.

Soit U un ouvert connexe de \mathbf{R}^n , et soient $x_0 \in U$ et X la composante connexe par arcs de x_0 . Soit $x \in X$. Comme U est ouvert, il existe $r > 0$ tel que $B(x, r)$ soit incluse dans U . Si $y \in B(x, r)$, le segment $[x, y]$ est inclus dans U , et comme il existe un chemin continu joignant x_0 à x dans U , il suffit de composer ce chemin avec le segment $[x, y]$ pour obtenir un chemin joignant x_0 à y dans U . On en déduit l'appartenance de y à X , et donc l'inclusion de $B(x, r)$ dans X , ce qui prouve que X est ouvert. Maintenant, soit x dans l'adhérence de X dans U , et soit $r > 0$ tel que $B(x, r)$ soit incluse dans U . Par définition de l'adhérence, il existe $y \in X \cap B(x, r)$, et comme le segment $[y, x]$ est contenu dans U , on déduit comme ci-dessus que $x \in X$, ce qui prouve que X est fermé. On a donc prouvé que X est à la fois ouvert et fermé dans U , et comme il est non vide et que U est supposé connexe, cela implique que $X = U$. Ceci permet de conclure.

- Un ouvert de \mathbf{R}^n est une réunion dénombrable d'ouverts connexes. Un ouvert de \mathbf{R} est une réunion dénombrable de segments ouverts.

⁽³⁴⁾C'est le principal intérêt de la connexité par arcs ; la connexité est d'utilisation nettement plus facile.

Soit U un ouvert de \mathbf{R}^n . Si $x \in U$, il existe $r > 0$ tel que $B(x, r) \subset U$, et comme $B(x, r)$ est connexe par arcs (et même par segments), la composante connexe de x contient $B(x, r)$. On en déduit que les composantes connexes de U sont des ouverts. Maintenant, un ouvert de \mathbf{R}^n contient un point dont toutes les coordonnées sont rationnelles, et comme les composantes connexes de U sont disjointes, on obtient une injection de l'ensemble de ces composantes connexes dans \mathbf{Q}^n , en choisissant un point à coordonnées rationnelles dans chacune d'entre elles. Comme \mathbf{Q}^n est dénombrable, cela montre que l'ensemble des composantes connexes de U est dénombrable. On en déduit le premier énoncé. Le second en est une conséquence immédiate puisqu'un ouvert connexe de \mathbf{R} est un segment ouvert.

Exercice 8.1. — Montrer que si $n \geq 2$, et si U est un ouvert connexe de \mathbf{R}^n , alors $U - \{x\}$ est connexe, quel que soit $x \in U$.

Exercice 8.2. — (i) Montrer que \mathbf{R} et \mathbf{R}^2 ne sont pas homéomorphes ; que $[0, 1]$ et $[0, 1]^2$ ne sont pas homéomorphes.

(ii) Montrer que $[0, 1]$ et le cercle $C = \{z \in \mathbf{C}, |z| = 1\}$ ne sont pas homéomorphes.

Exercice 8.3. — Montrer que $[0, 1]$ et $]0, 1[$ ne sont pas homéomorphes.

Exercice 8.4. — Soit X le sous-ensemble de \mathbf{R} constitué de trois cercles de rayon 1 dont les centres forment les trois sommets d'un triangle équilatéral dont la longueur des côtés est 2 (chacun des cercles est donc tangent aux deux autres). Soit Y formé de trois cercles de rayon 1 centrés en $(0, 0)$, $(2, 0)$ et $(4, 0)$. Montrer que X et Y ne sont pas homéomorphes.

Exercice 8.5. — (difficile, sa solution utilise des notions introduites plus tard dans le cours) Montrer que le cylindre et la bande de Moebius ne sont pas homéomorphes.

9. Complétude

9.1. Suites de Cauchy

Soit (X, d) un espace métrique. Une suite $(x_n)_{n \in \mathbf{N}}$ est *de Cauchy* (ou vérifie le *critère de Cauchy*) si le diamètre de $\{x_k, k \geq n\}$ tend vers 0 quand $n \rightarrow +\infty$, ce qui se traduit, au choix, par :

- quel que soit $\varepsilon > 0$, il existe $N \in \mathbf{N}$, tel que $d(x_{n+p}, x_n) < \varepsilon$ si $n \geq N$ et $p \in \mathbf{N}$;
- $\lim_{n \rightarrow +\infty} \left(\sup_{p \in \mathbf{N}} d(x_{n+p}, x_n) \right) = 0$.

On remarquera qu'une suite de Cauchy est en particulier bornée.

Exercice 9.1. — (i) Montrer que si d est ultramétrique, alors $(x_n)_{n \in \mathbf{N}}$ est de Cauchy si et seulement si $d(x_{n+1}, x_n) \rightarrow 0$.

(ii) Construire une suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de \mathbf{R} , vérifiant $d(x_{n+1}, x_n) \rightarrow 0$, et qui n'est pas de Cauchy.

- Une suite de Cauchy ayant au moins une valeur d'adhérence a une limite.

Soit $(x_n)_{n \in \mathbf{N}}$ une suite de Cauchy. Supposons que a en soit une valeur d'adhérence. Comme X est un espace métrique, il existe une suite $(x_{\varphi(n)})_{n \in \mathbf{N}}$ extraite de $(x_n)_{n \in \mathbf{N}}$ ayant a pour limite. Soit alors $\varepsilon > 0$. Comme $(x_n)_{n \in \mathbf{N}}$ est de Cauchy, il existe $N_0 \in \mathbf{N}$ tel que $d(x_{m+p}, x_m) < \varepsilon$,

si $m \geq N_0$ et $p \in \mathbf{N}$. Comme $\varphi(n)$ tend vers $+\infty$, il existe $N_1 \in \mathbf{N}$ tel que $\varphi(n) \geq N_0$, si $n \geq N_1$, et comme $x_{\varphi(n)} \rightarrow a$, il existe $N_2 \geq N_1$ tel que $d(x_{\varphi(n)}, a) < \varepsilon$, si $n \geq N_2$. Alors $d(x_{\varphi(n)+p}, a) < 2\varepsilon$, si $n \geq N_2$ et $p \in \mathbf{N}$, et donc $d(x_m, a) < 2\varepsilon$, si $m \geq \varphi(N_2)$. On en déduit que $x_n \rightarrow a$, ce qui permet de conclure.

L'espace (X, d) est *complet* si toute suite de Cauchy admet une valeur d'adhérence ou, ce qui revient au même, une limite. Le critère qui suit permet de ne considérer que des suites convergeant "normalement".

- (X, d) est complet si et seulement si toute suite $(x_n)_{n \in \mathbf{N}}$ telle que $\sum_{n=0}^{+\infty} d(x_{n+1}, x_n) < +\infty$ a une limite.

Si $\sum_{n=0}^{+\infty} d(x_{n+1}, x_n) < +\infty$, alors $\sup_{p \in \mathbf{N}} d(x_n, x_{n+p}) \leq \sum_{k=0}^{+\infty} d(x_{n+k+1}, x_{n+k})$ tend vers 0 quand $n \rightarrow +\infty$ puisque majoré par le reste d'une série convergente. On en déduit que la suite $(x_n)_{n \in \mathbf{N}}$ est de Cauchy, et donc converge si (X, d) est complet.

Réciproquement, si toute suite $(x_n)_{n \in \mathbf{N}}$ telle que $\sum_{n=0}^{+\infty} d(x_{n+1}, x_n) < +\infty$ a une limite, et si $(y_n)_{n \in \mathbf{N}}$ est une suite de Cauchy, on peut en extraire une sous-suite $(y_{\varphi(n)})_{n \in \mathbf{N}}$ telle que $\sup_{p \in \mathbf{N}} d(y_{\varphi(n)+p}, y_{\varphi(n)}) \leq 2^{-n}$ quel que soit $n \in \mathbf{N}$. Il suffit de définir $\varphi(n)$ comme le N correspondant à $\varepsilon = 2^{-n}$ dans la définition d'une suite de Cauchy. La suite $x_n = y_{\varphi(n)}$ vérifie $\sum_{n=0}^{+\infty} d(x_{n+1}, x_n) < +\infty$; elle converge donc, et comme elle est extraite de $(y_n)_{n \in \mathbf{N}}$, cela prouve que $(y_n)_{n \in \mathbf{N}}$ a une valeur d'adhérence et donc une limite puisqu'elle est de Cauchy. On en déduit la complétude de X .

- Si (X, d) est complet, et si Y est fermé dans X , alors (Y, d) est complet.

Si $(x_n)_{n \in \mathbf{N}}$ est une suite de Cauchy dans Y , alors c'est une suite de Cauchy dans X ; elle a donc une limite dans X qui appartient à Y puisque Y est fermé. D'où la complétude de Y .

- Un espace métrique compact est complet.

Si $(x_n)_{n \in \mathbf{N}}$ est de Cauchy dans un espace métrique compact X , alors $(x_n)_{n \in \mathbf{N}}$ admet une valeur d'adhérence puisque X est compact, et donc converge d'après le point ci-dessus, ce qui prouve que X est complet.

D'après le point précédent, un espace compact est complet quelle que soit la distance utilisée pour définir la topologie. Ce n'est pas le cas en général : *la complétude est une propriété métrique et pas topologique*.

Exercice 9.2. — (i) Montrer que $d'(x, y) = |f(y) - f(x)|$, avec $f(x) = \frac{x}{1-|x|}$ est une distance sur $] -1, 1[$ équivalente à la distance usuelle.

(ii) Montrer que $] -1, 1[$ est complet pour d' mais pas pour la distance usuelle.

- \mathbf{R} est complet.

Soit $(x_n)_{n \in \mathbf{N}}$ une suite de Cauchy d'éléments de \mathbf{R} . En particulier, la suite est bornée et il existe $M > 0$ telle que $(x_n)_{n \in \mathbf{N}}$ soit à valeurs dans $[-M, M]$. Comme $[-M, M]$ est compact, cela implique que $(x_n)_{n \in \mathbf{N}}$ a une valeur d'adhérence, et donc qu'elle a une limite puisqu'elle est de Cauchy. Ceci permet de conclure.

- Si X et Y sont complets, alors $X \times Y$ est complet.

Si $(x_n, y_n)_{n \in \mathbf{N}}$ est une suite de Cauchy dans $X \times Y$, alors $(x_n)_{n \in \mathbf{N}}$ est de Cauchy dans X et $(y_n)_{n \in \mathbf{N}}$ est de Cauchy dans Y , et si a et b désignent les limites respectives de $(x_n)_{n \in \mathbf{N}}$ et $(y_n)_{n \in \mathbf{N}}$, alors $(x_n, y_n)_{n \in \mathbf{N}}$ tend vers (a, b) . On en déduit la complétude de $X \times Y$.

9.2. Principales propriétés des espaces complets

L'intérêt principal de travailler dans un espace complet est que les problèmes d'existence sont nettement plus faciles. Le théorème du point fixe ci-dessous a de multiples applications à l'existence d'objets (solutions d'équations différentielles, racines de polynômes à coefficients réels, complexes, ou p -adiques, inversion locale de fonctions de classe $\mathcal{C}^1 \dots$). Le lemme de Baire est un autre de ces outils magiques fournissant l'existence d'une infinité de solutions à des problèmes pour lesquels on a du mal à en exhiber une⁽³⁵⁾; son utilisation nécessite nettement plus d'astuce que celle du théorème du point fixe.

- Dans un espace complet, une application strictement contractante admet un unique point fixe, et la suite des itérés de tout point tend vers ce point fixe (théorème du point fixe).

Soit (X, d) un espace métrique complet, soit $f : X \rightarrow X$ une application strictement contractante (i.e. il existe $\alpha < 1$ tel que $d(f(x), f(y)) \leq \alpha d(x, y)$ quels que soient $x, y \in X$), et soit $x \in X$. Définissons par récurrence une suite $(x_n)_{n \in \mathbf{N}}$ en posant $x_0 = x$ et $x_{n+1} = f(x_n)$, si $n \in \mathbf{N}$ (en notant f^n l'application $f \circ \dots \circ f$ composée n fois, on a aussi $x_n = f^n(x)$). Soit $a = d(x_0, x_1)$. Une récurrence immédiate montre que $d(x_n, x_{n+1}) \leq \alpha^n a$ quel que soit $n \in \mathbf{N}$. On a donc, si $p \in \mathbf{N}$, et $n \in \mathbf{N}$

$$d(x_{n+p}, x_n) \leq d(x_n, x_{n+1}) + \dots + d(x_{n+p-1}, x_{n+p}) \leq a(\alpha^n + \dots + \alpha^{n+p-1}) \leq \alpha^n \frac{a}{1 - \alpha}.$$

La suite $(x_n)_{n \in \mathbf{N}}$ est donc de Cauchy puisque α^n tend vers 0 quand n tend vers $+\infty$. Notons ℓ sa limite. Une application contractante étant en particulier continue, on a

$$f(\ell) = f\left(\lim_{n \rightarrow +\infty} x_n\right) = \lim_{n \rightarrow +\infty} f(x_n) = \lim_{n \rightarrow +\infty} x_{n+1} = \ell,$$

ce qui prouve que ℓ est un point fixe de f . On a donc prouvé que, si x est un point quelconque de X , alors la suite des itérés de x par f tend vers un point fixe de f . Maintenant, si x et y sont deux points fixes de f , on a $d(x, y) = d(f(x), f(y)) \leq \alpha d(x, y)$, et donc $d(x, y) = 0$, et $x = y$, ce qui prouve que f a un unique point fixe. Ceci permet de conclure.

- Dans un espace complet, l'intersection d'une suite de fermés emboîtés, non vides, dont le diamètre tend vers 0, est non vide et réduite à un point (théorème des fermés emboîtés).

Soit (X, d) un espace métrique complet, et soit $(F_n)_{n \in \mathbf{N}}$ une suite de fermés emboîtés (i.e. $F_{n+1} \subset F_n$ quel que soit $n \in \mathbf{N}$), non vides, dont le diamètre tend vers 0 (le diamètre d'un sous-ensemble Y de X est la borne supérieure de l'ensemble des $d(x, y)$, pour $x, y \in Y$).

Choisissons pour tout $n \in \mathbf{N}$ un élément x_n de F_n , et notons d_n le diamètre de F_n . Par hypothèse d_n tend vers 0 quand n tend vers $+\infty$. Par ailleurs, x_{n+p} et x_n sont deux éléments de F_n et donc $d(x_{n+p}, x_n) \leq d_n$ quels que soient $n, p \in \mathbf{N}$. La suite $(x_n)_{n \in \mathbf{N}}$ est donc de Cauchy. Comme X est supposé complet, cette suite admet une limite x . De plus, si on fixe m ,

⁽³⁵⁾On tombe alors sur le problème quasi-théologique de savoir si on peut vraiment prétendre avoir démontré qu'un ensemble est non vide si on est incapable d'en produire un élément.

alors $x_n \in F_n \subset F_m$, si $n \geq m$, et comme F_m est fermé, cela implique que $x \in F_m$. Ceci étant vrai pour tout $m \in \mathbf{N}$, on a $x \in F = \bigcap_{n \in \mathbf{N}} F_n$, ce qui prouve que F est non vide. Finalement, si x, y sont deux éléments de F , on a $x, y \in F_n$ pour tout $n \in \mathbf{N}$, et donc $d(x, y) \leq d_n$ quel que soit $n \in \mathbf{N}$. On en déduit la nullité de $d(x, y)$, ce qui implique $x = y$, et permet de conclure.

- Dans un espace complet, une intersection dénombrable d'ouverts denses est dense et donc, en particulier, est non vide (lemme de Baire).

Soit (X, d) un espace métrique complet, et soit $(U_n)_{n \in \mathbf{N}}$ une suite d'ouverts denses de X . Notre but est de prouver que, si $x_0 \in X$, et si $r_0 > 0$, alors $B(x_0, r_0^-) \cap (\bigcap_{n \in \mathbf{N}} U_n)$ est non vide. Pour cela, nous allons construire une suite $B(x_n, r_n)$ de boules fermées vérifiant :

$$0 < r_{n+1} \leq \frac{r_n}{2} \quad \text{et} \quad B(x_{n+1}, r_{n+1}) \subset U_{n+1} \cap B(x_n, r_n^-).$$

Supposons $B(x_n, r_n)$ construite. Comme U_{n+1} est dense dans X , $U_{n+1} \cap B(x_n, r_n^-)$ est non vide. Prenons $x_{n+1} \in U_{n+1} \cap B(x_n, r_n^-)$ quelconque. Comme $U_{n+1} \cap B(x_n, r_n^-)$ est un ouvert, il existe $r_{n+1} \in]0, \frac{r_n}{2}]$ tel que $B(x_{n+1}, 2r_{n+1}^-) \subset U_{n+1} \cap B(x_n, r_n^-)$, et donc $B(x_{n+1}, r_{n+1}) \subset U_{n+1} \cap B(x_n, r_n^-)$, ce qui permet de faire la construction à l'ordre $n + 1$.

Maintenant, par construction, les $B(x_n, r_n)$ forment une suite de fermés emboîtés (car on a imposé $B(x_{n+1}, r_{n+1}) \subset B(x_n, r_n^-)$) dont le diamètre tend vers 0 (car $r_{n+1} \leq \frac{r_n}{2}$), et $B(x_n, r_n) \subset B(x_0, r_0^-) \cap (\bigcap_{k \leq n} U_k)$, si $n \geq 1$, ce qui implique que $\bigcap_{n \in \mathbf{N}} B(x_n, r_n)$, qui est non vide d'après le théorème des fermés emboîtés, est inclus dans

$$\bigcap_{n \in \mathbf{N}} (B(x_0, r_0^-) \cap (\bigcap_{k \leq n} U_k)) = B(x_0, r_0^-) \cap (\bigcap_{n \in \mathbf{N}} U_n).$$

Ceci permet de conclure.

Le lemme de Baire s'utilise souvent en passant aux complémentaires.

- Dans un espace complet, une réunion dénombrable de fermés d'intérieur vide est d'intérieur vide ; autrement dit, si une réunion dénombrable de fermés est d'intérieur non vide, alors au moins un des fermés est d'intérieur non vide.

Exercice 9.3. — Montrer qu'une intersection dénombrable d'ouverts denses de \mathbf{R} est non dénombrable.

9.3. Complétion d'un espace métrique

Un espace métrique n'est pas forcément complet, mais il peut se compléter de manière unique. Plus précisément :

- Si (X, d) est un espace métrique, il existe, à isométrie près, un unique espace métrique complet (\widehat{X}, d) , contenant X comme sous-espace dense, qui vérifie la *propriété universelle* suivante : toute application uniformément continue f de X dans un espace métrique Y complet se prolonge de manière unique en une application continue de \widehat{X} dans Y .

Cet espace est le *complété* de X , et un espace complet est son propre complété ; plus généralement, si X est dense dans Y , et si Y est complet, alors Y est le complété de X .

L'unicité suit du résultat plus général (et très utile) suivant appliqué au cas où Y et Z sont deux complétés de X , et f est l'identité sur X , l'application $f : Y \rightarrow Z$ qu'on en tire est alors une isométrie puisque c'en est une sur X (cf. ex. 6.8).

• Soient (Y, d_Y) et (Z, d_Z) deux espaces complets. Si X est dense dans Y , et si $f : X \rightarrow Z$ est telle qu'il existe $\rho > 0$, tel que f soit uniformément continue sur $B_X(x, \rho)$, pour tout $x \in X$, alors f s'étend de manière unique en une application continue de Y dans Z .

Soit $y \in Y$, et soit $(x_n)_{n \in \mathbf{N}}$ une suite d'éléments de X tendant vers y quand n tend vers $+\infty$. La suite $(x_n)_{n \in \mathbf{N}}$ est alors de Cauchy, et il existe $n \in \mathbf{N}$ tel que $x_n \in B_X(x_{n_0}, \rho)$, quel que soit $n \geq n_0$. Comme on a supposé que f est uniformément continue sur $B_X(x_{n_0}, \rho)$, la suite $(f(x_n))_{n \in \mathbf{N}}$ est de Cauchy dans Z , et comme Z est complet, cette suite a une limite, et cette limite ne dépend pas de la suite $(x_n)_{n \in \mathbf{N}}$ de limite y (sinon on pourrait construire une telle suite de telle sorte que $(f(x_n))_{n \in \mathbf{N}}$ ait deux valeurs d'adhérence). Notons cette limite $f(y)$.

Maintenant, soit $\varepsilon > 0$ et soit $x_0 \in X$. Comme f est uniformément continue sur $B_X(x_0, \rho)$, il existe $\delta > 0$ tel que $d_Z(f(x), f(x')) \leq \varepsilon$, si $d_Y(x, x') < \delta$ et $x, x' \in B_X(x_0, \rho)$. Si $y_1, y_2 \in B_Y(x_0, \rho)$ vérifient $d_Y(y_1, y_2) < \delta$, et si $(x_{1,n})_{n \in \mathbf{N}}$ et $(x_{2,n})_{n \in \mathbf{N}}$ sont des suites d'éléments de X tendant vers y_1 et y_2 respectivement, alors $x_{1,n}, x_{2,n} \in B_X(x_0, \rho)$ et $d_Y(x_{1,n}, x_{2,n}) < \delta$ si n est assez grand. On a donc $d_Z(f(x_{1,n}), f(x_{2,n})) \leq \varepsilon$ pour tout n assez grand, et un passage à la limite montre que $d_Z(f(y_1), f(y_2)) \leq \varepsilon$, ce qui prouve que f est uniformément continue sur $B_Y(x_0, \rho)$. Comme les $B_Y(x_0, \rho)$, pour $x_0 \in X$, recouvrent Y , puisque X est dense dans Y , cela permet de conclure.

L'existence se démontre en rajoutant⁽³⁶⁾ de force les limites des suites de Cauchy.

Pour ce faire, notons $\text{Cauchy}(X)$ l'ensemble des suites de Cauchy à valeurs dans X . Si $\hat{x} = (x_n)_{n \in \mathbf{N}}$ et $\hat{y} = (y_n)_{n \in \mathbf{N}}$ sont deux éléments de $\text{Cauchy}(X)$, la suite $(d(x_n, y_n))_{n \in \mathbf{N}}$ est de Cauchy dans \mathbf{R} car

$$|d(x_{n+p}, y_{n+p}) - d(x_n, y_n)| = |d(x_{n+p}, y_{n+p}) - d(x_n, y_{n+p}) + d(x_n, y_{n+p}) - d(x_n, y_n)| \leq d(x_{n+p}, x_n) + d(y_{n+p}, y_n)$$

d'après l'inégalité triangulaire. Comme \mathbf{R} est complet, cette suite admet une limite que l'on note $\hat{d}(\hat{x}, \hat{y})$. De plus, si $\hat{x} = (x_n)_{n \in \mathbf{N}}, \hat{y} = (y_n)_{n \in \mathbf{N}}, \hat{z} = (z_n)_{n \in \mathbf{N}}$ sont trois éléments de $\text{Cauchy}(X)$, un passage à la limite dans l'inégalité triangulaire $d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n)$ montre que \hat{d} vérifie l'inégalité triangulaire $\hat{d}(\hat{x}, \hat{z}) \leq \hat{d}(\hat{x}, \hat{y}) + \hat{d}(\hat{y}, \hat{z})$. De même, \hat{d} vérifie la symétrie $\hat{d}(\hat{x}, \hat{y}) = \hat{d}(\hat{y}, \hat{x})$, mais elle ne vérifie pas la séparation de la distance (i.e. il n'est pas vrai que $\hat{d}(\hat{x}, \hat{y}) = 0$ implique $\hat{x} = \hat{y}$). De fait, il est assez clair que $\hat{d}(\hat{x}, \hat{y}) = 0$ équivaut au fait que \hat{x} et \hat{y} ont moralement la même limite. Cela nous conduit à introduire la relation \sim sur $\text{Cauchy}(X)$ définie par, $\hat{x} \sim \hat{y}$ si et seulement si $\hat{d}(\hat{x}, \hat{y}) = 0$, ce qui fait de \sim une relation d'équivalence, et nous permet de considérer le quotient \hat{X} de $\text{Cauchy}(X)$ par cette relation d'équivalence (ce qui revient à considérer comme égaux deux éléments \hat{x}, \hat{y} de $\text{Cauchy}(X)$ vérifiant $\hat{d}(\hat{x}, \hat{y}) = 0$).

Il n'y a plus qu'à vérifier que l'objet que l'on a construit est bien celui que l'on voulait.

L'inégalité triangulaire montre que $\hat{d}(\hat{x}, \hat{y}) = \hat{d}(\hat{x}', \hat{y}')$ si $\hat{x} \sim \hat{x}'$ et $\hat{y} \sim \hat{y}'$, ce qui montre que \hat{d} passe au quotient, et définit une distance sur \hat{X} puisque, par définition de \hat{X} , la condition $\hat{d}(\hat{x}, \hat{y}) = 0$ implique $\hat{x} = \hat{y}$.

Maintenant, on peut identifier $x \in X$, à la classe dans \hat{X} de la suite constante $\iota(x) = (x_n)_{n \in \mathbf{N}}$, avec $x_n = x$ quel que soit $n \in \mathbf{N}$. Si $x, y \in X$, on a $\hat{d}(x, y) = \hat{d}(\iota(x), \iota(y)) = \lim_{n \rightarrow +\infty} d(x, y) =$

⁽³⁶⁾Beaucoup d'objets mathématiques sont obtenus de cette manière, à commencer par \mathbf{R} qui est le complété de \mathbf{Q} pour la distance usuelle $d(x, y) = |x - y|$, où $|x - y|$ est la valeur absolue de $x - y$, et \mathbf{Q}_p qui est le complété de \mathbf{Q} pour la norme p -adique.

$d(x, y)$, ce qui montre que \hat{d} induit la distance d sur X . Par ailleurs, si $\hat{x} = (x_n)_{n \in \mathbf{N}}$ est un élément de $\text{Cauchy}(X)$, alors $\hat{d}(\hat{x}, \iota(x_k)) = \lim_{n \rightarrow +\infty} d(x_n, x_k) \leq \sup_{n \geq k} d(x_n, x_k)$, et comme la suite $(x_n)_{n \in \mathbf{N}}$ est de Cauchy, $\sup_{n \geq k} d(x_n, x_k)$ tend vers 0 quand k tend vers $+\infty$. On a donc $\hat{x} = \lim_{k \rightarrow +\infty} x_k$ dans \hat{X} , ce qui prouve que X est dense dans \hat{X} .

Il reste à prouver que \hat{X} est complet. Pour cela soit $(\hat{x}_n)_{n \in \mathbf{N}}$ une suite de Cauchy dans \hat{X} . Comme X est dense dans \hat{X} , on peut trouver, quel que soit $n \in \mathbf{N}$, un élément x_n de X tel que $\hat{d}(\hat{x}_n, x_n) \leq 2^{-n}$. Soit $\hat{x} = (x_n)_{n \in \mathbf{N}}$. On a

$$d(x_n, x_{n+p}) = \hat{d}(x_n, x_{n+p}) \leq \hat{d}(x_n, \hat{x}_n) + \hat{d}(\hat{x}_n, \hat{x}_{n+p}) + \hat{d}(\hat{x}_{n+p}, x_{n+p}) \leq 2^{1-n} + \hat{d}(\hat{x}_n, \hat{x}_{n+p}),$$

et comme la suite $(\hat{x}_n)_{n \in \mathbf{N}}$ est de Cauchy, on en déduit que $\hat{x} \in \text{Cauchy}(X)$. De plus,

$$\hat{d}(\hat{x}_n, \hat{x}) \leq \hat{d}(\hat{x}_n, x_n) + \hat{d}(x_n, \hat{x}) \leq 2^{-n} + \lim_{m \rightarrow +\infty} d(x_n, x_m) \leq 2^{-n} + \sup_{p \in \mathbf{N}} d(x_n, x_{n+p}),$$

et comme $(x_n)_{n \in \mathbf{N}}$ est de Cauchy, $\sup_{p \in \mathbf{N}} d(x_n, x_{n+p}) \rightarrow 0$. Autrement dit, $\hat{x}_n \rightarrow \hat{x}$ dans \hat{X} . On en déduit la complétude de \hat{X} .

10. Convergence de fonctions

10.1. Convergence simple

Si X et Y sont deux espaces topologiques, une suite de fonctions $f_n : X \rightarrow Y$ converge simplement vers f si pour tout $x \in X$, la suite $f_n(x)$ a pour limite $f(x)$ dans Y . Si c'est le cas, on dit que f est la *limite simple* de la suite f_n .

Il est, en pratique, largement inutile de savoir quelle topologie se cache derrière la convergence simple. Cette topologie n'a rien de mystérieux : c'est la topologie produit sur l'espace des fonctions Y^X de X dans Y . En effet, les conditions suivantes sont équivalentes :

- $f_n(x) \rightarrow f(x)$, pour tout $x \in X$;
- $(f_n(x))_{x \in I} \rightarrow (f(x))_{x \in I}$ pour tout $I \subset X$ fini ;
- pour tout $I \subset X$ fini, et tout ouvert de Y^I de la forme $U = \prod_{x \in I} U_x$ qui contient $(f(x))_{x \in I}$, il existe $N \in \mathbf{N}$, tel que $(f_n(x))_{x \in I} \in U$, si $n \geq N$;
- pour tout $I \subset X$ fini, et tout ouvert de Y^X de la forme $U = (\prod_{x \in I} U_x) \times (\prod_{x \notin I} Y)$ qui contient $(f(x))_{x \in X}$, il existe $N \in \mathbf{N}$, tel que $(f_n(x))_{x \in X} \in U$, si $n \geq N$;
- $f_n \rightarrow f$ dans Y^X .

D'après l'exercice ci-dessous, les fonctions continues sont denses dans l'ensemble des fonctions de \mathbf{R} dans \mathbf{C} pour la topologie produit sur l'espace $\mathbf{C}^{\mathbf{R}}$ des fonctions de \mathbf{R} dans \mathbf{C} . Or, Baire a montré qu'une limite simple de fonctions continues de \mathbf{R} dans \mathbf{C} est continue en au moins un point. Donc il existe des éléments de $\mathbf{C}^{\mathbf{R}}$ qui ne sont pas limite simple d'une suite de fonctions continues, ce qui n'est possible que si la topologie ci-dessus sur $\mathbf{C}^{\mathbf{R}}$ n'est pas définissable par une distance. Cela explique qu'il existe des fonctions qui sont limites simples de limites simples de fonctions continues, mais qui ne sont pas limites simples de fonctions continues.

Exercice 10.1. — Montrer que l'ensemble des fonctions continues de \mathbf{R} dans \mathbf{C} est dense dans $\mathbf{C}^{\mathbf{R}}$ (muni de la topologie produit).

10.2. Convergence uniforme

Soient X un ensemble et Y un espace métrique (par exemple $Y = \mathbf{C}$). Soient f et f_n , pour $n \in \mathbf{N}$, des fonctions de X dans Y . On dit que f_n converge uniformément vers f sur X ou que f est la *limite uniforme* des f_n , si $\lim_{n \rightarrow +\infty} \left(\sup_{x \in X} d_Y(f(x), f_n(x)) \right) = 0$. Ceci peut

se réécrire sous la forme : pour tout $\varepsilon > 0$, il existe $N = N(\varepsilon)$ tel que $d_Y(f(x), f_n(x)) < \varepsilon$, pour tout $n \geq N$ et tout $x \in X$.

La différence avec la convergence simple est que $N(\varepsilon)$ est le même pour tout $x \in X$; en particulier, la convergence uniforme⁽³⁷⁾ implique la convergence simple.

• Si X est un espace topologique, si $f_n \rightarrow f$ uniformément sur X , et si f_n est continue en x_0 , pour tout $n \in \mathbf{N}$, alors f est continue en x_0 . Si les f_n sont continues sur X , il en est de même de f .

Soit $\varepsilon > 0$, et soit $n \in \mathbf{N}$ tel que $\sup_{x \in X} d_Y(f(x), f_n(x)) < \varepsilon$. Comme f_n est continue en x_0 , il existe V ouvert de X contenant x_0 tel que $d_Y(f_n(x), f_n(x_0)) < \varepsilon$, pour tout $x \in V$. On a alors

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f_n(x)) + d_Y(f_n(x), f_n(x_0)) + d_Y(f_n(x_0), f(x_0)) < 3\varepsilon,$$

pour tout $x \in V$. On en déduit la continuité de f en x_0 . Le second énoncé en étant une conséquence immédiate, cela permet de conclure.

Exercice 10.2. — Soient $u = (u_k)_{k \in \mathbf{N}}$ et $u^{(n)} = (u_k^{(n)})_{k \in \mathbf{N}}$, pour $n \in \mathbf{N}$, des suites à valeurs dans \mathbf{C} . On suppose que $u^{(n)} \rightarrow u$ uniformément sur \mathbf{N} et que $\lim_{k \rightarrow +\infty} u_k^{(n)} = 0$, pour tout n . Montrer que $\lim_{k \rightarrow +\infty} u_k = 0$.

Si X est un ensemble et si Y est un espace métrique, une suite de fonctions $f_n : X \rightarrow Y$ vérifie le *critère de Cauchy uniforme* sur X si $\lim_{n \rightarrow +\infty} \left(\sup_{x \in X, p \in \mathbf{N}} d_Y(f_n(x), f_{n+p}(x)) \right) = 0$.

• Si X est un espace topologique, si Y est un espace métrique complet, et si $(f_n)_{n \in \mathbf{N}}$ est une suite de fonctions continues de X dans Y vérifiant le critère de Cauchy uniforme, alors $(f_n)_{n \in \mathbf{N}}$ a une limite simple f qui est continue, et f_n converge uniformément vers f sur X .

Si $x \in X$, la suite $(f_n(x))_{n \in \mathbf{N}}$ est de Cauchy, et donc admet une limite $f(x)$, puisque Y est complet. Soit $\delta_n = \sup_{p \in \mathbf{N}} d_Y(f_{n+p}(x), f_n(x))$; par hypothèse, on a $\delta_n \rightarrow 0$. Un passage à la limite montre que $d_Y(f(x), f_n(x)) \leq \delta_n$, pour tout x , et comme $\delta_n \rightarrow 0$, cela prouve que $f_n \rightarrow f$ uniformément sur X , ce qui permet de conclure puisqu'une limite uniforme de fonctions continues est continue.

Exercice 10.3. — Soit $(E, \|\cdot\|)$ un espace vectoriel normé (sur \mathbf{R} ou \mathbf{C}). On dit que $f : E \rightarrow \mathbf{C}$ tend vers ℓ à l'infini, si pour tout $\varepsilon > 0$, il existe $M > 0$ tel que $|f(x) - \ell| < \varepsilon$ pour tout x vérifiant $\|x\| > M$. Soient f et f_n , pour $n \in \mathbf{N}$, des fonctions de E dans \mathbf{C} . On suppose que $f_n \rightarrow f$ uniformément sur E , et que f_n tend vers ℓ_n à l'infini. Montrer que $(\ell_n)_{n \in \mathbf{N}}$ a une limite $\ell \in \mathbf{C}$, et que f tend vers ℓ à l'infini.

⁽³⁷⁾En filière PC, la convergence uniforme (concept universellement reconnu) a été remplacée par les deux demi-concepts que constituent la convergence normale (pour une norme qui n'est autre que celle de la convergence uniforme), et par l'approximation uniforme d'une fonction par des fonctions d'un certain ensemble. J'avoue avoir du mal à saisir la subtile différence.

11. Espaces vectoriels normés

11.1. Normes et applications linéaires continues

Si E est un espace vectoriel sur \mathbf{K} , avec $\mathbf{K} = \mathbf{R}$ ou $\mathbf{K} = \mathbf{C}$, une *norme* $\| \cdot \|$ sur E est une application $x \mapsto \|x\|$ de E dans \mathbf{R}_+ vérifiant les propriétés suivantes :

- (i) $\|x\| = 0$ si et seulement si $x = 0$;
- (ii) $\|\lambda x\| = |\lambda| \cdot \|x\|$, si $x \in E$ et $\lambda \in \mathbf{K}$;
- (iii) $\|x + y\| \leq \|x\| + \|y\|$, si $x, y \in E$.

Si $\| \cdot \|$ est une norme sur E , alors $d : E \times E \rightarrow \mathbf{R}_+$ définie par $d(x, y) = \|x - y\|$ est une distance sur E , ce qui permet de voir un espace vectoriel normé $(E, \| \cdot \|)$ comme un cas particulier d'espace métrique.

• Si $(E, \| \cdot \|_E)$ et $(F, \| \cdot \|_F)$ sont deux espaces vectoriels normés, et si $u : E \rightarrow F$ est une application *linéaire*, les conditions suivantes sont équivalentes :

- (i) u est continue ;
- (ii) u est uniformément continue ;
- (iii) il existe $M \in \mathbf{R}_+$ tel que $\|u(x)\|_F \leq M \cdot \|x\|_E$, quel que soit $x \in E$.

Si u est continue, l'image inverse de la boule unité ouverte de F contient un voisinage de 0 dans E , et donc une boule ouverte $B(0, r^-)$, avec $r > 0$. Autrement dit, $\|x\|_E < r$ implique $\|u(x)\|_F < 1$, et donc, quel que soit $x \in E - \{0\}$,

$$\|u(x)\|_F = \frac{\|x\|_E}{r} \cdot \left\| \frac{r}{\|x\|_E} u(x) \right\|_F \leq \frac{\|x\|_E}{r}.$$

On en déduit l'implication (i) \Rightarrow (iii) (avec $M = \frac{1}{r}$). Maintenant, si $\|u(x)\|_F \leq M \cdot \|x\|_E$, quel que soit $x \in E$, alors u est lipschitzienne de rapport M , et donc uniformément continue. On en déduit l'implication (iii) \Rightarrow (ii), et comme l'implication (ii) \Rightarrow (i) est une évidence, cela permet de conclure.

• Si $(E, \| \cdot \|_E)$ et $(F, \| \cdot \|_F)$ sont deux espaces vectoriels normés avec F complet, et si $u : E \rightarrow F$ est linéaire continue, alors u se prolonge, par continuité, en une application linéaire continue du complété \widehat{E} de E dans F .

C'est une conséquence de la propriété universelle vérifiée par \widehat{E} .

11.2. La norme d'un opérateur

Si $(E, \| \cdot \|_E)$ et $(F, \| \cdot \|_F)$ sont deux espaces vectoriels normés, et si $u : E \rightarrow F$ est une application linéaire continue, la *norme d'opérateur* $\|u\|$ de u est la borne supérieure de l'ensemble des $\|x\|_E^{-1} \|u(x)\|_F$, pour $x \in E - \{0\}$. On a donc $\|u(x)\|_F \leq \|u\| \cdot \|x\|_E$, quel que soit $x \in E$, et $\|u\|$ est le plus petit réel ayant cette propriété.

• La norme d'opérateur est une norme sur l'espace vectoriel $\text{Hom}(E, F)$ des applications linéaires continues de E dans F .

Si $\|u\| = 0$, alors $u(x) = 0$, pour tout x , et donc $u = 0$. Si $u \in \text{End}(E, F)$ et $\lambda \in \mathbf{K}$, alors

$$\|\lambda u\| = \sup_{x \in E - \{0\}} \|x\|_E^{-1} \|\lambda u(x)\|_F = \sup_{x \in E - \{0\}} |\lambda| \cdot \|x\|_E^{-1} \|u(x)\|_F = |\lambda| \sup_{x \in E - \{0\}} \|x\|_E^{-1} \|u(x)\|_F = |\lambda| \cdot \|u\|.$$

Si $u, v \in \text{End}(E, F)$, alors

$$\|u + v\| = \sup_{x \in E - \{0\}} \|x\|_E^{-1} \|u(x) + v(x)\|_F \leq \sup_{x \in E - \{0\}} \|x\|_E^{-1} (\|u(x)\|_F + \|v(x)\|_F) = \|u\| + \|v\|.$$

Ceci permet de conclure.

- La norme d'opérateur est une norme d'algèbre sur l'anneau $\text{End}(E)$ des endomorphismes linéaires continus de E .

Compte-tenu du point précédent, il ne reste plus que l'inégalité $\|u \circ v\| \leq \|u\| \cdot \|v\|$ à vérifier.

Or $\|u \circ v(x)\|_E \leq \|u\| \cdot \|v(x)\|_E \leq \|u\| \cdot \|v\| \cdot \|x\|_E$, pour tout $x \in E$, par définition de $\|u\|$ et $\|v\|$. On en déduit l'inégalité cherchée.

11.3. Normes équivalentes

Deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur E sont *équivalentes*, si l'application identité de $(E, \|\cdot\|_1)$ dans $(E, \|\cdot\|_2)$ est un homéomorphisme (i.e. est continue ainsi que son inverse). D'après l'alinéa précédent, cela équivaut à l'existence de $C > 0$ tel que $C^{-1}\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$, quel que soit $x \in E$.

- Soit E un espace vectoriel de dimension finie sur \mathbf{K} . Alors toutes les normes sur E sont équivalentes et E est complet pour n'importe laquelle d'entre elles.

Soit (e_1, \dots, e_n) une base de E sur \mathbf{K} . Comme \mathbf{K} est complet, il suffit de prouver que toute norme sur E est équivalente à la norme $\|\cdot\|_\infty$ définie par

$$\|x_1 e_1 + \dots + x_n e_n\|_\infty = \sup(|x_1|, \dots, |x_n|),$$

ce qui se fait par récurrence sur la dimension de E . Si cette dimension est 1, il n'y a rien à faire. Sinon, soit $\|\cdot\|$ une norme sur E . On déduit de l'inégalité triangulaire que

$$\|x_1 e_1 + \dots + x_n e_n\| \leq (\|e_1\| + \dots + \|e_n\|) \sup(|x_1|, \dots, |x_n|),$$

d'où l'une des deux inégalités à vérifier. Pour démontrer l'autre, raisonnons par l'absurde. Supposons qu'il existe une suite $x_1^{(k)} e_1 + \dots + x_n^{(k)} e_n$ qui tende vers 0 pour la norme $\|\cdot\|$ mais pas pour la norme $\|\cdot\|_\infty$. Il existe alors $C > 0$, $i \in \{1, \dots, n\}$ et une sous-suite infinie telle que l'on ait $|x_i^{(k)}| \geq C$, et donc la suite de terme général $v_k = \frac{x_1^{(k)}}{x_i^{(k)}} e_1 + \dots + \frac{x_n^{(k)}}{x_i^{(k)}} e_n$ tend encore vers 0 pour $\|\cdot\|$. On en déduit le fait que e_i est dans l'adhérence de $W = \text{Vect}(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n)$, qui est complet d'après l'hypothèse de récurrence, ce qui implique $e_i \in W$ et est absurde puisque les e_i forment une base de E .

- L'énoncé précédent devient *totalement faux en dimension infinie* : les normes sur un espace E de dimension infinie ne sont pas toutes équivalentes⁽³⁸⁾, et E peut être complet pour certaines d'entre elles, mais il y en a "beaucoup plus" pour lesquelles ce n'est pas le cas.

⁽³⁸⁾Un des problèmes de base en analyse fonctionnelle est précisément de choisir la bonne norme en fonction du problème à résoudre.

Exercice 11.1. — Soit $E = \mathcal{C}([0, 1])$ l'espace des fonctions continues de $[0, 1]$ dans \mathbf{C} .

(i) Montrer que, si $\phi \in E$, alors $\|\phi\|_\infty = \sup_{x \in [0, 1]} |\phi(x)|$ est fini et que $\|\cdot\|_\infty$ est une norme sur E pour laquelle E est complet.

(ii) Montrer que $\|\cdot\|_1$ définie par $\|\phi\|_1 = \int_0^1 |\phi(t)| dt$ est une norme sur E pour laquelle E n'est pas complet.

(iii) Les normes $\|\cdot\|_\infty$ et $\|\cdot\|_1$ sont-elles équivalentes ?

Exercice 11.2. — (i) Montrer que, si \mathcal{T}_1 et \mathcal{T}_2 sont des topologies sur X , alors \mathcal{T}_1 est plus fine que \mathcal{T}_2 si et seulement si $\text{id} : (X, \mathcal{T}_1) \rightarrow (X, \mathcal{T}_2)$ est continue.

(ii) Soit \mathcal{T}_1 la topologie sur $\mathcal{C}(\mathbf{R})$ définie par la norme $\|\cdot\|_1$ et \mathcal{T}_∞ celle définie par la norme $\|\cdot\|_\infty$. Montrer qu'aucune des deux topologies \mathcal{T}_1 et \mathcal{T}_∞ n'est plus fine que l'autre.

11.4. La boule unité d'un espace vectoriel normé

- Si E est de dimension finie, la boule unité fermée est compacte.

Par définition, la boule unité fermée est bornée, et comme elle est fermée, et que l'on est en dimension finie, elle est compacte.

- Soit E un espace vectoriel normé. Si la boule unité fermée $B(0, 1)$ est compacte, alors E est de dimension finie (théorème de Riesz, 1918).

Si $B(0, 1)$ est compacte, on peut extraire un recouvrement fini du recouvrement de $B(0, 1)$ par les $B(x, (\frac{1}{2})^-)$, pour $x \in B(0, 1)$. Autrement dit, on peut trouver un sous-ensemble fini $\{e_i, i \in I\}$ d'éléments de E tels que $B(0, 1) \subset \cup_{i \in I} B(e_i, \frac{1}{2})$. Nous allons montrer que le sous-espace E' engendré par les $(e_i)_{i \in I}$ est égal à E , ce qui permettra de conclure. Comme E' est fermé, puisque complet, car de dimension finie (n° II.1.1), il suffit de montrer que E' est dense dans E . Soit donc $x \in E$, et soient $a \in \mathbf{Z}$ et $y \in E'$ tels que $\|x - y\| \leq 2^{-a}$ (un tel couple existe : il suffit de prendre $y = 0$ et a assez petit pour que $\|x\| \leq 2^{-a}$). On a $2^a(x - y) \in B(0, 1)$ et, par définition de la famille $(e_i)_{i \in I}$, il existe $i \in I$ tel que $\|2^a(x - y) - e_i\| \leq \frac{1}{2}$. Mais alors $y' = y + 2^{-a}e_i \in E'$ et $\|x - y'\| \leq 2^{-a-1}$. Ceci permet de construire, par récurrence, une suite $(y_n)_{n \in \mathbf{N}}$ d'éléments de E' vérifiant $\|x - y_n\| \leq 2^{-n-a}$, ce qui prouve que x est dans l'adhérence de E' , et permet de conclure.

11.5. Applications bilinéaires continues

Si $(E_1, \|\cdot\|_1)$ et $(E_2, \|\cdot\|_2)$ sont deux espaces vectoriels normés, l'espace topologique $E_1 \times E_2$ est aussi un espace vectoriel normé, la topologie produit étant celle associée à la norme $\|(x_1, x_2)\| = \sup(\|x_1\|_1, \|x_2\|_2)$ ou à toute autre norme équivalente comme par exemple $\|(x_1, x_2)\| = (\|x_1\|_1^2 + \|x_2\|_2^2)^{1/2}$.

- Soient $(E_1, \|\cdot\|_1)$, $(E_2, \|\cdot\|_2)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés, et soit $b : E_1 \times E_2 \rightarrow F$ une application bilinéaire. Alors

(i) b est continue si et seulement si il existe $C > 0$ tel que $\|b(x_1, x_2)\|_F \leq C \cdot \|x_1\|_1 \cdot \|x_2\|_2$ quels que soient $x_1 \in E_1$ et $x_2 \in E_2$;

(ii) si F est complet et b continue, alors b s'étend par continuité en une application bilinéaire du complété $\widehat{E}_1 \times \widehat{E}_2$ de $E_1 \times E_2$ dans F .

Si b est continue, il existe $r_1 > 0$ et $r_2 > 0$ tels que $b^{-1}(B_F(0, 1^-))$ contienne $B_{E_1}(0, r_1^-) \times B_{E_2}(0, r_2^-)$. Autrement dit, on a $\|b(x_1, x_2)\|_F < 1$ si $\|x_1\|_1 < r_1$ et $\|x_2\|_2 < r_2$. Par linéarité, cela implique que

$$\|b(x_1, x_2)\|_F = \frac{\|x_1\|_1 \cdot \|x_2\|_2}{r_1 r_2} \|b(\frac{r_1}{\|x_1\|_1} x_1, (\frac{r_2}{\|x_2\|_2} x_2))\|_F \leq \frac{\|x_1\|_1 \cdot \|x_2\|_2}{r_1 r_2}.$$

Réciproquement, s'il existe $C > 0$ tel que $\|b(x_1, x_2)\|_F \leq C \cdot \|x_1\|_1 \cdot \|x_2\|_2$ quels que soient $x_1 \in E_1$ et $x_2 \in E_2$, alors

$$\|b(x_1 + h_1, x_2 + h_2) - b(x_1, x_2)\|_F \leq C(\|x_1\|_1 \cdot \|h_2\|_2 + \|h_1\|_1 \cdot \|x_2\|_2 + \|h_1\|_1 \cdot \|h_2\|_2),$$

ce qui prouve que b est lipschitzienne de rapport $C \cdot (\|x_1\|_1 + \|x_2\|_2 + 1)$ sur $B_{E_1}(x_1, 1^-) \times B_{E_2}(x_2, 1^-)$. Ceci prouve que b est continue (et donc termine la démonstration du (i)), et permet de déduire le (ii) du deuxième point du n° 9.3.

12. Tératologie

Ce § rassemble un certain nombre de monstres mathématiques.

12.1. Fonctions continues dérivables nulle part

Jusqu'au début du XIX^e siècle (au moins), il était évident pour tout le monde qu'une fonction continue de \mathbf{R} dans \mathbf{R} était dérivable, et même somme de sa série de Taylor, sauf en des points isolés. C'est malheureusement loin d'être le cas puisque Weierstrass a construit une fonction continue dérivable nulle part, et Banach a montré que l'ensemble de ces fonctions était dense dans celui des fonctions continues.

Soit $E = \mathcal{C}^0([0, 1], \|\cdot\|_\infty)$. Nous nous proposons de construire un sous-ensemble X , dense dans E , constitué de fonctions dérivables nulle part. Pour ce faire, fixons $a \in]\frac{1}{2}, 1[$. Si $n \in \mathbf{N}$, et si $k \in \{0, 1, \dots, 2^n - 1\}$, soit

$$U_{n,k} = \left\{ \phi \in E, \left| \phi\left(\frac{k+1}{2^n}\right) - \phi\left(\frac{k}{2^n}\right) \right| > a^n \right\}.$$

• $U_{n,k}$ est un ouvert de E : en effet $\phi \mapsto \left| \phi\left(\frac{k+1}{2^n}\right) - \phi\left(\frac{k}{2^n}\right) \right|$ est continue sur E comme composée de l'application linéaire continue $\phi \mapsto \Lambda_{n,k}(\phi) = \phi\left(\frac{k+1}{2^n}\right) - \phi\left(\frac{k}{2^n}\right)$ (la continuité de $\Lambda_{n,k}$ suit de la majoration $|\Lambda_{n,k}(\phi)| \leq 2\|\phi\|_\infty$), et de la valeur absolue.

On en déduit que $U_n = \bigcap_{k=0}^{2^n-1} U_{n,k}$ et $V_n = \bigcup_{m \geq n} U_m$ sont des ouverts de E .

• V_n est dense dans E . En effet, soit $\phi \in E$, et soit $\varepsilon > 0$. Comme $[0, 1]$ est compact, ϕ est uniformément continue, et il existe $n_0 \in \mathbf{N}$ tel que $\left| \phi\left(\frac{k+1}{2^n}\right) - \phi\left(\frac{k}{2^n}\right) \right| \leq \varepsilon$, quels que soient $n \geq n_0$ et $k \in \{0, 1, \dots, 2^n - 1\}$. Soit $m \geq \sup(n_0, n)$ tel que $a^m < \varepsilon$, et soit $\psi \in E$ définie par $\psi(x) = \phi(x) + \varepsilon \sin(2^m \pi x)$. Si $k \in \{0, 1, \dots, 2^m - 1\}$, on a $\|\psi - \phi\|_\infty \leq \varepsilon$ et

$$\left| \psi\left(\frac{k+1}{2^m}\right) - \psi\left(\frac{k}{2^m}\right) \right| = \left| \pm 2\varepsilon + \phi\left(\frac{k+1}{2^m}\right) - \phi\left(\frac{k}{2^m}\right) \right| \geq 2\varepsilon - \varepsilon > a^m,$$

ce qui prouve que $\psi \in U_m \subset V_n$. On en déduit que, pour tout $\phi \in E$, on peut trouver un élément de V_n dans tout voisinage de ϕ , et donc que V_n est effectivement dense dans E .

Comme E est complet, il résulte du lemme de Baire que $X = \bigcap_{n \in \mathbf{N}} V_n$ est dense dans E , et pour conclure, il suffit donc de prouver que, si $\phi \in X$, et si $x_0 \in [0, 1]$, alors ϕ n'est pas dérivable en x_0 . Pour cela, remarquons que $\phi \in X$ signifie que ϕ appartient à une infinité de U_n , et donc qu'il existe $b : \mathbf{N} \rightarrow \mathbf{N}$ tendant vers $+\infty$ en $+\infty$, telle que

$$\left| \phi\left(\frac{k+1}{2^{b(n)}}\right) - \phi\left(\frac{k}{2^{b(n)}}\right) \right| > a^{b(n)},$$

pour tout $n \in \mathbf{N}$ et tout $k \in \{0, 1, \dots, 2^{b(n)} - 1\}$. Soient k_n la partie entière de $2^{b(n)}x_0$, et $u_n = \frac{k_n}{2^{b(n)}}$, $v_n = \frac{k_n+1}{2^{b(n)}}$ (si $x_0 = 1$, on pose $u_n = 1 - \frac{1}{2^{b(n)}}$ et $v_n = 1$). Par construction, $u_n \leq x_0 \leq v_n$ et $v_n - u_n = \frac{1}{2^{b(n)}}$; en particulier, $u_n \rightarrow x_0$ et $v_n \rightarrow x_0$. Par ailleurs, pour tout $n \in \mathbf{N}$, on a $\left| \frac{\phi(v_n) - \phi(u_n)}{v_n - u_n} \right| > (2a)^{b(n)}$, et comme $2a > 1$, cela montre que $\left| \frac{\phi(v_n) - \phi(u_n)}{v_n - u_n} \right|$ tend vers $+\infty$, et donc que ϕ n'est pas dérivable en x_0 (si elle l'était, on aurait $\frac{\phi(v_n) - \phi(u_n)}{v_n - u_n} \rightarrow \phi'(x_0)$).

Ceci permet de conclure.

Exercice 12.1. — Adapter la dernière partie de l'argument pour montrer que $\sum_{n \geq 1} \frac{\sin(10^n \pi x)}{2^n}$ est continue sur \mathbf{R} , mais n'est dérivable nulle part.

12.2. L'escalier du diable

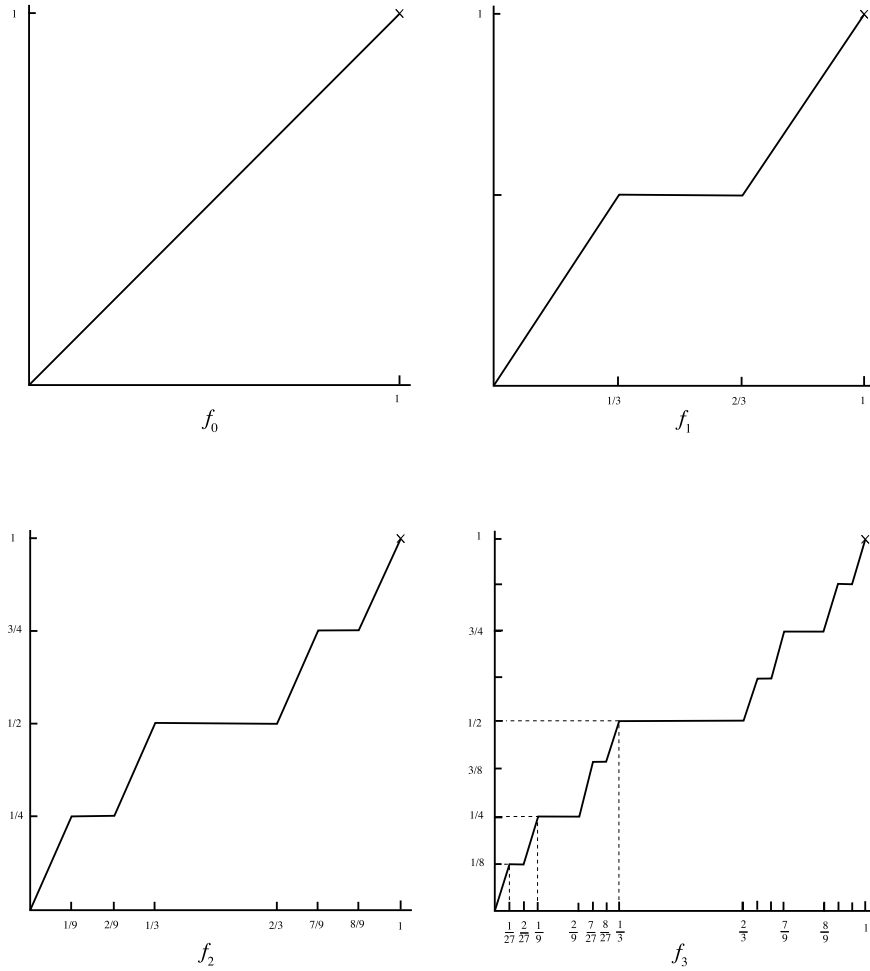
Il s'agit d'une fonction $f : [0, 1] \rightarrow \mathbf{R}$, continue, croissante, valant 0 en 0 et 1 en 1, mais qui croît subrepticement : il existe une famille de segments ouverts $]a_n, b_n[$ disjoints, pour $n \in \mathbf{N}$, tels que f soit constante sur chacun des segments $]a_n, b_n[$, et tels que la somme totale $\sum_{n \in \mathbf{N}} (b_n - a_n)$ des longueurs des segments soit égale à 1. La fonction f représente un contreexemple assez frappant à une extension naturelle du théorème fondamental de l'analyse ($\int_a^b f'(t) dt = f(b) - f(a)$).

On construit f , par un procédé fractal, comme la limite de $f_n : [0, 1] \rightarrow [0, 1]$, continues, croissantes, affines sur chaque intervalle $I_{n,i} = [\frac{i}{3^n}, \frac{i+1}{3^n}]$, pour $0 \leq i \leq 3^n - 1$, construites par récurrence à partir de $f_0(x) = x$ en utilisant la recette suivante : l'image de $I_{n,i}$ par f_{n+1} est la même que par f_n , mais le graphe de f_{n+1} sur cet intervalle est obtenu en coupant en trois le segment constituant le graphe de f_n , et en introduisant un palier horizontal au milieu.

De manière plus précise, si on note $a_{n,i}$ et $b_{n,i}$ les valeurs de f_n en $\frac{i}{3^n}$ et $\frac{i+1}{3^n}$, alors les fonctions f_n et f_{n+1} sont données par les formules suivantes sur $I_{n,i}$:

$$f_n(x) = a_{n,i} + (b_{n,i} - a_{n,i})(3^n x - i)$$

$$f_{n+1}(x) = \begin{cases} a_{n,i} + \frac{3}{2}(b_{n,i} - a_{n,i})(3^n x - i) & \text{si } x \in I_{n+1,3i}, \\ \frac{b_{n,i} + a_{n,i}}{2} & \text{si } x \in I_{n+1,3i+1}, \\ b_{n,i} + \frac{3}{2}(b_{n,i} - a_{n,i})(3^n x - i - 1) & \text{si } x \in I_{n+1,3i+2}. \end{cases}$$

FIG. 1. Graphes de f_0 , f_1 , f_2 et f_3 .

En particulier, si f_n est constante sur $I_{n,i}$, alors $f_{n+1} = f_n$ sur $I_{n,i}$, et dans le cas général, on a

$$b_{n+1,i} - a_{n+1,i} = \begin{cases} \frac{b_{n,i} - a_{n,i}}{2} & \text{si le chiffre des unités dans l'écriture de } i \text{ en base 3 est 0 ou 2,} \\ 0 & \text{si le chiffre des unités dans l'écriture de } i \text{ en base 3 est un 1.} \end{cases}$$

$$|f_{n+1}(x) - f_n(x)| \leq \frac{b_{n,i} - a_{n,i}}{6}, \quad \text{si } x \in I_{n,i}.$$

Une récurrence immédiate permet d'en déduire que $\|f_{n+1} - f_n\|_\infty \leq \frac{1}{6 \cdot 2^n}$, et

$$b_{n,i} - a_{n,i} = \begin{cases} \frac{1}{2^n} & \text{si tous les chiffres de l'écriture de } i \text{ en base 3 sont des 0 ou des 2,} \\ 0 & \text{si un des chiffres de l'écriture de } i \text{ en base 3 est un 1.} \end{cases}$$

Comme $\sum_{n \in \mathbf{N}} \frac{1}{6 \cdot 2^n} < +\infty$, la série $f_0 + \sum_{n=0}^{+\infty} (f_{n+1} - f_n)$ converge normalement, et sa somme f , qui est aussi la limite de la suite $(f_n)_{n \in \mathbf{N}}$, est continue. Chaque f_n étant croissante, il en est de même de la limite f . Finalement, f est constante sur $I_{n,i}$, si un des chiffres de i dans le développement en base 3 est un 1. Il y a $3^n - 2^n$ tels i , ce qui fait que la réunion F_n des $I_{n,i}$, pour i vérifiant la condition précédente, est de longueur totale égale à $1 - \frac{2^n}{3^n}$. Comme f est constante sur (chacun des intervalles composant) F_n , un passage à la limite montre que f est constante sur la réunion des F_n qui est de longueur totale égale à 1. D'un autre côté, on a $f_n(0) = 0$ et $f_n(1) = 1$, pour tout n , et donc $f(0) = 0$ et $f(1) = 1$ par passage à la limite. On a donc bien construit une fonction continue qui croît subrepticement de 0 à 1.

12.3. L'ensemble triadique de Cantor

C'est un fermé K de \mathbf{R} inclus dans $[0, 1]$, de mesure nulle, mais quand même assez gros pour qu'il existe une surjection de K sur $[0, 1]$. C'est l'ensemble des points de $[0, 1]$ en lesquels l'escalier du diable croît.

On construit par récurrence une suite $(K_n)_{n \in \mathbf{N}}$ de fermés de $[0, 1]$, chaque K_n étant la réunion de 2^n segments fermés. On part de $K_0 = [0, 1]$, et si K_n est construit, on obtient K_{n+1} en coupant chacun des segments fermés constituant K_n en 3 segments de même longueur et en enlevant le morceau du milieu (ouvert pour que K_{n+1} soit fermé). On a donc

$$K_1 = [0, 1] -]\frac{1}{3}, \frac{2}{3}[= [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \quad K_2 = K_1 - \left(]\frac{1}{9}, \frac{2}{9}[\cup]\frac{7}{9}, \frac{8}{9}[\right) = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1].$$

On note K l'intersection des K_n ; c'est un fermé de $[0, 1]$ comme intersection de fermés. La somme des longueurs des segments constituant K_n est $(\frac{2}{3})^n$ qui tend vers 0 quand $n \rightarrow +\infty$, ce qui fait que K est de mesure nulle, puisque $K \subset K_n$ pour tout n .

Par ailleurs, K est l'ensemble des $x \in [0, 1]$ dont un des développements en base 3 ne comporte que des 0 et des 2 (les seuls nombres ayant deux développements sont ceux de la forme $\frac{k}{3^n}$, avec $k \in \mathbf{Z}$ et $n \in \mathbf{N}$). En effet, les nombres que l'on retire pour passer de K_n à K_{n+1} sont précisément ceux dont tous les développements en base 3 ont un 1 en n -ième position et pas de 1 avant. L'application $(a_n)_{n \geq 1} \mapsto \sum_{n=1}^{+\infty} \frac{a_n}{3^n}$ induit donc une bijection de l'ensemble $\{0, 2\}^{\mathbf{N}-\{0\}}$ sur K , ce qui nous permet de définir une surjection $f : K \rightarrow [0, 1]$, en passant de la base 3 à la base 2, c'est-à-dire en envoyant $\sum_{n=1}^{+\infty} \frac{a_n}{3^n}$ sur $\sum_{n=1}^{+\infty} \frac{b_n}{2^n}$, où $b_n = a_n/2 \in \{0, 1\}$.

Exercice 12.2. — (i) Adapter la construction ci-dessus pour construire un fermé de $[0, 1]$ d'intérieur vide, mais de mesure non nulle.

(ii) Montrer qu'un tel ensemble est totalement discontinu.

12.4. La courbe de Peano

Il s'agit d'une courbe fractale qui remplit tout le carré, ce qui montre que la notion de dimension est plus problématique que ce qu'on pourrait croire (un probabiliste dirait que pour obtenir une courbe ayant cette propriété, il suffit de lancer un mouvement brownien qui se chargera de remplir le plan tout seul, et même de repasser une infinité de fois par chaque point).

On construit la courbe de Peano $f : [0, 1] \rightarrow [0, 1]^2$ comme une limite de fonctions f_n , affines par morceaux, construites par récurrence. La fonction f_0 est juste $t \mapsto (t, t)$; son image est donc la diagonale du carré $[0, 1]^2$. La fonction f_n est une fonction affine sur chaque intervalle de la forme $I_{n,i} = [\frac{i}{9^n}, \frac{i+1}{9^n}]$, et le passage de f_{n+1} à f_n se fait en remplaçant chacun des 9^n segments qui constituent l'image de f_n par 9 segments par le procédés indiqués à la figure 2. La figure 3 montre ce que cela donne pour f_2 (les fonctions f_0 et f_1 sont représentées sur la figure 2).

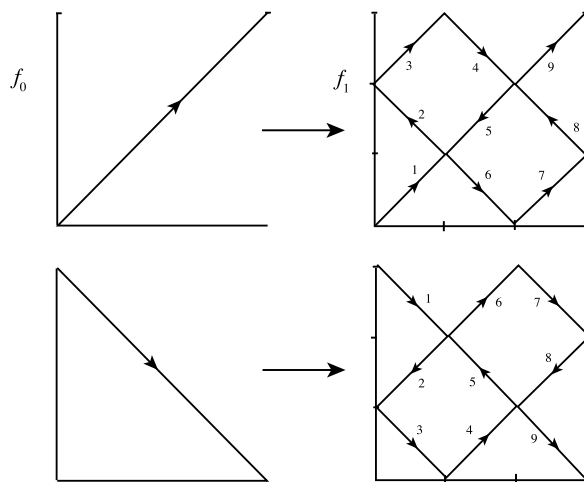


FIG. 2. Procédé d'obtention de f_{n+1} à partir de f_n ; pour un segment allant dans l'autre sens, on renverse juste le sens de parcours.

Par construction, les fonctions f_{n+1} et f_n ont une image incluse dans le même sous-carré de côté de longueur $\frac{1}{3^n}$, sur chacun des segments $I_{n,i}$, pour $0 \leq i \leq 9^n - 1$. On a donc $\|f_{n+1} - f_n\|_\infty \leq \frac{1}{3^n}$, si on munit \mathbf{R}^2 de la norme $\|(x, y)\| = \sup(|x|, |y|)$. On en déduit que f_n converge uniformément sur $[0, 1]$, et comme les f_n sont continues, il en est de même de la limite f .

On a $f_{n+1}(\frac{i}{9^n}) = f_n(\frac{i}{9^n})$, si $0 \leq i \leq 9^n$, et donc $f(\frac{i}{9^n}) = f_n(\frac{i}{9^n})$, si $n \in \mathbf{N}$ et $0 \leq i \leq 9^n$. Or l'image de $\{\frac{i}{9^n}, 0 \leq i \leq 9^n - 1\}$ par f_n est l'ensemble A_n des couples $(\frac{a}{3^n}, \frac{b}{3^n})$, avec a, b entiers, $0 \leq a, b \leq 3^n$, et $a + b$ pair. La réunion des A_n est dense dans $[0, 1]^2$, et est contenue dans l'image de f d'après ce qui précède; l'image de f est donc dense dans $[0, 1]^2$. Pour montrer que f remplit tout le carré $[0, 1]^2$, il n'y a plus qu'à remarquer que $[0, 1]$

étant compact et f continue, $f([0, 1])$ est compacte et donc fermée dans $[0, 1]^2$, et comme elle est dense, c'est $[0, 1]^2$ tout entier !

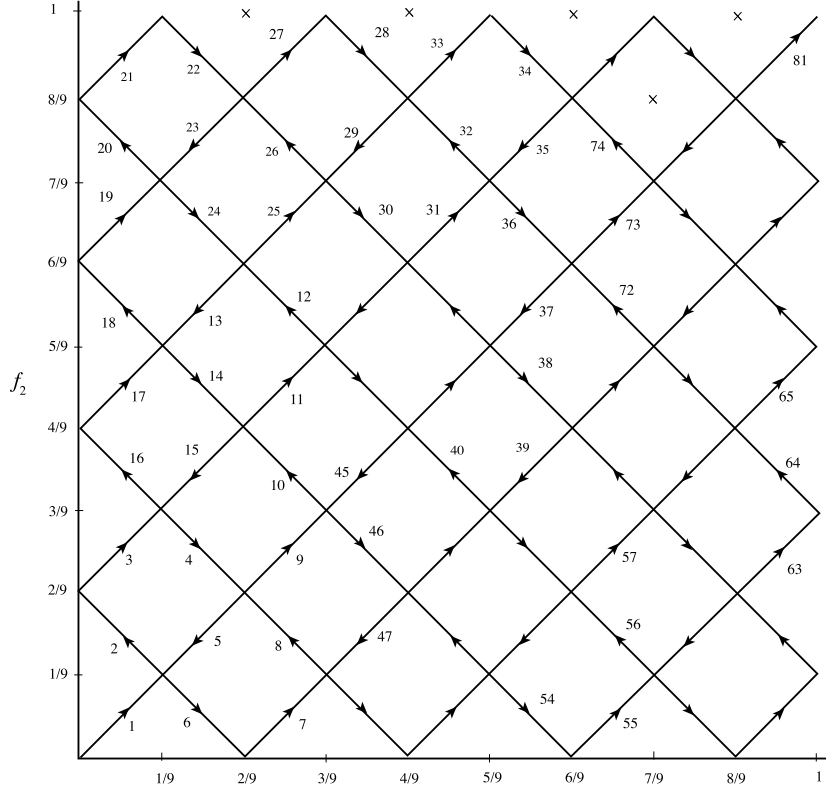


FIG. 3. La fonction f_2 : les nombres apparaissant sur la figure correspondent à l'ordre dans lequel les $9^2 = 81$ segments sont parcourus.

12.5. Ensembles connexes non connexes par arcs

12.5.1. Le graphe de $\sin \frac{1}{x}$

Soit X le graphe de la fonction $x \mapsto \phi(x) = \sin \frac{1}{x}$, pour $x > 0$. L'ensemble X est connexe par arcs, vu que c'est un arc en tant qu'image de \mathbf{R}_+^* par $x \mapsto (x, \phi(x))$ qui est une application continue de \mathbf{R}_+^* dans \mathbf{R}^2 . Son adhérence \overline{X} dans X est donc connexe ; nous allons montrer qu'elle n'est pas connexe par arc.

Commençons par montrer que $\overline{X} = X \cup I$, où I est le segment vertical $I = \{(0, y), y \in [-1, 1]\}$.

- Comme \mathbf{R}^2 est métrique, un point (a, b) est dans l'adhérence de X , s'il existe une suite $(x_n, y_n)_{n \in \mathbf{N}}$ d'éléments de X convergeant vers (a, b) dans \mathbf{R}^2 . Or $y_n = \phi(x_n)$ et ϕ est continue sur \mathbf{R}_+^* , ce qui fait que, si $a > 0$, on doit avoir $b = \phi(a)$ par continuité. L'intersection de \overline{X} avec $\mathbf{R}_+^* \times \mathbf{R}$ est donc réduite à X .
- Comme X est contenu dans le fermé $\mathbf{R}_+ \times [-1, 1]$, il en est de même de son adhérence ; on en déduit l'inclusion $\overline{X} \subset X \cup I$.

• Si $b \in [-1, 1]$, alors $(0, b)$ est la limite de $(\frac{1}{2n\pi + \arcsin(b)}, b) \in X$, quand $n \rightarrow +\infty$, ce qui montre que $(0, b) \in \overline{X}$, et donc que $I \subset \overline{X}$. On en déduit l'égalité $\overline{X} = X \cup I$ que l'on cherchait à établir.

Démontrons, par l'absurde, que \overline{X} n'est pas connexe par arcs. Supposons donc que \overline{X} est connexe par arcs; il existe alors $u : [0, 1] \rightarrow \overline{X}$, continue, telle que $u(0) = (0, 0)$, et $u(1) = (\pi^{-1}, 0)$. Soit $A = \{t, u(t) \in I\}$, et soit $a \in [0, 1]$ la borne supérieure de A . Alors $u(a) \in I$ car I est fermé et u est continue, et $u(t) \notin I$, si $t > a$. On a donc $u(a) = (0, b)$, et $u(t) = (x(t), y(t))$, avec $x(t) > 0$, si $t > a$. On peut supposer, sans nuire à la généralité, que $b \neq 1$ (sinon, on remplace 1 par -1 dans ce qui suit). Comme u est continue, il existe $\delta > 0$ tel que $y(t) \neq 1$, si $t \in [a, a + \delta]$. Comme $y(t) = \phi(x(t))$, cela implique que $x(t)$ n'est pas de la forme $\frac{1}{2n\pi + (\pi/2)}$, si $t \in [a, a + \delta]$. Or le seul intervalle de \mathbf{R}_+ contenant 0 et ne contenant aucun point de la forme $\frac{1}{2n\pi + (\pi/2)}$, pour $n \in \mathbf{N}$, est $\{0\}$. Comme $t \mapsto x(t)$ est continue sur $[a, a + \delta]$, cela implique $x(t) = 0$, si $t \in [a, a + \delta]$, ce qui est contraire à la définition de a . Ceci permet de conclure.

12.5.2. Le tipi de Cantor

C'est un sous-ensemble T du plan qui défie un peu l'entendement car il est connexe, et il existe $S \in T$ tel que, si on retire S à T , le résultat est totalement discontinu (ce qui signifie, rappelons-le, que les composantes connexes de $T - S$ sont réduites à des points).

Pour construire T , on part de l'ensemble triadique de Cantor K que l'on partitionne en un ensemble K_1 dénombrable et dense⁽³⁹⁾ et son complémentaire K_2 .

On identifie K à un sous-ensemble de \mathbf{R}^2 par $t \mapsto (t, 0)$, ce qui permet de voir K comme un sous-ensemble du segment horizontal $L = [0, 1] \times \{0\}$. On note S le point $(0, 1)$, et si $P = (t, 0)$, avec $t \in K_1$ (resp. $t \in K_2$), on définit le *rayon* T_P comme l'ensemble des (x, y) appartenant au segment $[P, S[$, avec $y \in \mathbf{Q}$ (resp. $y \notin \mathbf{Q}$). On définit le *tipi de Cantor* T comme la réunion des T_P , pour $P \in K$, auquel on rajoute le *sommet* S de T . Nous allons montrer que T est connexe, mais que T privé de S est totalement discontinu.

Pour montrer que T est connexe, considérons une partition de T en deux ouverts U_1 et U_2 , et supposons que $S \in U_1$. Comme U_1 est non vide, il s'agit de montrer que U_2 l'est. Comme il est plus confortable de travailler dans un carré que dans un triangle, on remarque que $(x, y) \mapsto ((1 - y)x, y)$ induit un homéomorphisme de $[0, 1] \times [0, 1[$ sur le triangle de sommets $A = (0, 0)$, $B = (1, 0)$ et $S = (0, 1)$, privé de son sommet S ; l'homéomorphisme réciproque étant $(x, y) \mapsto (\frac{x}{1-y}, y)$. Via cet homéomorphisme, le rayon T_P devient $T'_P = \{P\} \times ([0, 1] \cap \mathbf{Q})$, si $P \in K_1$, et $T'_P = \{P\} \times ([0, 1] \cap (\mathbf{R} - \mathbf{Q}))$, si $P \in K_2$, et $T - S$ devient la réunion T' de $K_1 \times ([0, 1] \cap \mathbf{Q})$, et de $K_2 \times ([0, 1] \cap (\mathbf{R} - \mathbf{Q}))$. L'ouvert $U_1 - S$ devient un ouvert U'_1 de T' contenant $([0, 1] \times]1 - \delta, 1]) \cap T'$, si $\delta > 0$ est assez petit, U_2 devient un ouvert U'_2 de T' , et U'_1 et U'_2 forment une partition de T' .

On est alors ramené à prouver que U'_2 est vide. On définit une fonction $h : K \rightarrow [0, 1]$, par $h(P) = 0$, si $T'_P \cap U'_2 = \emptyset$, et $h(P) = \sup\{y, (P, y) \in T'_P \cap U'_2\}$, si $T'_P \cap U'_2 \neq \emptyset$. Comme U'_2 est ouvert, sa vacuité est équivalente à $h = 0$ sur K ; on va donc s'intéresser aux points où $h \neq 0$.

• $h(P) < 1$ pour tout $P \in K$, car $U'_1 \cap U'_2 = \emptyset$ et U'_1 contient $([0, 1] \times]1 - \delta, 1]) \cap T'$, si $\delta > 0$ est assez petit.

⁽³⁹⁾ On peut, par exemple, prendre pour K_1 l'ensemble des éléments de K dont le développement en base 3 est limité, i.e. l'ensemble des nombres de la forme $\sum_{i=1}^n \frac{a_i}{3^n}$, avec $n \in \mathbf{N}$, et $a_i \in \{0, 2\}$, si $1 \leq i \leq n$.

• $h(P) \in \mathbf{Q}$ si $P \in K_2$, car sinon le point $(P, h(P))$ de T'_P appartiendrait à U'_i , pour $i = 1$ ou $i = 2$, et comme U'_i est ouvert, il existerait un segment ouvert $J \subset]0, 1[$, contenant $h(P)$, tel que $\{(P, t), t \in J\} \cap T'$, soit contenu dans U'_i . Dans les deux cas $i = 1$ et $i = 2$, on obtient une contradiction avec la définition de $h(P)$.

• Si $q \in]0, 1[\cap \mathbf{Q}$, et si $P \in K_1$, il existe un ouvert I de K contenant P tel que $h(Q) \neq q$ pour tout $Q \in I$. En effet, le point (P, q) appartient à T'_P par construction, et donc appartient à U'_i , pour $i = 1$ ou $i = 2$. Comme U'_i est ouvert et contient (P, q) , il contient un ouvert de la forme $(I \times J) \cap T'$, où I est un ouvert de K contenant P , et J est un ouvert de $]0, 1[$, contenant q ; la définition de h montre que l'on a $h(Q) \notin J$, si $Q \in I$.

Si $q \in]0, 1[\cap \mathbf{Q}$, soit F_q l'adhérence de $\{P \in K, h(P) = q\}$. C'est un fermé de K par construction, et il ne rencontre pas K_1 d'après le point précédent. Il est donc d'intérieur vide puisque K_1 est dense dans K . Comme K est un compact métrique, il est complet, et le lemme de Baire implique que la réunion X de K_1 et des F_q , pour $q \in \mathbf{Q} \cap]0, 1[$, est d'intérieur vide, puisque c'est une réunion dénombrable de fermés d'intérieur vide ($\mathbf{Q} \cap]0, 1[$ est dénombrable et K_1 est dénombrable et donc est une réunion dénombrable de singletons). L'ensemble $K - X$ est donc dense dans K . Or $P \in K - X$ implique $h(P) = 0$, et donc aussi $\{P\} \times (]0, 1[\cap (\mathbf{R} - \mathbf{Q})) \subset U'_1$. Donc U'_1 contient $(K - X) \times (]0, 1[\cap (\mathbf{R} - \mathbf{Q}))$ qui est dense dans T' car il l'est dans $K \times]0, 1[$ qui contient T' . Son complémentaire U'_2 est donc d'intérieur vide, et comme il est ouvert, il est vide. On en déduit la connexité de T .

Il reste à montrer que $T - S$ est totalement discontinu, et comme $T - S$ est homéomorphe à T' , il suffit de prouver que T' l'est. Pour cela considérons deux points distincts (P_1, y_1) et (P_2, y_2) de T' . Si $P_1 \neq P_2$, il existe $Q \notin K$ dans l'intervalle ouvert d'extrémités P_1 et P_2 puisque K est d'intérieur vide. La droite verticale $\{Q\} \times \mathbf{R}$ ne rencontre pas T' , et les deux demi-plans ouverts qu'elle délimite partitionnent T' en deux ouverts, l'un contenant (P_1, y_1) , l'autre (P_2, y_2) . On en déduit que (P_2, y_2) n'est pas dans la composante connexe de (P_1, y_1) . Une composante connexe de T' est donc incluse dans un rayon T'_P , or un tel ensemble est totalement discontinu puisqu'il est homéomorphe à $\mathbf{Q} \cap [0, 1[$ ou à $(\mathbf{R} - \mathbf{Q}) \cap [0, 1[$. Les composantes connexes de T' sont donc des points, ce qui permet de conclure.

13. Corrigé des exercices

Exercice 1.1. (i) Si $a = 0$ ou $b = 0$, on a $a\mathbf{Z} \cap b\mathbf{Z} = \{0\}$ et $\text{ppcm}(a, b) = 0$ puisque le seul multiple de 0 est 0. Si $a \neq 0$ et $b \neq 0$, alors $a\mathbf{Z} \cap b\mathbf{Z}$ est un sous-groupe de \mathbf{Z} comme intersection de deux sous-groupes, qui n'est pas réduit à 0 puisqu'il contient ab . Il existe donc $m \in \mathbf{N}$ tel que $a\mathbf{Z} \cap b\mathbf{Z} = m\mathbf{Z}$. Alors m est un multiple de a (car $a \in m\mathbf{Z}$) et de b (car $b \in m\mathbf{Z}$). Donc $\text{ppcm}(a, b) \mid m$. Réciproquement, si c est un multiple de a et b , alors $c \in a\mathbf{Z}$ et $c \in b\mathbf{Z}$ et donc $c \in m\mathbf{Z}$ et $m \mid c$. En particulier, $m \mid \text{ppcm}(a, b)$, et donc $m = \text{ppcm}(a, b)$, ce qu'il fallait démontrer.

(ii) On a $a \mid c$ (resp. $b \mid c$) si et seulement si $v_p(a) \leq v_p(c)$ (resp. $v_p(b) \leq v_p(c)$), pour tout $p \in \mathcal{P}$. Donc c est un multiple de a et b si et seulement si $v_p(c) \geq \sup(v_p(a), v_p(b))$, pour tout $p \in \mathcal{P}$. Le plus petit entier multiple de a et b est donc $\prod_{p \in \mathcal{P}} p^{\sup(v_p(a), v_p(b))}$, ce qu'il fallait démontrer.

Exercice 1.2. (i) Si $a = 0$ ou $b = 0$, on a $v_p(ab) = v_p(a) + v_p(b)$ car les deux membres valent $+\infty$. Si $a \neq 0$ et $b \neq 0$, on a $a = \text{sign}(a) \prod_{p \in \mathcal{P}} p^{v_p(a)}$, $b = \text{sign}(b) \prod_{p \in \mathcal{P}} p^{v_p(b)}$ et

$$ab = \text{sign}(ab) \prod_{p \in \mathcal{P}} p^{v_p(ab)} = \text{sign}(a)\text{sign}(b) \prod_{p \in \mathcal{P}} p^{v_p(a)+v_p(b)}.$$

On déduit de l'unicité de la décomposition en produit de facteurs premiers que $\text{sign}(ab) = \text{sign}(a)\text{sign}(b)$ et $v_p(ab) = v_p(a) + v_p(b)$, pour tout $p \in \mathcal{P}$.

Maintenant, si $m = \inf(v_p(a), v_p(b))$, alors $p^m \mid a$ et $p^m \mid b$, ce qui implique $p^m \mid a + b$ et donc $v_p(a + b) \geq m$, ce qu'on cherchait à démontrer.

(ii) Si $x = \frac{a}{b}$, avec $a \in \mathbf{Z}$ et $b \in \mathbf{Z} - \{0\}$, on doit avoir $v_p(x) = v_p(a) - v_p(b)$, et il faut vérifier que cela ne dépend pas de l'écriture choisie. Or, si $\frac{a'}{b'} = \frac{a}{b}$, on a $ab' = ba'$ et donc $v_p(a) + v_p(b') = v_p(b) + v_p(a')$ et $v_p(a) - v_p(b) = v_p(a') - v_p(b')$, ce qui prouve que $v_p(x)$ est bien défini. De plus, si $y = \frac{c}{d}$, alors $v_p(xy) = v_p(\frac{ac}{bd}) = v_p(ac) - v_p(bd) = v_p(a) + v_p(c) - v_p(b) - v_p(d) = v_p(x) + v_p(y)$. Enfin, si $x, y \in \mathbf{Q}$ et si $c \in \mathbf{N} - \{0\}$ est tel que $cx, cy \in \mathbf{Z}$, on a $v_p(c(x + y)) \geq \inf(v_p(cx), v_p(cy))$ et donc

$$v_p(c) + v_p(x + y) \geq \inf(v_p(c) + v_p(x), v_p(c) + v_p(y)) = v_p(c) + \inf(v_p(x), v_p(y)),$$

et comme $v_p(c)$ est fini, cela permet de conclure.

Exercice 1.3. On a $n! = \prod_{k=1}^n k$ et donc $v_p(n!) = \sum_{k=1}^n v_p(k)$. Or il y a exactement $\lfloor \frac{n}{p^i} \rfloor - \lfloor \frac{n}{p^{i+1}} \rfloor$ entiers $\leq n$ vérifiant $v_p(k) = i$ (les multiples de p^i privés des multiples de p^{i+1}). On en déduit que $v_p(n!) = \sum_{i=1}^{+\infty} i(\lfloor \frac{n}{p^i} \rfloor - \lfloor \frac{n}{p^{i+1}} \rfloor) = \sum_{i=1}^{+\infty} \lfloor \frac{n}{p^i} \rfloor (i - (i-1)) = \sum_{i=1}^{+\infty} \lfloor \frac{n}{p^i} \rfloor$.

Maintenant, si $n = a_0 + a_1p + \dots + a_r p^r$, où $a_i \in \{0, \dots, p-1\}$, pour tout i , alors $\lfloor \frac{n}{p^i} \rfloor = a_i + \dots + a_r p^{r-i}$ et donc

$$\begin{aligned} \sum_{i=1}^{+\infty} \lfloor \frac{n}{p^i} \rfloor &= \sum_{i=1}^r \left(\sum_{s=i}^r a_s p^{s-i} \right) = \sum_{s=1}^r \sum_{i=1}^s a_s p^{s-i} \\ &= \sum_{s=1}^r a_s p^{s-1} \left(\frac{1-p^{-s}}{1-p^{-1}} \right) = \sum_{s=1}^r a_s \left(\frac{p^s-1}{p-1} \right) = \sum_{s=0}^r a_s \left(\frac{p^s-1}{p-1} \right) = \frac{n - S_p(n)}{p-1}. \end{aligned}$$

Exercice 1.4. L'ensemble $\mathcal{P}(\mathbf{N})$ est en bijection avec $\{0, 1\}^{\mathbf{N}}$ (on associe à $X \subset \mathbf{N}$ la suite $(x_k)_{k \in \mathbf{N}}$ définie par $x_k = 1$ si $k \in X$ et $x_k = 0$ si $k \notin X$) ; il n'est donc pas dénombrable.

L'ensemble des parties finies de \mathbf{N} est la réunion, pour $n \in \mathbf{N}$, de l'ensemble des parties de $\{0, \dots, n\}$; il est donc dénombrable en tant que réunion dénombrable d'ensembles finis.

Exercice 1.5. Si n est fixé, l'ensemble $\mathbf{Q}[X]^{(n)}$ des polynômes de $\mathbf{Q}[X]$ de degré n s'injecte dans \mathbf{Q}^{n+1} en envoyant $P = a_n X^n + \dots + a_0$ sur (a_n, \dots, a_0) ; il est donc dénombrable puisque \mathbf{Q} l'est. On en déduit que $\mathbf{Q}[X] = \bigcup_{n \in \mathbf{N}} \mathbf{Q}[X]^{(n)}$ est dénombrable comme réunion dénombrable d'ensembles dénombrables. Enfin,

un polynôme n'ayant qu'un nombre fini de racines dans \mathbf{C} , l'ensemble $\overline{\mathbf{Q}}$ est une réunion dénombrable (d'après ce qui précède) d'ensembles finis, et donc est dénombrable.

L'ensemble des nombres transcendants n'est pas dénombrable (sinon \mathbf{R} le serait comme réunion de deux ensembles dénombrables) ; en particulier, il est non vide.

Exercice 1.6. Si on choisit dans chaque disque un point de la forme $a + ib$, avec $a, b \in \mathbf{Q}$, on obtient une injection de I dans \mathbf{Q}^2 , et comme \mathbf{Q}^2 est dénombrable puisque \mathbf{Q} l'est, cela permet de conclure.

Exercice 1.7 (i) Soit $a = \inf_{x > x_0} f(x)$. Par définition de a , on a $f(x) \geq a$, si $x > x_0$, et pour tout $\varepsilon > 0$, il existe $x_\varepsilon > x_0$ tel que $f(x_\varepsilon) < a + \varepsilon$. Soit $\delta = x_\varepsilon - x_0$. Comme f est croissante, on a $a \leq f(x) < a + \varepsilon$, pour tout $x \in]x_0, x_0 + \delta[$, ce qui prouve que f a une limite à droite $f(x_0^+)$ en x_0 , égale à a . La limite à gauche s'étudie exactement de la même manière (ou peut se déduire de ce qu'on vient de faire en étudiant $g(x) = -f(-x)$ en $-x_0$).

Maintenant, comme f est croissante, on a $f(x_0^-) = \sup_{x < x_0} f(x) \leq f(x_0) \leq \inf_{x > x_0} f(x) = f(x_0^+)$. Comme f admet des limites à gauche et à droite en x_0 , elle est continue en x_0 si et seulement si $f(x_0^-) = f(x_0) = f(x_0^+)$, et donc si et seulement si $f(x_0^-) = f(x_0^+)$.

(ii) Comme f est croissante, on a $f(x_0^+) = \inf_{x > x_0} f(x) = \inf_{x_1 > x > x_0} f(x) \leq \sup_{x_0 < x < x_1} f(x) = \sup_{x < x_1} f(x) = f(x_1^-)$.

(iii) Soit $x \in D$ un point de discontinuité. On a alors $f(x^-) < f(x^+)$, ce qui permet de choisir un élément $r(x) \in \mathbf{Q}$ dans l'intervalle $]f(x^-), f(x^+)[$. Si $x_1 < x_2$ sont deux éléments de D , on a $r(x_1) < f(x_1^+) \leq f(x_2^-) < r(x_2)$, ce qui prouve que $x \mapsto r(x)$ est une injection de D dans \mathbf{Q} , et \mathbf{Q} étant dénombrable, cela implique que D est dénombrable.

Exercice 1.8. Soit $(H_i)_{i \in I}$ une famille de huit dans le plan, deux à deux disjoints. Si H_i est constitué des cercles $C_{i,1}$ et $C_{i,2}$, choisissons un point $P_{i,1}$ (resp. $P_{i,2}$) à coordonnées rationnelles dans le disque $D_{i,1}$ (resp. $D_{i,2}$) délimité par $C_{i,1}$ (resp. $C_{i,2}$). On obtient de la sorte une application de I dans \mathbf{Q}^4 . Soient $i \neq j$ deux éléments de I . Si $P_{i,1} = P_{j,1}$, alors l'un des disques $D_{i,1}$ et $D_{j,1}$ contient l'autre puisque les cercles $C_{i,1}$ et $C_{j,1}$ sont disjoints. Quitte à permuter i et j , on peut supposer que c'est $D_{i,1}$ qui contient $D_{j,1}$, mais alors $D_{i,1}$ contient aussi le point de contact entre $C_{j,1}$ et $C_{j,2}$, et donc aussi le cercle $C_{j,2}$ tout entier puisque $C_{j,2}$ et $C_{i,1}$ sont disjoints, et donc aussi le disque $D_{j,2}$ et le point $P_{j,2}$. Comme il ne contient pas $P_{i,2}$ par construction, on en déduit que $i \mapsto (P_{i,1}, P_{i,2})$ est injective, et comme \mathbf{Q}^4 est dénombrable, il en est de même de I .

Exercice 1.9. L'idée est de prouver que deux tripodes disjoints ne peuvent pas être trop proches. Soient donc Y et Y' deux tripodes de sommets respectifs (A, B, C) et (A', B', C') et de centres de gravité G et G' . Soit $r = d(G, A)$. Si $d(A, A')$, $d(B, B')$ et $d(C, C')$ sont toutes trois $< \frac{r}{2}$, on a aussi $d(G, G') < \frac{r}{2}$ et un petit dessin montre que suivant le tiers de plan dans lequel se trouve G' , l'un des segments $[G', A']$, $[G', B']$ ou $[G', C']$ rencontre Y . Maintenant, soit $(Y_i)_{i \in I}$ une famille de tripodes dans le plan, deux à deux disjoints. Si $i \in I$, soient A_i, B_i, C_i les sommets de Y_i , G_i le centre de gravité de (A_i, B_i, C_i) et $r_i = d(G_i, A_i)$. Choisissons pour tout i un triplet $(P_{i,1}, P_{i,2}, P_{i,3})$ de points à coordonnées rationnelles, avec $d(A_i, P_{i,1}) < \frac{r_i}{4}$, $d(B_i, P_{i,2}) < \frac{r_i}{4}$ et $d(C_i, P_{i,3}) < \frac{r_i}{4}$. Il résulte de la discussion préliminaire que l'on obtient ainsi une injection de I dans \mathbf{Q}^6 , ce qui prouve que I est dénombrable.

Exercice 2.1 Soit $m = \text{ppcm}(a, b)$. Comme a et b ne sont pas premiers entre eux, on a $m < |ab|$. Or m annule tout élément de $\mathbf{Z}/a\mathbf{Z}$ puisque c'est un multiple de a et tout élément de $\mathbf{Z}/b\mathbf{Z}$ puisque c'est un multiple de b ; on a donc $mx = 0$, pour tout $x \in (\mathbf{Z}/a\mathbf{Z}) \oplus (\mathbf{Z}/b\mathbf{Z})$. Or m n'annule pas 1 dans $\mathbf{Z}/ab\mathbf{Z}$ puisque $m < |ab|$ n'est pas un multiple de ab .

Exercice 2.2. 4 admet 16 comme inverse dans $\mathbf{Z}/21\mathbf{Z}$; l'équation $4x + 3 = 0$ est donc équivalente à $x + 48 = 0$, soit $x = -48 = 3 \times 21 - 48 = 15$.

$14x$ est multiple de 7 dans $\mathbf{Z}/21\mathbf{Z}$, ce que -2 n'est pas. L'équation $14x + 2 = 0$ n'a donc pas de solution dans $\mathbf{Z}/21\mathbf{Z}$.

$14x + 7 = 0$ dans $\mathbf{Z}/21\mathbf{Z}$ équivaut à $7(2x + 1) = 0$ dans $\mathbf{Z}/21\mathbf{Z}$, soit encore à $2x + 1$ multiple de 3 dans $\mathbf{Z}/21\mathbf{Z}$. Les solutions sont donc 1, 4, 7, 10, 13, 16 et 19 modulo 21.

Exercice 2.3. On a $91 = 7 \times 13$ et donc $\mathbf{Z}/91\mathbf{Z} = \mathbf{F}_7 \times \mathbf{F}_{13}$, ce qui nous ramène à trouver les solutions dans les corps \mathbf{F}_7 et \mathbf{F}_{13} . On remarque que 2 est racine dans \mathbf{F}_7 , et comme la somme des racines vaut -1 , l'autre est $-3 = 4$. De même 3 est racine dans \mathbf{F}_{13} , et donc l'autre est $-1 - 3 = -4 = 9$. On est alors confronté au problème de trouver quels sont les éléments de $\mathbf{Z}/91\mathbf{Z}$ correspondant aux couples $(2, 3)$, $(2, 9)$, $(4, 3)$ et $(4, 9)$ de $\mathbf{F}_7 \times \mathbf{F}_{13}$. Pour cela, on remarque que $1 = 2 \times 7 - 13$, et donc que $14 = 2 \times 7$ a pour image 0 dans \mathbf{F}_7 et pour image 1 dans \mathbf{F}_{13} , alors que -13 a pour image 1 dans \mathbf{F}_7 et pour image 0 dans \mathbf{F}_{13} . On en déduit, que si $(a, b) \in \mathbf{Z}$, alors $-13a + 14b \in \mathbf{Z}$ ne dépend modulo 91 que des réductions de a et b modulo 7 et 13 respectivement, et l'image de $-13a + 14b$ dans $\mathbf{F}_7 \times \mathbf{F}_{13}$ est (a, b) . Les solutions de l'équation $x^2 + x + 1 = 0$ dans $\mathbf{Z}/91\mathbf{Z}$ sont donc $-13 \cdot 2 + 14 \cdot 3 = 16$, $-13 \cdot 2 + 14 \cdot 9 = 100 = 9$, $-13 \cdot 4 + 14 \cdot 3 = -10$ et $-13 \cdot 4 + 14 \cdot 9 = 74 = -17$.

Exercice 2.4. (i) Si a est solution de l'équation $x^2 + x + 1 = 0$, alors $-1 - a$ aussi. Or le système d'équations $x^2 + x + 1 = 0$ et $2x = -1$ est équivalent à $2x = -1$ et $x(x - 1) = 0$. Comme \mathbf{F}_p est un corps, $x(x - 1) = 0$ équivaut à $x = 0$ ou $x = 1$, ce qui est incompatible avec $2x = -1$, sauf si $2 \cdot 1 = -1$, c'est-à-dire si $3 = 0$, et donc si $p = 3$. On en déduit que, si $p \neq 3$, l'équation $x^2 + x + 1 = 0$ a deux solutions dans \mathbf{F}_p si et seulement si elle en a au moins une.

(ii) D'après le (i), si $p \neq 3$, l'équation $x^2 + x + 1 = 0$ a deux solutions modulo p , s'il existe $n \in \mathbf{N}$ tel que p divise $n^2 + n + 1$. Supposons, par l'absurde, que l'ensemble des p vérifiant ceci est fini. Cela signifie qu'il existe des nombres premiers p_1, \dots, p_k tels que pour tout $n \in \mathbf{N}$, il existe $a_1, \dots, a_k \in \mathbf{N}$ tels que $n^2 + n + 1 = p_1^{a_1} \cdots p_k^{a_k}$. Si $n \leq X - 1$, cela implique que $n^2 + n + 1 \leq X^2$, et donc que chacun des a_i vérifie $a_i \leq \frac{\log X^2}{\log p_i} \leq \frac{2}{\log 2} \log X$; on en déduit que $n^2 + n + 1$ peut prendre au plus $(\frac{2}{\log 2} \log X)^k$ valeurs pour $n \leq X - 1$, ce qui est absurde pour X tendant vers $+\infty$, les valeurs de $n^2 + n + 1$ étant toutes distinctes pour $n \geq 0$.

(iii) Il existe un ensemble infini $\{p_1, p_2, \dots\}$ de nombres premiers tels que l'équation $x^2 + x + 1 = 0$ ait deux solutions dans \mathbf{F}_p . Soit $D_k = p_1 \cdots p_k$. D'après le théorème des restes chinois, $\mathbf{Z}/D_k\mathbf{Z} = \prod_{i=1}^k \mathbf{F}_{p_i}$, et comme l'équation $x^2 + x + 1 = 0$ a deux solutions dans \mathbf{F}_{p_i} , pour tout i , elle en a 2^k dans $\mathbf{Z}/D_k\mathbf{Z}$. Comme 2^k peut être rendu arbitrairement grand, cela permet de conclure.

Exercice 2.5 Les éléments inversibles de $\mathbf{Z}/p^n\mathbf{Z}$ sont en bijection avec les éléments de $\{0, 1, \dots, p^n - 1\}$ qui sont premiers à p^n . Or être premier à p^n est équivalent à être premier à p d'après le lemme de Gauss et donc aussi à ne pas être divisible par p , puisque p est premier. Comme il y a p^{n-1} multiples de p dans $\{0, 1, \dots, p^n - 1\}$, on en déduit que $|\mathbf{Z}/p^n\mathbf{Z}^*| = p^n - p^{n-1}$.

Maintenant, si $D \geq 2$ est quelconque, on peut factoriser D sous la forme $D = \prod_{p|D} p^{n_p}$, avec $n_p \geq 1$, et le théorème des restes chinois nous dit que l'anneau $\mathbf{Z}/D\mathbf{Z}$ est isomorphe à $\prod_{p|D} (\mathbf{Z}/p^{n_p}\mathbf{Z})$. On a donc $(\mathbf{Z}/D\mathbf{Z})^* = \prod_{p|D} (\mathbf{Z}/p^{n_p}\mathbf{Z})^*$, ce qui nous donne

$$\varphi(D) = \prod_{p|D} (p^{n_p} - p^{n_p-1}) = D \prod_{p|D} \left(1 - \frac{1}{p}\right).$$

Exercice 2.7 (i) Si $v_1 = (x_1, y_1)$ et $v_2 = (x_2, y_2)$ engendrent la même droite, il existe $\alpha \in K^*$ tel que $v_2 = \alpha v_1$, et on a $\lambda(v_2) = \frac{x_2}{y_2} = \frac{\alpha x_1}{\alpha y_1} = \frac{x_1}{y_1} = \lambda(v_1)$, ce qui prouve que $\lambda(v)$ ne dépend que de la droite engendrée par v , et donc que λ induit une application de $\mathbf{P}^1(K)$ dans $K \cup \{\infty\}$. Cette application est injective car « $\lambda(v_1) = \lambda(v_2)$ » équivaut à « $x_1 y_2 = x_2 y_1$ », et donc « v_1 et v_2 colinéaires ». Elle est surjective car $(1, 0)$ s'envoie sur ∞ et $(z, 1)$ sur z , si $z \in K$. C'est donc une bijection.

(ii) Soit $z \in K \cup \{\infty\}$, et soit $v = (x, y)$ tel que $\frac{x}{y} = \lambda(v) = z$. Alors, par définition, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \lambda\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot v\right) = \lambda(ax + by, cx + dy) = \frac{ax+by}{cx+dy} = \frac{az+b}{cz+d}$.

Exercice 2.9 (i) Une isométrie u du carré permute ses sommets et laisse fixe son centre de gravité O . En particulier u est linéaire et est déterminée par l'image de deux points non colinéaires avec O , par exemple A et B . L'image de A doit appartenir à $\{A, B, C, D\}$, et comme l'angle $\{u(A), O, u(B)\}$ doit être un angle droit, cela ne laisse que deux possibilités pour $u(B)$ pour chaque choix de $u(A)$. On en déduit que D_4 a au plus 8 éléments. Comme il contient l'identité id , la symétrie $-\text{id}$ par rapport à O , les rotations ρ^+ et ρ^- de centre O et d'angles respectifs $\frac{\pi}{2}$ et $-\frac{\pi}{2}$, les symétries $\sigma_{A,C}$ et $\sigma_{C,D}$ par rapport aux deux diagonales, et les symétries σ_H et σ_V par rapport aux droites horizontale et verticale, on voit que D_4 a exactement 8 éléments qui sont ceux que nous venons d'énumérer.

(ii) S est de toute évidence stable par D_4 , et il y a deux orbites :

- O est fixe par tout élément de D_4 ; son orbite est donc $\{O\}$ et son stabilisateur est D_4 ;
- on passe de A à B, C et D en itérant ρ^+ , ce qui montre que l'orbite de A est $\{A, B, C, D\}$ (elle ne peut contenir O puisque les orbites sont distinctes), et on détermine par inspection que le stabilisateur de A est le groupe à 2 éléments $\{\text{id}, \sigma_{A,C}\}$.

(iii) Les orbites de T sous l'action de D_4 sont au nombre de 3 :

- l'orbite de $\{O, A\}$ consiste en les 4 paires contenant O (on passe de $\{O, A\}$ aux autres en itérant ρ^+), et le stabilisateur de $\{O, A\}$ est $\{\text{id}, \sigma_{A,C}\}$;
- l'orbite de $\{A, B\}$ consiste en les 4 paires de sommets consécutifs (on passe de $\{A, B\}$ aux autres en itérant ρ^+), et le stabilisateur de $\{A, B\}$ est $\{\text{id}, \sigma_V\}$;
- l'orbite de $\{A, C\}$ consiste en les 2 paires de sommets opposés $\{A, C\}$ et $\{B, D\}$, et le stabilisateur de $\{A, C\}$ est $\{\text{id}, -\text{id}, \sigma_{A,C}, \sigma_{B,D}\}$.

(iv) On remarque que dans tous les cas, le produit du cardinal de l'orbite par celui du stabilisateur d'un de ses éléments est $8 = |D_4|$; il s'agit d'un cas particulier d'un théorème général (si G opère sur X , si $x \in X$, et si G_x est le stabilisateur de x , alors l'orbite O_x est isomorphe à G/G_x , et donc $|O_x| = |G|/|G_x|$).

Exercice 3.2. (i) On a $|x + y|_p \leq |x|_p$. Si $|x + y|_p < |x|_p$, alors $x = (x + y) - y$ et donc $|x|_p \leq \sup(|x + y|_p, |y|_p) < |x|_p$, ce qui est absurde. Donc $|x + y|_p = |x|_p$.

(ii) Comme $u_n \rightarrow 0$, la série $\sum_{n=1}^{+\infty} u_n$ converge et si on note y sa somme, alors $|y|_p \leq \sup_{n \geq 1} |u_n|_p$. Comme on a supposé $|u_0|_p > |u_n|_p$, pour tout $n \geq 1$, on en déduit $|y|_p < |u_0|_p$, puis $|u_0 + y|_p = |u_0|_p$; en particulier, $u_0 + y = \sum_{n \in \mathbb{N}} u_n \neq 0$.

Exercice 4.3 $108 = 2^2 \times 3^3$, et donc $(\mathbb{Z}/108\mathbb{Z})^* \cong (\mathbb{Z}/4\mathbb{Z})^* \oplus (\mathbb{Z}/27\mathbb{Z})^*$. Or $(\mathbb{Z}/4\mathbb{Z})^* = \{\pm 1\}$ est isomorphe à $\mathbb{Z}/2\mathbb{Z}$, et $(\mathbb{Z}/27\mathbb{Z})^*$ est un groupe de cardinal $\varphi(27) = 2 \cdot 9$ qui est donc isomorphe à $(\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/9\mathbb{Z})$ ou à $(\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/3\mathbb{Z}) \oplus (\mathbb{Z}/3\mathbb{Z})$. Dans le second cas, tout élément de $(\mathbb{Z}/27\mathbb{Z})^*$ vérifierait $x^6 = 1$, or $2^6 = 64 \neq 1$ dans $(\mathbb{Z}/27\mathbb{Z})^*$. On a donc $(\mathbb{Z}/27\mathbb{Z})^* \cong (\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/9\mathbb{Z})$ et $(\mathbb{Z}/108\mathbb{Z})^* \cong (\mathbb{Z}/2\mathbb{Z})^2 \oplus (\mathbb{Z}/9\mathbb{Z})$.

$200 = 2^3 \cdot 5^2$, et donc $(\mathbb{Z}/200\mathbb{Z})^* \cong (\mathbb{Z}/8\mathbb{Z})^* \oplus (\mathbb{Z}/25\mathbb{Z})^*$. Or $(\mathbb{Z}/8\mathbb{Z})^*$ est un groupe d'ordre 4 dans lequel tout les éléments sont d'ordre 2 (en effet, $1^2, 3^2, 5^2$ et 7^2 sont congrus à 1 modulo 8) ; il est donc isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$. Par ailleurs, $(\mathbb{Z}/25\mathbb{Z})^*$ est un groupe de cardinal $\varphi(25) = 4 \cdot 5$ qui est donc isomorphe à $(\mathbb{Z}/4\mathbb{Z}) \oplus (\mathbb{Z}/5\mathbb{Z})$ ou à $(\mathbb{Z}/2\mathbb{Z})^2 \oplus (\mathbb{Z}/5\mathbb{Z})$. Dans le second cas, toute puissance 5-ième serait d'ordre 2, or $2^5 = 32 \equiv 7 \pmod{25}$ a un carré égal à $49 \equiv -1 \pmod{25} \neq 1$, et donc $(\mathbb{Z}/25\mathbb{Z})^* \cong (\mathbb{Z}/4\mathbb{Z}) \oplus (\mathbb{Z}/5\mathbb{Z})$ et $(\mathbb{Z}/200\mathbb{Z})^* \cong (\mathbb{Z}/2\mathbb{Z})^2 \oplus (\mathbb{Z}/4\mathbb{Z}) \oplus (\mathbb{Z}/5\mathbb{Z})$.

La solution ci-dessus est un peu artisanale ; on peut aller plus vite en utilisant les résultats de l'ex. 4.6.

Exercice 4.4. Comme $|\mathbb{F}_p^*| = p - 1$, on a $x^{p-1} = 1$ pour tout $x \in \mathbb{F}_p^*$ d'après le théorème de Lagrange. On en déduit que $x^p = x$ pour tout $x \in \mathbb{F}_p$, ce qui se traduit, en remontant dans \mathbb{Z} , par $p|n^p - n$, pour tout $n \in \mathbb{Z}$.

Exercice 4.5. Soit $\oplus_{p \in \mathcal{P}} (\oplus_i (\mathbf{Z}/p^{a_{p,i}} \mathbf{Z}))$ la décomposition de K^* fournie par le th. 4.1. Si K^* n'est pas cyclique, il existe p tel que $a_{p,2} \neq 0$; en effet, sinon on aurait $K^* \cong \mathbf{Z}/D\mathbf{Z}$, où $D = \prod_p p^{a_{p,1}}$, d'après le théorème des restes chinois, et K^* serait cyclique. Mais alors l'équation $x^p = 1$ a au moins p^2 solutions dans K [les éléments de $(p^{a_{p,1}-1} \mathbf{Z}/p^{a_{p,1}} \mathbf{Z}) \oplus (p^{a_{p,2}-1} \mathbf{Z}/p^{a_{p,2}} \mathbf{Z})$], ce qui est impossible dans un corps commutatif.

Exercice 4.6 (i) On a $(1+p^k a)^p = 1 + p^{k+1} a + \frac{p(p-1)}{2} p^{2k} a^2 + p^{3k} a^3 \left(\sum_{i=3}^p \binom{p}{i} (p^k a)^{p-i} \right)$. Dans cette somme, tous les termes sauf les deux premiers sont divisibles par p^{k+2} , si $k \geq 1$ (ou si $k \geq 2$, dans le cas $p = 2$, où $\frac{p(p-1)}{2}$ n'est pas divisible par p). On a donc bien $x \equiv 1 + p^{k+1} a \pmod{p^{k+2}}$, dans les cas considérés, et une récurrence immédiate montre que $(1+p)^{p^{n-2}} = 1 + p^{n-1} \neq 1$ dans $\mathbf{Z}/p^n \mathbf{Z}$, si $p \neq 2$ et $n \geq 2$, et que $(1+4)^{p^{n-3}} = 1 + 2^{n-1} \neq 1$ dans $\mathbf{Z}/2^n \mathbf{Z}$, si $n \geq 3$.

(ii) Supposons p impair. Alors N est le sous-groupe image de $1 + p\mathbf{Z}$ dans $(\mathbf{Z}/p^n \mathbf{Z})^*$, qui est de cardinal p^{n-1} (car $x \mapsto 1 + px$ induit une bijection de $\mathbf{Z}/p^{n-1} \mathbf{Z}$ sur $1 + p\mathbf{Z}$ modulo $p^n \mathbf{Z}$). Comme $(1+p)^{p^{n-2}} \neq 1$ dans $(\mathbf{Z}/p^n \mathbf{Z})^*$, dans la décomposition $\oplus_i (\mathbf{Z}/p^{a_i} \mathbf{Z})$ du groupe N (dont le cardinal est une puissance de p), au moins un des a_i est $\geq n-1$, et donc $N \cong \mathbf{Z}/p^{n-1} \mathbf{Z}$. Le cas $p = 2$ se traite de la même manière.

(iii) La réduction modulo p fournit une surjection $\pi : G = (\mathbf{Z}/p^n \mathbf{Z})^* \rightarrow \mathbf{F}_p^*$, et \mathbf{F}_p^* est un groupe isomorphe à $(\mathbf{Z}/(p-1)\mathbf{Z})$ d'après l'ex. 4.5. Comme $p-1$ et p^{n-1} sont premiers entre eux, il résulte du th. 4.1, que $G_p = N \cong \mathbf{Z}/p^{n-1} \mathbf{Z}$ et que G est de la forme $(\mathbf{Z}/p^{n-1} \mathbf{Z}) \oplus G'$, avec $G' = \oplus_{\ell \neq p} G_\ell$. Alors $G/N \cong G'$, et comme $G/N \cong \mathbf{F}_p^*$ par définition de N et surjectivité de π , cela permet de conclure.

(iv) Le groupe $(\mathbf{Z}/2^n \mathbf{Z})^*$ est de cardinal $\mathbf{Z}/2^{n-1} \mathbf{Z}$, et contient les sous-groupes N et $\{\pm 1\}$ dont l'intersection est nulle. Ceci implique que N et $\{\pm 1\}$ sont en somme directe, et comme $|N| \cdot |\{\pm 1\}| = |(\mathbf{Z}/2^n \mathbf{Z})^*|$, cela prouve que $(\mathbf{Z}/2^n \mathbf{Z})^* = N \oplus \{\pm 1\}$, ce qui permet de conclure puisque $N \cong \mathbf{Z}/2^{n-2} \mathbf{Z}$ et $\{\pm 1\} \cong \mathbf{Z}/2\mathbf{Z}$.

Exercice 4.7. On obtient le 5-cycle $(1, 2, 3, 4, 5)$.

Exercice 4.8. La démonstration se fait par récurrence sur n . Le résultat est trivial si $n = 2$. Soit $n \geq 3$, et soient $\sigma \in S_n$, et $a = \sigma(n)$. Si $a \neq n$, alors $\sigma' = (n-1, n) \cdots (a, a+1) \sigma$ fixe n , et est dans le sous-groupe engendré par $(1, 2), (2, 3), \dots, (n-2, n-1)$ d'après l'hypothèse de récurrence. Donc $\sigma = (a, a+1) \cdots (n-1, n) \sigma'$ est dans le sous-groupe engendré par $(1, 2), (2, 3), \dots, (n-1, n)$. Si $a = n$, alors σ est déjà dans le sous-groupe engendré par $(1, 2), (2, 3), \dots, (n-2, n-1)$, ce qui prouve que le sous-groupe engendré par $(1, 2), (2, 3), \dots, (n-1, n)$ est S_n .

Exercice 4.9. Comme les τ_i commutent deux à deux, on a $\sigma^n = \tau_1^n \cdots \tau_s^n$, et comme les τ_i^n sont à supports disjoints, on a $\sigma^n = 1$ si et seulement si $\tau_i^n = 1$ pour tout i . On en déduit que l'ordre de σ est le ppcm des ordres des τ_i , et comme τ_i est d'ordre ℓ_i , l'ordre de σ est le ppcm des ℓ_i .

Exercice 4.10 (i) Choisir un cycle de longueur k revient à choisir les k éléments (n choix pour le premier, \dots , $n-k+1$ pour le dernier), en tenant compte du fait que les k permutations circulaires des éléments donnent le même cycle; il y a donc $\frac{1}{k}(n(n-1) \cdots (n-k+1))$ cycles de longueur k .

(ii) Soit $\tau = (i_1, \dots, i_k)$ un cycle de longueur k . Alors τ apparaît dans la décomposition de σ si et seulement si la restriction de σ à $\{i_1, \dots, i_k\}$ est τ , et σ peut permuter les autres éléments comme il veut, et donc τ apparaît dans la décomposition de $(n-k)!$ permutations.

Maintenant, le nombre total de cycles apparaissant dans les permutations de S_n est aussi la somme pour chaque cycle du nombre de permutations dans lequel il apparaît. Ce nombre total est donc, d'après ce qui précède, égal à $\sum_{k=1}^n \frac{1}{k}(n(n-1) \cdots (n-k+1)) \cdot (n-k)! = n!(1 + \frac{1}{2} + \cdots + \frac{1}{n})$, et le nombre moyen de cycles est $1 + \frac{1}{2} + \cdots + \frac{1}{n}$ qui tend bien vers $+\infty$.

Exercice 4.12 Si $\tau_1 \dots \tau_r$ est la décomposition de σ en cycles (en incluant les cycles de longueur 1), et si τ_i est de longueur ℓ_i , alors $\omega(\sigma) = r$, $\sum_{i=1}^r \ell_i = n$ et

$$\text{sign}(\sigma) = \prod_{i=1}^r \text{sign}(\tau_i) = \prod_{i=1}^r (-1)^{\ell_i-1} = (-1)^{n-r} = (-1)^{n-\omega(\sigma)}.$$

Exercice 4.13 (i) On a $u_{\sigma\tau}(e_i) = e_{\sigma\tau(i)} = e_{\sigma(\tau(i))} = u_\sigma(e_{\tau(i)}) = u_\sigma(u_\tau(e_i))$, ce qui prouve que les endomorphismes $u_{\sigma\tau}$ et $u_\sigma u_\tau$ coïncident sur la base canonique, et donc sont égaux. De plus, l'image de la base canonique est une base (vu que c'est la base canonique à l'ordre près); u_σ est donc élément de $\text{GL}(\mathbf{C}^n)$ et $\sigma \mapsto u_\sigma$ est un morphisme de groupes de S_n dans $\text{GL}(\mathbf{C}^n)$.

(ii) Si τ est la transposition (i, j) , alors u_τ est la symétrie par rapport à l'hyperplan engendré par $\frac{e_i+e_j}{2}$ et les e_ℓ , pour $\ell \notin \{i, j\}$, de direction la droite engendrée par $\frac{e_i-e_j}{2}$. Ceci implique que u_τ a $n-1$ valeurs propres égales à 1 et une égale à -1 et donc $\det u_\tau = -1$.

(iii) Comme $\det : \text{GL}(\mathbf{C}^n) \rightarrow \mathbf{C}^*$ est un morphisme de groupes, l'application $\sigma \mapsto \det u_\sigma$ est un morphisme de groupes. Par ailleurs, il ressort du (ii) que l'on a $\det u_\sigma = -1 = \text{sign}(\sigma)$, si σ est une transposition, et comme les transpositions engendrent S_n , cela implique que les deux morphismes de groupes $\sigma \mapsto \det u_\sigma$ et $\sigma \mapsto \text{sign}(\sigma)$ coïncident sur S_n .

Exercice 4.14. (i) D'après le théorème de structure, G est isomorphe à une somme directe $\oplus_{i \in I} (\mathbf{Z}/p_i^{a_i}\mathbf{Z})$, où les p_i sont des nombres premiers (pas forcément distincts). On a alors $|G| = \prod_{i \in I} p_i^{a_i}$, et si d divise $|G|$, on peut trouver des entiers b_i , avec $b_i \leq a_i$, tels que $d = \prod_{i \in I} p_i^{b_i}$. Comme $\mathbf{Z}/p_i^{a_i}\mathbf{Z}$ est cyclique, et comme $p^{b_i} | p^{a_i}$, le groupe $\mathbf{Z}/p_i^{a_i}\mathbf{Z}$ contient un sous-groupe H_i d'ordre p^{b_i} , et $\oplus_{i \in I} H_i$ est un sous-groupe de G de cardinal d .

(ii) Comme $|A_5| = 60 > 6 = |S_3|$, la restriction de f à A_5 n'est pas injective, et comme A_5 est simple, cela implique que $f(A_5) = \{\text{id}\}$, et donc que f se factorise à travers S_5/A_5 . Comme le cardinal de S_5/A_5 est 2, l'image de f a 1 ou 2 éléments.

(iii) Soit H un sous-groupe de S_5 d'ordre 40, et soit $X = S_5/H$. Alors $|X| = |S_5|/|H| = 3$. Par ailleurs, S_5 agit sur X par translation à droite, et permute les éléments de X . On en déduit l'existence d'un morphisme de groupes de S_5 dans $\text{Perm}(X) \cong S_3$ dont l'image a au moins 3 éléments. Ceci étant en contradiction avec le (ii), cela prouve que H n'existe pas.

Exercice 5.1 C'est un cas particulier de la prop. 5.10, mais on peut en donner une démonstration plus directe. Soit $d = \deg Q$. Alors $K[T]/Q$ est un K -espace vectoriel de dimension d (de base $(1, \dots, X^{d-1})$), et si $P \in K[T]$ n'est pas divisible par Q , la multiplication par P est injective sur $K[T]/Q$ (si R est dans le noyau, alors PR est divisible par Q , et comme Q est irréductible et P est premier à Q , cela implique que R est divisible par Q , et donc est nul dans $K[T]/Q$), et donc est surjective, ce qui prouve que tout élément non nul de $K[T]/Q$ a un inverse.

Exercice 6.2. La vérification de ce que d est une distance ne pose pas de problème, et comme les singletons sont ouverts puisque $\{x\} = B(x, (1/2)^-)$, la topologie associée est la topologie discrète.

Exercice 6.3. Si $d'(x, y) = 0$, on a $f(x) = f(y)$ et donc $x = y$ car f est injective (strictement croissante). La symétrie est évidente et l'inégalité triangulaire suit de ce que $d'(x, z) = |f(x) - f(z)| \leq |f(x) - f(y)| + |f(y) - f(z)| = d'(x, y) + d'(y, z)$. Il reste à prouver que si $x \in \mathbf{R}$ et si $\varepsilon > 0$, il existe $\delta > 0$ tel que $d(x, y) < \delta$ implique $d'(x, y) < \varepsilon$ et $d'(x, y) < \delta$ implique $d(x, y) < \varepsilon$, ce qui suit de la continuité de f et de sa réciproque $g(x) = \frac{x}{1-|x|}$, si $x \in]-1, 1[$.

Exercice 6.5 C'est la topologie grossière : si $x \in \mathbf{R}$ et si U est un ouvert non vide de \mathbf{R} , alors U contient un élément de la forme $x + r$, avec $r \in \mathbf{Q}$, et donc tout ouvert non vide de \mathbf{R}/\mathbf{Q} contient l'image de x , pour tout x , et donc est égal à \mathbf{R}/\mathbf{Q} .

Exercice 6.6. Soient $a \neq b$ deux points de X . Comme f est injective, on a $f(a) \neq f(b)$, et comme Y est séparé, on peut trouver des ouverts disjoints U et V de Y tels que $f(a) \in U$ et $f(b) \in V$. Maintenant, comme f est continue, $f^{-1}(U)$ et $f^{-1}(V)$ sont des ouverts de X , qui sont disjoints car U et V le sont, et qui contiennent respectivement a et b . Ceci permet de conclure.

Exercice 6.7. Il suffit de passer aux complémentaires.

Exercice 6.8. (i) Soit $U \neq \emptyset$ un ouvert de $X_1 \times X_2$. Il existe alors $U_1 \neq \emptyset$ ouvert de X_1 et $U_2 \neq \emptyset$ ouvert de X_2 tels que U contienne $U_1 \times U_2$. Comme Y_1 est dense, $Y_1 \cap U_1$ est non vide et comme Y_2 est dense, il en est de même de $Y_2 \cap U_2$, ce qui montre que $(Y_1 \times Y_2) \cap U$ qui contient $(Y_1 \times Y_2) \cap (U_1 \times U_2) = (Y_1 \cap U_1) \times (Y_2 \cap U_2)$ est non vide. On en déduit la densité de $Y_1 \times Y_2$.

(ii) Soient $g : Y \times Y \rightarrow \mathbf{R}_+$ définie par $g(x, x') = d_Y(x, x')$ et $h : Y \times Y \rightarrow \mathbf{R}_+$ définie par $g(x, x') = d_Z(f(x'), f(x'))$. On cherche à prouver que g et h sont égales. Or elles sont égales sur $X \times X$ par hypothèse, et comme $X \times X$ est dense dans $Y \times Y$, et Z est séparé car métrique, on peut en conclure qu'elles sont égales sur $Y \times Y$, en utilisant le point ci-dessus (ou l'ex. 6.11).

Exercice 6.9. (i) Comme \bar{U} contient U , son intérieur, qui est le plus grand ouvert contenu dans \bar{U} contient U . Si U est l'ouvert $]0, 1[\cup]1, 2[$ de \mathbf{R} , alors $\bar{U} = [0, 2]$ et l'intérieur de \bar{U} est $]0, 2[$ qui contient strictement U . Revenons au cas d'un ouvert général U et notons V l'intérieur de son adhérence. Comme $U \subset V$, on a $\bar{U} \subset \bar{V}$, et comme \bar{U} est un fermé qui contient V , on a $\bar{V} \subset \bar{U}$, et donc $\bar{V} = \bar{U}$, ce qui termine la démonstration du (i).

Le (ii) se déduit du (i) en passant aux complémentaires.

Exercice 6.10 Si A n'est pas dense, son adhérence n'est pas \mathbf{C}^2 , et il existe un polynôme $P \in \mathbf{C}[X, Y]$ non nul s'annulant sur A . Soit donc $P \in \mathbf{C}[X, Y]$ tel que $P(n, e^n) = 0$ pour tout $n \in \mathbf{N}$. On écrit P sous la forme $P(X, Y) = P_d(X)Y^d + \dots + P_0(X)$, avec $P_0, \dots, P_d \in \mathbf{C}[X]$. On a donc $P_d(n)e^{dn} + \dots + P_0(n) = 0$ pour tout n , et en divisant par e^{dn} , on en déduit que $P_d(n) \rightarrow 0$ quand $n \rightarrow +\infty$. Ceci n'est possible que si $P_d = 0$. On en déduit que $P = 0$; d'où la densité de A dans \mathbf{C}^2 .

A n'est pas dense dans \mathbf{C}^2 pour la topologie usuelle car A ne contient aucun point de l'ouvert $\{z = (z_1, z_2), \sup(|z_1|, |z_2|) < 1\}$. En fait, il n'est pas difficile de voir que A est fermé dans \mathbf{C}^2 pour la topologie usuelle.

Exercice 6.11. Si X est métrisable, la topologie peut être définie par une métrique d , ce qui permet de supposer que (X, d) est métrique dans tout ce qui suit.

(i) Soit $a \in X$. Comme les $B(a, 2^{-n})$ forment une base de voisinages de a , on voit que si $a \in \bar{Z}$, alors, pour tout $n \in \mathbf{N}$, il existe $x_n \in Z$ avec $d(a, x_n) \leq 2^{-n}$; la suite $(x_n)_{n \in \mathbf{N}}$ a alors a comme limite. Réciproquement, si $(x_n)_{n \in \mathbf{N}}$ est une suite d'éléments de Z ayant a pour limite, et si U est un voisinage de a , alors $x_n \in U$, pour tout n assez grand, ce qui prouve que U contient des éléments de Z , et permet de montrer que $a \in \bar{Z}$ (noter que ce sens n'a pas utilisé le fait que X est métrique).

(ii) Z est dense dans X si et seulement si $\bar{Z} = X$, et donc le résultat suit du (i).

(iii) Si $x \in X$, il existe une suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de Z tendant vers x . Mais alors $f(x_n)$ tend vers $f(x)$ et $g(x_n)$ tend vers $g(x)$ puisque f et g sont continues, et comme $f(x_n) = g(x_n)$ pour tout n , cela implique que $f(x)$ et $g(x)$ sont des limites de la suite $(f(x_n))_{n \in \mathbf{N}}$. Comme Y est supposé métrique et donc séparé, il y a unicité de la limite d'une suite et donc $f(x) = g(x)$.

Exercice 7.1. (i) Soit X un compact métrique, et soient $(x_n)_{n \in \mathbf{N}}$ ayant une unique valeur d'adhérence a , et U un ouvert contenant a . Alors $X - U$ ne contient qu'un nombre fini de termes de la suite, sinon on pourrait extraire une sous-suite $(x_{\varphi(n)})_{n \in \mathbf{N}}$ de $(x_n)_{n \in \mathbf{N}}$, dont tous les termes sont dans $X - U$, et comme $X - U$ est compact puisque fermé dans un compact, cela implique que $(x_{\varphi(n)})_{n \in \mathbf{N}}$ et donc aussi $(x_n)_{n \in \mathbf{N}}$, a une valeur d'adhérence dans $X - U$, contrairement à l'hypothèse. Il existe donc $N \in \mathbf{N}$ tel que $x_n \in U$, si $n \geq N$, ce qui prouve que a est la limite de la suite $(x_n)_{n \in \mathbf{N}}$.

(ii) La suite $(1 + (-1)^n)n$ admet 0 comme unique valeur d'adhérence dans \mathbf{R} , mais ne converge pas.

Exercice 7.2. Il suffit d'adapter la démonstration de la compacité de \mathbf{Z}_p (alinéa 3.4.2) en partant du développement décimal des éléments de $[0, 1]$.

Exercice 7.3. (i) L'ensemble A des rationnels appartenant à $[0, 1]$ est dénombrable; soit $n \mapsto r_n$ une bijection de \mathbf{N} sur A . Comme A est dense dans $[0, 1]$, l'adhérence de tout ouvert contenant A contient $[0, 1]$; il suffit donc de prendre $]a_n, b_n[=]r_n - \frac{\varepsilon}{2^{n+3}}, r_n + \frac{\varepsilon}{2^{n+3}}[$.

(ii) Comme $[0, 1]$ est compact, si les $]a_n, b_n[$, pour $n \in \mathbf{N}$, recouvrent $[0, 1]$, on peut en extraire un recouvrement fini, et le résultat suit du cas d'une famille finie. (Pour démontrer le résultat dans le cas d'une famille finie, on peut remarquer que $\sum_{n \in J} (b_n - a_n)$ est l'intégrale (de Riemann) de la fonction continue par morceaux $\phi = \sum_{n \in J} \mathbf{1}_{]a_n, b_n[}$. Or l'hypothèse $[0, 1] \subset \cup_{n \in J}]a_n, b_n[$ se traduit par $\phi(x) \geq 1$, si $x \in [0, 1]$, et donc l'intégrale de ϕ est supérieure ou égale à celle de $\mathbf{1}_{[0, 1]}$ qui vaut 1. L'exercice permet de montrer que l'intégrale de Lebesgue de $\mathbf{1}_{[0, 1]}$ est supérieure ou égale à 1 et donc aussi égale à 1, ce qui est rassurant...)

Exercice 7.4. Si f est identiquement nulle, il n'y a rien à démontrer. Sinon, il existe $x_0 \in E$ tel que $|f(x_0)| > 0$, et comme f tend vers 0 à l'infini, il existe $M > 0$, tel que $|f(x)| < \frac{|f(x_0)|}{2}$, si $\|x\| > M$. Mais alors la boule $B(0, M)$ contient x_0 et est compacte, puisque E est de dimension finie. Cela implique que $|f|$ atteint son maximum sur cette boule en un point x_1 , et on a $|f(x_1)| \geq |f(x_0)| > \frac{|f(x_0)|}{2}$, ce qui prouve que $|f(x_1)|$ est aussi le maximum de $|f|$ sur E tout entier. Ceci permet de conclure.

Exercice 7.5. (i) Si $x_1, x_2 \in X$, on a $d(x_1, y) \leq d(x_1, x_2) + d(x_2, y)$ pour tout $y \in F$. En passant à la borne inférieure sur $y \in F$, on en déduit que $d(x_1, F) \leq d(x_1, x_2) + d(x_2, F)$. Par symétrie, on a $d(x_2, F) \leq d(x_1, x_2) + d(x_1, F)$. On en déduit que $|d(x_1, F) - d(x_2, F)| \leq d(x_1, x_2)$, et donc que $d(x, F)$ est 1-lipschitzienne.

(ii) On a $d(x, F) = 0$, si $x \in F$, et donc, par continuité, $d(x, F) = 0$, si $x \in \overline{F}$. Réciproquement, si $x \in \overline{F}$, alors pour tout $n > 0$, il existe $x_n \in F$ avec $d(x, x_n) < 2^{-n}$, ce qui implique que $d(x, F) < 2^{-n}$, pour tout n , et donc $d(x, F) = 0$.

(iii) La fonction $f(x) = d(x, F_1) - d(x, F_2)$ est continue sur X , et donc $U_1 = f^{-1}(\mathbf{R}_+^*)$ et $U_2 = f^{-1}(\mathbf{R}_-^*)$ sont deux ouverts de X (en tant qu'images inverses d'ouverts de \mathbf{R} par une fonction continue) qui sont disjoints puisque \mathbf{R}_+^* et \mathbf{R}_-^* sont disjoints. Maintenant, si $x \in F_1$, alors $d(x, F_2) > 0$ puisque F_2 est fermé et $x \notin F_2$; donc $f(x) > 0$. On en déduit que $F_1 \subset U_1$. De même, $F_2 \subset U_2$, ce qui permet de conclure.

(iv) La fonction $(x, y) \mapsto d(x, y)$ est continue sur $X \times X$. Comme $F_1 \times F_2$ est compact comme produit de deux compacts, le minimum de $d(x, y)$ sur $F_1 \times F_2$ est atteint en (x_0, y_0) , et comme $F_1 \cap F_2 = \emptyset$, on a $d(x_0, y_0) \neq 0$, et donc $d(F_1, F_2) > 0$.

(v) La fonction $x \mapsto d(x, F_1)$ est continue sur F_2 et ne s'annule pas car $F_1 \cap F_2 = \emptyset$ et F_1 est fermé. Comme F_2 est compact, elle atteint son minimum qui, de ce fait est > 0 . Or ce minimum est $\inf_{x \in F_2} d(x, F_1) = \inf_{x \in F_2} \inf_{y \in F_1} d(x, y) = d(F_1, F_2)$, ce qui permet de conclure.

(vi) Dans \mathbf{R} , on peut prendre $F_1 = \mathbf{N}$ et $F_2 = \{n + 2^{-n-1}, n \in \mathbf{N}\}$. Dans \mathbf{R}^2 , on peut prendre $F_1 = \{(x, y), xy = 1\}$ et $F_2 = \{(x, y), xy = 0\}$.

Exercice 7.6. (i) Soit $g : X \rightarrow \mathbf{R}$ définie par $g(x) = d(x, f(x))$. Alors g est continue comme composée de $g_1 : X \rightarrow X \times X$ envoyant x sur $(x, f(x))$ et $g_2 : X \times X \rightarrow \mathbf{R}$ envoyant (x, y) sur $d(x, y)$. Elle atteint donc son minimum en un point x_0 , et on a $f(x_0) = x_0$, sinon $d(f(f(x_0)), f(x_0)) < d(f(x_0), x_0)$, ce qui est contraire à la définition de x_0 . La fonction f admet donc au moins un point fixe. Si elle en admet deux $x_1 \neq x_2$, on a $d(f(x_1), f(x_2)) < d(x_1, x_2)$, ce qui est contraire à l'hypothèse $f(x_1) = x_1$ et $f(x_2) = x_2$. Le point fixe de f est donc unique, ce qui permet de conclure.

(ii) Soit $\delta_n = d(f^n(x), x_0)$. Comme f est strictement contractante, $\delta_{n+1} = d(f(f^n(x)), f(x_0)) < \delta_n$, si $f^n(x) \neq x_0$. Maintenant, soit a une valeur d'adhérence de la suite $(f^n(x))_{n \in \mathbf{N}}$, et soit $f^{\varphi(n)}(x)$ une suite

extraite tendant vers a . Si $a \neq x_0$, on a

$$\delta_{\varphi(n_0)+1} \leq d(f^{\varphi(n_0)+1}(x), f(a)) + d(f(a), x_0) < d(f^{\varphi(n_0)}(x), a) + d(a, x_0) \leq d(a, x_0),$$

si n_0 est assez grand. On aboutit à une contradiction car $\delta_m \leq \delta_{\varphi(n_0)} < d(a, x_0)$ pour tout $m > \varphi(n_0)$, et la suite extraite $\delta_{\varphi(n)}$ tend vers $d(a, x_0)$ quand n tend vers $+\infty$. On en déduit que $a = x_0$ et donc que $f^n(x)$ a x_0 comme une unique valeur d'adhérence dans X . Comme X est compact, cela implique que $f^n(x) \rightarrow x_0$.

(iii) Soit $\delta_n = \sup_{x \in X} d(f^n(x), f(x_0))$. Il s'agit de prouver que $\delta_n \rightarrow 0$. Comme f est compact et $x \mapsto d(f^n(x), f(x_0))$ est continue puisque f est continue, il existe $x_n \in X$ tel que $d(f^n(x_n), x_0) = \delta_n$. On a alors $\delta_{n+1} = d(f^{n+1}(x), x_0) = d(f^n(f(x_{n+1})), x_0) \leq \delta_n$, ce qui montre que la suite δ_n est décroissante. Il suffit donc d'exhiber une suite extraite de $(\delta_n)_{n \in \mathbf{N}}$ tendant vers 0.

Soit a une valeur d'adhérence de la suite x_n , et soit $x_{\varphi(n)}$ une suite extraite tendant vers a . On a alors $\delta_{\varphi(n)} = d(f^{\varphi(n)}(x_{\varphi(n)}), x_0) \leq d(f^{\varphi(n)}(x_{\varphi(n)}), f^{\varphi(n)}(a)) + d(f^{\varphi(n)}(a), x_0) \leq d(x_{\varphi(n)}, a) + d(f^{\varphi(n)}(a), x_0)$, ce qui montre que $\delta_{\varphi(n)} \rightarrow 0$ car $d(x_{\varphi(n)}, a) \rightarrow 0$ par construction, et $d(f^{\varphi(n)}(a), x_0) \rightarrow 0$ d'après le (ii). Ceci permet de conclure.

Exercice 7.7. La démonstration se fait par l'absurde. Supposons X non compact, et construisons une fonction continue $\phi : X \rightarrow \mathbf{R}$ non bornée. Il existe une suite $(x_n)_{n \in \mathbf{N}}$ n'ayant pas de valeur d'adhérence dans X , ce qui se traduit, pour tout $a \in X$, par l'existence de $\delta_a > 0$ tel que $B(a, 2\delta_a^-)$ contienne au plus un x_n , à savoir a si l'un des x_n vaut a . Soit $\phi_n(x) = \sup(n - n^2 d(x, x_n), 0)$. C'est une fonction continue sur X , nulle en dehors de $B(x_n, \frac{1}{n})$ et valant n en x_n . Si $a \in X$, la restriction de ϕ_n à $B(a, \delta_a^-)$ est identiquement nulle, si $\frac{1}{n} < \delta_a$ et si $x_n \neq a$. Comme il n'y a qu'un nombre fini de n ne vérifiant pas ces conditions, cela montre que $\phi(x) = \sum_{n \in \mathbf{N}} \phi_n(x)$ est la somme d'un nombre fini de fonctions continues sur $B(a, \delta_a^-)$, pour tout a ; c'est donc une fonction continue sur X . Par ailleurs, on a $\phi(x_n) \geq n$, pour tout n , et donc ϕ est non bornée. Ceci permet de conclure.

Exercice 8.1. On sait que U est connexe par arcs, et il suffit de prouver qu'il en est de même de $V = U - \{x\}$. Soient donc $y_1, y_2 \in V$, et soit $u : [0, 1] \rightarrow U$ un chemin continu joignant y_1 à y_2 dans U . Si u ne passe pas par x , il n'y a rien à faire. Sinon, il existe $r < \inf(d(x, y_1), d(x, y_2))$ tel que $B(x, r) \subset U$, et l'ensemble des t tels que $d(x, u(t)) \leq r$ admet un plus petit (resp. grand) élément t_1 (resp. t_2). Alors u permet de joindre y_1 à $u(t_1)$ et $u(t_2)$ à y_2 dans V , et on peut passer de $u(t_1)$ à $u(t_2)$ en restant sur la sphère de rayon r (il suffit de prendre l'arc de cercle délimité par le cône de sommet x et dont les bords sont les demi-droites $[x, u(t_1))$ et $[x, u(t_2))$).

Exercice 8.2. Si f est un homéomorphisme de X sur Y , alors la restriction de f à $X - \{x\}$ est encore un homéomorphisme de $X - \{x\}$ sur $Y - \{f(x)\}$ pour tout $x \in X$. Il ne peut donc pas y avoir d'homéomorphisme de \mathbf{R} sur \mathbf{R}^2 puisque \mathbf{R} privé d'un point est non connexe, alors que \mathbf{R}^2 privé d'un point est connexe. Les autres cas se traitent de la même manière en enlevant à $[0, 1]$ n'importe quel élément différent de 0 et 1.

Exercice 8.3. Si f est une bijection de $[0, 1]$ sur $]0, 1[$, alors $f([0, 1]) =]0, 1[- \{f(0)\}$ est non connexe, tandis que $]0, 1]$ est connexe, ce qui prouve que f ne peut pas être continue.

Exercice 8.4. Si on enlève de Y les deux points de contacts, on obtient un ensemble avec 4 composantes connexes, alors que si on enlève deux points à X , le mieux que l'on puisse obtenir est 3 composantes connexes.

Exercice 8.5. Définissons le bord du cylindre et de la bande de Moebius comme l'image de $\{0, 1\} \times [0, 1]$. Dans le cylindre, on obtient deux lacets disjoints, alors que dans la bande de Moebius on n'obtient qu'un seul lacet, car $(0, 0)$ est identifié à $(1, 1)$. Maintenant, si x est sur le bord, alors x admet une base de voisinages constituée de demi-disques de centre x , et si on prive un de ces demi-disques de x , on obtient

un ensemble contractile. Si x n'est pas sur le bord, alors tout voisinage de x contient un disque de centre x , et si on le prive de x , on obtient un ensemble non contractile. On en déduit qu'un homéomorphisme du cylindre sur la bande de Moebius induit un homéomorphisme entre les bords, mais ce n'est pas possible car le bord du cylindre n'est pas connexe, alors que celui de la bande de Moebius l'est.

Exercice 9.1. (i) Si d est ultramétrique, $d(x_m, x_{m+p}) \leq \sup_{0 \leq i \leq p-1} d(x_{m+i}, x_{m+i+1}) \leq \sup_{m \geq n} d(x_m, x_{m+1})$. On en déduit que si $d(x_{n+1}, x_n) \rightarrow 0$, et donc si $\lim_{n \rightarrow +\infty} (\sup_{m \geq n} d(x_{m+1}, x_m)) = 0$, alors $\lim_{m \rightarrow +\infty} (\sup_{p \in \mathbf{N}} d(x_{m+p}, x_m)) = 0$, et la suite est de Cauchy.

(ii) Si $n \geq 1$, soit $i = \left\lfloor \frac{\log n}{\log 2} \right\rfloor$, de telle sorte que $n = 2^i + j$, avec $0 \leq j \leq 2^i - 1$. Posons alors $x_n = \frac{j}{2^i}$, si i est pair et $x_n = 1 - \frac{j}{2^i}$, si i est impair. On vérifie que $x_{n+1} - x_n = \frac{1}{2^i}$ tend vers 0, mais que la suite $(x_n)_{n \in \mathbf{N}}$ balaie consciencieusement l'intervalle $[0, 1]$ et que l'ensemble de ses valeurs d'adhérence est $[0, 1]$. Elle n'est donc pas de Cauchy. On aurait aussi pu prendre $x_n = \log(n+1)$ qui tend vers $+\infty$, et donc n'est pas de Cauchy.

Exercice 9.3. Soit $(U_n)_{n \in \mathbf{N}}$ une famille d'ouverts denses de \mathbf{R} . Supposons que $X = \bigcap_{n \in \mathbf{N}} U_n$ est dénombrable, et choisissons une surjection $n \mapsto x_n$ de \mathbf{N} sur X . Alors $V_n = U_n - \{x_n\}$ est un ouvert dense de \mathbf{R} pour tout n et $\bigcap_{n \in \mathbf{N}} V_n = \emptyset$, ce qui est contraire au lemme de Baire.

Exercice 10.1 En revenant à la définition de la topologie produit, on voit qu'il suffit de prouver qu'on peut toujours construire une fonction continue de \mathbf{R} dans \mathbf{C} prenant des valeurs prescrites en un nombre fini de points. Ceci ne pose pas de problème (on peut par exemple prendre un polynôme d'interpolation de Lagrange).

Exercice 10.2. On peut prolonger $u^{(n)}$ en une fonction continue sur $\overline{\mathbf{N}}$ en posant $u^{(n)}(+\infty) = 0$. On prolonge aussi u en posant $u(+\infty) = 0$. Alors $u^{(n)} \rightarrow u$ uniformément sur $\overline{\mathbf{N}}$ et donc u est continue en $+\infty$, ce qui se traduit par $\lim_{k \rightarrow +\infty} u_k = 0$.

On peut aussi se passer de $\overline{\mathbf{N}}$, en recopiant la démonstration du point précédent l'exercice. Soit $\varepsilon > 0$. Comme $u^{(n)} \rightarrow u$ uniformément sur \mathbf{N} , il existe $N_0 \in \mathbf{N}$ tel que $|u_k^{(n)} - u_k| < \varepsilon$, quels que soient $n \geq N_0$ et $k \in \mathbf{N}$. Choisissons $n \geq N_0$. Comme $\lim_{k \rightarrow +\infty} u_k^{(n)} = 0$, il existe $N \in \mathbf{N}$ tel que $|u_k^{(n)}| < \varepsilon$, pour tout $k \geq N$, et on a $|u_k| \leq |u_k^{(n)} - u_k| + |u_k^{(n)}| < 2\varepsilon$, pour tout $k \geq N$. On en déduit que $\lim_{k \rightarrow +\infty} u_k = 0$.

Exercice 10.3 Comme $f_n \rightarrow f$ uniformément sur E , elle vérifie le critère de Cauchy uniforme, et on a $\lim_{n \rightarrow +\infty} \left(\sup_{x \in E, p \in \mathbf{N}} |f_n(x) - f_{n+p}(x)| \right) = 0$. Or $|\ell_n - \ell_{n+p}| \leq \sup_{x \in E} |f_n(x) - f_{n+p}(x)|$, et donc $\lim_{n \rightarrow +\infty} \left(\sup_{p \in \mathbf{N}} |\ell_n - \ell_{n+p}| \right) = 0$, ce qui prouve que $(\ell_n)_{n \in \mathbf{N}}$ est de Cauchy et comme \mathbf{C} est complet, elle admet une limite ℓ .

Soit maintenant $\varepsilon > 0$. Comme $f_n \rightarrow f$ uniformément sur E , il existe $N_0 \in \mathbf{N}$ tel que l'on ait $|f_n(x) - f(x)| < \varepsilon$, quels que soient $n \geq N_0$ et $x \in E$. Choisissons $n \geq N_0$. En passant à la limite, on en déduit que $|\ell_n - \ell| \leq \varepsilon$. Par ailleurs, il existe $M > 0$ tel que $|f_n(x) - \ell_n| < \varepsilon$, si $\|x\| > M$; on a donc

$$|f(x) - \ell| \leq |f(x) - f_n(x)| + |f_n(x) - \ell_n| + |\ell_n - \ell| < 3\varepsilon,$$

si $\|x\| > M$, ce qui prouve que f tend vers ℓ à l'infini.

Exercice 11.1. (i) Si $\phi \in E$, alors $\|\phi\|_\infty$ est fini car $[0, 1]$ est compact et une fonction continue sur un compact est bornée. Que $\|\cdot\|_\infty$ soit une norme sur E est alors immédiat. Maintenant, une suite $(\phi_n)_{n \in \mathbf{N}}$ est de Cauchy pour $\|\cdot\|_\infty$ si et seulement si elle vérifie le critère de Cauchy uniforme sur $[0, 1]$, et \mathbf{C} étant complet, on sait (alinéa 10.2) que $(\phi_n)_{n \in \mathbf{N}}$ admet une limite simple ϕ qui est continue sur $[0, 1]$, et que $\phi_n \rightarrow \phi$ uniformément sur $[0, 1]$, ce qui signifie exactement que $\phi_n \rightarrow \phi$ pour $\|\cdot\|_\infty$. On en déduit la complétude de $(E, \|\cdot\|_\infty)$.

(ii) Que $\|\cdot\|_1$ soit une norme est immédiat à part peut-être le fait que « $\|\phi\|_1 = 0$ » implique « $\phi = 0$ ». Mais si $\phi \neq 0$, il existe $x_0 \in [0, 1]$ avec $\phi(x_0) \neq 0$, et comme ϕ est continue, il existe un intervalle I de

longueur non nulle ℓ sur lequel $|\phi(x)| \geq |\phi(x_0)/2|$. On a alors $\|\phi\|_1 \geq \ell |\phi(x_0)/2| > 0$. Maintenant, soit $\phi_n = x^{-1/2} \mathbf{1}_{[1/n, 1]}$. La suite $(\phi_n)_{n \geq 1}$ est de Cauchy car

$$\|\phi_{n+p} - \phi_n\|_1 = \int_{1/(n+p)}^{1/n} x^{-1/2} dx = 2 \left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+p}} \right) \leq \frac{2}{\sqrt{n}}.$$

Si cette suite avait une limite ϕ dans E , on aurait $\lim_{n \rightarrow +\infty} \int_0^1 |\phi - \phi_n| = 0$. Or, pour tous $a > 0$ et $n \geq 1/a$, on a $\int_0^1 |\phi - \phi_n| \geq \int_a^1 |\phi - \phi_n| = \int_a^1 |\phi(x) - x^{-1/2}| dx$. On devrait donc avoir $\int_a^1 |\phi(x) - x^{-1/2}| dx = 0$, quel que soit $a > 0$, et ϕ étant continue, cela implique que $\phi(x) = x^{-1/2}$, pour tout $x > a$ et tout $a > 0$, et donc que $\phi(x) = x^{-1/2}$ si $x \in]0, 1]$. Ceci n'est pas possible car cette fonction n'est pas la restriction à $]0, 1]$ d'une fonction continue sur $[0, 1]$. En résumé $(\phi_n)_{n \in \mathbf{N}}$ n'a pas de limite dans E , et E n'est pas complet pour $\|\cdot\|_1$.

(iii) Si les normes étaient équivalentes, les suites de Cauchy seraient les mêmes dans les deux cas, et donc E serait simultanément complet ou non pour les deux normes, ce qui n'est pas le cas. On peut aussi remarquer que $\|\phi_n\|_1 \leq 2$ pour tout n , alors que $\|\phi_n\|_\infty \rightarrow +\infty$

Exercice 11.2. (i) $\text{id} : (X, \mathcal{T}_1) \rightarrow (X, \mathcal{T}_2)$ est continue si et seulement si l'image réciproque de tout ouvert de (X, \mathcal{T}_2) par id est un ouvert de (X, \mathcal{T}_1) , et donc si et seulement si tout élément de \mathcal{T}_2 est élément de \mathcal{T}_1 .

(ii) Si $\phi_n(x) = \mathbf{1}_{[1/n, 1]}(x) \cdot x^{-1}$, alors $\|\phi_n\|_\infty = 1$, alors que $\|\phi_n\|_1 = \log n$ tend vers $+\infty$, ce qui prouve que $\text{id} : (\mathcal{C}(\mathbf{R}), \|\cdot\|_\infty) \rightarrow (\mathcal{C}(\mathbf{R}), \|\cdot\|_1)$ n'est pas continue et donc que \mathcal{T}_∞ n'est pas plus fine que \mathcal{T}_1 .

De même, si $\phi_n = \mathbf{1}_{[1/n, 1]}(x) \cdot x^{-1/2}$, alors $\|\phi_n\|_1 = \frac{1}{2}(1 - 1/\sqrt{n}) \leq \frac{1}{2}$, alors que $\|\phi_n\|_\infty = \sqrt{n}$ tend vers $+\infty$, ce qui prouve que $\text{id} : (\mathcal{C}(\mathbf{R}), \|\cdot\|_1) \rightarrow (\mathcal{C}(\mathbf{R}), \|\cdot\|_\infty)$ n'est pas continue et donc que \mathcal{T}_1 n'est pas plus fine que \mathcal{T}_∞ .

Index du chapitre

- adhérence, 55
- anneau
 - intègre, 44
 - noethérien, 45
 - principal, 44
- boule
 - fermée, 50
 - ouverte, 50
- catégorie, 12
- Cauchy
 - critère, 68
 - critère uniforme, 74
 - suite, 68
- centralisateur, 21
- centre, 22
- classe
 - d'équivalence, 13
 - de conjugaison, 21
 - formule des, 32
- compacité, 57
- complétion, 71
- complétude, 69
- congruence, 15
- conjugaison, 21
- connexité, 65
 - composante connexe, 66
 - composantes connexes par arcs, 67
 - par arcs, 67
- continuité, 51
 - uniforme, 52
- convergence
 - simple, 73
 - uniforme, 73
- coupures de Dedekind, 25
- cycle, 33
- dénombrable, 9
- densité, 55
- distance, 50
 - équivalence, 50
 - p -adique, 26
 - triviale, 51
 - ultramétrique, 50
- domaine fondamental, 21
- endomorphisme, 38
- diagonalisable, 39
- trace, 39
- équivalence
 - classe, 13
 - quotient par une relation, 14
 - relation, 13
- espace
 - caractéristique, 40
 - métrique, 50
 - métrisable, 50
 - propre, 39
 - topologique, 48
- fermé, 48
 - de Zariski, 49
- fonction
 - continue, 51
 - φ indicatrice d'Euler, 15
 - lipschitzienne, 52
 - uniformément continue, 52
- groupe
 - abélien, 29
 - alterné, 36
 - cyclique, 30
 - distingué, 22
 - orthogonal, 21
 - p -groupe, 37
 - p -Sylow, 37
 - simple, 23
 - sous-groupe, 30
 - symétrique, 33
 - symplectique, 21
 - unitaire, 21
- homéomorphisme, 51
- homothétie, 38
- idéal, 18
 - maximal, 44
 - premier, 44
 - principal, 44
- inégalité
 - triangulaire, 50
 - ultramétrique, 26
- intérieur, 55
- Jordan

- bloc, 40
- forme, 40
- limite
 - inférieure, 63
 - simple, 73
 - supérieure, 63
 - uniforme, 73
- module, 40
 - de torsion, 41
 - de type fini, 41
 - engendré, 41
- nombre
 - algébrique, 10
 - complexe, 26
 - entier, 23
 - p -adique, 26
 - premier, 7
 - réel, 25
 - rationnel, 24
 - transcendant, 10
- norme
 - équivalence, 76
 - espace vectoriel, 75
 - opérateur, 75
 - p -adique, 26
- orbite, 20
- ordre
 - d'un élément, 31
 - d'un groupe, 37
- ouvert, 48
 - base, 49
- partition
 - d'un ensemble, 13
 - d'un entier, 34
- permutation, 20, 33
 - signature, 35
 - support, 33
- polynôme
 - caractéristique, 38
 - irréductible, 41
 - minimal, 39
- propre
 - espace, 39
 - valeur, 39
 - vecteur, 39
- propriété universelle, 11, 71
- réduction
 - des endomorphismes, 38
 - modulo D , 15
- somme directe
 - d'espaces vectoriels, 12
 - de groupes, 12
- spectre, 39
- stabilisateur, 21
- suite
 - convergente, 56
 - de Cauchy, 68
 - extraite, 56
- théorème
 - Bézout, 7, 46
 - Borel-Lebesgue, 59
 - Cayley-Hamilton, 39
 - de Fermat (petit), 32
 - du point fixe, 70
 - fermés emboîtés, 70
 - Lagrange, 32
 - lemme de Baire, 71
 - lemme de Gauss, 7, 45
 - restes chinois, 16
 - Riesz, 77
 - structure des groupes abéliens finis, 31
 - structure des modules de torsion sur un anneau principal, 47
 - Sylow, 37
 - valeurs intermédiaires, 66
 - Wedderburn, 32
 - Weierstrass, 55
- théorème fondamental
 - algèbre, 26
 - analyse, 79
 - arithmétique, 8
- topologie, 48
 - discrète, 49
 - grossière, 49
 - induite, 53
 - produit, 53
 - quotient, 53
 - séparée, 54
 - totalelement discontinue, 66
 - Zariski, 49
- transposition, 34
- valuation p -adique, 8
- voisinage, 49
 - base, 49

CHAPITRE I

REPRÉSENTATIONS DES GROUPE FINIS

Si G est un groupe, une représentation V de G est un espace vectoriel sur un corps K (ou, plus généralement, un module sur un anneau A) muni d'une action linéaire de G (i.e. on demande que $x \mapsto g \cdot x$ soit une application linéaire de V dans V , quel que soit $g \in G$). Il arrive souvent que G et V soient munis de topologies, et on demande alors, en général, que $(g, x) \mapsto g \cdot x$ soit continue de $G \times V$ dans V .

Les représentations de groupes interviennent de multiples façons en mathématique, en physique, ou en chimie. Par exemple, une des motivations initiales de la théorie des représentations des groupes finis, dont il sera question dans ce chapitre, est venue de la cristallographie. La physique des particules utilise grandement les représentations des groupes de Lie⁽¹⁾ comme le groupe $\mathbf{SU}(2)$ des isométries de déterminant 1 de \mathbf{C}^2 (muni du produit scalaire usuel $\langle (x_1, x_2), (y_1, y_2) \rangle = \overline{x_1} y_1 + \overline{x_2} y_2$), ou le groupe d'Heisenberg des matrices 3×3 unipotentes supérieures (i.e. triangulaires supérieures avec des 1 sur la diagonale), à coefficients réels, ou encore ceux de Lorentz (i.e. $\mathbf{O}(1, 3)$) et Poincaré.

Si G est un groupe, la connaissance des représentations de G fournit des tas d'informations sur G , et certains groupes ne sont accessibles qu'à travers leurs représentations. Par exemple, l'existence du *monstre*, le plus grand des groupes finis simples sporadiques⁽²⁾, de cardinal

$$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71,$$

⁽¹⁾Voir le cours "Groupes et représentations" de D. Renard, en troisième année.

⁽²⁾Un groupe G est *simple* si ses seuls sous-groupes distingués sont $\{1\}$ et G . Si un groupe fini n'est pas simple, il possède un sous-groupe distingué non trivial H , et H et G/H sont deux groupes plus petits que G à partir desquels G est construit. En réitérant ce procédé, cela permet de casser n'importe quel groupe fini en une famille finie de groupes simples. La classification des groupes finis se ramène donc à celle des groupes finis simples, et à comprendre comment on peut composer ces groupes simples pour fabriquer des groupes plus gros. La classification des groupes finis simples s'est achevée au début des années 1980 (elle court sur quelques milliers de pages, et personne n'en maîtrise vraiment la totalité...). Il y a un certain nombre de familles infinies comme les $\mathbf{Z}/p\mathbf{Z}$, pour p premier, les groupes alternés A_n , pour $n \geq 5$, les quotients de $\mathbf{SL}_n(\mathbf{F}_q)$ par leur centre (\mathbf{F}_q est le corps à q élément), et quelques autres découvertes par C. Chevalley en 1954. A côté de ces familles, il y a 26 groupes isolés, dit sporadiques.

n'a été démontrée en 1982, par R. Griess, que grâce à la construction d'une de ses représentations⁽³⁾ (de dimension 196883), alors que l'existence du monstre avait été prédite en 1973 par R. Griess et B. Fischer (il y a une ressemblance certaine avec la chasse aux particules élémentaires).

De même, on n'a de prise sur $\mathcal{G}_{\mathbf{Q}}$, groupe des automorphismes du corps $\overline{\mathbf{Q}}$ des nombres algébriques (cf. annexe D) qu'à travers ses représentations. Celles-ci fournissent de précieuses informations sur $\overline{\mathbf{Q}}$, et permettent de résoudre des problèmes classiques de théorie des nombres. Par exemple, la démonstration du théorème de Fermat par A. Wiles (1994) consiste à relier, de manière à en tirer une contradiction, deux types de représentations de $\mathcal{G}_{\mathbf{Q}}$: d'une part celles provenant des solutions (dans $\overline{\mathbf{Q}}$) de l'équation $y^2 = x(x - a^p)(x + b^p)$, où $a^p + b^p = c^p$ est un contreexemple potentiel au théorème de Fermat, et d'autre part, des représentations provenant des formes modulaires.

L'exemple de $\mathcal{G}_{\mathbf{Q}}$ agissant sur $\overline{\mathbf{Q}}$ est en fait assez typique. Si on a un ensemble X sur lequel un groupe G agit, et si on connaît bien les représentations de G , alors on peut espérer en tirer des informations fines sur X .

La théorie des représentations présente deux aspects. Le premier de ces aspects (décomposition d'une représentation en représentations irréductibles) est une généralisation de la théorie de la réduction des endomorphismes (valeurs propres, espaces propres, diagonalisation), qui correspond, modulo un petit exercice de traduction (ex. I.1.1 et I.2.1, rem. I.2.2), au cas du groupe \mathbf{Z} . Le second aspect (théorie des caractères) est une première approche de l'analyse de Fourier dans un cadre non commutatif (ou commutatif, cf. alinéa 5.1 du n° I.2). A part le cas de \mathbf{Z} qui permet de faire le lien avec l'algèbre linéaire classique, nous ne considérerons essentiellement que les représentations complexes des groupes finis dans ce cours. Ce cas présente l'avantage d'être à la fois simple (il n'y a pas à se battre avec les problèmes de convergence ou autres subtilités analytiques que l'on rencontre par exemple dans l'étude des séries de Fourier qui correspondent au groupe

⁽³⁾ C'est la plus petite des représentations non triviales du monstre ; le début de la liste des dimensions des représentations irréductibles du monstre est le suivant :

$$f_1 = 1, f_2 = 196883, f_3 = 21296876, f_4 = 842609326, f_5 = 18538750076, f_6 = 19360062527, \dots$$

J. McKay a remarqué en 1977, que 196883 avait un rapport avec les coefficients de Fourier de la fonction modulaire j de l'ex. VII.6.10 : si on écrit $j(z)$ sous la forme $j(z) = \frac{1}{q} + 744 + \sum_{n \geq 1} c_n q^n$, avec $q = e^{2i\pi z}$, alors $c_1 = f_2 + f_1$, $c_2 = f_3 + f_2 + f_1$, $c_3 = f_4 + f_3 + 2f_2 + 2f_1 \dots$ Vu la taille des nombres en présence, il y avait peu de chance que ceci soit une coïncidence fortuite. Ce mystère, connu sous le nom de "monster's moonshine", a été résolu par R. Borcherds en 1992, en utilisant des objets venant de la physique mathématique, ce qui lui a valu la médaille Fields (1998). L'expression "monster's moonshine" est moins poétique que ce qu'elle suggère, car "moonshine" doit être pris dans le sens de "bêtise, faribole", comme dans la citation suivante de E. Rutherford : « The energy produced by the breaking down of the atom is a very poor kind of thing. Anyone who expects a source of power from the transformations of these atoms is talking moonshine. ».

\mathbf{R}/\mathbf{Z}), et tout à fait représentatif du genre d'énoncés que l'on peut espérer dans d'autres situations.

I.1. Représentations et caractères

1. Représentations de groupes, exemples

Le lecteur est renvoyé au chapitre « Vocabulaire mathématique », n° 2.6.1, § 4 et n° 5.1, pour le vocabulaire et les résultats de base d'algèbre linéaire et de théorie des groupes.

Soit G un groupe de loi de groupe $(g, h) \mapsto gh$. Une *représentation* V de G est un \mathbf{C} -espace vectoriel muni d'une action (à gauche) de G agissant de manière linéaire. Une telle représentation est équivalente à la donnée d'un morphisme de groupes ρ_V de G dans $\mathbf{GL}(V)$: si $g \in G$, l'application $v \mapsto g \cdot v$ est linéaire bijective et donc nous définit un élément $\rho_V(g)$ de $\mathbf{GL}(V)$, et l'identité $g \cdot (h \cdot v) = gh \cdot v$, valable quels que soient $g, h \in G$ et $v \in V$, se traduit par l'identité $\rho_V(gh) = \rho_V(g)\rho_V(h)$. Dans la suite on parlera indifféremment de la représentation V de G ou de la représentation ρ_V de G , suivant qu'on veut mettre l'accent sur l'espace vectoriel de la représentation ou sur le morphisme de G dans $\mathbf{GL}(V)$. On notera aussi parfois $\rho_{V,g}$ l'élément $\rho_V(g)$ de $\mathbf{GL}(V)$, de manière à pouvoir écrire $\rho_{V,g}(v)$ au lieu de $\rho_V(g)(v)$ l'image $g \cdot v$ de $v \in V$ sous l'action de $g \in G$.

Exemple I.1.1. — (Représentations de \mathbf{Z})

(i) Si $\lambda \in \mathbf{C}^*$, alors $n \mapsto \lambda^n$ est un morphisme de groupes de \mathbf{Z} dans \mathbf{C}^* , ce qui nous fabrique une représentation de \mathbf{Z} que nous noterons $\mathbf{C}(\lambda)$; l'action de $n \in \mathbf{Z}$ sur $z \in \mathbf{C}$ étant donnée par $\rho_{\mathbf{C}(\lambda),n}(z) = \lambda^n z$ (ce qu'on peut aussi écrire sous la forme $n \cdot z = \lambda^n z$).

(ii) Si V est un \mathbf{C} -espace vectoriel, et si $u : V \rightarrow V$ est un isomorphisme linéaire, l'application $n \mapsto u^n$ est un morphisme de groupes de \mathbf{Z} dans $\mathbf{GL}(V)$, ce qui fait de V une représentation du groupe additif \mathbf{Z} , l'action de $n \in \mathbf{Z}$ sur $v \in V$ étant donnée par $n \cdot v = u^n(v)$. Réciproquement, si V est une représentation de \mathbf{Z} , alors $u = \rho_V(1) \in \mathbf{GL}(V)$, et on a $\rho_V(n) = u^n$ pour tout $n \in \mathbf{Z}$, et donc $n \cdot v = u^n(v)$, si $n \in \mathbf{Z}$ et $v \in V$. En d'autres termes, une représentation de \mathbf{Z} n'est rien d'autre que la donnée d'un \mathbf{C} -espace vectoriel V et d'un élément u de $\mathbf{GL}(V)$.

Exemple I.1.2. — (Représentations de $\mathbf{Z}/D\mathbf{Z}$)

Si V est un \mathbf{C} -espace vectoriel muni d'un isomorphisme linéaire u vérifiant $u^D = 1$, l'application $n \mapsto u^n$ est un morphisme de groupes de \mathbf{Z} dans $\mathbf{GL}(V)$ dont le noyau contient $D\mathbf{Z}$; il induit donc un morphisme de $\mathbf{Z}/D\mathbf{Z}$ dans $\mathbf{GL}(V)$, ce qui fait de V une représentation de $\mathbf{Z}/D\mathbf{Z}$, l'action de $n \in \mathbf{Z}$ sur $v \in V$ étant donnée par $n \cdot v = u^n(v)$. Réciproquement, si V est une représentation de $\mathbf{Z}/D\mathbf{Z}$, et si $u = \rho_V(1) \in \mathbf{GL}(V)$, alors $u^D = \rho_V(D) = \rho_V(0) = 1$, car $D = 0$ dans $\mathbf{Z}/D\mathbf{Z}$. En d'autres termes, une représentation de $\mathbf{Z}/D\mathbf{Z}$ n'est rien d'autre que la donnée d'un \mathbf{C} -espace vectoriel V et d'un élément u de $\mathbf{GL}(V)$ vérifiant $u^D = 1$.

Remarque I.1.3. — Dans les deux cas ci-dessus, on dispose d'une présentation du groupe à partir de générateurs (dans les deux cas G est engendré par 1) et de relations entre les générateurs (pas de relation dans le cas de \mathbf{Z} , une relation $D = 0$ pour $\mathbf{Z}/D\mathbf{Z}$). Ceci permet de décrire une représentation de G en disant ce que fait chaque générateur, les relations entre les générateurs imposant des relations entre leurs actions. Ce type de description est très efficace quand on dispose d'une présentation relativement simple du groupe G .

Par exemple, le groupe \mathbf{Z}^2 est engendré par $e_1 = (1, 0)$ et $e_2 = (0, 1)$, et est décrit par la relation de commutation $e_1 + e_2 = e_2 + e_1$. Une représentation de \mathbf{Z}^2 est donc la donnée d'un \mathbf{C} -espace vectoriel V et de deux éléments de $\mathbf{GL}(V)$ commutant entre eux.

Le groupe $\mathbf{SL}_2(\mathbf{Z})$ est engendré par les matrices $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, et toute relation entre S et T est conséquence des relations $S^4 = I$, $S^2T = TS^2$ et $(ST)^3 = S^2$; une représentation de $\mathbf{SL}_2(\mathbf{Z})$ est donc la donnée d'un \mathbf{C} -espace vectoriel V et de deux éléments u et v de $\mathbf{GL}(V)$ vérifiant $u^4 = 1$, $u^2v = vu^2$ et $(uv)^3 = u^2$.

De même, l'exercice ci-dessous montre qu'une représentation du groupe S_3 des permutations de $\{1, 2, 3\}$ est juste un \mathbf{C} -espace vectoriel V muni de deux symétries s_1, s_2 vérifiant $s_1s_2s_1 = s_2s_1s_2$.

Exercice I.1.4. — Soient $\sigma_1, \sigma_2 \in S_3$ les permutations $(1, 2)$ et $(2, 3)$.

- (i) Vérifier que $\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2$ est la permutation $(1, 3)$.
- (ii) Montrer que σ_1 et σ_2 engendrent S_3 .
- (iii) Montrer que toute relation entre σ_1 et σ_2 dans S_3 est conséquence des relations $\sigma_1^2 = \sigma_2^2 = 1$ et $\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2$. (On montrera qu'une relation de longueur $n \geq 4$ peut toujours se ramener à une relation de longueur $n - 2$.)

La *dimension* $\dim V$ d'une représentation V est juste la dimension du \mathbf{C} -espace vectoriel V . Par exemple, $\dim \mathbf{C}(\lambda) = 1$, pour tout $\lambda \in \mathbf{C}^*$.

Dans tout ce qui suit, les représentations sont implicitement supposées de dimension finie. Si $\dim V = d$ et si (e_1, \dots, e_d) est une base de V , on note $R_V(g)$ ou $R_{V,g}$ la matrice de $\rho_V(g)$ dans la base (e_1, \dots, e_d) (qui dépend du choix de la base bien que ça n'apparaisse pas dans la notation). Alors $R_V : G \rightarrow \mathbf{GL}_d(\mathbf{C})$ est un morphisme de groupes.

Exemple I.1.5. — (Construction d'une représentation de dimension 2 de S_3)

Soient $A = (1, 0)$, $B = (-\frac{1}{2}, \frac{\sqrt{3}}{2})$ et $C = (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$. Les points A, B, C sont les sommets d'un triangle équilatéral de centre de gravité $O = (0, 0)$. Les isométries du plan laissant stable ce triangle fixent O et donc sont linéaires; elles forment donc un sous-groupe⁽⁴⁾ D_3 de $\mathbf{O}(2) \subset \mathbf{GL}_2(\mathbf{C})$. L'injection de D_3 dans $\mathbf{GL}_2(\mathbf{C})$ fait de \mathbf{C}^2 une représentation du groupe D_3 , et nous allons montrer que ce groupe est isomorphe à S_3 pour construire notre représentation de S_3 . Un élément de D_3 laisse stable l'ensemble $\{A, B, C\}$, et fournit un morphisme de groupes f de D_3 dans le groupe des permutations $S_{\{A, B, C\}}$ de $\{A, B, C\}$. Comme A, B et C ne sont pas alignés, un élément de D_3 est uniquement déterminé par les images de A, B et C , ce qui signifie que f est injectif. Par ailleurs, f est surjectif car D_3 contient les symétries par rapport aux droites (OA) , (OB) et (OC) qui s'envoient respectivement sur les transpositions (B, C) , (A, C) et (A, B) , et les rotations d'angles 0 , $\frac{2\pi}{3}$ et $\frac{-2\pi}{3}$ dont les images respectives sont l'identité et les cycles (A, B, C) et (A, C, B) . En résumé, $f : D_3 \rightarrow S_{\{A, B, C\}}$ est un isomorphisme de groupes. La bijection $1 \mapsto A$, $2 \mapsto B$, $3 \mapsto C$ de $\{1, 2, 3\}$ sur $\{A, B, C\}$ fournit un isomorphisme $g : S_3 \cong S_{\{A, B, C\}}$. On obtient un morphisme de

⁽⁴⁾Plus généralement, on note D_n le groupe des isométries du plan fixant un polygone régulier à n côtés.

groupes de S_3 dans $\mathbf{GL}_2(\mathbf{C})$ en composant $f^{-1} \circ g : S_3 \rightarrow D_3$ avec l'injection de D_3 dans $\mathbf{GL}_2(\mathbf{C})$. Ce morphisme fait de \mathbf{C}^2 une représentation de S_3 .

Remarque I.1.6. — Soit G un groupe fini ; tout élément de G est alors d'ordre fini. Soit V une représentation de G . Si $g \in G$ est d'ordre n , on a $\rho_V(g)^n = \rho_V(g^n) = 1$. Comme le polynôme $X^n - 1$ n'a que des racines simples, cela prouve que $\rho_V(g)$ est diagonalisable.

Exercice I.1.7. — (i) Soit $V = \mathbf{C}^2$ la représentation de S_3 construite à l'exemple I.1.5. Montrer qu'il n'existe pas de base de V dans laquelle les matrices des actions de tous les éléments de S_3 sont simultanément diagonales.

(ii) Soient V un espace vectoriel de dimension finie sur \mathbf{C} , et u_1, u_2 deux endomorphismes diagonalisables de V commutant l'un à l'autre. Montrer que tout espace propre de u_1 est stable par u_2 . En déduire qu'il existe une base de V dans laquelle les matrices de u_1 et u_2 sont toutes les deux diagonales.

(iii) Soit G un groupe commutatif fini, et soit V une représentation de G . Montrer qu'il existe une base de V dans laquelle les matrices $R_V(g)$, pour $g \in G$, sont toutes diagonales. En déduire qu'il existe une décomposition de V en somme directe de droites stables par l'action de G .

2. Caractère d'une représentation, exemples

Le caractère χ_V de V est l'application de G dans \mathbf{C} définie par $\chi_V(g) = \text{Tr}(\rho_V(g))$, où $\text{Tr}(\rho_V(g))$ désigne la trace de l'endomorphisme $\rho_V(g)$; c'est aussi la trace de la matrice $R_V(g)$ dans n'importe quelle base de V , et c'est aussi la somme des valeurs propres de $\rho_V(g)$ comptées avec multiplicité.

On a en particulier $\chi_V(1) = \text{Tr}(1) = \dim V$; la valeur de χ_V en l'élément neutre est donc un entier ; cet entier est appelé le *degré* du caractère χ_V ; d'après ce qui précède, c'est aussi la dimension de la représentation V ; cette observation est d'usage constant. De plus, comme $\text{Tr}(AB) = \text{Tr}(BA)$, si A et B sont deux éléments de $\mathbf{M}_d(\mathbf{C})$, on a

$$\text{Tr}(\rho_V(hgh^{-1})) = \text{Tr}(\rho_V(h)\rho_V(g)\rho_V(h)^{-1}) = \text{Tr}(\rho_V(g)),$$

ce qui montre que χ_V est une *fonction centrale* sur G (i.e. χ_V est constante sur chacune des classes de conjugaison de G : on a $\chi_V(hgh^{-1}) = \chi_V(g)$ quels que soient $h, g \in G$).

Remarque I.1.8. — Si G est fini, et si $g \in G$, les valeurs propres de $\rho_V(g)$ sont des racines de l'unité. En particulier, elles sont de module 1, et donc $\lambda^{-1} = \bar{\lambda}$, si λ est une valeur propre de $\rho_V(g)$. Comme les valeurs propres de $\rho_V(g^{-1}) = \rho_V(g)^{-1}$ sont les inverses de celles de $\rho_V(g)$, et comme la trace est la somme des valeurs propres, on en déduit que $\chi_V(g^{-1}) = \overline{\chi_V(g)}$, quel que soit $g \in G$.

2.1. Caractères linéaires

Si V est de dimension 1, les endomorphismes de V sont les homothéties, et l'application qui à une homothétie associe son rapport induit un isomorphisme de $\mathbf{GL}(V)$ sur \mathbf{C}^* . Une représentation de dimension 1 n'est donc rien d'autre qu'un morphisme de groupes $\chi : G \rightarrow \mathbf{C}^*$; un tel morphisme est aussi souvent appelé un *caractère linéaire* de G . On note \widehat{G} l'ensemble de ces caractères linéaires.

Si V est une représentation de dimension 1 correspondant au caractère linéaire χ , on a $\chi_V = \chi$ de manière évidente. Autrement dit, *le caractère d'un caractère linéaire est le caractère linéaire lui-même*.

La *représentation triviale*, notée $\mathbf{1}$, est la représentation de dimension 1 correspondant au caractère trivial $\chi : G \rightarrow \mathbf{C}^*$, défini par $\chi(g) = 1$, pour tout $g \in G$.

Si V est une représentation de G , et si $\chi \in \widehat{G}$, on note $V(\chi)$ ou $V \otimes \chi$ la *tordue de V par le caractère linéaire χ* : c'est la représentation définie par $\rho_{V(\chi)}(g) = \chi(g)\rho_V(g)$ (l'espace vectoriel de $V(\chi)$ est V , mais l'action est tordue par χ ; la matrice $R_{V(\chi)}(g)$ est le produit de $R_V(g)$ par $\chi(g)$). On a $\chi_{V(\chi)}(g) = \chi(g)\chi_V(g)$, si $g \in G$.

Exercice I.1.9. — On définit le produit $\chi_1\chi_2$ de deux caractères linéaires par la formule $(\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g)$. Montrer que \widehat{G} , muni de ce produit, est un groupe commutatif.

Exercice I.1.10. — (Orthogonalité des caractères linéaires)

(i) Soit G un groupe fini et soit $\chi \in \widehat{G}$. Montrer que $\frac{1}{|G|} \sum_{g \in G} \chi(g)$ vaut 1 si χ est le caractère trivial, et 0 sinon. (Multiplier la somme par $\chi(h)$, pour un $h \in G$ bien choisi.)

(ii) En déduire que, si $\chi_1, \chi_2 \in \widehat{G}$, et si $\langle \chi_1, \chi_2 \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_1(g)} \chi_2(g)$, alors $\langle \chi_1, \chi_2 \rangle = 1$ si $\chi_1 = \chi_2$, et $\langle \chi_1, \chi_2 \rangle = 0$ si $\chi_1 \neq \chi_2$.

2.2. Sommes directes

Si V_1 et V_2 sont deux représentations de G , on peut munir la somme directe $V_1 \oplus V_2$ des espaces vectoriels V_1 et V_2 d'une action de G . Rappelons que $V_1 \oplus V_2$ est un espace vectoriel dont V_1 et V_2 sont des sous-espaces vectoriels qui sont en somme directe dans $V_1 \oplus V_2$. Comme on fait la somme directe d'un nombre fini d'espaces, on a une identification naturelle de $V_1 \oplus V_2$ avec le produit $V_1 \times V_2$, où $v_1 \in V_1$ s'identifie à $(v_1, 0) \in V_1 \times V_2$ et $v_2 \in V_2$ à $(0, v_2) \in V_1 \times V_2$. En utilisant cette identification, l'action de $g \in G$ sur $(v_1, v_2) \in V_1 \oplus V_2$ est donnée par $g \cdot (v_1, v_2) = (g \cdot v_1, g \cdot v_2)$. La représentation de G ainsi obtenue est encore notée $V_1 \oplus V_2$, et appelée *somme directe de V_1 et V_2* . Si on choisit une base e_1, \dots, e_m de V_1 et une base f_1, \dots, f_n de V_2 , alors $(e_1, 0), \dots, (e_m, 0), (0, f_1), \dots, (0, f_n)$ est une base de V , et la matrice $R_V(g)$ dans cette base est la matrice diagonale par blocs $\begin{pmatrix} R_{V_1}(g) & 0 \\ 0 & R_{V_2}(g) \end{pmatrix}$, dont la trace est la somme des traces de $R_{V_1}(g)$ et $R_{V_2}(g)$. On a donc

$$\chi_{V_1 \oplus V_2} = \chi_{V_1} + \chi_{V_2}.$$

Le cas $V_1 = V_2$ n'est pas exclus, et $V \oplus V$ est une représentation de G contenant deux copies de V d'intersection nulle et dont la somme est tout. Par exemple, la représentation $\mathbf{C}(\lambda) \oplus \mathbf{C}(\lambda)$ de \mathbf{Z} est \mathbf{C}^2 muni de l'homothétie λ , les deux copies de $\mathbf{C}(\lambda)$ obtenue en identifiant la somme au produit étant $\mathbf{C} \times \{0\}$ et $\{0\} \times \mathbf{C}$ (il y a beaucoup d'autres copies de $\mathbf{C}(\lambda)$ dans $\mathbf{C}(\lambda) \oplus \mathbf{C}(\lambda)$ puisque toute droite en est une). Plus généralement, si $m \in \mathbf{N}$, on note mV la somme directe de m copies de V (pour $m = 0$, on obtient l'espace

vectorel 0). Si les V_i , pour $i \in I$ fini, sont des représentations de G , et si $m_i \in \mathbf{N}$, pour tout $i \in I$, alors $\oplus_{i \in I} m_i V_i$ est une représentation de G de caractère

$$\chi_{\oplus_{i \in I} m_i V_i} = \sum_{i \in I} m_i \chi_{V_i}.$$

2.3. Représentations de permutation, représentation régulière

Si X est un ensemble fini muni d'une action (à gauche) de G donnée par $(g, x) \mapsto g \cdot x$, on définit la *représentation de permutation* V_X , associée à X , comme l'espace vectoriel V_X de dimension $|X|$, de base $(e_x)_{x \in X}$, muni de l'action linéaire de G donnée, sur les vecteurs de la base, par $g \cdot e_x = e_{g \cdot x}$. Si $g_1, g_2 \in G$, et si $x \in X$, on a $g_1 \cdot (g_2 \cdot e_x) = g_1 \cdot e_{g_2 \cdot x} = e_{g_1 g_2 \cdot x} = g_1 g_2 \cdot e_x$, ce qui prouve que la formule précédente définit bien une action de G sur V_X . Dans la base $(e_x)_{x \in X}$, la matrice de g est une *matrice de permutation* (i.e. a exactement un 1 par ligne et par colonne, et tous les autres coefficients sont nuls), et le terme diagonal $a_{x,x}$ est égal à 1 si et seulement si $g \cdot x = x$ (i.e. si x est un point fixe de g), sinon, il vaut 0. On en déduit que la trace de la matrice de g est le nombre de points fixes de g agissant sur X . Autrement dit, on a

$$\chi_{V_X}(g) = |\{x \in X, g \cdot x = x\}|.$$

Un cas particulier intéressant est celui où G est fini, $X = G$, et l'action de G est donnée par la multiplication à gauche (i.e. $g \cdot h = gh$). La représentation V_G ainsi obtenue est la *représentation régulière* de G . Comme $gh = h$ implique $g = 1$, on voit que le caractère de la représentation régulière est donné par la formule

$$\chi_{V_G}(1) = |G|, \quad \text{et} \quad \chi_{V_G}(g) = 0, \quad \text{si } g \in G - \{1\}.$$

3. Morphismes de représentations

3.1. La représentation $\text{Hom}(V_1, V_2)$

Soient V_1 et V_2 deux représentations de G , et soit $u : V_1 \rightarrow V_2$ une application linéaire. Si $g \in G$, on définit $g \cdot u : V_1 \rightarrow V_2$ par la formule $(g \cdot u)(v) = g \cdot u(g^{-1} \cdot v)$, quel que soit $v \in V_1$. Si $g_1, g_2 \in G$, et si $v \in V_1$, on a

$$\begin{aligned} (g_1 \cdot (g_2 \cdot u))(v) &= g_1 \cdot ((g_2 \cdot u)(g_1^{-1} \cdot v)) \\ &= g_1 \cdot (g_2 \cdot u(g_2^{-1} \cdot g_1^{-1} \cdot v)) = g_1 g_2 \cdot u((g_1 g_2)^{-1} \cdot v) = (g_1 g_2 \cdot u)(v), \end{aligned}$$

et donc $g_1 \cdot (g_2 \cdot u) = g_1 g_2 \cdot u$, ce qui prouve que l'on a défini de la sorte une action de G sur l'espace $\text{Hom}(V_1, V_2)$ des applications linéaires de V_1 dans V_2 .

Si $g \in G$, l'endomorphisme $\rho_{\text{Hom}(V_1, V_2), g}$ de $\text{Hom}(V_1, V_2)$ est alors donné par la formule :

$$\rho_{\text{Hom}(V_1, V_2), g}(u) = \rho_{V_2, g} \circ u \circ \rho_{V_1, g}^{-1}, \quad \text{si } u \in \text{Hom}(V_1, V_2).$$

Proposition I.1.11. — *Si G est fini, et si $g \in G$, alors*

$$\chi_{\text{Hom}(V_1, V_2)}(g) = \overline{\chi_{V_1}(g)} \chi_{V_2}(g).$$

Démonstration. — Si g est fixé, on peut choisir une base $(e_i)_{i \in I}$ de V_1 et une base $(f_j)_{j \in J}$ de V_2 dans lesquelles les actions de g sont diagonales. Il existe donc des racines de l'unité α_i , pour $i \in I$, et β_j , pour $j \in J$, tels que $g \cdot e_i = \alpha_i e_i$, si $i \in I$, et $g \cdot f_j = \beta_j f_j$, si $j \in J$. On a alors $\chi_{V_1}(g) = \sum_{i \in I} \alpha_i$ et $\chi_{V_2}(g) = \sum_{j \in J} \beta_j$.

Si $(i, j) \in I \times J$, soit $u_{i,j} : V_1 \rightarrow V_2$ l'application linéaire définie par $u_{i,j}(e_i) = f_j$, et $u_{i,j}(e_{i'}) = 0$, si $i' \neq i$. Les $u_{i,j}$, pour $(i, j) \in I \times J$ forment une base de $\text{Hom}(V_1, V_2)$, et on a $g \cdot u_{i,j} = \alpha_i^{-1} \beta_j u_{i,j} = \overline{\alpha_i} \beta_j u_{i,j}$. On a donc

$$\chi_{\text{Hom}(V_1, V_2)}(g) = \sum_{(i,j) \in I \times J} \overline{\alpha_i} \beta_j = \left(\sum_{i \in I} \overline{\alpha_i} \right) \left(\sum_{j \in J} \beta_j \right) = \overline{\chi_{V_1}(g)} \chi_{V_2}(g).$$

Ceci permet de conclure.

Remarque I.1.12. — Si $V_1 = V$ et si V_2 est la représentation triviale, la représentation $\text{Hom}(V_1, V_2) = \text{Hom}(V, \mathbf{C})$ est la *représentation duale* V^* de V . On a $\chi_{V^*} = \overline{\chi_V}$, d'après la prop. I.1.11.

3.2. Opérateurs d'entrelacement, représentations isomorphes

Notons $\text{Hom}_G(V_1, V_2)$ l'ensemble des applications linéaires de V_1 dans V_2 commutant à l'action de G . C'est un sous-espace vectoriel de $\text{Hom}(V_1, V_2)$ et un élément u de $\text{Hom}(V_1, V_2)$ est dans $\text{Hom}_G(V_1, V_2)$, si et seulement si on a $g \cdot u(v) = u(g \cdot v)$, quel que soit $v \in V_1$. Appliqué à $g^{-1} \cdot v$, ceci peut aussi se réécrire sous la forme $g \cdot u(g^{-1} \cdot v) = u(v)$, quel que soit $v \in V_1$, ou encore, sous la forme $g \cdot u = u$. Autrement dit, $\text{Hom}_G(V_1, V_2)$ est le sous-espace vectoriel de $\text{Hom}(V_1, V_2)$ des éléments fixes sous l'action de G . Les éléments de $\text{Hom}_G(V_1, V_2)$ sont souvent appelés des *opérateurs d'entrelacement*.

Exemple I.1.13. — Si V est une représentation de G , l'ensemble V^G des éléments de V fixes sous l'action de G est un sous-espace vectoriel de V (c'est l'intersection des noyaux des $g - 1$, pour $g \in G$) qui est stable sous l'action de G , et sur lequel G agit trivialement par construction ; c'est donc une représentation de G . Maintenant, si G est fini, on peut considérer l'*opérateur de moyenne* $M : V \rightarrow V$ défini par $M(v) = \frac{1}{|G|} \sum_{g \in G} g \cdot v$. Alors M est un opérateur d'entrelacement entre V et V^G . En effet, si $h \in G$, et si $v \in V$, on a

$$\begin{aligned} h \cdot M(v) &= h \cdot \left(\frac{1}{|G|} \sum_{g \in G} g \cdot v \right) = \frac{1}{|G|} \sum_{g \in G} h \cdot (g \cdot v) = \frac{1}{|G|} \sum_{g \in G} hg \cdot v \\ M(h \cdot v) &= \frac{1}{|G|} \sum_{g \in G} g \cdot (h \cdot v) = \frac{1}{|G|} \sum_{g \in G} gh \cdot v \end{aligned}$$

et comme $g \mapsto hg$ et $g \mapsto gh$ sont des bijections de G dans G , les deux quantités sont égales à $M(v)$. Cela prouve à la fois que $M(v) \in V^G$ et que $M : V \rightarrow V^G$ commute à l'action de G .

Remarque I.1.14. — L'idée, selon laquelle il suffit de faire la moyenne sous l'action du groupe pour obtenir quelque chose de fixe par tout le groupe, joue un rôle très important dans la théorie.

On dit que deux représentations V_1 et V_2 de G sont *isomorphes*, s'il existe un isomorphisme linéaire $u : V_1 \rightarrow V_2$ commutant à l'action de G , (autrement dit, s'il existe $u \in \text{Hom}_G(V_1, V_2)$ bijectif, ce qui implique, en particulier, que V_1 et V_2 ont la même dimension). Traduit en termes des morphismes $\rho_{V_1} : G \rightarrow \mathbf{GL}(V_1)$ et $\rho_{V_2} : G \rightarrow \mathbf{GL}(V_2)$ attachés à V_1 et V_2 , cette relation devient $u \circ \rho_{V_1}(g) = \rho_{V_2}(g) \circ u$, quel que soit $g \in G$. Traduit en termes matriciels (après avoir choisi des bases de V_1 et V_2), cela se traduit par l'existence de $T \in \mathbf{GL}_d(\mathbf{C})$, tel que $T R_{V_1}(g) = R_{V_2}(g) T$, quel que soit $g \in G$, ce qui peut encore se mettre sous la forme $R_{V_2}(g) = T R_{V_1}(g) T^{-1}$. En particulier, $\chi_{V_1}(g) = \chi_{V_2}(g)$ quel que soit $g \in G$.

Remarque I.1.15. — (i) On verra plus loin que, si G est fini, la réciproque est vraie : si V_1 et V_2 ont mêmes caractères, alors elles sont isomorphes, ce qui peut sembler un peu surprenant, le caractère ne permettant, a priori, que de calculer la trace des endomorphismes. L'exercice I.1.16 ci-dessous rend le résultat un peu plus envisageable.

(ii) Dans le cas de \mathbf{Z} , en notant R_i , pour $i = 1, 2$, la matrice de $\rho_{V_i}(1)$, on voit que V_1 et V_2 sont isomorphes si et seulement s'il existe T inversible telle que $R_2 = T R_1 T^{-1}$ (i.e. si et seulement si R_2 et R_1 sont des matrices semblables). Il en résulte que la classification des représentations de \mathbf{Z} à isomorphisme près est équivalente à celle des matrices à similitude près, ce qui se fait en utilisant la forme de Jordan. Si on impose à $R = R_V(1)$ d'être diagonalisable, alors V est donnée, à isomorphisme près, par les valeurs propres de R avec leurs multiplicités.

Exercice I.1.16. — Soit V une représentation d'un groupe fini G , et soit $g \in G$. Montrer que, si $P_g(T) = \det(1 - T \rho_V(g))$, on a l'identité des séries formelles

$$-\sum_{n=0}^{+\infty} \chi_V(g^{n+1}) T^n = \frac{P'_g(T)}{P_g(T)}.$$

En déduire que χ_V permet de déterminer $\rho_V(g)$, pour tout $g \in G$, à conjugaison près.

I.2. Décomposition des représentations

1. Décomposition en somme directe de représentations irréductibles

Soient G un groupe et V une représentation de G . Une *sous-représentation* de V est un sous-espace vectoriel de V stable par G . On dit que V est *irréductible* si V ne possède pas de sous-représentation autre que 0 ou V . De manière équivalente, V est irréductible si, quel que soit $v \in V - \{0\}$, le sous-espace vectoriel de V engendré par les $g \cdot v$, pour $g \in G$, est égal à V .

Exemple I.2.1. — Soit V une représentation de \mathbf{Z} , et soit $u = \rho_V(1)$. Comme \mathbf{C} est algébriquement clos, u admet une valeur propre λ , non nulle car u est inversible. Soit $e_\lambda \in V$ un vecteur propre pour la valeur propre λ . On a alors $n \cdot e_\lambda = u^n(e_\lambda) = \lambda^n e_\lambda$ pour tout $n \in \mathbf{Z}$, ce qui prouve que la droite $\mathbf{C}e_\lambda$ est stable sous l'action de \mathbf{Z} et est une sous-représentation de \mathbf{Z} isomorphe à la représentation $\mathbf{C}(\lambda)$ de l'ex. I.1.1. En particulier, si V est de dimension ≥ 2 , alors V n'est pas irréductible, et donc toute représentation irréductible de \mathbf{Z} est de dimension 1, isomorphe à $\mathbf{C}(\lambda)$, pour un $\lambda \in \mathbf{C}^*$ uniquement déterminé.

Supposons maintenant que u est diagonalisable. Soit v_1, \dots, v_d une base de V constituée de vecteurs propres de u , et soit λ_i la valeur propre associée à e_i . Alors V est la somme directe $\bigoplus_{i=1}^d \mathbf{C}e_i$ des droites $\mathbf{C}e_i$ qui sont des sous-représentations de V , chaque $\mathbf{C}e_i$ étant isomorphe à $\mathbf{C}(\lambda_i)$ en tant que représentation de \mathbf{Z} . On en déduit que V est, en tant que représentation de \mathbf{Z} , isomorphe à $\bigoplus_{i=1}^d \mathbf{C}(\lambda_i)$.

Remarque I.2.2. — (i) Dire que V est isomorphe à $\bigoplus_{i=1}^d \mathbf{C}(\lambda_i)$ signifie juste que $u = \rho_V(1)$ est diagonalisable, et que son polynôme caractéristique est $\prod_{i=1}^d (X - \lambda_i)$, ce qui est nettement moins précis que d'exhiber une base de vecteurs propres, et donc un isomorphisme de $\bigoplus_{i=1}^d \mathbf{C}(\lambda_i)$ sur V entre représentations de \mathbf{Z} .

(ii) Si u est diagonalisable, si les valeurs propres de u sont $\lambda_1, \dots, \lambda_r$, avec $\lambda_i \neq \lambda_j$ si $i \neq j$, et si la multiplicité de λ_i est m_i , alors $V \cong \bigoplus_{i=1}^r m_i \mathbf{C}(\lambda_i)$.

(iii) Si u n'est pas diagonalisable, la représentation V ne se décompose pas comme une somme directe de représentations irréductibles.

Nous allons prouver que, si G est fini, toute représentation de G est somme directe de représentations irréductibles. Cela revient, en choisissant une base de chacune de ces représentations irréductibles, à exhiber une base de V dans laquelle les $\rho_V(g)$, pour $g \in G$, se mettent simultanément sous une forme diagonale par blocs⁽⁵⁾, la taille des blocs étant la plus petite possible. (C'est un peu analogue à la forme minimale d'une matrice de rotation dans \mathbf{R}^n .) Nous aurons besoin du résultat suivant.

Théorème I.2.3. — *Soit V une représentation de G . Il existe sur V un produit scalaire hermitien $\langle \cdot, \cdot \rangle_V$ invariant sous l'action de G .*

Démonstration. — Partons d'un produit scalaire hermitien quelconque $\langle \cdot, \cdot \rangle$ sur V , et définissons $\langle \cdot, \cdot \rangle_V$ comme la moyenne des transformés de $\langle \cdot, \cdot \rangle$ sous l'action de G . Autrement

⁽⁵⁾C'est assez particulier aux représentations des groupes finis sur un corps de caractéristique 0. Même dans le cas des groupes finis, si on considère des représentations sur un corps de caractéristique > 0 , le mieux que l'on puisse espérer est une mise sous forme triangulaire supérieure par blocs. Par exemple, si V est la représentation de dimension 2 de $G = \mathbf{Z}/p\mathbf{Z}$ sur le corps $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, la matrice de $n \in \mathbf{Z}$ dans une base (e_1, e_2) étant $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$, alors le sous-espace V_1 , engendré par e_1 , est stable (et même fixe) par G , mais il est facile de voir que c'est le seul sous-espace propre de V ayant cette propriété. La représentation V n'est donc pas irréductible, mais n'est pas somme directe de représentations irréductibles.

dit, on a

$$\langle v_1, v_2 \rangle_V = \frac{1}{|G|} \sum_{g \in G} \langle g \cdot v_1, g \cdot v_2 \rangle.$$

L'action de G étant linéaire, le résultat est bien linéaire par rapport à v_2 et sesquilinéaire par rapport à v_1 . De plus, $\langle v, v \rangle_V \geq \frac{1}{|G|} \langle v, v \rangle$, ce qui prouve que $\langle \cdot, \cdot \rangle_V$ est défini positif. Finalement, si $h \in G$, on a

$$\langle h \cdot v_1, h \cdot v_2 \rangle_V = \frac{1}{|G|} \sum_{g \in G} \langle g \cdot (h \cdot v_1), g \cdot (h \cdot v_2) \rangle = \frac{1}{|G|} \sum_{g \in G} \langle gh \cdot v_1, gh \cdot v_2 \rangle = \langle v_1, v_2 \rangle_V,$$

car $g \mapsto gh$ induit une bijection de G sur lui-même. Ceci permet de conclure.

Théorème I.2.4. — (Maschke, 1899) *Toute représentation de G est somme directe de représentations irréductibles.*

Démonstration. — La démonstration se fait par récurrence sur la dimension. Si V est de dimension 1 ou est irréductible, il n'y a rien à faire. Si V est de dimension ≥ 2 et n'est pas irréductible, alors V possède une sous-représentation V_1 distincte de 0 et V . Si $\langle \cdot, \cdot \rangle_V$ est un produit scalaire hermitien sur V , invariant sous l'action de G , le supplémentaire orthogonal V_2 de V_1 est lui-aussi stable par G puisque que « v orthogonal à V_1 » équivaut à « $g \cdot v$ orthogonal à $g \cdot V_1 = V_1$ » par invariance du produit scalaire. On a alors $V = V_1 \oplus V_2$, et V_1 et V_2 sont de dimensions strictement inférieures à celle de V . L'hypothèse de récurrence permet de les décomposer comme des sommes directes de représentations irréductibles, ce qui prouve qu'on peut en faire autant de V .

Remarque I.2.5. — Si G est cyclique engendré par g , la décomposition de V en somme de représentations irréductibles est équivalente à une décomposition de V en droites invariantes sous l'action de g . On sait bien que si g a une valeur propre de multiplicité > 1 , cette décomposition n'est pas unique. Par contre, la décomposition en sous-espaces propres est, elle, parfaitement canonique. On verra plus loin (cor. I.2.16) que la situation est la même en ce qui concerne la décomposition en somme de représentations irréductibles d'une représentation d'un groupe fini quelconque.

2. Le lemme de Schur et ses conséquences immédiates

Théorème I.2.6. — (Lemme de Schur, 1905) *Soient G un groupe et V_1, V_2 des représentations irréductibles de G .*

- (i) *Si V_1 et V_2 ne sont pas isomorphes, alors $\text{Hom}_G(V_1, V_2) = 0$.*
- (ii) *Si $V_1 = V_2$, alors $\text{Hom}_G(V_1, V_2)$ est la droite des homothéties.*

Démonstration. — Soit $u \in \text{Hom}_G(V_1, V_2)$. Le fait que u commute à l'action de G , montre que $\text{Ker}(u) \subset V_1$ et $\text{Im}(u) \subset V_2$ sont stables par G . Comme par hypothèse, V_1 et V_2 sont irréductibles, on a soit $\text{Ker}(u) = V_1$, auquel cas $u = 0$, soit $\text{Ker}(u) = 0$, auquel cas

$\text{Im}(u) \neq 0$ et donc $\text{Im}(u) = V_2$. On en déduit que, si $u \neq 0$, alors u est à la fois injective (puisque $\text{Ker}(u) = 0$) et surjective, et donc est un isomorphisme. Cela démontre le (i).

Passons au (ii). Comme on travaille avec des \mathbf{C} -espaces vectoriels, u admet une valeur propre λ . Donc $u - \lambda$, qui commute à l'action de G puisque u le fait et qu'une homothétie commute à tout, a un noyau non nul. Le même raisonnement qu'au (i) montre que ce noyau doit donc être égal à V_1 , ce qui se traduit par le fait que u est une homothétie de rapport λ . Ceci permet de conclure.

Exercice I.2.7. — Soit G un groupe commutatif. Montrer que toute représentation irréductible de G est de dimension 1.

Proposition I.2.8. — Soient G un groupe fini et V_1, V_2 des représentations de G .

(i) Si V_1 et V_2 sont irréductibles, non isomorphes, et si $u \in \text{Hom}(V_1, V_2)$, alors $M(u) = \frac{1}{|G|} \sum_{g \in G} g \cdot u = 0$.

(ii) Si V est irréductible, et si $u \in \text{Hom}(V, V)$, alors $M(u) = \frac{1}{|G|} \sum_{g \in G} g \cdot u$ est l'homothétie de rapport $\frac{1}{\dim V} \text{Tr}(u)$.

(iii) Si V est irréductible, et si ϕ est une fonction centrale sur G , alors $\sum_{g \in G} \phi(g) \rho_V(g)$ est l'homothétie de rapport $\frac{1}{\dim V} \sum_{g \in G} \phi(g) \chi(g)$.

Démonstration. — Si V_1 et V_2 sont deux représentations de G , si $u \in \text{Hom}(V_1, V_2)$, et si $h \in G$, on a $h \cdot (\sum_{g \in G} g \cdot u) = \sum_{g \in G} hg \cdot u$. Comme $g \mapsto hg$ est une bijection de G sur lui-même, cette dernière quantité est aussi égale à $\sum_{g \in G} g \cdot u$. On en déduit que $M(u) = \frac{1}{|G|} \sum_{g \in G} g \cdot u$ appartient à $\text{Hom}_G(V_1, V_2)$.

Le (i) est donc une conséquence du (i) du lemme de Schur. Le (ii) du lemme de Schur montre, quant à lui, que $M(u)$ est une homothétie, si $u \in \text{Hom}(V, V)$ et V est irréductible. Pour déterminer le rapport de cette homothétie, il suffit d'en calculer la trace et de diviser par $\dim V$. Or on a $M(u) = \frac{1}{|G|} \sum_{g \in G} \rho_V(g) u \rho_V(g)^{-1}$, et donc $M(u)$ est la moyenne de $|G|$ termes dont chacun a pour trace $\text{Tr}(u)$, puisque la trace est invariante par conjugaison. On a donc $\text{Tr}(M(u)) = \text{Tr}(u)$, ce qui permet d'en déduire le (ii).

Finalement, si ϕ est une fonction centrale (i.e., si $\phi(hgh^{-1}) = \phi(g)$ quels que soient $h, g \in G$), si $u_\phi = \sum_{g \in G} \phi(g) \rho_V(g) \in \text{Hom}(V, V)$, et si $h \in G$, on a

$$\begin{aligned} h \cdot u_\phi &= \rho_V(h) \left(\sum_{g \in G} \phi(g) \rho_V(g) \right) \rho_V(h)^{-1} = \sum_{g \in G} \phi(g) \rho_V(hgh^{-1}) \\ &= \sum_{g \in G} \phi(hgh^{-1}) \rho_V(hgh^{-1}) = \sum_{g \in G} \phi(g) \rho_V(g) = u_\phi. \end{aligned}$$

On conclut comme ci-dessus que u_ϕ est l'homothétie de rapport

$$\frac{1}{\dim V} \text{Tr}(u_\phi) = \frac{1}{\dim V} \sum_{g \in G} \phi(g) \text{Tr}(\rho_V(g)) = \frac{1}{\dim V} \sum_{g \in G} \phi(g) \chi_V(g),$$

ce qui permet de conclure.

3. Orthogonalité des caractères

Soit G un groupe fini. On note $R_{\mathbf{C}}(G)$ l'espace vectoriel des fonctions centrales. Cet espace contient l'ensemble $R^+(G)$ des caractères des représentations de G qui lui-même, contient l'ensemble $\text{Irr}(G)$ des *caractères irréductibles* de G (i.e. les caractères des représentations irréductibles de G). Finalement, on note $R_{\mathbf{Z}}(G)$ le groupe des *caractères virtuels* de G ; c'est le sous-groupe (additif) de $R_{\mathbf{C}}(G)$ engendré par $R^+(G)$.

Exercice I.2.9. — Montrer que $R^+(G)$ est stable par addition et que $R_{\mathbf{Z}}(G)$ est l'ensemble des $\chi_1 - \chi_2$, avec $\chi_1, \chi_2 \in R^+(G)$.

On munit $R_{\mathbf{C}}(G)$ du produit scalaire hermitien $\langle \cdot, \cdot \rangle$ défini par

$$\langle \phi_1, \phi_2 \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\phi_1(g)} \phi_2(g).$$

Théorème I.2.10. — (Frobenius, 1897) *Les caractères irréductibles forment une base orthonormale de l'espace des fonctions centrales.*

Démonstration. — Soient χ_1 et χ_2 deux caractères, et soient V_1 et V_2 des représentations de G dont les caractères sont χ_1 et χ_2 . En utilisant la prop. I.1.11, on peut réécrire $\langle \chi_1, \chi_2 \rangle$ sous la forme

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_1(g)} \chi_2(g) = \frac{1}{|G|} \sum_{g \in G} \chi_{\text{Hom}(V_1, V_2)}(g).$$

Comme $\chi_{\text{Hom}(V_1, V_2)}(g)$ est, par définition, la trace de g agissant sur $\text{Hom}(V_1, V_2)$, cela permet de voir $\langle \chi_1, \chi_2 \rangle$ comme la trace de l'application linéaire $u \mapsto \frac{1}{|G|} \sum_{g \in G} g \cdot u = M(u)$ définie dans la prop. I.2.8. On déduit alors des (i) et (ii) de cette proposition les faits suivants.

- Si χ_1 et χ_2 sont irréductibles et distincts, alors M est identiquement nul, et donc $\langle \chi_1, \chi_2 \rangle = \text{Tr}(M) = 0$.

- Si χ est irréductible, et si V est une représentation de caractère χ , alors M est l'application associant à $u \in \text{Hom}(V, V)$ l'homothétie de rapport $\frac{1}{\dim V} \text{Tr}(u)$. On en déduit que M admet comme valeurs propres 1 avec multiplicité 1, l'espace propre correspondant étant la droite des homothéties, et 0 avec multiplicité $(\dim V)^2 - 1$, le noyau de M étant l'hyperplan des endomorphismes de trace nulle. La trace de M est donc 1, ce qui se traduit par $\langle \chi, \chi \rangle = 1$.

Il résulte des deux points ci-dessus que les caractères irréductibles forment une famille orthonormale. Il reste à vérifier qu'ils forment une base de $R_{\mathbf{C}}(G)$, et pour cela, il suffit de vérifier qu'une fonction centrale ϕ , qui est orthogonale à tous les éléments de $\text{Irr}(G)$, est nulle. Pour cela, considérons la représentation régulière V_G de G , que l'on décompose en somme directe $V_1 \oplus \cdots \oplus V_r$ de représentations irréductibles. Si ϕ est une

fonction centrale orthogonale à χ_{V_i} , il résulte du (iii) de la prop. I.2.8, que l'endomorphisme $\sum_{g \in G} \overline{\phi(g)} \rho_{V_i}(g)$ de V_i est nul. Donc, si ϕ est orthogonale à tous les caractères irréductibles, l'endomorphisme $\sum_{g \in G} \overline{\phi(g)} \rho_{V_G}(g)$ de V_G est nul. En faisant agir cet endomorphisme sur $e_1 \in V_G$, on en déduit que $0 = \sum_{g \in G} \overline{\phi(g)} g \cdot e_1 = \sum_{g \in G} \overline{\phi(g)} e_g$. Or les e_g , pour $g \in G$, forment une base de V_G ; la nullité de $\sum_{g \in G} \overline{\phi(g)} e_g$ implique donc celle de $\overline{\phi(g)}$, quel que soit $g \in G$, et donc aussi celle de ϕ . Ceci permet de conclure.

Remarque I.2.11. — Si V_1 et V_2 sont irréductibles et non isomorphes, l'application M est identiquement nulle et donc $\text{Tr}(M) = 0$. Or il résulte de la démonstration du th. I.2.10 que $\text{Tr}(M) = \langle \chi_{V_1}, \chi_{V_2} \rangle$. On en déduit en particulier que $\chi_{V_1} \neq \chi_{V_2}$. Autrement dit, l'application $W \mapsto \chi_W$ est une injection de l'ensemble des représentations irréductibles de G (à isomorphisme près) dans $\text{Irr}(G)$. Comme, par définition de $\text{Irr}(G)$, cette application est surjective, c'est une bijection, ce qui permet de voir $\text{Irr}(G)$ aussi comme l'ensemble des représentations irréductibles de G . C'est cette interprétation de $\text{Irr}(G)$ qui est utilisée dans la suite.

4. Applications du théorème principal

Le théorème I.2.10 a des tas de conséquences agréables.

4.1. Nombre des représentations irréductibles

Corollaire I.2.12. — *Le nombre de représentations irréductibles de G est égal au nombre $|\text{Conj}(G)|$ de classes de conjugaison dans G . En particulier, il est fini.*

Démonstration. — D'après le th. I.2.10, le nombre de représentations irréductibles de G est égal à la dimension de l'espace $R_G(G)$ des fonctions centrales. Or une fonction est centrale si et seulement si elle est constante sur chaque classe de conjugaison; une fonction centrale ϕ peut donc s'écrire de manière unique sous la forme $\phi = \sum_{C \in \text{Conj}(G)} \lambda_C \mathbf{1}_C$, où $\mathbf{1}_C$ est la fonction indicatrice de C , et $\lambda_C \in \mathbf{C}$ (on a $\lambda_C = \phi(g)$, où g est n'importe quel élément de C). Les $\mathbf{1}_C$, pour $C \in \text{Conj}(G)$ forment donc une base de $R_G(G)$ qui, de ce fait, est de dimension $|\text{Conj}(G)|$. Ceci permet de conclure.

Remarque I.2.13. — L'ensemble des représentations irréductibles de G et celui des classes de conjugaison dans G ont le même cardinal mais il n'y a, en général, aucune bijection naturelle entre ces deux ensembles. Les groupes symétriques constituent une exception remarquable (cf. n° B.2).

4.2. La décomposition canonique d'une représentation

Le résultat suivant est un peu magique et très utile (par exemple pour décomposer une représentation obtenue par des procédés tensoriels comme au n° 1 du § I.3, ce qui sert beaucoup en physique des particules); on peut le voir comme une généralisation du calcul des valeurs propres d'un endomorphisme à partir du polynôme caractéristique. Dans les

deux cas, on n'a pas besoin d'exhiber des vecteurs ayant le bon comportement ; on se contente de prouver qu'ils existent.

Corollaire I.2.14. — Si V est une représentation de G , si $V = W_1 \oplus \cdots \oplus W_k$ est une décomposition de V en somme directe de représentations irréductibles, et si $W \in \text{Irr}(G)$, alors le nombre m_W de W_i qui sont isomorphes à W est égal à $\langle \chi_W, \chi_V \rangle$. En particulier, il ne dépend pas de la décomposition, et $V \cong \bigoplus_{W \in \text{Irr}(G)} \langle \chi_W, \chi_V \rangle W$.

Démonstration. — On a $\chi_V = \chi_{W_1} + \cdots + \chi_{W_k}$, et donc

$$\langle \chi_W, \chi_V \rangle = \langle \chi_W, \chi_{W_1} \rangle + \cdots + \langle \chi_W, \chi_{W_k} \rangle.$$

Or $\langle \chi_W, \chi_{W_i} \rangle$ est égal à 1 ou 0 suivant que W_i est ou n'est pas isomorphe à W ; on a donc $\langle \chi_W, \chi_V \rangle = m_W$. Ceci permet de conclure.

Corollaire I.2.15. — Deux représentations V_1 et V_2 de G ayant même caractère χ sont isomorphes.

Démonstration. — Elles sont toutes les deux isomorphes à $\bigoplus_{W \in \text{Irr}(G)} \langle \chi_W, \chi \rangle W$, d'après le cor. I.2.14.

Corollaire I.2.16. — (Décomposition d'une représentation en composantes isotypiques) Si V est une représentation de G , et si $W \in \text{Irr}(G)$, alors

$$p_W = \frac{\dim W}{|G|} \sum_{g \in G} \overline{\chi_W(g)} \rho_V(g),$$

est un projecteur commutant à l'action de G . De plus, toutes les représentations irréductibles de G apparaissant dans la décomposition de $p_W(V)$ sont isomorphes à W , et V est la somme directe des $p_W(V)$, pour $W \in \text{Irr}(G)$.

Démonstration. — Soit $V = W_1 \oplus \cdots \oplus W_k$ une décomposition de V en somme directe de représentations irréductibles. D'après le (iii) de la proposition I.2.8, la restriction de p_W à W_i est l'homothétie de rapport

$$\frac{\dim W}{|G| \dim W_i} \sum_{g \in G} \overline{\chi_W(g)} \chi_{W_i}(g) = \frac{\dim W}{\dim W_i} \langle \chi_W, \chi_{W_i} \rangle.$$

D'après les relations d'orthogonalité des caractères, cela implique que p_W est l'identité sur W_i , si $W_i \cong W$, et est nulle dans le cas contraire. Le résultat s'en déduit.

4.3. Un critère d'irréductibilité

Corollaire I.2.17. — Une représentation V de G est irréductible, si et seulement si $\langle \chi_V, \chi_V \rangle = 1$

Démonstration. — Si $V \cong \bigoplus_{W \in \text{Irr}(G)} m_W W$, alors

$$\langle \chi_V, \chi_V \rangle = \left\langle \sum_{W \in \text{Irr}(G)} m_W \chi_W, \sum_{W \in \text{Irr}(G)} m_W \chi_W \right\rangle = \sum_{W \in \text{Irr}(G)} m_W^2.$$

Comme les m_W sont des entiers naturels, on en déduit que $\langle \chi_V, \chi_V \rangle = 1$ si et seulement si tous les m_W sont égaux à 0, sauf un qui est égal à 1. Ceci permet de conclure.

Exercice I.2.18. — Soit $\chi \in R_{\mathbf{Z}}(G)$. Montrer que les conditions suivantes sont équivalentes.

- (i) $\chi \in \text{Irr}(G)$.
- (ii) $\langle \chi, \chi \rangle = 1$ et $\chi(1) \geq 0$.

4.4. La décomposition de la représentation régulière

Corollaire I.2.19. — (i) Si W est irréductible, alors W apparaît dans la représentation régulière avec la multiplicité $\dim W$.

(ii) On a $\sum_{W \in \text{Irr}(G)} (\dim W)^2 = |G|$ (formule de Burnside⁽⁶⁾).

(iii) Si $g \neq 1$, alors $\sum_{W \in \text{Irr}(G)} \dim W \chi_W(g) = 0$.

Démonstration. — Le caractère χ_{V_G} de la représentation régulière est donné (alinéa 2.3 du § I.1) par $\chi_{V_G}(1) = |G|$ et $\chi_{V_G}(g) = 0$, si $g \neq 1$. Or la multiplicité de W dans V_G est, d'après le cor. I.2.14, égale à

$$\langle \chi_W, \chi_{V_G} \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_W(g)} \chi_{V_G}(g) = \frac{1}{|G|} \overline{\chi_W(1)} |G| = \overline{\chi_W(1)} = \dim W,$$

ce qui démontre le (i). On en déduit que $\chi_{V_G} = \sum_{W \in \text{Irr}(G)} \dim W \chi_W$. En appliquant cette identité à $g = 1$, on en déduit le (ii), et à $g \neq 1$, on en déduit le (iii).

5. Le cas des groupes commutatifs

5.1. La transformée de Fourier

Théorème I.2.20. — Si G est commutatif, toute représentation irréductible de G est de dimension 1. Autrement dit $\text{Irr}(G)$ coïncide avec l'ensemble \widehat{G} des caractères linéaires de G .

Démonstration. — Si G est commutatif, les classes de conjugaison sont réduites à un élément, et donc $|\text{Conj}(G)| = |G|$. Comme $|\text{Irr}(G)| = |\text{Conj}(G)|$ d'après le cor. I.2.12, comme $\sum_{W \in \text{Irr}(G)} (\dim W)^2 = |G|$, d'après le (ii) du cor. I.2.19, et comme $\dim W \geq 1$, quel

⁽⁶⁾Dans le cas de S_n , on dispose d'une démonstration directe de cette formule, cf. note 2 de l'annexe B

que soit $W \in \text{Irr}(G)$, on en déduit que $\dim W = 1$, quel que soit $W \in \text{Irr}(G)$, ce que l'on voulait démontrer⁽⁷⁾.

Corollaire I.2.21. — *Si G est commutatif, toute fonction de G dans \mathbf{C} est combinaison linéaire de caractères linéaires.*

Démonstration. — D'après le th. I.2.10, toute fonction centrale (et donc toute fonction puisque G est commutatif) est combinaison linéaire de caractères irréductibles. Le th. I.2.20 permet de conclure.

Comme les caractères linéaires d'un groupe commutatif G forment une base orthonormale des fonctions de G dans \mathbf{C} , il est très facile de décomposer une fonction quelconque comme une combinaison linéaire de caractères linéaires. Si ϕ est une fonction sur G , on définit la *transformée de Fourier* $\hat{\phi}$ comme la fonction définie sur \hat{G} par

$$\hat{\phi}(\chi) = \langle \chi, \phi \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\chi(g)} \phi(g) = \frac{1}{|G|} \sum_{g \in G} \chi(g)^{-1} \phi(g).$$

La *formule d'inversion de Fourier* s'exprime alors sous la forme

$$\phi = \sum_{\chi \in \hat{G}} \hat{\phi}(\chi) \chi;$$

c'est une conséquence immédiate du fait que les χ , pour $\chi \in \hat{G}$, forment une famille orthonormale. Par exemple, si on applique ce qui précède à la fonction $\phi_a : G \rightarrow \mathbf{C}$ valant 1 en a et 0 ailleurs, on a $\hat{\phi}_a(\chi) = \frac{1}{|G|} \overline{\chi(a)}$, et on obtient :

$$\frac{1}{|G|} \sum_{\chi \in \hat{G}} \overline{\chi(a)} \chi(x) = \begin{cases} 1 & \text{si } x = a, \\ 0 & \text{sinon.} \end{cases}$$

Un caractère linéaire de $(\mathbf{Z}/D\mathbf{Z})^*$ est appelé un *caractère de Dirichlet modulo D* . On note $\text{Dir}(D)$ l'ensemble de ces caractères. Le résultat suivant est un des ingrédients de la démonstration de Dirichlet du théorème de la progression arithmétique (cf. annexe A).

Proposition I.2.22. — *Si a est premier à D , alors*

$$\frac{1}{\varphi(D)} \sum_{\chi \in \text{Dir}(D)} \overline{\chi(a)} \chi(n) = \begin{cases} 1 & \text{si } n \equiv a \pmod{D}, \\ 0 & \text{sinon.} \end{cases}$$

Démonstration. — Il suffit d'appliquer ce qui précède au groupe $(\mathbf{Z}/D\mathbf{Z})^*$, dont le cardinal est $\varphi(D)$, et à la fonction $\phi_a : (\mathbf{Z}/D\mathbf{Z})^* \rightarrow \mathbf{C}$ valant 1 en a et 0 ailleurs.

⁽⁷⁾Des démonstrations plus terre-à-terre sont proposées dans les ex. I.1.7 et I.2.7.

5.2. Le groupe dual

Remarque I.2.23. — Si G est un groupe, alors \widehat{G} est un groupe commutatif pour la multiplication des caractères linéaires ($\chi_1\chi_2(x) = \chi_1(x)\chi_2(x)$, pour tout $x \in G$, si $\chi_1, \chi_2 \in \widehat{G}$), et on peut donc considérer le groupe $\widehat{\widehat{G}}$ de ses caractères linéaires. La formule de multiplication ci-dessus montre que, si $x \in G$, alors $\chi \mapsto \chi(x)$ est un caractère linéaire de \widehat{G} ; d'où une application naturelle $\iota : G \rightarrow \widehat{\widehat{G}}$, définie par $(\iota(x))(\chi) = \chi(x)$. Cette application est un morphisme de groupes puisque, si $x, y \in G$, on a

$$(\iota(xy))(\chi) = \chi(xy) = \chi(x)\chi(y) = (\iota(x))(\chi) (\iota(y))(\chi),$$

pour tout $\chi \in \widehat{G}$, et donc $\iota(xy) = \iota(x)\iota(y)$.

Proposition I.2.24. — Si G est un groupe commutatif fini, alors $\iota : G \rightarrow \widehat{\widehat{G}}$ est un isomorphisme de groupes.

Démonstration. — Compte-tenu de ce qui précède, il suffit de vérifier la bijectivité de ι . Si H est un groupe commutatif fini, on a $|\text{Conj}(H)| = |H|$, et $|\text{Irr}(H)| = |\widehat{H}|$ d'après le cor. I.2.21. On en déduit, en utilisant le cor. I.2.12, que $|H| = |\widehat{H}|$. En utilisant ce résultat pour G et \widehat{G} , on en déduit que G et $\widehat{\widehat{G}}$ ont le même cardinal. Il suffit donc de vérifier que ι est injective. Or la décomposition de Fourier de la fonction ϕ_a , introduite dans le paragraphe précédant la prop. I.2.22, montre que, si $\chi(a) = \chi(b)$ pour tout $\chi \in \widehat{G}$, alors $\phi_a = \phi_b$, et donc $a = b$. Ceci permet de conclure.

Exercice I.2.25. — Montrer que le groupe dual de $\mathbf{Z}/n\mathbf{Z}$ est isomorphe à $\mathbf{Z}/n\mathbf{Z}$ (en fait à μ_n).

Lemme I.2.26. — Soit G un groupe commutatif fini.

(i) Si $x \in G$ est d'ordre a , si $y \in G$ est d'ordre b , et si a et b sont premiers entre eux, alors xy est d'ordre ab .

(ii) Si $a, b \in \mathbf{N} - \{0\}$, et si G contient des éléments d'ordre a et b , alors il contient un élément d'ordre $\text{ppcm}(a, b)$.

(iii) Soit N le maximum des ordres des éléments de G . Alors $x^N = 1$ pour tout $x \in G$.

Démonstration. — (i) Comme x et y commutent, on a $(xy)^n = x^n y^n$, pour tout $n \in \mathbf{N}$. En particulier, $(xy)^{ab} = x^{ab} y^{ab} = 1$, et donc l'ordre de xy divise ab . Réciproquement, si $(xy)^n = 1$, alors $1 = (xy)^{an} = y^{an}$ et $1 = (xy)^{bn} = x^{bn}$, et donc an est un multiple de b et bn est un multiple de a . Comme a et b sont premiers entre eux, cela implique que n est un multiple de a et b et donc aussi de ab ; autrement dit l'ordre de xy est un multiple de ab . On en déduit le (i).

(ii) Soit \mathcal{P}_1 (resp. \mathcal{P}_2) l'ensemble des nombres premiers tels que $v_p(a) > 0$ et $v_p(a) \geq v_p(b)$ (resp. $v_p(b) > v_p(a)$). Alors \mathcal{P}_1 et \mathcal{P}_2 sont des ensembles disjoints, ce qui fait que $k = \prod_{p \in \mathcal{P}_1} p^{v_p(a)}$ et $\ell = \prod_{p \in \mathcal{P}_2} p^{v_p(b)}$ sont premiers entre eux. De plus, on a $v_p(k\ell) = v_p(a)$, si $v_p(a) \geq v_p(b)$, et $v_p(k\ell) = v_p(b)$, si $v_p(a) < v_p(b)$, et donc $k\ell = \text{ppcm}(a, b)$. Maintenant, soient $x \in G$ d'ordre a et $y \in G$ d'ordre b . Comme k divise a , cela implique que $x' = x^{a/k}$ est d'ordre k . De même $y' = y^{b/\ell}$ est d'ordre ℓ , et le (i) montre que $x'y'$ est d'ordre $k\ell = \text{ppcm}(a, b)$. D'où le (ii).

(iii) Il résulte du (ii) que si $x \in G$ est d'ordre a , alors G contient un élément d'ordre $\text{ppcm}(a, N)$. Comme $\text{ppcm}(a, N) \geq N$, cela implique, par définition de N , que $\text{ppcm}(a, N) = N$, et donc que a divise N . On en déduit le (iii).

Remarque I.2.27. — Le cas de S_3 montre que les trois résultats du lemme I.2.26 peuvent se trouver en défaut, si G n'est pas commutatif.

Si G est un groupe commutatif fini, l'entier N dont le (iii) du lemme I.2.26 décrit les propriétés est appelé *l'exposant* de G .

Lemme I.2.28. — Si G est un groupe commutatif fini, alors G et \widehat{G} ont même exposant.

Démonstration. — Si H est un groupe commutatif fini, notons $N(H)$ son exposant. Si $\chi \in \widehat{H}$, on a

$$\chi^{N(H)}(x) = \chi(x)^{N(H)} = \chi(x^{N(H)}) = \chi(1) = 1, \quad \text{pour tout } x \in G,$$

et donc $\chi^{N(H)} = 1$. On en déduit que l'exposant de \widehat{H} divise celui de H . En utilisant ce résultat pour $H = G$ et $H = \widehat{G}$, et l'isomorphisme $\widehat{\widehat{G}} \cong G$, on en déduit les inégalités $N(G) = N(\widehat{\widehat{G}}) \leq N(\widehat{G}) \leq N(G)$, qui permettent de conclure.

5.3. *Le théorème de structure des groupes finis commutatifs*

Théorème I.2.29. — Si G est un groupe fini commutatif, il existe $r \in \mathbf{N}$, et des entiers N_1, \dots, N_r , où N_1 est l'exposant de G et $N_{i+1} | N_i$, si $i \leq r-1$, tels que $G \cong \bigoplus_{i=1}^r \mathbf{Z}/N_i \mathbf{Z}$.

Démonstration. — La démonstration se fait par récurrence sur $|G|$, le résultat étant évident (avec $r = 0$), si $|G| = 1$. Supposons donc $|G| > 1$, et notons $N = N_1$ l'exposant de G . Alors $\chi(x)$ est une racine N -ième de l'unité, pour tous $\chi \in \widehat{G}$ et $x \in G$. De plus, comme N est aussi l'exposant de \widehat{G} (lemme I.2.28), il existe χ_1 d'ordre N , et comme $\chi_1(G)$ est un sous-groupe du groupe cyclique μ_N , c'est μ_N tout entier. Il existe donc $x_1 \in G$ tel que $\chi_1(x_1) = e^{2i\pi/N}$. Comme l'ordre de x_1 divise N , par définition de l'exposant d'un groupe, il en résulte que x_1 est d'ordre N , et donc que le sous-groupe H_1 de G engendré par x_1 est isomorphe à $\mathbf{Z}/N\mathbf{Z}$. Montrons que G est la somme directe de H_1 et $G_1 = \ker \chi_1$, ce qui permettra de conclure en appliquant l'hypothèse de récurrence à G_1 , l'exposant d'un sous-groupe divisant celui du groupe de manière évidente. Pour cela, constatons que χ_1 induit un isomorphisme de H_1 sur μ_N , puisqu'il est surjectif et que les deux groupes ont le même cardinal N ; notons $\alpha : \mu_N \rightarrow H_1$ son inverse. Si $x \in G$, alors $a = \alpha(\chi_1(x)) \in H_1$ et $b = a^{-1}x$ vérifie $\chi_1(b) = \chi_1(a)^{-1}\chi_1(x) = 1$, et donc $b \in G_1$. On peut donc écrire tout élément x de G sous la forme $x = ab$, avec $a \in H_1$ et $b \in G_1$. Enfin $H_1 \cap G_1 = \{1\}$ puisque χ_1 est injectif sur H_1 . Ceci montre que $G = H_1 \oplus G_1$, et permet de conclure.

Exercice I.2.30. — Soit G un groupe commutatif fini. Montrer que \widehat{G} et G sont isomorphes.

6. Table des caractères d'un groupe fini

Soit G un groupe fini, et soit $c = |\text{Conj}(G)|$. La *table des caractères* de G est un tableau $c \times c$ dont les entrées sont les valeurs des caractères irréductibles sur les classes de conjugaison de G , le coefficient à l'intersection de la colonne correspondant au caractère χ et de la ligne correspondant à la classe de conjugaison C , étant $\chi(C)$. C'est en quelque sorte la carte du groupe G .

Remarque I.2.31. — Notons T_G la matrice $c \times c$ définie par la table des caractères. Notons aussi K la matrice diagonale dont la coordonnée diagonale sur la ligne correspondant à une classe de conjugaison C est $\frac{|C|}{|G|}$. Alors les relations d'orthogonalité des caractères s'expriment de manière compacte par la relation $T_G^* K T_G = I$. On en déduit que $K = (T_G^*)^{-1} T_G^{-1}$ et $K^{-1} = T_G T_G^*$. En particulier, les lignes de T_G forment une famille orthogonale, ce qui permet de remplir la table des caractères en n'en connaissant qu'une partie.

Par exemple, le groupe $\{\pm 1\}$ a deux classes de conjugaison 1 et -1 , et deux caractères irréductibles 1 et χ (de dimension 1 puisque $\{\pm 1\}$ est commutatif) ; sa table des caractères est très facile à établir :

	1	χ
1	1	1
-1	1	-1

FIG. 1. Table des caractères de $\{\pm 1\}$

L'exemple du groupe $\{\pm 1\}$ est un peu trop trivial pour donner une idée de la manière dont on peut construire la table des caractères d'un groupe. L'exemple de A_4 , traité ci-dessous, est nettement plus riche. Le lecteur trouvera dans les exercices d'autres techniques pour établir des tables de caractères de petits groupes. L'appendice B contient des exemples un peu plus sophistiqués. Le contraste entre la simplicité et la puissance de la théorie générale et le côté artisanal du traitement des cas particuliers est assez saisissant.

Rappelons que A_4 est le sous-groupe des permutations de S_4 de signature 1. Comme S_4 a 24 éléments, on a $|A_4| = 12$, et les éléments de A_4 sont :

- l'élément neutre id ,
- les trois produits de deux transpositions $s_2 = (12)(34)$, $s_3 = (13)(24)$ et $s_4 = (14)(23)$, qui sont d'ordre 2,
- les huit 3-cycles (123) , (234) , (341) , (412) et (132) , (243) , (314) , (421) , qui sont d'ordre 3.

Nous nous proposons d'établir la table des caractères de A_4 . Il y a plusieurs manières d'arriver au résultat. La manière la plus systématique consiste à déterminer les classes de conjugaison de A_4 , construire toutes les représentations irréductibles de A_4 , et calculer la valeur de leurs caractères sur les classes de conjugaison. C'est celle que nous explorons en premier⁽⁸⁾ ; ensuite nous montrons un certain nombre de raccourcis possibles qui utilisent les théorèmes du cours.

(a) Soit t le 3-cycle (123) . On a $t^2 = (132)$, et comme t est d'ordre 3, le sous-groupe $T = \{1, t, t^2\}$ de A_4 engendré par t est d'ordre 3.

(b) $H = \{\text{id}, s_2, s_3, s_4\}$ est un sous-groupe commutatif distingué de A_4 .

Cela peut se vérifier par un calcul un peu fastidieux. On peut aussi remarquer qu'un 2-Sylow de A_4 est de cardinal 4, et comme H est de cardinal 4 et contient tous les éléments de A_4 d'ordre divisant 4, cela prouve qu'il n'y a qu'un seul 2-Sylow (qui est donc distingué puisque

⁽⁸⁾Nous n'avons pas cherché la solution la plus courte ; au contraire, nous avons essayé d'employer un maximum de techniques de base de la théorie des groupes finis. On pourrait aller plus vite en utilisant ce qu'on sait des classes de conjugaison de S_4 (n° 4.5 du chap. « Vocabulaire mathématique »).

la conjugaison transforme un 2-Sylow en un 2-Sylow), et que ce 2-Sylow est H . De plus, tous les éléments de H sont d'ordre divisant 2, et un groupe ayant cette propriété est commutatif car $(xy)^2 = 1 = x^2y^2$ implique $xy = yx$.

(c) Tout élément de A_4 peut s'écrire de manière unique sous la forme $t^a h$, avec $a \in \{0, 1, 2\}$ et $h \in H$.

Les sous-groupes T et H de A_4 ont une intersection réduite à $\{\text{id}\}$. On en déduit que $(c, h) \mapsto ch$ est une injection de $T \times H$ dans A_4 ; en effet, si $c_1 h_1 = c_2 h_2$, alors $c_2^{-1} c_1 = h_2 h_1^{-1}$, et comme $c_2^{-1} c_1 \in T$ et $h_2 h_1^{-1} \in H$, on a $c_2 = c_1$ et $h_2 = h_1$. Comme $T \times H$ et A_4 ont le même cardinal, une injection est une bijection, ce qui permet de conclure.

(d) t et t^2 ne commutent à aucun élément de $H - \{\text{id}\}$.

Cela peut se vérifier par un calcul un peu pénible. On peut aussi remarquer que si t et $s \in H - \{\text{id}\}$ commutent, le sous-groupe G de A_4 engendré par s et t est commutatif et isomorphe à $(\mathbf{Z}/2\mathbf{Z}) \oplus (\mathbf{Z}/3\mathbf{Z}) \cong \mathbf{Z}/6\mathbf{Z}$, car les sous-groupes $\{\text{id}, s\}$ et T engendrés par s et t sont en somme directe. Ceci n'est pas possible car A_4 ne contient pas d'élément d'ordre 6. L'argument est le même pour n'importe quel 3-cycle, et donc en particulier pour t^2 .

(e) Les classes de conjugaison de A_4 sont $B_1 = \{\text{id}\}$, $B_2 = H - \{\text{id}\}$, $B_3 = tH$ et $B_4 = t^2H$.

Un calcul particulièrement ennuyeux mènerait au résultat... Comme dans tout groupe, la classe de conjugaison de l'élément neutre n'a qu'un élément, et donc $B_1 \in \text{Conj}(A_4)$. Si $s \in B_2$, et si $t^a h$, avec $a \in \{0, 1, 2\}$ et $h \in H$ commute à s , on a $t^a h s = s t^a h$, et donc $t^a h s h = s t^a h^2$, et comme H est commutatif et $h^2 = \text{id}$, on obtient $t^a s = s t^a$, ce qui implique $a = 0$. Le centralisateur de s est donc H , et le cardinal de la classe de conjugaison de s est égal à $\frac{|A_4|}{|H|} = 3$. Comme un conjugué de s est d'ordre 2, cette classe de conjugaison est incluse dans B_2 , et donc lui est égale pour des raisons de cardinal. Enfin, le centralisateur de t et t^2 est T (si $t^a h t = t t^a h$, on a $h t = t h$, et donc $h = \text{id}$), ce qui fait que le cardinal de la classe de conjugaison de t est $\frac{|A_4|}{|T|} = 4$. Or on a $(t^a h) t (t^a h)^{-1} = t^a h t h^{-1} t^{-a} = t(t^{a-1} h t^{1-a})(t^a h^{-1} t^{-a}) \in tH$, car H est distingué et donc $t^{a-1} h t^{1-a} \in H$ et $t^a h^{-1} t^{-a} \in H$. La classe de conjugaison de t est donc incluse dans B_3 et lui est égale pour des raisons de cardinal. De même, la classe de conjugaison de t^2 est B_4 . Ceci permet de conclure.

(f) Soit $\rho = e^{\frac{2i\pi}{3}}$ une racine primitive 3-ième de l'unité. Si $i \in \{0, 1, 2\}$, on définit $\eta^i : A_4 \rightarrow \mu_3$ par $\eta^i(t^a h) = \rho^{ia}$, si $a \in \{0, 1, 2\}$ et si $h \in H$. Alors $\eta^0 = 1$, η et η^2 sont des caractères linéaires distincts de A_4 .

Si $a, b \in \{0, 1, 2\}$, et si $h, g \in H$, alors $t^a h t^b g = t^{a+b} (t^{-b} h t^b) g$, et comme H est distingué, on a $t^{-b} h t^b \in H$, et donc $(t^{-b} h t^b) g \in H$ et $\eta^i(t^a h t^b g) = \rho^{i(a+b)} = \rho^{ia} \rho^{ib} = \eta^i(t^a h) \eta^i(t^b g)$.

(g) Soit V la représentation de permutation associée à l'action naturelle de A_4 sur $\{1, 2, 3, 4\}$. Rappelons que cette représentation est \mathbf{C}^4 muni de l'action de A_4 définie, dans la base canonique e_1, \dots, e_4 , par $g(e_i) = e_{g(i)}$. L'hyperplan W d'équation $x_1 + \dots + x_4 = 0$ est stable par A_4 , et la représentation que l'on obtient est irréductible de caractère donné par $\chi_W(\text{id}) = 3$, $\chi_W(g) = -1$, si $g \in H - \{\text{id}\}$, et $\chi_W(g) = 0$, si $g \notin H$.

La représentation V se décompose sous la forme $V' \oplus W$, où V' est la droite engendré par $e_1 + \dots + e_4$ (isomorphe à $\mathbf{1}$ car $e_1 + \dots + e_4$ est fixe par A_4). Comme V est une représentation de permutation, $\chi_V(g)$ est le nombre de points fixes de g agissant sur $\{1, 2, 3, 4\}$. On a donc $\chi_V(\text{id}) = 4$, $\chi_V(g) = 0$, si $g \in H - \{\text{id}\}$ et $\chi_V(g) = 1$, si $g \notin H$. On en déduit le caractère de W , car on a $\chi_V = \chi_{V'} + \chi_W$, et $\chi_{V'}(g) = 1$ pour tout $g \in A_4$, puisque $V' \cong \mathbf{1}$. On a donc $\chi_W(\text{id}) = 3$, $\chi_W(g) = -1$, si $g \in H - \{\text{id}\}$, et $\chi_W(g) = 0$, si $g \notin H$. On en déduit que $\langle \chi_W, \chi_W \rangle = \frac{1}{12}(3^2 + 3 \cdot (-1)^2 + 8 \cdot 0) = 1$, ce qui prouve que W est irréductible.

(h) La table des caractères de A_4 est celle de la figure 2.

	$\mathbf{1}$	η	η^2	χ_W
B_1	1	1	1	3
B_2	1	1	1	-1
B_3	1	ρ	ρ^2	0
B_4	1	ρ^2	ρ	0

FIG. 2. Table des caractères de A_4

En effet, A_4 ayant 4 classes de conjugaison, il a aussi 4 représentations irréductibles à isomorphisme près, qui sont donc les 3 caractères linéaires 1 , η et η^2 , et la représentation W de dimension 3. Les valeurs des caractères de ces représentations ont été calculées ci-dessus ; ce sont les valeurs reportées dans la table. Ceci permet de conclure.

- *Premier raccourci.* Imaginons que l'on ait construit des représentations 1 , η , η^2 et W dont les caractères prennent les valeurs de la table sur B_1 , B_2 , B_3 et B_4 , mais qu'on ne sache pas quelles sont les classes de conjugaison de A_4 . Alors on peut en déduire que ces classes sont exactement B_1 , B_2 , B_3 et B_4 , ce qui permet de se passer des points (d) et (e) ci-dessus. En effet, comme $1^2 + 1^2 + 1^2 + 3^2 = 12$, la formule de Burnside ((ii) du cor. I.2.19) montre que 1 , η , η^2 et χ_W sont les éléments de $\text{Irr}(A_4)$, et donc (cor. I.2.12) que A_4 a 4 classes de conjugaison. Or on remarque que, si $i \neq j$, il existe $\chi \in \text{Irr}(A_4)$ prenant des valeurs distinctes sur B_i et B_j . Comme un élément de $\text{Irr}(A_4)$ est constant sur une classe de conjugaison, on en déduit que si $C \in \text{Conj}(A_4)$, il existe $i(C) \in \{1, 2, 3, 4\}$ tel que $C \subset B_{i(C)}$. Les éléments de C formant une partition de A_4 , l'application $C \mapsto i(C)$ est surjective, et comme les deux ensembles ont le même nombre d'éléments, elle est bijective ; de plus, on a $B_{i(C)} = C$, sinon un élément de $B_{i(C)} - C$ ne serait pas dans la réunion des classes de conjugaison. En résumé, les classes de conjugaison de A_4 sont les B_i .

- *Second raccourci.* Supposons W construite. La formule de Burnside ((ii) du cor. I.2.19) nous fournit alors l'identité $12 = |A_4| = 9 + \sum_{W' \in \text{Irr}(A_4) - \{W\}} (\dim W')^2$, et comme il y a une seule manière d'écrire 3 comme une somme de carrés, on en déduit que A_4 a trois caractères linéaires distincts. Autrement dit, le groupe \hat{A}_4 (cf. ex. I.1.9) est de cardinal 3 et donc isomorphe à $\mathbf{Z}/3\mathbf{Z}$; en particulier, il est cyclique et si on note η un générateur, les éléments de \hat{A}_4 sont η , η^2 et le caractère trivial qui est aussi égal à η^3 . Comme η est d'ordre 3, il est à valeurs dans le groupe μ_3 des racines 3-ièmes de l'unité, et son image étant un sous-groupe de μ_3 non réduit à 1, c'est μ_3 tout entier. En particulier, l'image de η est de cardinal 3, et donc le noyau est de cardinal $|A_4|/3 = 4$. Par ailleurs, on a $H \subset \ker \chi$ car le seul élément de μ_3 d'ordre divisant 2 est 1. On en déduit que $\ker \chi = H$, ce qui permet de redémontrer le point (b). Enfin, on a

$\eta(t) \neq 1$ puisque $t \notin H$, et donc $\eta(t) = \rho$ ou $\eta(t) = \rho^2$; quitte à remplacer η par η^2 , on peut supposer que $\eta(t) = \rho$. On a alors $\eta(g) = 1$ si $g \in H = B_1 \cup B_2$, $\eta(g) = \rho$ si $g \in B_3 = tH$, et $\eta(g) = \rho^2$ si $g \in B_4 = t^2H$. Ceci permet, en utilisant le premier raccourci, de compléter la table des caractères de A_4 sans avoir utilisé un seul des points (a)-(e) au sujet de la structure de A_4 ni le point (f).

• *Troisième raccourci.* On suppose ce coup-ci que l'on a construit η , ce qui utilise les points (a)-(c) et (f), mais pas les (d), (e) et (g). La formule de Burnside montre alors, qu'il y a a priori quatre possibilités pour les caractères irréductibles distincts des caractères linéaires 1, η et η^2 :

- un unique caractère χ_W de degré 3,
- deux caractères de degré 2 et un de degré 1, ou un de degré 2 et cinq de degré 1,
- neuf caractères de degré 1.

Si on est dans le premier cas, on a gagné car $1 + \eta + \eta^2 + 3\chi_W$ est le caractère de la représentation régulière, ce qui permet de calculer χ_W , et donc de compléter la table en utilisant le premier raccourci. Il suffit donc d'éliminer les deux autres possibilités. La dernière implique que $|\text{Irr}(A_4)| = 12$, ce qui implique que A_4 a 12 classes de conjugaison (cor. I.2.12), et donc que celles-ci sont des singletons, et que A_4 est commutatif, ce qui n'est pas. Si A_4 a au moins une représentation irréductible V de dimension 2, alors χ_V , $\chi_V\eta$ et $\chi_V\eta^2$ sont des caractères irréductibles de degré 2 (cf. alinéa 2.1), et comme il y a au plus deux tels caractères, il existe $\eta_1 \neq \eta_2 \in \{1, \eta, \eta^2\}$ tels que $\chi_V\eta_1 = \chi_V\eta_2$. Or la condition $\eta_1 \neq \eta_2$ implique que $\eta_1(t) \neq \eta_2(t)$, et la relation $\chi_V\eta_1 = \chi_V\eta_2$ entraîne donc $\chi_V(t) = 0$. Ceci n'est pas possible, car t est d'ordre 3, ce qui fait que les deux valeurs propres de $\rho_V(t)$ sont des racines 3-ièmes de l'unité, et la somme de deux racines 3-ièmes n'est jamais nulle. L'existence d'une représentation irréductible de dimension 2 est donc exclue, ce qui permet de conclure.

I.3. Construction de représentations

1. Constructions tensorielles de représentations

1.1. Produit tensoriel d'espaces vectoriels de dimension finie

Soient V_1, V_2 deux espaces vectoriels de dimension finie, et soient (e_1, \dots, e_n) une base de V_1 et (f_1, \dots, f_m) une base de V_2 . Soit $V_1 \otimes V_2$ le *produit tensoriel* de V_1 et V_2 : c'est l'espace vectoriel de base⁽⁹⁾ les $e_i \otimes f_j$, pour $1 \leq i \leq n$ et $1 \leq j \leq m$. Si $x = \sum_{i=1}^n \lambda_i e_i \in V_1$ et $y = \sum_{j=1}^m \mu_j f_j \in V_2$, on note $x \otimes y$ l'élément de $V_1 \otimes V_2$ défini par la formule

$$x \otimes y = \sum_{i=1}^n \sum_{j=1}^m \lambda_i \mu_j e_i \otimes f_j.$$

Le produit tensoriel $V_1 \otimes V_2$ est en général un objet nouveau, mais il arrive qu'il puisse se décrire de manière plus explicite.

Exemple I.3.1. — (i) Si X est un ensemble fini, on note \mathbf{C}^X l'ensemble des fonctions $\phi : X \rightarrow \mathbf{C}$. C'est un espace vectoriel dont une base est l'ensemble des ϕ_x , pour $x \in X$, où $\phi_x(y) = 1$, si $y = x$ et $\phi_x(y) = 0$, si $y \neq x$. Il est facile de vérifier que, si I et J sont deux ensembles finis, alors $\phi_i \otimes \phi_j \mapsto \phi_{(i,j)}$, pour $i \in I$ et $j \in J$, induit un isomorphisme de $\mathbf{C}^I \otimes \mathbf{C}^J$ sur $\mathbf{C}^{I \times J}$.

⁽⁹⁾On aurait pu noter $g_{i,j}$ la base de $V_1 \otimes V_2$, mais la notation $e_i \otimes f_j$ est plus parlante pour la suite.

(ii) Si V_1, V_2 sont deux espaces vectoriels, et si V_1^* et V_2^* sont leur duals (V_i^* est l'espace des formes linéaires sur V_i), alors $V_1^* \otimes V_2^*$ est l'espace des formes bilinéaires sur $V_1 \times V_2$. Si $\lambda_1 \in V_1^*$ et $\lambda_2 \in V_2^*$, alors $\lambda_1 \otimes \lambda_2$ est la forme bilinéaire $(x, y) \mapsto \lambda_1(x)\lambda_2(y)$.

Exercice I.3.2. — Montrer que $V_1^* \otimes V_2 = \text{Hom}(V_1, V_2)$.

Par construction, $(x, y) \mapsto x \otimes y$ est une application bilinéaire de $V_1 \times V_2$ dans $V_1 \otimes V_2$. Le lemme suivant montre que $V_1 \otimes V_2$ est *universel* pour les applications bilinéaires⁽¹⁰⁾ sur $V_1 \times V_2$.

Lemme I.3.3. — Si W est un espace vectoriel, et si $u : V_1 \times V_2 \rightarrow W$ est bilinéaire, alors il existe une unique application linéaire $\tilde{u} : V_1 \otimes V_2 \rightarrow W$, telle que $\tilde{u}(x \otimes y) = u(x, y)$, quels que soient $x \in V_1$ et $y \in V_2$.

Démonstration. — On définit \tilde{u} par ses valeurs sur les éléments de la base des $e_i \otimes f_j$, en posant $\tilde{u}(e_i \otimes f_j) = u(e_i, f_j)$. Un calcul immédiat montre alors que la bilinéarité de u est équivalente à la relation $\tilde{u}(x \otimes y) = u(x, y)$, quels que soient $x \in V_1$ et $y \in V_2$. Ceci permet de conclure.

Maintenant, si $u_1 \in \text{End}(V_1)$ et $u_2 \in \text{End}(V_2)$, alors $(x, y) \mapsto u_1(x) \otimes u_2(y)$ est bilinéaire de $V_1 \times V_2$ dans $V_1 \otimes V_2$. D'après le lemme I.3.3, il existe $u_1 \otimes u_2 \in \text{End}(V_1 \otimes V_2)$ unique, tel que $(u_1 \otimes u_2)(x \otimes y) = u_1(x) \otimes u_2(y)$, quels que soient $x \in V_1$ et $y \in V_2$. Si $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbf{M}_n(\mathbf{C})$ est la matrice de u_1 , et $B = (b_{i',j'})_{1 \leq i',j' \leq m} \in \mathbf{M}_m(\mathbf{C})$ est la matrice de u_2 , alors la matrice $A \otimes B \in \mathbf{M}_{nm}(\mathbf{C})$ de $u_1 \otimes u_2$, dans la base des $e_i \otimes f_{i'} = g_{(i-1)m+i', (j-1)m+j'}$, est la matrice des $c_{(i-1)m+i', (j-1)m+j'}$, avec $1 \leq i, j \leq n$ et $1 \leq i', j' \leq m$ et $c_{(i-1)m+i', (j-1)m+j'} = a_{i,j}b_{i',j'}$. En particulier, on a

$$\text{Tr}(u_1 \otimes u_2) = \sum_{1 \leq k \leq nm} c_{k,k} = \sum_{i=1}^n \sum_{i'=1}^m a_{i,i} b_{i',i'} = \left(\sum_{i=1}^n a_{i,i} \right) \left(\sum_{i'=1}^m b_{i',i'} \right) = \text{Tr}(u_1) \text{Tr}(u_2).$$

Finalement, on déduit du lemme I.3.3 que, si $u'_1, u''_1 \in \text{End}(V_1)$, et si $u'_2, u''_2 \in \text{End}(V_2)$, alors

$$(u'_1 \circ u''_1) \otimes (u'_2 \circ u''_2) = (u'_1 \otimes u'_2) \circ (u''_1 \otimes u''_2).$$

(Il suffit de comparer l'image de $x \otimes y$ par les endomorphismes dans les deux membres.)

⁽¹⁰⁾On peut en donner une construction naturelle, sans choisir de base, en prenant le quotient de l'espace de base les $e_{x,y}$, pour $(x, y) \in V_1 \times V_2$, par les relations $e_{x,y_1+y_2} = e_{x,y_1} + e_{x,y_2}$, $e_{x_1+x_2,y} = e_{x_1,y} + e_{x_2,y}$ et $e_{\lambda x,y} = e_{x,\lambda y} = \lambda e_{x,y}$. Alors $x \otimes y$ est l'image de $e_{x,y}$ dans le quotient. La construction du texte est moins canonique mais plus concrète... La construction naturelle a l'avantage de marcher aussi en dimension infinie où une base n'est pas toujours facile à exhiber.

1.2. Produit tensoriel de représentations

Soit maintenant G un groupe fini, et soient V_1, V_2 deux représentations de G . D'après ce qui précède, $V_1 \otimes V_2$ est une représentation de G , l'action de G étant telle que $g \cdot (x \otimes y) = (g \cdot x) \otimes (g \cdot y)$. La formule ci-dessus pour la trace de $u_1 \otimes u_2$ montre que

$$\chi_{V_1 \otimes V_2}(g) = \chi_1(g) \chi_2(g).$$

Si V_2 est de dimension 1, on retrouve la construction de la torsion d'une représentation par un caractère linéaire (alinéa 2.1 du § I.1).

Remarque I.3.4. — (i) Si G_1 et G_2 sont deux groupes finis, et si V_1 et V_2 sont des représentations de G_1 et G_2 respectivement, on peut définir de la même manière une représentation $V_1 \boxtimes V_2$ de $G_1 \times G_2$, en faisant agir $g = (g_1, g_2) \in G_1 \times G_2$ sur l'espace vectoriel $V_1 \otimes V_2$ par $g \cdot (x \otimes y) = (g_1 \cdot x) \otimes (g_2 \cdot y)$.

(ii) Si $G_1 = G_2 = G$, la représentation $V_1 \otimes V_2$ de G définie précédemment est la restriction à G , vu comme ensemble des couples (g, g) de $G \times G$, de la représentation $V_1 \boxtimes V_2$ de $G \times G$ (c'est pour pouvoir faire la distinction que $V_1 \boxtimes V_2$ n'est pas notée $V_1 \otimes V_2$).

Exercice I.3.5. — Montrer que $R^+(G)$ est stable par produit, et que $R_{\mathbf{Z}}(G)$ est un anneau.

1.3. Carré symétrique et carré extérieur d'une représentation

Si V est une représentation d'un groupe fini G , la représentation $V \otimes V$ n'est pas irréductible. En effet, les *tenseurs symétriques* (i.e. les expressions de la forme $xy = \frac{1}{2}(x \otimes y + y \otimes x)$, avec $x, y \in V$, et donc $xy = yx$, si $x, y \in V$) et les *tenseurs alternés* (les $x \wedge y = x \otimes y - y \otimes x$, avec $x, y \in V$, et donc $x \wedge y = -y \wedge x$, si $x, y \in V$) sont stables sous l'action de G ; il en est donc de même des sous-espaces de $V \otimes V$ qu'ils engendrent.

On note $\text{Sym}^2 V$ le *carré symétrique* de V ; c'est le sous-espace de $V \otimes V$ engendré par les tenseurs symétriques. Si V est de dimension d , de base (e_1, \dots, e_d) , alors $\text{Sym}^2 V$ est un espace de dimension $\frac{d(d+1)}{2}$ dont une base est constituée des $e_i e_j$, pour $1 \leq i \leq j \leq d$.

On note $\wedge^2 V$ le *carré extérieur* de V ; c'est le sous-espace de $V \otimes V$ engendré par les tenseurs alternés. C'est un espace de dimension $\frac{d(d-1)}{2}$ dont une base est constituée des $e_i \wedge e_j$, pour $1 \leq i < j \leq d$.

De plus, on a $V \otimes V = \text{Sym}^2 V \oplus \wedge^2 V$.

Exemple I.3.6. — Si V^* est le dual de V , alors $\text{Sym}^2 V^*$ est l'espace des formes bilinéaires symétriques sur V et $\wedge^2 V^*$ est celui des formes bilinéaires alternées.

Proposition I.3.7. — Si V est une représentation de G , et si $g \in G$, alors

$$\chi_{\text{Sym}^2 V}(g) = \frac{1}{2}(\chi_V(g)^2 + \chi_V(g^2)) \quad \text{et} \quad \chi_{\wedge^2 V}(g) = \frac{1}{2}(\chi_V(g)^2 - \chi_V(g^2)).$$

Démonstration. — Choisissons une base (e_1, \dots, e_d) de V formée de vecteurs propres de g . On a alors $g \cdot e_i = \lambda_i e_i$, $g^2 \cdot e_i = \lambda_i^2 e_i$, et $g \cdot e_i e_j = \lambda_i \lambda_j e_i e_j$, $g \cdot e_i \wedge e_j = \lambda_i \lambda_j e_i \wedge e_j$. On en déduit que

$$\chi_{\text{Sym}^2 V}(g) = \sum_{i \leq j} \lambda_i \lambda_j \quad \text{et} \quad \chi_{\wedge^2 V}(g) = \sum_{i < j} \lambda_i \lambda_j,$$

et comme

$$\chi_V(g^2) = \sum_i \lambda_i^2 \quad \text{et} \quad \chi_V(g)^2 = \left(\sum_i \lambda_i \right)^2 = \sum_i \lambda_i^2 + 2 \sum_{i < j} \lambda_i \lambda_j,$$

le résultat s'en déduit.

Remarque I.3.8. — Ce que l'on a fait avec deux copies de la même représentation V de G peut se généraliser à n copies de V . On note $\otimes^n V$ le produit tensoriel de n copies de V (avec une définition évidente). On note S_n le groupe des permutations de $\{1, \dots, n\}$, et $\text{sign} : S_n \rightarrow \{\pm 1\}$ la signature. Un tenseur symétrique est un tenseur de la forme

$$x_1 \cdots x_n = \frac{1}{n!} \sum_{\sigma \in S_n} x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)},$$

et un tenseur alterné est un tenseur de la forme

$$x_1 \wedge \cdots \wedge x_n = \sum_{\sigma \in S_n} \text{sign}(\sigma) x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)}.$$

La *puissance symétrique n -ième* de V est le sous-espace $\text{Sym}^n V$ de $\otimes^n V$ engendré par les tenseurs symétriques, et la *puissance extérieure n -ième* de V est le sous-espace $\wedge^n V$ de $\otimes^n V$ engendré par les tenseurs alternés. Alors $\text{Sym}^n V$ et $\wedge^n V$ sont des représentations de G de dimensions respectives $\binom{d+n-1}{n}$ et $\binom{d}{n}$. ($\text{Sym}^n V \oplus \wedge^n V$ est un sous-espace strict de $\otimes^n V$, dès que $n \geq 3$.)

En particulier, $\wedge^n V = 0$ si $n > d$, et $\wedge^d V$ est de dimension 1 ; cette représentation est souvent notée $\det V$, l'action de g sur $\det V$ étant la multiplication par $\det \rho_V(g)$.

2. Représentations induites

2.1. Caractère d'une représentation induite

Soit H un sous-groupe de G , et soit V une représentation de H . On définit l'espace vectoriel $\text{Ind}_H^G V$ par

$$\text{Ind}_H^G V = \{ \varphi : G \rightarrow V, \varphi(hx) = h \cdot \varphi(x), \text{ quels que soient } h \in H \text{ et } x \in G \}.$$

Soit $S \subset G$ un système de représentants de $H \backslash G$. Si $x \in G$, il existe alors un unique $h_x \in H$ tel que $h_x^{-1}x \in S$. Ceci permet d'établir un isomorphisme de $\text{Ind}_H^G V$ sur l'espace V^S des applications de S dans V , en envoyant φ sur $(\varphi(s))_{s \in S}$; la bijection réciproque envoie $(v_s)_{s \in S} \in V^S$ sur l'application $\varphi : G \rightarrow V$ définie par $\varphi(x) = h_x \cdot v_{h_x^{-1}x}$. Pour vérifier que φ est bien un élément de $\text{Ind}_H^G V$, il suffit de constater que, si $h \in H$, alors $h_{hx} = hh_x$, et donc

$$\varphi(hx) = hh_x \cdot v_{(hh_x)^{-1}hx} = hh_x \cdot v_{h_x^{-1}x} = h \cdot (h_x \cdot v_{h_x^{-1}x}) = h \cdot \varphi(x).$$

On munit $\text{Ind}_H^G V$ d'une action de G en définissant $g \cdot \varphi$ comme la fonction $x \mapsto \varphi(xg)$. Si $h \in H$, on a

$$(g \cdot \varphi)(hx) = \varphi(hxg) = h \cdot \varphi(xg) = h \cdot ((g \cdot \varphi)(x)),$$

ce qui prouve que $g \cdot \varphi$ est bien élément de $\text{Ind}_H^G V$. De plus, si $g_1, g_2 \in G$, alors

$$(g_1 \cdot (g_2 \cdot \varphi))(x) = (g_2 \cdot \varphi)(xg_1) = \varphi(xg_1g_2) = (g_1g_2 \cdot \varphi)(x),$$

ce qui prouve que l'on a bien défini une action de groupe de G sur $\text{Ind}_H^G V$. La représentation de G ainsi obtenue est la *représentation induite de H à G de la représentation V* . L'isomorphisme de $\text{Ind}_H^G V$ sur V^S montre que la dimension de $\text{Ind}_H^G V$ est $|S| \cdot \dim V = \frac{|G|}{|H|} \dim V$.

Par exemple, si $H = 1$, et $V = \mathbf{1}$ est la représentation triviale, la représentation $\text{Ind}_H^G V$ est l'espace des fonctions $\varphi : G \rightarrow \mathbf{C}$. Il admet comme base les fonctions φ_h , pour $h \in G$, définies par $\varphi_h(x) = 1$, si $xh = 1$, et $\varphi_h(x) = 0$, si $xh \neq 1$. Si $g \in G$, on a alors $(g \cdot \varphi_h)(x) = \varphi_h(xg) = \varphi_{gh}(x)$. On en déduit que $\text{Ind}_{\{1\}}^G \mathbf{1}$ est la *représentation régulière de G* .

Remarque I.3.9. — Les représentations induites à partir de représentations de sous-groupes sont la principale source de représentations d'un groupe G . L'un de leurs intérêts est que leur caractère se calcule très facilement (cf. annexe B pour un certain nombre d'applications).

Théorème I.3.10. — Soient $H \subset G$ deux groupes finis, $S \subset G$ un système de représentants de $H \backslash G$, V une représentation de H , et $W = \text{Ind}_H^G V$. Alors, quel que soit $g \in G$, on a

$$\chi_W(g) = \sum_{\substack{s \in S \\ sgs^{-1} \in H}} \chi_V(sgs^{-1}) = \frac{1}{|H|} \sum_{\substack{s \in G \\ sgs^{-1} \in H}} \chi_V(sgs^{-1}).$$

Démonstration. — On utilise l'isomorphisme de W avec $V^S = \bigoplus_{s \in S} V_s$. Dans cet isomorphisme, si φ est l'image de $(v_s)_{s \in S}$, on a $\varphi(x) = h_x \cdot v_{h_x^{-1}x}$, et $g \cdot \varphi$ est l'image de $((g \cdot \varphi)(s))_{s \in S} = (\varphi(sg))_{s \in S}$, ce qui fait que l'on obtient

$$g \cdot (v_s)_{s \in S} = (h_{sg} \cdot v_{h_{sg}^{-1}sg})_{s \in S}.$$

Choisissons une base $(e_i)_{i \in I}$ de V , et notons $e_{i,s}$ l'élément $(v_t)_{t \in S}$ de V^S défini par $v_s = e_i$, et $v_t = 0$, si $t \neq s$. Les $e_{i,s}$, pour $i \in I$, forment une base de V_s , et les $e_{i,s}$, pour $(i, s) \in I \times S$, forment une base de V^S . La matrice de g dans cette base est constituée de blocs indexés par $(s, s') \in S \times S$, le bloc correspondant à (s, s') étant nul sauf si $s' = h_{sg}^{-1}sg$. En particulier, les seuls blocs qui vont contribuer à la trace sont ceux pour lesquels $s = h_{sg}^{-1}sg$, ce qui peut se réécrire sous la forme $h_{sg} = sgs^{-1}$. L'action de g sur le V_s correspondant coïncide alors avec celle de $h_{sg} = sgs^{-1}$, et sa contribution à la trace est donc $\chi_V(sgs^{-1})$. On en déduit la première égalité du théorème. La seconde s'en déduit en remarquant que

$\chi_V(hs g (hs)^{-1}) = \chi_V(h(sgs^{-1})h^{-1}) = \chi_V(sgs^{-1})$, si $h \in H$ et $sgs^{-1} \in H$, et en écrivant $s \in G$ sous la forme $h_s^{-1}s$.

Exercice I.3.11. — Retrouver la formule $\dim(\text{Ind}_H^G V) = \frac{|G|}{|H|} \dim V$ en utilisant le th. I.3.10.

2.2. La formule de réciprocité de Frobenius

Soient $H \subset G$ deux groupes finis. On définit des applications linéaires

$$\text{Res}_G^H : R_C(G) \rightarrow R_C(H) \quad \text{et} \quad \text{Ind}_H^G : R_C(H) \rightarrow R_C(G),$$

de la manière suivante. Si $\phi \in R_C(G)$, alors $\text{Res}_G^H \phi$ est juste la fonction centrale sur H , restriction de ϕ à H , et si $\phi \in R_C(H)$, alors $\text{Ind}_H^G \phi$ est la fonction centrale sur G donnée par la formule

$$(\text{Ind}_H^G \phi)(g) = \frac{1}{|H|} \sum_{\substack{s \in G \\ sgs^{-1} \in H}} \phi(sgs^{-1}).$$

Il est immédiat que, si W est une représentation de G , alors $\text{Res}_G^H \chi_W$ est le caractère de la représentation de H obtenue en ne considérant que l'action du sous-groupe H de G . Dans l'autre sens, si V est une représentation de H , alors $\text{Ind}_H^G \chi_V$ est, d'après le th. I.3.10, le caractère de la représentation induite $\text{Ind}_H^G V$.

Pour les distinguer, on note \langle , \rangle_H et \langle , \rangle_G les produits scalaires sur $R_C(H)$ et $R_C(G)$. On a alors le résultat suivant.

Théorème I.3.12. — (formule de réciprocité de Frobenius) *Si $\phi_1 \in R_C(H)$ et $\phi_2 \in R_C(G)$, alors*

$$\langle \phi_1, \text{Res}_G^H \phi_2 \rangle_H = \langle \text{Ind}_H^G \phi_1, \phi_2 \rangle_G.$$

Démonstration. — Par définition, on a

$$\begin{aligned} \langle \text{Ind}_H^G \phi_1, \phi_2 \rangle_G &= \frac{1}{|G|} \sum_{g \in G} \overline{\text{Ind}_H^G \phi_1(g)} \phi_2(g) \\ &= \frac{1}{|G|} \sum_{g \in G} \left(\frac{1}{|H|} \sum_{\substack{s \in G \\ sgs^{-1} \in H}} \overline{\phi_1(sgs^{-1})} \right) \phi_2(g). \end{aligned}$$

En posant $h = sgs^{-1}$, et donc $g = s^{-1}hs$, on peut réécrire la somme ci-dessus sous la forme

$$\frac{1}{|G| |H|} \sum_{h \in H, s \in G} \overline{\phi_1(h)} \phi_2(s^{-1}hs),$$

et comme ϕ_2 est une fonction centrale sur G , on a $\phi_2(s^{-1}hs) = \phi_2(h)$, quel que soit $s \in G$. On obtient donc

$$\langle \text{Ind}_H^G \phi_1, \phi_2 \rangle_G = \frac{1}{|H|} \sum_{h \in H} \overline{\phi_1(h)} \phi_2(h) = \langle \phi_1, \text{Res}_G^H \phi_2 \rangle_H,$$

ce qui permet de conclure.

Exercice I.3.13. — (i) Montrer que, si V, V_1, V_2 sont des représentations de G , alors

$$\begin{aligned}\mathrm{Hom}_G(V, V_1 \oplus V_2) &= \mathrm{Hom}_G(V, V_1) \oplus \mathrm{Hom}_G(V, V_2) \\ \mathrm{Hom}_G(V_1 \oplus V_2, V) &= \mathrm{Hom}_G(V_1, V) \oplus \mathrm{Hom}_G(V_2, V).\end{aligned}$$

(ii) En déduire que, si V et V' sont deux représentations de G , alors $\dim(\mathrm{Hom}_G(V, V')) = \langle \chi_V, \chi_{V'} \rangle$.

(iii) Soit H un sous-groupe de G , et soient W une représentation de H et V une représentation de G . Si $u \in \mathrm{Hom}_H(W, \mathrm{Res}_G^H V)$, on note $\alpha_u : \mathrm{Ind}_H^G W \rightarrow V$ l'application qui à un élément $\phi : G \rightarrow W$ de $\mathrm{Ind}_H^G W$, associe $\alpha_u(\phi) = \frac{1}{|G|} \sum_{g \in G} g^{-1} u(\phi(g))$. Montrer que $\alpha_u \in \mathrm{Hom}_G(\mathrm{Ind}_H^G W, V)$.

(iv) Montrer que $u \mapsto \alpha_u$, de $\mathrm{Hom}_H(W, \mathrm{Res}_G^H V)$ dans $\mathrm{Hom}_G(\mathrm{Ind}_H^G W, V)$, est une injection linéaire. En déduire que c'est un isomorphisme (réciprocité de Frobenius pour les représentations).

2.3. Transitivité des inductions

Proposition I.3.14. — Soient $K \subset H \subset G$ des groupes finis.

- (i) Si $\phi \in R_G(K)$, alors $\mathrm{Ind}_H^G(\mathrm{Ind}_K^H \phi) = \mathrm{Ind}_K^G \phi$.
- (ii) Si W est une représentation de K , alors $\mathrm{Ind}_H^G(\mathrm{Ind}_K^H W) = \mathrm{Ind}_K^G W$.

Démonstration. — Le (ii) est, modulo le fait qu'une représentation est déterminée par son caractère (cor. I.2.15), une conséquence du (i) appliqué à $\phi = \chi_W$. Pour démontrer le (i), on part de la formule

$$\begin{aligned}\mathrm{Ind}_H^G(\mathrm{Ind}_K^H \phi)(g) &= \frac{1}{|H|} \sum_{\substack{s \in G \\ sgs^{-1} \in H}} (\mathrm{Ind}_K^H \phi)(sgs^{-1}) \\ &= \frac{1}{|H|} \sum_{\substack{s \in G \\ sgs^{-1} \in H}} \frac{1}{|K|} \sum_{\substack{h \in H \\ hsgs^{-1}h^{-1} \in K}} \phi(hsgs^{-1}h^{-1}).\end{aligned}$$

On pose alors $hs = t$ de telle sorte que $s = h^{-1}t$, ce qui permet de réécrire la somme ci-dessus sous la forme

$$\frac{1}{|H| \cdot |K|} \sum_{\substack{h \in H, t \in G \\ tgt^{-1} \in K, h^{-1}tgt^{-1}h \in H}} \phi(tgt^{-1}).$$

Comme la condition $h^{-1}tgt^{-1}h \in H$ est automatique, si $tgt^{-1} \in K$ et $h \in H$, la somme ci-dessus se simplifie et devient

$$\frac{1}{|K|} \sum_{\substack{t \in G \\ tgt^{-1} \in K}} \phi(tgt^{-1}) = (\mathrm{Ind}_K^G \phi)(g),$$

ce qui permet de conclure.

2.4. Les théorèmes d'Artin et de Brauer

Théorème I.3.15. — (Artin, 1930) Soit G un groupe fini, et soit V une représentation de G , alors il existe un entier non nul d_V , et une famille finie de couples (C_i, χ_i) , pour $i \in I$, où C_i est un sous-groupe cyclique de G , et $\chi_i \in \widehat{C}_i$ est un caractère linéaire de C_i , tels que l'on ait

$$d_V \chi_V = \sum_{i \in I} n_i \mathrm{Ind}_{C_i}^G \chi_i, \quad \text{avec } n_i \in \mathbf{Z}, \text{ si } i \in I.$$

Démonstration. — Commençons par démontrer que les $\mathrm{Ind}_C^G \chi$, où C décrit les sous-groupes cycliques de G , et χ les éléments de \widehat{C} , forment une famille génératrice de $R_G(G)$. Dans le cas contraire, il existe $\phi \in R_G(G)$ non nulle, orthogonale à tous les $\mathrm{Ind}_C^G \chi$. En utilisant la formule de réciprocité de Frobenius, on en déduit que $\langle \phi, \chi \rangle_C = 0$, quel que soit $\chi \in \widehat{C}$. Soit alors $c \in G$. Le sous-groupe C de G engendré par c est cyclique par définition. Comme un groupe cyclique est en particulier commutatif, les $\chi \in \widehat{C}$ engendrent l'espace vectoriel des fonctions de C dans \mathbf{C} (cor. I.2.21). Si $\phi_c : C \rightarrow \mathbf{C}$ est la fonction valant

1 en c , et 0 ailleurs, on a donc $0 = \langle \phi, \phi_c \rangle_C = \frac{1}{|C|} \overline{\phi(c)}$, et donc $\phi(c) = 0$. On en déduit le fait que ϕ est identiquement nulle, ce qui permet de prouver notre affirmation selon laquelle les $\text{Ind}_C^G \chi$ forment une famille génératrice de $R_C(G)$.

Extrayons-en une base $e_i = \text{Ind}_{C_i}^G \chi_i$, pour $i \in I$. Si $\chi \in \text{Irr}(G)$, on a $\langle \chi, e_i \rangle \in \mathbf{N}$, puisque $\langle \chi, e_i \rangle$ est la multiplicité de la représentation correspondant à χ dans la décomposition de $\text{Ind}_{C_i}^G \chi_i$ en représentations irréductibles. De plus, on a $e_i = \sum_{\chi \in \text{Irr}(G)} \langle \chi, e_i \rangle \chi$. La matrice de passage de la base des e_i , pour $i \in I$, à celle des χ , pour $\chi \in \text{Irr}(G)$, est donc à coefficients rationnels, et comme χ_V a, pour la même raison que précédemment, des coordonnées entières dans la base des χ , pour $\chi \in \text{Irr}(G)$, cela implique que χ_V a des coordonnées rationnelles dans la base des e_i , pour $i \in I$. Il suffit alors de prendre pour d_V le p.p.c.m. des dénominateurs des coordonnées de χ_V dans la base des e_i , pour $i \in I$, pour obtenir la décomposition voulue. Ceci permet de conclure.

Théorème I.3.16. — (R. Brauer, 1947) *Soit G un groupe fini, et soit V une représentation de G , alors il existe une famille finie de couples (H_i, χ_i) , pour $i \in I$, où H_i est un sous-groupe de G , et $\chi_i \in \widehat{H_i}$ est un caractère linéaire de H_i , tels que l'on ait*

$$\chi_V = \sum_{i \in I} n_i \text{Ind}_{H_i}^G \chi_i, \quad \text{avec } n_i \in \mathbf{Z}, \text{ si } i \in I.$$

La principale différence avec le théorème d'Artin est la disparition de l'entier d_V , ce qui a des conséquences assez remarquables (une de ces conséquences est évoquée au n° 3 du § D.1). L'autre différence est que l'on ne peut pas se restreindre aux groupes cycliques (R. Brauer montre que l'on peut se restreindre aux *groupes élémentaires* : un groupe fini H est dit élémentaire s'il existe un nombre premier p tel que H soit le produit d'un p -groupe (un groupe d'ordre une puissance de p) par un groupe cyclique d'ordre premier à p). La démonstration demande d'utiliser des propriétés d'intégralité des $\chi_V(g)$, et déborde un peu du cadre de ce cours. Signalons que ces propriétés d'intégralité permettent aussi de démontrer le résultat suivant.

Proposition I.3.17. — *Si G est un groupe fini et si V est une représentation irréductible de G , alors $\dim V$ divise $|G|$.*

3. Exercices

On rappelle que S_n (resp. A_n) désigne le groupe symétrique (resp. alterné), cf. alinéa 4.5 du § 1 du chapitre « Vocabulaire mathématique ».

Exercice I.3.18. — Soit n un entier ≥ 1 . Quelles sont les représentations irréductibles de $\mathbf{Z}/n\mathbf{Z}$?

Exercice I.3.19. — (i) Montrer qu'un groupe non commutatif d'ordre 6 a deux représentations irréductibles de dimension 1 et une de dimension 2.

(ii) Dresser la table des caractères de S_3 .

Exercice I.3.20. — (i) Montrer qu'un groupe non commutatif d'ordre 8 a quatre représentations irréductibles de dimension 1 et une de dimension 2.

(ii) Soit D_4 le groupe des symétries du carré. Montrer que D_4 est d'ordre 8, et dresser la table des caractères de D_4 .

(iii) Soit H_4 le groupe des quaternions. C'est l'ensemble des matrices $\begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$, avec $\{a, b\} \subset \{0, 1, -1, i, -i\}$, et $a = 0$ ou $b = 0$. Montrer que H_4 est un groupe d'ordre 8 non isomorphe à D_4 , et dresser sa table de caractères.

Exercice I.3.21. — Soient G_1, G_2 deux groupes finis, et soit $G = G_1 \times G_2$.

(i) Montrer que, si V_1 et V_2 sont des représentations irréductibles de G_1 et G_2 , alors $V_1 \boxtimes V_2$ (cf. rem. I.3.4) est une représentation irréductible de G .

(ii) Montrer que toute représentation irréductible de G est obtenue de cette manière.

Exercice I.3.22. — (o) Quelles sont les classes de conjugaison de S_4 ?

(i) Montrer que $C = \{1, (12)(34), (13)(24), (14)(23)\}$ est un sous-groupe distingué de S_4 , et que $S_4/C = S_3$. (On pourra faire agir S_4 sur $C - \{1\}$ par conjugaison.)

(ii) En déduire l'existence d'une représentation irréductible de S_4 de dimension 2 et de deux de dimension 1.

(iii) On fait agir S_4 sur \mathbf{C}^4 par permutation des éléments de la base canonique. Montrer que l'hyperplan $V = \{x_1 + x_2 + x_3 + x_4 = 0\}$ est stable par S_4 , et calculer le caractère χ_V .

(iv) Montrer que V est irréductible et non isomorphe à $V \otimes \text{sign}$. En déduire la table des caractères de S_4 .

Exercice I.3.23. — (o) Montrer que S_5 a 7 classes de conjugaison, et calculer le cardinal de chaque classe.

(i) On note U la représentation de S_5 sur l'hyperplan $\sum_{i=1}^5 x_i = 0$ de \mathbf{C}^5 . Calculer χ_U , et montrer que U et $U \otimes \text{sign}$ sont irréductibles non isomorphes.

(ii) Calculer $\chi_{\wedge^2 U}$ et montrer que $\wedge^2 U$ est irréductible.

(iii) Calculer $\chi_{\text{Sym}^2 U}$ et montrer que $\text{Sym}^2 U = \mathbf{1} \oplus U \oplus V$, où V est irréductible.

(iv) Dresser la table des caractères de S_5 .

Exercice I.3.24. — Soit $n \geq 3$, et soient D_n le groupe des symétries d'un polygone régulier à n sommets et $C_n \subset D_n$ le sous-groupe des rotations.

(i) Montrer que C_n est un groupe cyclique d'indice 2 dans D_n . Montrer que, si n est pair (resp. impair), les symétries forment deux (resp. une) classes de conjugaison, et les rotations $\frac{n}{2} + 1$ (resp. $\frac{n+1}{2}$).

(ii) Montrer que, si on identifie la rotation d'angle α avec la multiplication par $e^{i\alpha}$ dans le plan complexe, les représentations irréductibles de C_n sont les χ_a , pour $a \in \mathbf{Z}/n\mathbf{Z}$, définies par $\chi_a(e^{i\alpha}) = e^{ai\alpha}$.

(iii) Si $a \in \mathbf{Z}/n\mathbf{Z}$, calculer le caractère ϕ_a de $\text{Ind}_{C_n}^{D_n} \chi_a$.

(iv) Calculer $\langle \phi_a, \phi_a \rangle$; en déduire dans quel cas $\text{Ind}_{C_n}^{D_n} \chi_a$ est irréductible.

(v) Dresser la table des caractères de D_n .

Exercice I.3.25. — Soit G un groupe fini, et soit V une représentation fidèle de G (i.e. $\rho_V(g) \neq 1$ si $g \neq 1$).

(i) Montrer que $\chi_V(g) \neq \dim V$, si $g \neq 1$.

- (ii) Soit W une représentation irréductible de G . Montrer que $\sum_{n=0}^{+\infty} \langle \chi_W, \chi_V^n \rangle T^n$ est une fraction rationnelle que l'on explicitera, mais n'est pas un polynôme.
- (iii) En déduire que W apparaît dans la décomposition en représentations irréductibles d'une infinité de $\otimes^n V$.

Exercice I.3.26. — Soit G un groupe fini, soit $H \neq G$ un sous-groupe de G , et soit V la représentation de permutation associée à l'action de G sur G/H .

- (i) Montrer que $V = \text{Ind}_H^G \mathbf{1}$. En déduire que $\sum_{g \in G} \chi_V(g) = |G|$.
- (ii) Montrer que V n'est pas irréductible; en déduire que $\frac{1}{|G|} \sum_{g \in G} \chi_V(g)^2 \geq 2$. (On remarquera que χ_V est à valeurs réelles.)
- (iii) Soit Y l'ensemble des $g \in G$ vérifiant $\chi_V(g) = 0$. Montrer que

$$\sum_{g \in G} (\chi_V(g) - 1)(\chi_V(g) - |G/H|) \leq |G/H| \cdot |Y|.$$

- (iv) En déduire que $|Y| \geq |H|$.
- (v) Soit X , avec $|X| \geq 2$, un ensemble sur lequel G agit transitivement (i.e., quels que soient $x, y \in X$, il existe $g \in G$, tel que $y = g \cdot x$). Montrer que la proportion des $g \in G$ agissant sans point fixe sur X est supérieure⁽¹¹⁾ ou égale à $1/|X|$.

Exercice I.3.27. — Soit G un groupe fini.

- (i) Si $\sigma \in S_n$, soit M_σ la matrice de l'isomorphisme u_σ de \mathbf{C}^n , envoyant l'élément e_i de la base canonique sur $e_{\sigma(i)}$. Quelles sont, en fonction de la décomposition de σ en cycles, les valeurs propres de M_σ (avec multiplicité). En déduire que, si M_σ et M_τ sont semblables, alors σ et τ ont le même nombre de points fixes.
- (ii) Montrer que la matrice T_G , définie par la table des caractères de G , est inversible.
- (iii) Montrer que, si $C \in \text{Conj}(G)$, alors $C^{-1} = \{g^{-1}, g \in C\}$ appartient à $\text{Conj}(G)$, et que, si $\chi \in \text{Irr}(G)$, alors $\chi(C^{-1}) = \overline{\chi(C)}$.
- (iv) Montrer que le nombre de classes de conjugaison symétriques de G (i.e. vérifiant $C = C^{-1}$) est égal au nombre de caractères irréductibles réels de G (i.e. $\chi(C) \in \mathbf{R}$, pour tout $C \in \text{Conj}(G)$).

⁽¹¹⁾En théorie algébrique des nombres, ce résultat de Jordan permet de montrer que si $P \in \mathbf{Z}[T]$, de degré ≥ 2 , est irréductible dans $\mathbf{Q}[T]$, alors il existe une infinité de nombres premiers p tels que P n'ait aucune solution dans \mathbf{F}_p .

CHAPITRE II

ESPACES DE BANACH

La théorie des espaces vectoriels normés complets (appelés « espaces de Banach » en raison du rôle joué par ce dernier dans sa mise en forme) est issue des travaux du 19-ième siècle sur les équations différentielles, les équations aux dérivées partielles ou les équations intégrales du type $u(x) + \int_a^b K(x, y)u(y) dy = f(x)$, où u est une fonction inconnue. Il s'est écoulé une vingtaine d'années entre l'introduction par D. Hilbert de l'espace qui porte son nom (l'espace ℓ^2 des suites de carré sommable), la réalisation l'année suivante, par E. Fischer et F. Riesz, de ce que l'espace des fonctions de carré sommable lui était isomorphe, et la forme définitive de la théorie par l'école polonaise (S. Banach, H. Hahn, H. Steinhaus). En retour, cette théorie a permis de nombreuses avancées sur les problèmes qui l'ont motivée. Le problème de la classification des espaces de Banach est toujours d'actualité⁽¹⁾.

II.1. Espaces de Banach

Dans tout ce qui suit, \mathbf{K} désigne soit le corps \mathbf{R} des nombres réels, soit le corps \mathbf{C} des nombres complexes, et « espace vectoriel » est une abréviation pour « espace vectoriel sur \mathbf{K} ». Le lecteur est renvoyé au § 11 du chapitre « Vocabulaire mathématique » pour le vocabulaire et les propriétés élémentaires des espaces vectoriels normés.

1. Convergence normale, séries sommables

Soit $(E, \| \cdot \|)$ un espace vectoriel normé. On rappelle que, si $x \in E$ et si $r \in \mathbf{R}_+$, on note $B(x, r)$ ou $B_E(x, r)$ (resp. $B(x, r^-)$ ou $B_E(x, r^-)$) la boule fermée (resp. ouverte) de centre x et de rayon r .

Une série $\sum_{n \in \mathbf{N}} u_n$ d'éléments de E est *normalement convergente*, si $\sum_{n \in \mathbf{N}} \|u_n\| < +\infty$.

⁽¹⁾Un problème qui est resté ouvert pendant longtemps était de savoir si un endomorphisme continu d'un \mathbf{C} -espace de Banach possède toujours un sous-espace fermé invariant (c'est le cas en dimension finie ≥ 2). Un contreexemple a été construit par P. Enflo vers 1981, et C. Read (1984) en a construit un dans l'espace ℓ^1 des suites sommables, mais on ne sait pas s'il existe des contreexemples dans ℓ^2 ou, plus généralement, dans des espaces réflexifs (isomorphes au dual de leur dual), ce que ℓ^1 n'est pas.

Un espace vectoriel normé $(E, \|\cdot\|)$, qui est complet (pour la distance associée à la norme), est appelé un *espace de Banach*. D'après le n° 9.1 du chap. « Vocabulaire mathématique », $(E, \|\cdot\|)$ est un espace de Banach si et seulement si toute série normalement convergente converge dans E . Comme un sous-espace fermé d'un espace complet est complet, un sous-espace vectoriel fermé d'un espace de Banach est un espace de Banach.

Les exemples les plus simples d'espaces de Banach sont les espaces de dimension finie, mais ceux-ci ont des propriétés très spéciales. On a en particulier les résultats classiques suivants (cf. n°s II.1.1 et II.1.4 du chap. « Vocabulaire mathématique »).

Proposition II.1.1. — (i) Si V est un espace vectoriel de dimension finie, alors toutes les normes sur V sont équivalentes et V est complet pour n'importe laquelle d'entre elles.

(ii) Soit E un espace de Banach. La boule unité fermée $B(0, 1)$ est compacte si et seulement si E est de dimension finie.

Remarque II.1.2. — Insistons sur le fait que le (i) de la prop. II.1.1 devient *totale*ment faux en dimension infinie : les normes sur un espace E de dimension infinie ne sont pas toutes équivalentes, et E peut être complet pour certaines d'entre elles, mais il y en a “beaucoup plus” pour lesquelles ce n'est pas le cas.

Un espace est dit *séparable* s'il contient un sous-ensemble dénombrable dense⁽²⁾ (i.e. s'il est « pas trop gros »).

Exercice II.1.3. — (i) Montrer qu'un espace vectoriel de dimension finie est séparable.

(ii) Montrer qu'un sous-espace d'un espace séparable est un espace séparable.

Soit E un espace de Banach. Une série $\sum_{i \in I} x_i$, avec I dénombrable, est dite *sommable* si elle vérifie le critère de Cauchy non ordonné suivant : pour tout $\varepsilon > 0$, il existe $I(\varepsilon) \subset I$ fini, tel que pour tout $J \subset I$ fini avec $J \cap I(\varepsilon) = \emptyset$, on ait $\|\sum_{i \in J} x_i\| \leq \varepsilon$.

- Si I est fini, toute série $\sum_{i \in I} x_i$ est sommable.
- Si I est infini, si $\sum_{i \in I} x_i$ est sommable, alors pour toute bijection $n \mapsto i(n)$ de \mathbf{N} sur I , les sommes partielles de la série $\sum_{n=0}^{+\infty} x_{i(n)}$ vérifient le critère de Cauchy usuel, et comme E est complet, la série converge et la limite ne dépend pas du choix de la bijection $n \mapsto i(n)$; c'est la somme de la série $\sum_{i \in I} x_i$.

La sommabilité et la convergence normale sont des notions assez proches. L'exercice ci-dessous (dans lequel on s'intéresse aux séries $\sum_{i \in I} x_i$, où I est un ensemble dénombrable infini) explore leurs liens.

Exercice II.1.4. — (i) Montrer qu'une série normalement convergente est sommable.

⁽²⁾La plupart des espaces de Banach de l'analyse fonctionnelle sont séparables ; une exception notable étant l'espace $\mathcal{C}_b(\mathbf{R})$ des fonctions continues bornées sur \mathbf{R} (cf. Ex. II.1.7) ou l'espace ℓ^∞ des suites bornées.

(ii) Montrer que dans \mathbf{R} une série est sommable si et seulement si elle est absolument convergente.

(iii) En déduire que dans un espace vectoriel normé, de dimension finie sur \mathbf{R} ou \mathbf{C} , une série est sommable si et seulement si elle est normalement convergente.

(iv) Soit e_i la suite $(e_{i,j})_{j \in \mathbf{J}}$ d'éléments de \mathbf{C} , définie par $e_{i,i} = 1$ et $e_{i,j} = 0$, si $i \neq j$.

(a) Montrer que $\sum_{i \in \mathbf{I}} a_i e_i$ est sommable dans $\ell^\infty(\mathbf{I})$ si et seulement si la suite $(a_i)_{i \in \mathbf{I}}$ tend vers 0 à l'infini (pour tout $\varepsilon > 0$, il existe $\mathbf{I}(\varepsilon) \subset \mathbf{I}$ fini tel que $|a_i| \leq \varepsilon$, si $i \notin \mathbf{I}(\varepsilon)$).

(b) Montrer que $\sum_{i \in \mathbf{I}} a_i e_i$ est sommable dans $\ell^2(\mathbf{I})$ si et seulement si $\sum_{i \in \mathbf{I}} |a_i|^2 < +\infty$.

(c) Quelle est la somme dans ces deux cas ?

2. Espaces de suites

Exemple II.1.5. — (i) On note ℓ^∞ l'ensemble des suites bornées $(x_n)_{n \in \mathbf{N}}$. Muni de la norme $\| \cdot \|_\infty$ définie par $\|(x_n)_{n \in \mathbf{N}}\|_\infty = \sup_{n \in \mathbf{N}} |x_n|$, l'espace ℓ^∞ est un espace de Banach. L'espace ℓ_0^∞ , sous-espace de ℓ^∞ des suites tendant vers 0 quand n tend vers $+\infty$ est un sous-espace fermé de ℓ^∞ , et donc aussi un Banach.

(ii) On note ℓ^1 l'ensemble des suites $(x_n)_{n \in \mathbf{N}}$, telles que $\sum_{n \in \mathbf{N}} |x_n| < +\infty$. Muni de la norme $\| \cdot \|_1$ définie par $\|(x_n)_{n \in \mathbf{N}}\|_1 = \sum_{n \in \mathbf{N}} |x_n|$, l'espace ℓ^1 est un espace de Banach.

(iii) On note ℓ^2 l'ensemble des suites $(x_n)_{n \in \mathbf{N}}$, telles que $\sum_{n \in \mathbf{N}} |x_n|^2 < +\infty$. Muni de la norme $\| \cdot \|_2$ définie par $\|(x_n)_{n \in \mathbf{N}}\|_2 = (\sum_{n \in \mathbf{N}} |x_n|^2)^{1/2}$, l'espace ℓ^2 est un espace de Banach⁽³⁾.

(iv) Si \mathbf{I} est un ensemble dénombrable infini, on définit de même les espaces $\ell^\infty(\mathbf{I})$, $\ell_0^\infty(\mathbf{I})$, $\ell^1(\mathbf{I})$ et $\ell^2(\mathbf{I})$; ce sont aussi des espaces de Banach.

Démonstration. — Le cas \mathbf{I} dénombrable se déduit du cas de \mathbf{N} en choisissant une bijection entre \mathbf{I} et \mathbf{N} ; il suffit donc de démontrer les (i), (ii) et (iii). Soit \mathbf{E} un des espaces ℓ^∞ , ℓ^1 ou ℓ^2 , et soit $\| \cdot \|$ la norme correspondante. Pour prouver que \mathbf{E} est un espace de Banach, il s'agit de vérifier que toute série normalement convergente d'éléments de \mathbf{E} admet une limite dans \mathbf{E} . Soit donc $(x^{(k)})_{k \in \mathbf{N}}$ une suite d'éléments de \mathbf{E} vérifiant $\sum_{k \in \mathbf{N}} \|x^{(k)}\| < +\infty$. Chaque $x^{(k)}$ est une suite $(x_n^{(k)})_{n \in \mathbf{N}}$ d'éléments de \mathbf{K} , et dans les trois cas, on $|x_n^{(k)}| \leq \|x^{(k)}\|$ pour tout $n \in \mathbf{N}$, ce qui fait que, quel que soit $n \in \mathbf{N}$, la série $\sum_{k \in \mathbf{N}} x_n^{(k)}$ est normalement convergente dans \mathbf{K} , et donc converge dans \mathbf{K} (puisque \mathbf{K} est complet); nous noterons y_n la somme de cette série et y la suite $(y_n)_{n \in \mathbf{N}}$. Pour conclure, il suffit de vérifier que $\|y\| \leq \sum_{k \in \mathbf{N}} \|x^{(k)}\|$: en effet, ceci prouve que $y \in \mathbf{E}$, que $\|y - \sum_{i \leq k} x^{(i)}\| \leq \sum_{i \geq k+1} \|x^{(i)}\|$ tend vers 0 quand $k \rightarrow +\infty$, puisque majoré par le reste d'une série convergente, et donc que y est somme de la série $\sum_{k \in \mathbf{N}} x^{(k)}$ dans \mathbf{E} .

• Si $\mathbf{E} = \ell^\infty$, on a $|y_n| \leq \sum_{k \in \mathbf{N}} |x_n^{(k)}| \leq \sum_{k \in \mathbf{N}} \|x^{(k)}\|_\infty$, et donc $\|y\|_\infty \leq \sum_{k \in \mathbf{N}} \|x^{(k)}\|_\infty$. Pour la fermeture de ℓ_0^∞ dans ℓ^∞ , cf. chap. « Vocabulaire mathématique », ex. 10.2.

⁽³⁾Comme la norme $\| \cdot \|_2$ est définie par un produit scalaire hermitien, c'est même un espace de Hilbert.

- Si $E = \ell^1$, on a $|y_n| \leq \sum_{k \in \mathbf{N}} |x_n^{(k)}|$, et donc

$$\|y\|_1 = \sum_{n \in \mathbf{N}} |y_n| \leq \sum_{n \in \mathbf{N}} \sum_{k \in \mathbf{N}} |x_n^{(k)}| = \sum_{k \in \mathbf{N}} \sum_{n \in \mathbf{N}} |x_n^{(k)}| = \sum_{k \in \mathbf{N}} \|x^{(k)}\|_1.$$

- Si $E = \ell^2$, on a

$$\begin{aligned} \sum_{n \in \mathbf{N}} |y_n|^2 &\leq \sum_{n \in \mathbf{N}} \left(\sum_{k \in \mathbf{N}} |x_n^{(k)}| \right)^2 = \sum_{n \in \mathbf{N}} \sum_{k_1, k_2 \in \mathbf{N}} |x_n^{(k_1)}| |x_n^{(k_2)}| \\ &= \sum_{k_1, k_2 \in \mathbf{N}} \sum_{n \in \mathbf{N}} |x_n^{(k_1)}| |x_n^{(k_2)}| = \sum_{k_1, k_2 \in \mathbf{N}} \langle |x^{(k_1)}|, |x^{(k_2)}| \rangle \\ &\leq \sum_{k_1, k_2 \in \mathbf{N}} \|x^{(k_1)}\|_2 \|x^{(k_2)}\|_2 = \left(\sum_{k \in \mathbf{N}} \|x^{(k)}\|_2 \right)^2 \end{aligned}$$

où l'on a utilisé le fait que l'on pouvait réordonner les termes comme on le voulait dans une série à termes positifs, la notation $|x^{(k)}|$ pour désigner la suite $(|x_n^{(k)}|)_{n \in \mathbf{N}}$, et l'inégalité de Cauchy-Schwartz (cf. n° 1 du § II.2). On en déduit que $\|y\|_2 \leq \sum_{k \in \mathbf{N}} \|x^{(k)}\|_2$.

Ceci permet de conclure.

3. Espaces de fonctions continues

Si X est un espace topologique, on note $\mathcal{C}(X)$ l'espace des fonctions continues de X dans \mathbf{C} .

Exemple II.1.6. — (i) Si X est un espace métrique (ou plus généralement un espace topologique), on peut munir l'espace $\mathcal{C}_b(X)$ des fonctions continues bornées de X dans \mathbf{C} de la norme $\|\cdot\|_\infty$ de la *convergence uniforme* définie par $\|\phi\|_\infty = \sup_{x \in X} |\phi(x)|$. Alors $(\mathcal{C}_b(X), \|\cdot\|_\infty)$ est un espace de Banach. En effet, la complétude de $\mathcal{C}_b(X)$ est une traduction du théorème de taupe, selon lequel une limite uniforme de fonctions continues est continue.

(ii) On note $\mathcal{C}_c(X)$ l'espace des fonctions continues sur X , nulles en dehors d'un compact (le « c » en indice signifie « à support compact »). Comme une fonction continue sur un compact a une image compacte et donc bornée, on a $\mathcal{C}_c(X) \subset \mathcal{C}_b(X)$, et cette inclusion est stricte sauf si X est compact auquel cas $\mathcal{C}_c(X) = \mathcal{C}_b(X) = \mathcal{C}(X)$. On note $\mathcal{C}_0(X)$ l'adhérence de $\mathcal{C}_c(X)$ dans $\mathcal{C}_b(X)$; c'est l'espace des fonctions continues sur X tendant vers 0 à l'infini.

Exercice II.1.7. — Si $\lambda \in \mathbf{R}$, on note e_λ la fonction $t \mapsto e^{2i\pi\lambda t}$.

(i) Montrer que $e_\lambda \in \mathcal{C}_b(\mathbf{R})$, et que $\|e_\lambda - e_\mu\|_\infty = 2$ si $\lambda \neq \mu$.

(ii) En déduire que $\mathcal{C}_b(\mathbf{R})$ n'est pas séparable.

Théorème II.1.8. — (de Stone-Weierstrass, Stone (1948)) Si X est un espace compact, et si \mathcal{A} est une sous-algèbre de $\mathcal{C}(X)$ qui contient les fonctions constantes, sépare les points, et est stable par $f \mapsto \bar{f}$, alors \mathcal{A} est dense dans $\mathcal{C}(X)$.

Démonstration. — Avant de faire la démonstration de cet important théorème, explicitons la condition « \mathcal{A} sépare les points » : elle signifie que, si $x \neq y$, on peut trouver $f \in \mathcal{A}$, avec $f(x) \neq f(y)$.

Maintenant, quitte à remplacer \mathcal{A} par son adhérence dans $\mathcal{C}(X)$, qui est encore une algèbre vérifiant les conditions du théorème, on peut supposer que \mathcal{A} est complète et on doit démontrer qu'alors $\mathcal{A} = \mathcal{C}(X)$.

Soit $\mathcal{A}_{\mathbf{R}} = \mathcal{A} \cap \mathcal{C}(X, \mathbf{R})$. C'est une sous-algèbre fermée (et donc complète) de $\mathcal{C}(X, \mathbf{R})$, ensemble des fonctions continues sur X , à valeurs dans \mathbf{R} , et la condition « \mathcal{A} est stable par $f \mapsto \bar{f}$ » entraîne que

$\mathcal{A}_{\mathbf{R}}$ sépare les points puisque $\mathcal{A}_{\mathbf{R}}$ contient $\operatorname{Re}(f) = \frac{1}{2}(f + \bar{f})$ et $\operatorname{Im}(f) = \frac{1}{2i}(f - \bar{f})$, si $f \in \mathcal{A}$. Comme $\mathcal{C}(X) = \mathcal{C}(X, \mathbf{R}) + i\mathcal{C}(X, \mathbf{R})$, il suffit de prouver que $\mathcal{A}_{\mathbf{R}} = \mathcal{C}(X, \mathbf{R})$. Nous aurons besoin du lemme suivant.

Lemme II.1.9. — *Il existe une suite $(P_n(t))_{n \in \mathbf{N}}$ de polynômes à coefficients réels tendant vers $|t|$ uniformément sur $[-1, 1]$.*

Démonstration. — La formule de Taylor avec reste intégral

$$f(x) = f(0) + xf'(0) + \cdots + \frac{x^n}{n!}f^{(n)}(0) + \int_0^x \frac{(x-t)^{n-1}}{(n-1)!}f^{(n)}(t) dt$$

permet de montrer que, si $\alpha > 0$, la série $\sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n$ tend vers $(1+x)^\alpha$ uniformément sur $x \in [-1, 1]$. En particulier, pour $\alpha = \frac{1}{2}$ et $x = t^2 - 1$, les sommes partielles de cette série fournissent une suite de polynômes tendant, uniformément sur $[-1, 1]$, vers $(1+t^2-1)^{1/2} = |t|$.

Revenons à la démonstration du théorème.

• Si $f \in \mathcal{A}_{\mathbf{R}}$, il existe $a \in \mathbf{R}_+^*$ tel que f prenne ses valeurs dans $[-a, a]$. Mais alors $aP_n(a^{-1}f)$ est une suite d'éléments de $\mathcal{A}_{\mathbf{R}}$ tendant uniformément vers $|f|$, et comme $\mathcal{A}_{\mathbf{R}}$ est complète, on en déduit que, si $f \in \mathcal{A}_{\mathbf{R}}$, alors $|f| \in \mathcal{A}_{\mathbf{R}}$.

• Maintenant, si $f, g \in \mathcal{A}_{\mathbf{R}}$, on déduit de ce qui précède, que $\sup(f, g) = \frac{f+g}{2} + \frac{|f-g|}{2}$ et $\inf(f, g) = \frac{f+g}{2} - \frac{|f-g|}{2}$ appartiennent toutes les deux à $\mathcal{A}_{\mathbf{R}}$.

• Soit alors $h \in \mathcal{C}(X, \mathbf{R})$. Fixons $x \in X$ et $\varepsilon > 0$. Comme $\mathcal{A}_{\mathbf{R}}$ sépare les points et contient les constantes, on peut trouver, quel que soit $y \in X$, une fonction $f_y \in \mathcal{A}_{\mathbf{R}}$ vérifiant $f_y(x) = h(x)$ et $f_y(y) = h(y)$. Comme $h - f_y$ est continue, il existe un ouvert U_y contenant y tel que $|h(z) - f_y(z)| < \varepsilon$, si $z \in U_y$. Comme X est compact, et comme les U_y recouvrent X , on peut trouver un sous-ensemble fini Y de X tel que $X = \cup_{y \in Y} U_y$. Alors $g_x = \inf_{y \in Y} f_y$ est un élément de $\mathcal{A}_{\mathbf{R}}$, d'après le point précédent, et g_x vérifie $g_x(x) = h(x)$ et $g_x(z) \leq h(z) + \varepsilon$, quel que soit $z \in X$, puisque z appartient à au moins l'un des U_y , avec $y \in Y$.

• Comme $g_x(x) = h(x)$ et comme g_x est continu, il existe un ouvert V_x contenant x tel que $|g_x(z) - h(z)| \leq \varepsilon$, si $z \in V_x$. Comme ci-dessus, on peut extraire du recouvrement de X par les V_x un sous-recouvrement $(V_x)_{x \in X'}$, avec X' fini. Comme ci-dessus, la fonction $g = \sup_{x \in X'} g_x$ appartient à $\mathcal{A}_{\mathbf{R}}$ et vérifie $g(z) \leq h(z) + \varepsilon$ puisque cette inégalité est vérifiée par tous les g_x , et $g(z) \geq h(z) - \varepsilon$ puisque z appartient à au moins l'un des U_x , avec $x \in X'$.

On a donc construit, quel que soit $\varepsilon > 0$, un élément g de $\mathcal{A}_{\mathbf{R}}$ vérifiant $\|g - h\|_\infty \leq \varepsilon$, ce qui prouve que h est dans l'adhérence de $\mathcal{A}_{\mathbf{R}}$, et donc dans $\mathcal{A}_{\mathbf{R}}$. Ceci permet de conclure.

Exemple II.1.10. — (i) Si I est un intervalle compact de \mathbf{R} , alors les polynômes sont denses dans $\mathcal{C}(I)$ (Weierstrass 1885).

(ii) Plus généralement, si K est un compact de \mathbf{R}^m , les polynômes (en x_1, \dots, x_m)⁽⁴⁾ sont denses dans $\mathcal{C}(K)$.

(iii) Les *polynômes trigonométriques* (i.e. les fonctions de la forme $\sum_{k \in I} a_k e^{2i\pi kt}$, où I décrit les sous-ensembles finis de \mathbf{Z}) sont denses dans l'espace $\mathcal{C}(\mathbf{R}/\mathbf{Z})$ des fonctions continues, périodiques de période 1 (Weierstrass). Ils ne sont pas denses dans $\mathcal{C}([0, 1])$ car un élément de l'adhérence doit vérifier $f(0) = f(1)$, les points 0 et 1 n'étant pas séparés par les polynômes trigonométriques.

⁽⁴⁾Attention au fait que, si D est le disque unité de \mathbf{C} , les polynômes en z ne sont pas denses dans $\mathcal{C}(D)$; en effet, un élément de l'adhérence est une fonction holomorphe, comme nous le verrons. Le problème vient de ce que les polynômes en z ne sont pas stables par $f \mapsto \bar{f}$.

Exercice II.1.11. — Soit $h : \mathbf{R} \rightarrow \mathbf{R}$ la fonction indicatrice de \mathbf{Q} .

(i) Construire une suite double $f_{n,k}$ de fonctions continues de \mathbf{R} dans \mathbf{R} telle que, si n est fixé, alors la suite $f_{n,k}$ tend simplement (i.e. $f_{n,k}(x) \rightarrow g_n(x)$, quel que soit $x \in \mathbf{R}$) vers une fonction g_n quand k tend vers $+\infty$, et la suite g_n tend simplement vers h quand n tend vers $+\infty$. (En bref, h est limite simple de limites simples de fonctions continues).

(ii) Montrer que h n'est pas une limite simple de fonctions continues. (Si h_n est une suite de fonctions continues tendant simplement vers h , construire une suite extraite $h_{\varphi(n)}$ et une suite de segments emboîtés $[a_n, b_n]$ tels que l'image de $[a_n, b_n]$ par $h_{\varphi(n)}$ soit incluse dans $[\frac{1}{4}, \frac{3}{4}]$ et en tirer une contradiction.)

4. Complétion d'espaces vectoriels normés

La manière la plus standard pour construire des espaces de Banach est de partir d'espaces vectoriels normés et de les compléter⁽⁵⁾. On renvoie au n° 9.3 du chapitre « Vocabulaire mathématique » pour tout ce qui a trait à la complétion d'un espace métrique. De manière générale, on a le résultat suivant.

Proposition II.1.12. — (i) Si $(E, \|\cdot\|)$ est un espace vectoriel normé, alors le complété \widehat{E} de E (pour la distance associée à $\|\cdot\|$) est un espace vectoriel. De plus, $\|\cdot\|$ s'étend par continuité en une norme sur \widehat{E} , et $(\widehat{E}, \|\cdot\|)$ est un espace de Banach.

(ii) Si F est un espace de Banach, et si $u : E \rightarrow F$ est linéaire continue, alors u admet un unique prolongement continu à \widehat{E} , et ce prolongement est linéaire.

Démonstration. — Le (i) est un petit exercice utilisant de manière répétée les résultats du n° 9.3 du chapitre « Vocabulaire mathématique ». Par exemple, pour montrer que l'addition $s : E \times E \rightarrow E$ s'étend par continuité en une addition $s : \widehat{E} \times \widehat{E} \rightarrow \widehat{E}$, on peut munir $E \times E$ de la norme $\|(x, y)\| = \sup(\|x\|, \|y\|)$. Alors $s : E \times E \rightarrow E \subset \widehat{E}$ est lipschitzienne de rapport 2, et donc s'étend par continuité à $\widehat{E} \times \widehat{E}$.

Pour le (ii), voir le n° 11.1 du chapitre « Vocabulaire mathématique ».

Exemple II.1.13. — Soit Ω un ouvert de \mathbf{R}^m .

(i) L'espace $L^1(\Omega)$ défini au n° 1 du § III.2 est un espace de Banach dans lequel $\mathcal{C}_c(\Omega)$ est dense ; on peut donc aussi le définir comme le complété de $\mathcal{C}_c(\Omega)$ pour la norme $\|\phi\|_1$ définie par $\|\phi\|_1 = \int_{\Omega} |\phi(x)| dx$.

(ii) De même, l'espace $L^2(\Omega)$ défini au n° 2 du § III.2 peut aussi être défini comme le complété de $\mathcal{C}_c(\Omega)$ pour la norme $\|\cdot\|_2$ définie par $\|\phi\|_2 = \left(\int_{\Omega} |\phi(x)|^2 dx \right)^{1/2}$.

(iii) Si $k \geq 1$, on définit l'espace de Sobolev $H^k(\mathbf{R}^m)$ comme le complété de l'espace $\mathcal{C}_c^k(\mathbf{R}^m)$ des fonctions de classe \mathcal{C}^k sur \mathbf{R}^m , nulles en dehors d'un compact, pour la norme $\|\cdot\|_{H^k}$ définie par

$$\|\phi\|_{H^k} = \left(\sum_{|\ell| \leq k} (\|\partial^\ell \phi\|_2)^2 \right)^{1/2} = \left(\sum_{|\ell| \leq k} \int_{\mathbf{R}^m} |\partial^\ell \phi(x)|^2 dx \right)^{1/2},$$

⁽⁵⁾Pour beaucoup de questions c'est très utile, car on obtient un espace dans lequel l'analyse devient plus facile ; en particulier, il est nettement plus aisé de démontrer des résultats d'existence dans un espace complet. Évidemment, le problème est qu'il est difficile de retrouver ses petits après complétion. Par exemple, il est impossible de montrer que deux nombres réels sont égaux (\mathbf{R} est obtenu en complétant \mathbf{Q}) sauf si on sait par ailleurs que leur différence est un entier (à multiplication près par un nombre réel explicite). De même, il est nettement plus facile de démontrer l'existence de solutions d'équations différentielles dans un espace de Sobolev H^k , mais si ce qui nous intéresse sont des solutions de classe \mathcal{C}^k , il y a un travail supplémentaire pour vérifier que les solutions obtenues conviennent.

où, si $\ell = (\ell_1, \dots, \ell_m) \in \mathbf{N}^m$, on a posé $|\ell| = \sum_{j=1}^m \ell_j$, et noté ∂^ℓ l'opérateur différentiel

$$\partial^\ell = \left(\frac{\partial}{\partial x_1}\right)^{\ell_1} \dots \left(\frac{\partial}{\partial x_m}\right)^{\ell_m}.$$

Par définition, si $\ell = (\ell_1, \dots, \ell_m) \in \mathbf{N}^m$ vérifie $|\ell| \leq k$, l'application linéaire

$$\partial^\ell = \left(\frac{\partial}{\partial x_1}\right)^{\ell_1} \dots \left(\frac{\partial}{\partial x_m}\right)^{\ell_m} : \mathcal{C}_c^k(\mathbf{R}^m) \rightarrow \mathcal{C}_c^{k-|\ell|}(\mathbf{R}^m)$$

est continue (et même 1-lipschitzienne) si on munit $\mathcal{C}_c^k(\mathbf{R}^m)$ de la norme $\|\cdot\|_{H^k}$ et $\mathcal{C}_c^{k-|\ell|}(\mathbf{R}^m)$ de la norme $\|\cdot\|_{H^{k-|\ell|}}$. Elle s'étend donc, par continuité, en une application linéaire, encore notée ∂^ℓ , de $H^k(\mathbf{R}^m)$ dans $H^{k-|\ell|}(\mathbf{R}^m)$.

Proposition II.1.14. — Si $k \in \mathbf{N}$, si $\ell \in \mathbf{N}^m$ vérifie $|\ell| \leq k$, si $\phi \in \mathcal{C}_c^k(\mathbf{R}^m)$, et si $f \in H^k(\mathbf{R}^m)$, alors⁽⁶⁾

$$\int_{\mathbf{R}^m} (\partial^\ell \phi) f = (-1)^{|\ell|} \int_{\mathbf{R}^m} (\partial^\ell f) \phi.$$

Démonstration. — Les deux membres sont bien définis car $\partial^\ell \phi$, f , $\partial^\ell f$ et ϕ sont de carré sommable. On en déduit que

$$f \mapsto L_\phi(f) = \int_{\mathbf{R}^m} (\partial^\ell \phi) f - (-1)^{|\ell|} \int_{\mathbf{R}^m} (\partial^\ell f) \phi$$

est une forme linéaire sur $H^k(\mathbf{R}^m)$, qui est continue car

$$|L_\phi(f)| \leq \|f\|_2 \|\partial^\ell \phi\|_2 + \|\partial^\ell f\|_2 \|\phi\|_2 \leq (\|\partial^\ell \phi\|_2 + \|\phi\|_2) \|f\|_{H^k}.$$

Par ailleurs, une (suite d') intégration par partie montre que $L_\phi(f) = 0$ si $f \in \mathcal{C}_c^k(\mathbf{R}^m)$, et comme $\mathcal{C}_c^k(\mathbf{R}^m)$ est dense dans $H^k(\mathbf{R}^m)$, cela implique que L_ϕ est identiquement nulle sur $H^k(\mathbf{R}^m)$. Ceci permet de conclure.

5. Applications linéaires continues entre espaces de Banach

Théorème II.1.15. — (Banach-Steinhaus, 1927) Soient E un espace de Banach, F un espace vectoriel normé, et $(u_n)_{n \in \mathbf{N}}$ une suite d'applications linéaires continues de E dans F . Alors, de deux choses l'une : soit la suite $(\|u_n\|)_{n \in \mathbf{N}}$ est bornée⁽⁷⁾, et donc $u_n(x)$ est bornée pour tout $x \in E$, soit $\{x \in E, \sup_{n \in \mathbf{N}} \|u_n(x)\|_F = +\infty\}$ est dense dans E .

Démonstration. — Il s'agit de prouver que, si $(\|u_n\|)_{n \in \mathbf{N}}$ n'est pas bornée, et si on définit $\varphi : E \rightarrow \mathbf{R}_+ \cup \{+\infty\}$ par $\varphi(x) = \sup_{n \in \mathbf{N}} \|u_n(x)\|_F$, alors $\{x \in E, \varphi(x) = +\infty\}$ est un G_δ dense. Pour cela, considérons, si $N \in \mathbf{N}$, l'ensemble $U_N = \{x \in E, \varphi(x) > N\}$. On a $U_N = \bigcup_{n \in \mathbf{N}} \{x \in E, \|u_n(x)\|_F > N\}$, et comme chaque u_n est continue, U_N est une réunion d'ouverts et donc est ouvert. Si U_N n'est pas dense, il existe $x_0 \in E$ et $r > 0$ tel que $\|u_n(x + x_0)\|_F \leq N$, quel que soient $x \in E$, avec $\|x\|_E < r$, et $n \in \mathbf{N}$. Mais alors $\|u_n(x)\|_F = \|u_n(x + x_0) - u_n(x_0)\|_F \leq 2N$, quel que soit $x \in E$, avec $\|x\|_E < r$, et quel que soit $n \in \mathbf{N}$. Autrement dit, on a $\|u_n\| \leq \frac{2N}{r}$, quel que soit $n \in \mathbf{N}$, contrairement à l'hypothèse. C'est donc que U_N est un ouvert dense, quel que soit $N \in \mathbf{N}$, et comme $\{x \in E, \varphi(x) = +\infty\} = \bigcap_{N \in \mathbf{N}} U_N$, le lemme de Baire (n° 9.2 du chap. « Vocabulaire mathématique ») montre qu'il est dense dans E , ce que l'on cherchait à démontrer.

⁽⁶⁾ Autrement dit, $\partial^\ell f$ est la dérivée ℓ -ième de f au sens des distributions.

⁽⁷⁾ $\|u_n\|$ est la norme d'opérateur de u_n . Rappelons qu'elle est définie par $\|u_n\| = \sup_{x \neq 0} \|x\|_E^{-1} \|u_n(x)\|_F$, cf. n° 11 du § 6.

Remarque II.1.16. — Il ressort de la démonstration que $\{x \in E, \sup_{n \in \mathbf{N}} \|u_n(x)\|_F = +\infty\}$ est une intersection dénombrable d'ouverts denses, si $\|u_n\|$ n'est pas bornée. Une intersection dénombrable d'ouverts est appelée un G_δ , et il résulte du lemme de Baire qu'un G_δ dense dans un espace de Banach E est non dénombrable. (Si $X = \bigcap_{n \in \mathbf{N}} U_n$ est dénombrable, alors en numérotant les éléments x_n de X , on voit que $\bigcap_{n \in \mathbf{N}} (U_n - \{x_n\}) = \emptyset$, ce qui contredit le lemme de Baire si chacun des U_n est dense car alors $U_n - \{x_n\}$ est encore un ouvert dense de E .)

Le théorème de Banach-Steinhaus admet comme corollaire le très utile résultat suivant, qui est un peu surprenant quand on pense à ce qui se passe pour une limite simple de fonctions continues⁽⁸⁾.

Corollaire II.1.17. — *Si E est un espace de Banach, et si F un espace vectoriel, alors une limite simple d'applications linéaires continues de E dans F est une application linéaire continue de E dans F .*

Démonstration. — Soit $(u_n)_{n \in \mathbf{N}}$ une suite d'applications linéaires continues de E dans F telle que la suite $(u_n(x))_{n \in \mathbf{N}}$ ait une limite $u(x) \in F$, quel que soit $x \in E$. Si $x \in E$ et $\lambda \in \mathbf{K}$, on a $u(\lambda x) = \lim_{n \rightarrow +\infty} u_n(\lambda x) = \lim_{n \rightarrow +\infty} \lambda u_n(x) = \lambda \lim_{n \rightarrow +\infty} u_n(x) = \lambda u(x)$, et de même, $u(x+y) = u(x) + u(y)$, quels que soient $x, y \in E$, ce qui prouve que u est linéaire. Maintenant, le fait que la suite $(u_n(x))_{n \in \mathbf{N}}$ a une limite quel que soit $x \in E$, implique, d'après le théorème de Banach-Steinhaus, l'existence de $M \in \mathbf{R}_+$ tel que $\|u_n\| \leq M$ quel que soit $n \in \mathbf{N}$. On a donc $\|u_n(x)\|_F \leq M \cdot \|x\|_E$ quels que soient $n \in \mathbf{N}$ et $x \in E$. On en déduit, en passant à la limite, que $\|u(x)\|_F \leq M \cdot \|x\|_E$ quel que soit $x \in E$, et donc que u est continue, ce qui permet de conclure.

Théorème II.1.18. — (de l'image ouverte, Banach (1929)) *Si E et F sont deux espaces de Banach, et si $u : E \rightarrow F$ est une application linéaire continue surjective, alors il existe $\rho > 0$ tel que $u(B_E(0, 1^-))$ contienne $B_F(0, \rho^-)$.*

Démonstration. — Si $n \in \mathbf{N}$, soit A_n l'adhérence dans F de $u(B_E(0, n^-))$. Comme E est la réunion des $B_E(0, n^-)$, pour $n \in \mathbf{N}$, et comme u est supposée surjective, on a $\bigcup_{n \in \mathbf{N}} A_n = F$. Le lemme de Baire implique donc l'existence de n tel que A_n soit d'intérieur non vide. Ceci se traduit par l'existence de $x_0 \in B_E(0, n^-)$ et de $r > 0$, tels que, si $\|y\|_F < r$, alors quel que soit $\varepsilon > 0$, il existe $x \in B_E(0, n^-)$, avec $\|u(x) - (u(x_0) + y)\|_F < \varepsilon$. Comme $x - x_0 \in B_E(0, 2n^-)$, quitte à faire une homothétie de rapport $\frac{1}{4n}$, on voit que l'on a démontré le résultat suivant (avec $\rho = \frac{r}{4n}$) : il existe $\rho > 0$ tel que, quel que soit $y \in B_F(0, \rho^-)$ et quel que soit $\varepsilon > 0$, il existe $x \in B_E(0, (\frac{1}{2})^-)$ avec $\|y - u(x)\|_F < \varepsilon$.

Ce n'est pas tout à fait le résultat cherché, mais presque. Si $y \in B_F(0, \rho^-)$, on peut construire par récurrence, en utilisant ce qui précède, une suite $(x_m)_{m \in \mathbf{N}}$ d'éléments de $B_E(0, (\frac{1}{2})^-)$, et une suite $(y_m)_{m \in \mathbf{N}}$ d'éléments de $B_F(0, \rho^-)$ vérifiant :

$$y_0 = y, \quad \|y_m - u(x_m)\|_F < \frac{\rho}{2}, \quad \text{et} \quad y_{m+1} = 2(y_m - u(x_m)).$$

On a alors $y = u(x_0) + 2^{-1}u(x_1) + \dots + 2^{-m}u(x_m) + 2^{-m-1}y_{m+1}$, et comme la série $\sum_{m=0}^{+\infty} 2^{-m}x_m$ converge dans E vers un élément x de $B_E(0, 1^-)$, un passage à la limite montre que $y = u(x)$. Ceci démontre l'inclusion $B_F(0, \rho^-) \subset u(B_E(0, 1^-))$ que l'on cherchait à obtenir.

⁽⁸⁾Bien que Cauchy ait réussi à "démontrer", dans son *Cours d'analyse*, qu'une limite simple de fonctions continues est continue, on sait bien, à l'heure actuelle, qu'il n'en est rien, en général. Baire (1904) a démontré qu'une fonction $f : \mathbf{R}^n \rightarrow \mathbf{R}$ est limite simple de fonctions continues si et seulement si la restriction de f à tout fermé non vide a au moins un point où elle est continue. C'est à cette occasion qu'il a introduit son fameux lemme.

Remarque II.1.19. — Si $x \in E$ et $r > 0$, alors $B_E(x, r^-) = x + rB_E(0, 1^-)$ et donc $u(B_E(x, r^-)) = u(x) + ru(B_E(0, 1^-))$. Le théorème ci-dessus montre donc que, si u est surjective, alors $u(B_E(x, r^-))$ contient un voisinage ouvert de $u(x)$. On en déduit le fait que, si U est un ouvert de E , alors $u(U)$ est voisinage ouvert de $u(x)$, quel que soit $x \in U$; autrement dit $u(U)$ est ouvert. Le théorème ci-dessus peut donc se reformuler sous la forme « l'image d'un ouvert par une application linéaire continue *surjective* entre deux espaces de Banach est un ouvert »; c'est ce qui explique son nom.

Corollaire II.1.20. — Si E et F sont deux espaces de Banach, et si $u : E \rightarrow F$ est une application linéaire continue bijective, alors $u^{-1} : F \rightarrow E$ est aussi continue.

Démonstration. — Si U est un ouvert de E , alors $(u^{-1})^{-1}(U) = u(U)$ est ouvert d'après la remarque ci-dessus. Ceci permet de conclure.

Exercice II.1.21. — (Théorème du graphe fermé)

Soient E et F deux espaces de Banach, et $u : E \rightarrow F$ une application linéaire. Soit $G \subset E \times F$ le graphe de u (i.e. l'ensemble des couples $(x, u(x))$, pour $x \in E$).

(i) Montrer que $E \times F$ muni de la norme $\|(x, y)\| = \sup(\|x\|_E, \|y\|_F)$ est un espace de Banach et que les deux projections $p_E : E \times F \rightarrow E$ et $p_F : E \times F \rightarrow F$ sont continues.

(ii) Montrer que G est un sous-espace vectoriel de $E \times F$ et que, si G est fermé dans $E \times F$, alors u est continue.

(iii) Construire $f : \mathbf{R} \rightarrow \mathbf{R}$ non continue dont le graphe est fermé.

6. Le dual d'un espace de Banach

Si E est un espace vectoriel normé, on note E^* le *dual* de E , c'est-à-dire l'ensemble des formes linéaires continues sur E . Si $\Lambda : E \rightarrow \mathbf{K}$ est une forme linéaire continue, on rappelle que l'on définit sa norme $\|\Lambda\|$ comme la borne inférieure de l'ensemble des $C \in \mathbf{R}_+$ tels que $|\Lambda(x)| \leq C\|x\|$ quel que soit $x \in E$.

Théorème II.1.22. — Si E est un espace vectoriel normé, alors $(E^*, \|\cdot\|)$ est un espace de Banach.

Démonstration. — Le fait que $\|\cdot\|$ est une norme d'espace vectoriel est un exercice. Maintenant, soit Λ_n une suite d'éléments de E^* telle que $\sum_{n \in \mathbf{N}} \|\Lambda_n\| = C < +\infty$. Si $x \in E$, la série $\sum_{n \in \mathbf{N}} \Lambda_n(x)$ est alors absolument convergente, et la somme $\Lambda(x)$ vérifie $|\Lambda(x)| \leq \sum_{n \in \mathbf{N}} |\Lambda_n(x)| \leq C\|x\|$. Comme par ailleurs, il est facile de voir que $x \mapsto \Lambda(x)$ est linéaire, la majoration ci-dessus montre que $x \mapsto \Lambda(x)$ est aussi continue. On en déduit que $\Lambda \in E^*$ et que $\sum_{n \in \mathbf{N}} \Lambda_n = \Lambda$, ce qui prouve que E est complet.

Théorème II.1.23. — (Hahn-Banach, 1927) Si E est un espace vectoriel normé, si F est un sous-espace vectoriel de E , et si f est une forme linéaire continue sur F , alors il existe $\Lambda \in E^*$ dont la restriction à F est f et qui vérifie $\|\Lambda\| = \|f\|$.

Le théorème de Hahn-Banach, que nous ne démontrerons pas, a un certain nombre de conséquences intéressantes, dont le fait que $E^* \neq 0$. On en trouvera d'autres dans les exercices suivants.

Exercice II.1.24. — Montrer que l'adhérence \bar{F} de F dans E est l'intersection des noyaux des formes linéaires, continues sur E , nulles sur F .

Exercice II.1.25. — (i) Montrer que, si $x_0 \in E$, il existe $\Lambda \in E^*$, avec $\|\Lambda\| = 1$ et $|\Lambda(x_0)| = \|x_0\|$.

(ii) Montrer que E^* sépare les points (si $x \neq y$, il existe $\Lambda \in E^*$ tel que $\Lambda(x) \neq \Lambda(y)$.)

(iii) Montrer que l'application $\Lambda \mapsto \Lambda(x)$ est une forme linéaire continue sur E^* et induit une isométrie de E dans $(E^*)^*$. (On dit qu'un espace de Banach E est *réflexif* si cette isométrie est bijective, et donc

si E s'identifie au dual de son dual ; il peut arriver que $(E^*)^*$ soit beaucoup plus gros que E , mais les espaces de Hilbert sont réflexifs d'après le théorème de Riesz (th. II.2.11)).

II.2. Espaces de Hilbert

Les espaces de Hilbert sont des espaces de Banach aux propriétés mathématiques particulièrement agréables : l'existence de bases hilbertiennes montre que tous ceux qu'on rencontre en pratique sont isomorphes, et l'existence de projecteurs orthogonaux permet très souvent de se ramener à la dimension finie. La nature étant bien faite, ce sont précisément ces espaces qui interviennent naturellement dans beaucoup de questions physiques (par exemple en mécanique quantique).

1. Espaces préhilbertiens

Soit E un espace vectoriel sur \mathbf{K} .

- Un *produit scalaire* $(x, y) \mapsto \langle x, y \rangle$ sur E est une application de $E \times E$ dans \mathbf{K} qui est :
 - *sesquilinéaire*, i.e. linéaire par rapport à la seconde variable ($\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle$ et $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$, si $\lambda \in \mathbf{K}$, $x, y, y_1, y_2 \in E$), et semi-linéaire⁽⁹⁾ par rapport à la première variable (i.e. $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$ et $\langle \lambda x, y \rangle = \bar{\lambda} \langle x, y \rangle$, si $\lambda \in \mathbf{K}$, $x, y, x_1, x_2 \in E$) ;
 - *symétrique*, i.e. $\langle y, x \rangle = \overline{\langle x, y \rangle}$, quels que soient $x, y \in E$;
 - *définie positive*, i.e. $\langle x, x \rangle \geq 0$, si $x \in E$, et $\langle x, x \rangle = 0$, si et seulement si $x = 0$.
- Un *espace préhilbertien* est un espace vectoriel muni d'un produit scalaire. Si E est préhilbertien, on définit $\| \cdot \| : E \rightarrow \mathbf{R}$ en posant $\|x\| = \langle x, x \rangle^{1/2}$. Alors $\| \cdot \|$ est une norme, et on a $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$ pour tous $x, y \in E$ (*inégalité de Cauchy-Schwarz*) : l'application \mathbf{R} -linéaire $(x, y) \mapsto \langle x, y \rangle$, de $E \times E$ dans \mathbf{K} , est continue.

$\|x + ty\|^2 = \|x\|^2 + 2t \operatorname{Re}(\langle x, y \rangle) + t^2 \|y\|^2$ est toujours ≥ 0 , pour $t \in \mathbf{R}$; son discriminant est donc ≤ 0 , ce qui se traduit par $|\operatorname{Re}(\langle x, y \rangle)| \leq \|x\| \cdot \|y\|$ pour tous $x, y \in E$. Choisissons alors $\theta \in \mathbf{R}$ tel que $e^{-i\theta} \langle x, y \rangle \in \mathbf{R}_+$. En utilisant la majoration précédente pour $e^{i\theta} x$ et y au lieu de x et y , on obtient $|\langle x, y \rangle| = \operatorname{Re}(\langle e^{i\theta} x, y \rangle) \leq \|e^{i\theta} x\| \cdot \|y\| = \|x\| \cdot \|y\|$, ce qui prouve l'inégalité de Cauchy-Schwarz. L'inégalité triangulaire s'en déduit car

$$\|x + y\|^2 = \|x\|^2 + 2\operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

L'identité $\|\lambda x\| = |\lambda| \|x\|$ étant immédiate, on en déduit que $\| \cdot \|$ est une norme, ce qui permet de conclure.

- On dit que $x, y \in E$ sont *orthogonaux* si $\langle x, y \rangle = 0$. Si x et y sont orthogonaux, ils vérifient la *relation de Pythagore*⁽¹⁰⁾ $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. Dans le cas général, ils

⁽⁹⁾Si $\mathbf{K} = \mathbf{R}$, on a $\bar{x} = x$, et donc la sesquilinearité n'est autre que la bilinéarité.

⁽¹⁰⁾Si $\mathbf{K} = \mathbf{R}$, la relation de Pythagore entraîne l'orthogonalité (« théorème » de Pythagore, pauvre Pythagore...) ; ce n'est plus le cas si $\mathbf{K} = \mathbf{C}$.

vérifient *l'identité de la médiane* $\|x\|^2 + \|y\|^2 = 2\left\|\frac{x+y}{2}\right\|^2 + \frac{1}{2}\|x-y\|^2$, qui se démontre sans problème en développant le membre de droite.

- Si F est un sous-espace vectoriel de E , et si $x \in E$, il existe au plus un élément $p_F(x)$ de F , appelé (s'il existe) *projection orthogonale de x sur F* , tel que $x - p_F(x)$ soit orthogonal à F tout entier. De plus, on a $p_F(x) = x$, si $x \in F$, et p_F est linéaire et 1-lipschitzien sur son ensemble de définition.

Si $y_1, y_2 \in F$ sont tels que $x - y_1$ et $x - y_2$ sont orthogonaux à F tout entier, alors $y_1 - y_2 = (x - y_1) - (x - y_2)$ est orthogonal à F , et comme $y_1 - y_2 \in F$, on a $\langle y_1 - y_2, y_1 - y_2 \rangle = 0$, ce qui implique $y_1 = y_2$. On en déduit l'unicité de p_F . La linéarité de p_F et la formule $p_F(x) = x$, si $x \in F$, en sont des conséquences immédiates. Enfin, $x - p_F(x)$ et $p_F(x)$ étant orthogonaux, on a $\|p_F(x)\|^2 = \|x\|^2 - \|x - p_F(x)\|^2$, et donc $\|p_F(x)\| \leq \|x\|$. Ceci permet de conclure.

- Une famille $(e_i)_{i \in I}$ d'éléments de E est dite *orthonormale*, si $\|e_i\| = 1$ pour tout i , et si e_i et e_j sont orthogonaux si $i \neq j$. On a alors $\|\sum_{i \in J} x_i e_i\|^2 = \sum_{i \in J} |x_i|^2$, pour toute famille *finie* $(x_i)_{i \in J}$ de nombres complexes.

- Si F est un sous-espace de dimension finie de E muni d'une base orthonormale (e_1, \dots, e_d) , alors p_F est partout définie, et $p_F(x) = \sum_{i=1}^d \langle e_i, x \rangle e_i$. En particulier, si $x \in F$, ses coordonnées dans la base (e_1, \dots, e_d) sont les $\langle e_i, x \rangle$, et $\|x\|^2 = \sum_{i=1}^d |\langle e_i, x \rangle|^2$.

Soit $y = \sum_{i=1}^d \langle e_i, x \rangle e_i$. Alors $\langle e_j, x - y \rangle = \langle e_j, x \rangle - \sum_{i=1}^d \langle e_i, x \rangle \langle e_j, e_i \rangle = 0$, pour tout j . On en déduit que $x - y$ est orthogonal à chacun des e_j , et donc à F tout entier par linéarité. De plus, $y \in F$ par construction, et donc $y = p_F(x)$. On en déduit le résultat.

- Le *procédé d'orthonormalisation de Schmidt*, décrit ci-dessous, permet, si $(f_i)_{i \in I}$ est une famille libre d'éléments de E , avec I dénombrable, de fabriquer une base orthonormale de l'espace F engendré par les f_i .

On se ramène, en numérotant les éléments de I , au cas où I est un intervalle de \mathbf{N} contenant 0.

On note F_n le sous-espace de F engendré par les f_i , pour $i \leq n$. On construit par récurrence une famille orthonormale e_i d'éléments de F telle que (e_0, \dots, e_n) soit une base (orthonormale) de F_n , pour tout n . Pour cela, on pose $e_0 = \frac{1}{\|f_0\|}$, et en supposant e_0, \dots, e_{n-1} construits (et donc F_{n-1} muni d'une base orthonormale), on note $g_n = f_n - p_{F_{n-1}}(f_n)$. On a $g_n \neq 0$ puisque $f_n \notin F_{n-1}$, la famille des f_j étant supposée libre. On pose $e_n = \frac{1}{\|g_n\|} g_n$. Par construction, g_n (et donc aussi e_n) est orthogonal à chacun des e_i , pour $i \leq n-1$, et comme $\|e_n\| = 1$, cela permet de faire marcher la récurrence.

- Tout sous-espace de dimension finie de E possède une base orthonormale.

Il suffit d'appliquer le procédé d'orthonormalisation de Schmidt à une base quelconque.

2. Espaces de Hilbert

Un *espace de Hilbert* est un espace préhilbertien complet ; c'est donc un cas particulier d'espace de Banach.

Exemple II.2.1. — (i) Un espace de dimension finie muni d'un produit scalaire est un espace de Hilbert.

- (ii) Si E est un espace préhilbertien, son complété \widehat{E} est un espace de Hilbert (le produit scalaire s'étendant par continuité).
- (iii) ℓ^2 et, plus généralement, $\ell^2(I)$ si I est dénombrable, sont des espaces de Hilbert.
- (iv) Si Ω est un ouvert non vide de \mathbf{R}^n , alors $L^2(\Omega)$ est un espace de Hilbert.
- (v) Le complété $L^2(\mathbf{R}/\mathbf{Z})$ de $\mathcal{C}(\mathbf{R}/\mathbf{Z})$ pour la norme $\| \cdot \|_2$ définie par $\|f\|_2 = \int_0^1 |f(t)|^2 dt$ est un espace de Hilbert (de produit scalaire $\langle f, g \rangle = \int_0^1 \overline{f(t)}g(t) dt$).
- (vi) Les espaces de Sobolev $H^k(\mathbf{R}^m)$ sont des espaces de Hilbert.

2.1. Bases hilbertiennes

Soit E un espace de Hilbert *séparable*⁽¹¹⁾ de dimension infinie. Une *base hilbertienne*⁽¹²⁾ de E est une famille orthonormale dénombrable $(e_i)_{i \in I}$ d'éléments de E telle que le sous-espace vectoriel de E engendré par les e_i , pour $i \in I$, soit dense dans E .

Exemple II.2.2. — (i) Si I est un ensemble dénombrable, alors $\ell^2(I)$ possède une base hilbertienne naturelle, à savoir celle constituée des e_i , pour $i \in I$, où e_i est la suite avec un 1 en i et des 0 partout ailleurs.

(ii) Les $e^{2i\pi nt}$, pour $n \in \mathbf{Z}$ forment une base hilbertienne de $L^2(\mathbf{R}/\mathbf{Z})$.

Démonstration. — (i) Que les e_i forment une famille orthonormale est immédiat. Maintenant, si $x = (x_i)_{i \in I} \in \ell^2(I)$, et si $J \subset I$ est fini, alors $\|x - \sum_{j \in J} x_j e_j\|_2^2 = \sum_{j \in I-J} |x_j|^2$. Comme $\sum_{j \in I} |x_j|^2 < +\infty$, on peut rendre $\|x - \sum_{j \in J} x_j e_j\|_2$ aussi petit que l'on veut en augmentant J , ce qui prouve que le sous-espace engendré par les e_i , pour $i \in I$, est dense dans $\ell^2(I)$. On en déduit le (i).

(ii) Un calcul immédiat montre que les $e^{2i\pi nt}$ forment une famille orthonormale ; l'espace qu'elle engendre est l'espace des polynômes trigonométriques. Si $f \in L^2(\mathbf{R}/\mathbf{Z})$, et si $\varepsilon > 0$, il existe, par définition de $L^2(\mathbf{R}/\mathbf{Z})$, une fonction continue g sur \mathbf{R}/\mathbf{Z} avec $\|f - g\|_2 < \frac{\varepsilon}{2}$. Par ailleurs, l'espace des polynômes trigonométriques est dense dans $\mathcal{C}(\mathbf{R}/\mathbf{Z})$ d'après le théorème de Stone-Weierstrass (cf. (iii) de l'ex. II.1.10) ; il existe donc P , polynôme trigonométrique, tel que $\|P - g\|_\infty < \frac{\varepsilon}{2}$. Comme $\|h\|_2 \leq \|h\|_\infty$, si $h \in \mathcal{C}(\mathbf{R}/\mathbf{Z})$, on a $\|f - P\|_2 < \varepsilon$, ce qui prouve que toute boule ouverte de $L^2(\mathbf{R}/\mathbf{Z})$ contient un polynôme trigonométrique, et donc que l'espace engendré par les $e^{2i\pi nt}$, pour $n \in \mathbf{Z}$, est dense dans $L^2(\mathbf{R}/\mathbf{Z})$. Ceci permet de conclure.

Proposition II.2.3. — *Un espace de Hilbert séparable admet des bases hilbertiennes.*

⁽¹¹⁾La théorie qui suit s'étend au cas des espaces de Hilbert non séparables, mais ceux-ci ne se rencontrent pas en pratique. La seule différence est qu'une base hilbertienne n'est plus de cardinal dénombrable, si E n'est pas séparable.

⁽¹²⁾Une base hilbertienne est aussi souvent appelée une *base orthonormale*. On fera attention au fait qu'une base hilbertienne n'est, en général, pas une base au sens algébrique. Plus précisément, une base hilbertienne est une base algébrique si et seulement si on est en dimension finie, ce qui est rarement le cas en analyse fonctionnelle.

Démonstration. — Soient E un espace de Hilbert séparable et $A \subset E$ un sous-ensemble dénombrable dense. Pour construire une base hilbertienne à partir de A , on numérote les éléments de A , on élimine ceux qui se trouvent dans l'espace vectoriel engendré par les éléments précédents, ce qui nous fournit une base $(b_i)_{i \in I}$, avec $I \subset \mathbf{N}$, de l'espace vectoriel E' engendré par A . Enfin, on utilise le procédé d'orthonormalisation de Schmidt pour construire une famille orthonormale d'éléments de E engendrant le sous-espace E' ; cette famille est une base hilbertienne de E puisque E' , qui contient A , est dense dans E .

2.2. Projection orthogonale sur un sous-espace fermé

Lemme II.2.4. — Soit E un espace de Hilbert, et soit $(e_i)_{i \in I}$ une famille orthonormale dénombrable d'éléments de E .

(i) Si $(x_i)_{i \in I}$ est une famille d'éléments de \mathbf{C} , alors la série $\sum_{i \in I} x_i e_i$ est sommable si et seulement si $(x_i)_{i \in I} \in \ell^2(I)$.

(ii) Si $(x_i)_{i \in I} \in \ell^2(I)$, et si $x \in E$ est la somme de la série $\sum_{i \in I} x_i e_i$, alors $\langle e_i, x \rangle = x_i$, pour tout $i \in I$, et $\|x\|^2 = \sum_{i \in I} |x_i|^2$.

Démonstration. — On a $\|\sum_{i \in J} x_i e_i\| = (\sum_{i \in J} |x_i|^2)^{1/2}$, pour tout $J \subset I$ fini, par orthonormalité de la famille $(e_i)_{i \in I}$. La sommabilité de $\sum_{i \in J} x_i e_i$ est donc équivalente à la condition $\sum_{i \in I} |x_i|^2 < +\infty$. On en déduit le (i).

Le (ii) est évident si I est fini. On peut donc supposer $I = \mathbf{N}$. Alors x est la limite de la suite de terme général $y_n = \sum_{j=0}^n x_j e_j$, et comme on a $\langle e_i, y_n \rangle = x_i$, pour tout $n \geq i$, et $\|y_n\|^2 = \sum_{j=0}^n |x_j|^2$, le (ii) s'en déduit par un passage à la limite, en utilisant la continuité de la norme et du produit scalaire.

Proposition II.2.5. — Soit F un sous-espace vectoriel fermé de E muni d'une base hilbertienne $(e_i)_{i \in I}$.

(i) Si $x \in E$, alors $(\langle e_i, x \rangle)_{i \in I} \in \ell^2(I)$.

(ii) La série $\sum_{i \in I} \langle e_i, x \rangle e_i$ est sommable; sa somme $p_F(x)$ appartient à F , et $\|p_F(x)\|^2 = \sum_{i \in I} |\langle e_i, x \rangle|^2$.

(iii) $p_F : E \rightarrow F$ est un projecteur, et $p_F(x)$ est l'unique élément de F tel que $x - p_F(x)$ soit orthogonal à F tout entier. De plus, p_F est 1-lipschitzien.

Démonstration. — Si F est de dimension finie, le résultat a déjà été démontré. On peut donc supposer que $I = \mathbf{N}$. Notons F_i le sous-espace vectoriel de E engendré par les e_j , pour $j \leq i$. Alors $y_i = \sum_{j=0}^i \langle e_j, x \rangle e_j$ est la projection orthogonale de x sur F_i . En particulier, $\sum_{j=0}^i |\langle e_j, x \rangle|^2 = \|y_i\|^2 \leq \|x\|^2$, quel que soit $i \in \mathbf{N}$. On en déduit l'appartenance de $(\langle e_i, x \rangle)_{i \in \mathbf{N}} \in \ell^2$, ce qui démontre le (i).

Le (ii) est une conséquence directe du lemme II.2.4, dont on déduit aussi que $x - p_F(x)$ est orthogonal à tous les e_i , et donc à F tout entier par linéarité et densité. Le reste du (iii) suit de l'unicité de la projection orthogonale sur un sous-espace (pas forcément fermé). Ceci permet de conclure.

Théorème II.2.6. — Si $(e_i)_{i \in I}$ est une base hilbertienne de E , l'application $x \mapsto (\langle e_i, x \rangle)_{i \in I}$ induit une isométrie⁽¹³⁾ de E sur $\ell^2(I)$. Autrement dit,

(a) si $x \in E$, alors $(\langle e_i, x \rangle)_{i \in I} \in \ell^2(I)$ et $\sum_{i \in I} |\langle e_i, x \rangle|^2 = \|x\|^2$ (identité de Bessel-Parseval) ;

(b) si $(x_i)_{i \in I} \in \ell^2(I)$, alors $\sum_{i \in I} x_i e_i$ converge dans E et sa somme x vérifie $\langle e_i, x \rangle = x_i$ quel que soit $i \in I$.

Démonstration. — Commençons par justifier le « Autrement dit » : le (a) peut se reformuler en disant que $x \mapsto (\langle e_i, x \rangle)_{i \in I}$ est une isométrie de E sur un sous-espace de $\ell^2(I)$, tandis que le (b) montre que $\ell^2(I)$ est dans l'image de $x \mapsto (\langle e_i, x \rangle)_{i \in I}$. Maintenant, le (a) suit des (i) et (ii) de la prop. II.2.5 utilisée pour $F = E$, et le (b) suit du lemme II.2.4.

Si on spécialise le th. II.2.6 à la base hilbertienne de $L^2(\mathbf{R}/\mathbf{Z})$ constituée des $e^{2i\pi nt}$, on obtient, en particulier, le résultat suivant.

Corollaire II.2.7. — Si $f \in L^2(\mathbf{R}/\mathbf{Z})$, soit $c_n(f)$ son n -ième coefficient de Fourier :

$$c_n(f) = \langle e^{2i\pi nt}, f \rangle = \int_{\mathbf{R}/\mathbf{Z}} f(t) e^{-2i\pi nt} dt = \int_0^1 f(t) e^{-2i\pi nt} dt.$$

Alors $f = \sum_{n \in \mathbf{Z}} c_n(f) e^{2i\pi nt}$ dans⁽¹⁴⁾ $L^2(\mathbf{R}/\mathbf{Z})$ et $\sum_{n \in \mathbf{Z}} |c_n(f)|^2 = (\|f\|_2)^2 = \int_0^1 |f(t)|^2 dt$ (Bessel-Parseval).

Corollaire II.2.8. — (critère de totalité). Si $(e_i)_{i \in I}$ est une famille orthonormale dénombrable d'éléments de E , les conditions suivantes sont équivalentes :

- (i) $(e_i)_{i \in I}$ est une base hilbertienne de E ;
- (ii) l'ensemble des $x \in E$, orthogonaux à tous les e_i , est réduit à $\{0\}$.

Démonstration. — L'implication (i) \Rightarrow (ii) est une conséquence directe du (a) du th. II.2.6. Pour montrer (ii) \Rightarrow (i), introduisons l'espace vectoriel F , adhérence dans E de l'espace engendré par les e_i . Alors $(e_i)_{i \in I}$ est une base hilbertienne de F , et la condition (ii) implique que l'on a $p_F(x) = x$, pour tout $x \in E$, et donc que $F = E$. Ceci permet de conclure.

⁽¹³⁾Un sous-espace fermé de dimension infinie de ℓ^2 est un espace de Hilbert séparable, et donc, d'après le théorème, isomorphe à ℓ^2 . T. Gowers a reçu la médaille Fields en 1998, en grande partie pour avoir démontré que ceci caractérise ℓ^2 : un espace de Banach séparable, qui est isomorphe à tous ses sous-espaces fermés de dimension infinie, est isomorphe à ℓ^2 .

⁽¹⁴⁾On prendra garde au fait que cette convergence dans $L^2(\mathbf{R}/\mathbf{Z})$ (convergence en moyenne quadratique) n'implique la convergence en aucun point ; de fait, rien n'empêche a priori que tout réarrangement de la série diverge en tout point sauf celui où on s'est débrouillé pour le faire converger.

3. Le dual d'un espace de Hilbert

Dans ce n° , E est un espace de Hilbert. Si $x \in E$, on note Λ_x la forme linéaire définie par $\Lambda_x(y) = \langle x, y \rangle$.

Lemme II.2.9. — Si $x \in E$, la forme linéaire Λ_x est continue et $\|\Lambda_x\| = \|x\|$.

Démonstration. — L'inégalité de Cauchy-Schwarz qui devient $|\Lambda_x(y)| \leq \|x\| \cdot \|y\|$, nous donne la continuité de Λ_x ainsi que l'inégalité $\|\Lambda_x\| \leq \|x\|$. L'inégalité $\|\Lambda_x\| \geq \|x\|$, se déduit de ce que $|\Lambda_x(x)| = \|x\| \cdot \|x\|$.

Proposition II.2.10. — Soit E un espace de Hilbert séparable.

(i) Si A est une partie de E , alors l'orthogonal A^\perp de A (i.e. l'ensemble des $x \in E$ tels que $\langle x, a \rangle = 0$, quel que soit $a \in A$) est un sous-espace vectoriel fermé de E .

(ii) Si F est un sous-espace vectoriel fermé de E , et si $x \in E$, alors $x - p_F(x) \in F^\perp$, et on a⁽¹⁵⁾ $E = F \oplus F^\perp$ et $(F^\perp)^\perp = F$.

Démonstration. — (i) On a $\langle x, a \rangle = 0$ si et seulement si $\Lambda_a(x) = 0$, et comme Λ_a est continue, son noyau H_a est un hyperplan fermé de E , ce qui démontre le (i) puisque $A^\perp = \bigcap_{a \in A} H_a$.

(ii) On a $x - p_F(x) \in F^\perp$ par définition. Maintenant, si $x \in F \cap F^\perp$, alors $\langle x, x \rangle = 0$, et donc $x = 0$; on en déduit que $F \cap F^\perp = \{0\}$. Par ailleurs, si $x \in E$, alors $x = (x - p_F(x)) + p_F(x)$, avec $x - p_F(x) \in F^\perp$ et $p_F(x) \in F$. On en déduit que $E = F + F^\perp$ et donc que $E = F \oplus F^\perp$. Enfin, l'unicité de la projection sur un convexe fermé montre que $p_{F^\perp}(x) = x - p_F(x)$, et donc que $x = p_F(x) + p_{F^\perp}(x)$. Comme on a aussi $x = p_{F^\perp}(x) + p_{(F^\perp)^\perp}(x)$, on en déduit que $p_{(F^\perp)^\perp}(x) = p_F(x)$, quel que soit $x \in E$, et finalement que $(F^\perp)^\perp = F$.

Théorème II.2.11. — (Théorème de Riesz) Si E est un espace de Hilbert séparable, alors l'application qui associe, à $x \in E$, la forme linéaire Λ_x , définie par $\Lambda_x(y) = \langle x, y \rangle$, est une isométrie de E sur son dual E^* . Autrement dit :

(i) si $x \in E$, la forme linéaire Λ_x est continue et $\|\Lambda_x\| = \|x\|$;

(ii) si $\Lambda : E \rightarrow \mathbf{K}$ est une forme linéaire continue, il existe (un unique) $x \in E$ tel que $\Lambda(y) = \langle x, y \rangle$ quel que soit $y \in E$.

Démonstration. — Le (i) a déjà été démontré (c'est le contenu du lemme II.2.9). Passons à la démonstration du (ii)⁽¹⁶⁾.

Première démonstration. Supposons Λ non nulle sinon il n'y a qu'à prendre $x = 0$. Soit H le noyau de Λ . C'est un hyperplan de E , qui est fermé puisque Λ est continue. Soit $h \in H^\perp$, non nul. Comme $H \cap H^\perp = \{0\}$, et comme H est le noyau de Λ , on a $\Lambda(h) \neq 0$. Soit $y \in E$ et soit $z = y - \frac{\Lambda(y)}{\Lambda(h)}h$. On a $\Lambda(z) = 0$, et donc $z \in H$, ce qui implique $\langle h, z \rangle = 0$ et se traduit par $\langle h, y \rangle = \frac{\|h\|^2}{\Lambda(h)}\Lambda(y)$ quel que soit $y \in E$. Autrement dit, si on pose $x = \frac{\bar{\Lambda}(h)}{\|h\|^2}h$, on a $\Lambda(y) = \Lambda_x(y)$ quel que soit $y \in E$. Ceci permet de conclure.

⁽¹⁵⁾Il ressort de ce théorème que tout sous-espace vectoriel fermé d'un espace de Hilbert admet un supplémentaire fermé. Réciproquement, J. Lindenstrauss et L. Tzafriri (1971) ont démontré qu'un espace de Banach séparable ayant cette propriété est isomorphe à ℓ^2 .

Une question ouverte concernant la classification des espaces de Banach séparables est de savoir si ℓ^2 est le seul pour lequel le groupe des isométries (applications linéaires bijectives, vérifiant $\|u(z)\| = \|z\|$, quel que soit z) agit transitivement sur la sphère unité (i.e. si, quels que soient x, y de norme 1, il existe une isométrie u , avec $u(x) = y$). En dimension finie, l'énoncé analogue est vrai, mais pas totalement évident.

⁽¹⁶⁾C'est la partie non triviale du théorème et celle qui a les conséquences les plus spectaculaires en analyse fonctionnelle. On en déduit sans effort des tas de théorèmes d'existence.

Seconde démonstration. Comme E est supposé séparable, il est isométrique à ℓ^2 , et on peut donc supposer que $E = \ell^2$. On note e_n , pour $n \in \mathbf{N}$, la base hilbertienne standard de ℓ^2 (i.e. e_n est la suite dont tous les termes sont nuls sauf le n -ième qui est égal à 1), et on pose $a_n = \Lambda(e_n)$. Notons $\pi_n : \ell^2 \rightarrow \ell^2$, l'application $(y_i)_{i \in \mathbf{N}} \mapsto (z_i)_{i \in \mathbf{N}}$, avec $z_i = y_i$, si $i \leq n$, et $z_i = 0$, si $i > n$. Alors π_n est linéaire, continue car $\|\pi_n(y)\|_2 \leq \|y\|_2$, et on a $\pi_n(y) \rightarrow y$, pour tout $y \in \ell^2$. Soit $\Lambda_n = \Lambda \circ \pi_n$. Par linéarité, on a

$$\Lambda_n(y) = \Lambda\left(\sum_{i \leq n} y_i e_i\right) = \sum_{i \leq n} a_i y_i = \langle x^{(n)}, y \rangle,$$

où $x^{(n)} \in \ell^2$ est définie par $x_i^{(n)} = \overline{a_i}$, si $i \leq n$, et $x_i^{(n)} = 0$, si $i > n$. On déduit du (i) que $\|\Lambda_n\| = \|x^{(n)}\|_2$. Par ailleurs, si $y \in \ell^2$, on a $\Lambda_n(y) \rightarrow \Lambda(y)$, puisque $\pi_n(y) \rightarrow y$ et Λ est continue. Il résulte du théorème de Banach-Steinhaus que $\|\Lambda_n\|$ est bornée et donc que $\|x^{(n)}\|_2$ est majorée. Ceci implique que $x = (\overline{a_n})_{n \in \mathbf{N}} \in \ell^2$. Par ailleurs, $y \mapsto \Lambda'(y) = \Lambda(y) - \langle x, y \rangle$ est une forme linéaire continue sur ℓ^2 , nulle sur e_n pour tout n , et donc identiquement nulle puisque les e_n engendrent un sous-espace dense de ℓ^2 . Ceci permet de conclure.

4. Le théorème de projection sur un convexe

Rappelons que, si E est un espace vectoriel sur \mathbf{K} , un sous-ensemble C de E est *convexe* si C contient le segment $[x, y]$ quels que soient $x, y \in C$. En particulier, si C est convexe, et si $x, y \in C$, alors le milieu $\frac{x+y}{2}$ de x et y appartient à C . Le théorème suivant joue un rôle fondamental en analyse fonctionnelle.

Théorème II.2.12. — Soient E un espace de Hilbert et $C \neq \emptyset$ un convexe fermé de E .

- (i) Quel que soit $x \in E$, il existe $p_C(x) \in C$ unique (appelé la projection de x sur C) tel que $d(x, p_C(x)) \leq d(x, y)$ quel que soit $y \in C$.
- (ii) $p_C(x)$ est l'unique point y de C tel que, quel que soit $z \in C$, l'angle $(x - y, z - y)$ soit obtus (i.e. $\operatorname{Re}(\langle x - y, z - y \rangle) \leq 0$).
- (iii) L'application $x \mapsto p_C(x)$ est 1-lipschitzienne.

Démonstration. — Notons $d(x, C)$ la borne inférieure des $d(x, y)$, pour $y \in C$. Si y_1, y_2 réalisent cette borne inférieure, et si $z = \frac{y_1 + y_2}{2}$, alors $z \in C$, et l'identité de la médiane nous donne

$$\|y_1 - y_2\|^2 = 2\|y_1 - x\|^2 + 2\|y_2 - x\|^2 - 4\|z - x\|^2 = 4(d(x, C)^2 - d(x, z)^2) \leq 0.$$

On a donc $y_1 = y_2$, d'où l'unicité de la projection.

Passons à l'existence⁽¹⁷⁾. Par définition de $d(x, C)$, il existe une suite $(y_n)_{n \in \mathbf{N}}$ d'éléments de C , telle que $d(x, y_n)$ tende vers $d(x, C)$ quand n tend vers $+\infty$. L'identité de la médiane se traduit, en notant $z_{n,p}$ le milieu de y_n et y_p , par

$$\begin{aligned} \|y_n - y_{n+p}\|^2 &= 2\|y_n - x\|^2 + 2\|y_{n+p} - x\|^2 - 4\|z_{n,n+p} - x\|^2 \\ &\leq 2(\|y_n - x\|^2 + \|y_{n+p} - x\|^2 - 2d(x, C)^2). \end{aligned}$$

Comme par hypothèse, le membre de droite tend vers 0 quand $n \rightarrow +\infty$ et $p \in \mathbf{N}$, la suite $(y_n)_{n \in \mathbf{N}}$ est de Cauchy, et comme on a supposé C fermé dans un espace complet, elle converge vers un élément $p_C(x)$ appartenant à C . Par continuité de la norme, on a $d(x, p_C(x)) = \lim_{n \rightarrow +\infty} d(x, y_n) = d(x, C)$, ce qui démontre le (i).

Soit $z \in C$. Si $0 < t < 1$, le point $y_t = (1 - t)p_C(x) + tz$ appartient à C , et donc

$$\|p_C(x) - x\|^2 \leq \|y_t - x\|^2 = \|p_C(x) - x\|^2 + t^2\|z - p_C(x)\|^2 + 2t\operatorname{Re}\langle p_C(x) - x, z - p_C(x) \rangle.$$

⁽¹⁷⁾En dimension finie, un petit argument de compacité permettrait de la démontrer (exercice).

En faisant tendre t vers 0, on en déduit l'inégalité $\operatorname{Re}\langle x - p_C(x), z - p_C(x) \rangle \leq 0$. Réciproquement, si $\operatorname{Re}\langle x - y, z - y \rangle \leq 0$, alors

$$\|z - x\|^2 = \|y - x\|^2 + \|z - y\|^2 - 2\operatorname{Re}\langle x - y, z - y \rangle \geq \|y - x\|^2,$$

ce qui montre que, si $\operatorname{Re}\langle x - y, z - y \rangle \leq 0$ quel que soit $z \in C$, alors y vérifie la propriété définissant $p_C(x)$. On en déduit le (ii).

Finalement, si $x, y \in E$, on a

$$\operatorname{Re}\langle x - p_C(x), p_C(y) - p_C(x) \rangle \leq 0 \quad \text{et} \quad \operatorname{Re}\langle y - p_C(y), p_C(x) - p_C(y) \rangle \leq 0.$$

On en déduit, en faisant la somme, l'inégalité

$$\operatorname{Re}\langle x - y, p_C(y) - p_C(x) \rangle \geq \operatorname{Re}\langle p_C(y) - p_C(x), p_C(y) - p_C(x) \rangle = \|p_C(y) - p_C(x)\|^2.$$

On conclut la démonstration du (iii) en utilisant l'inégalité de Cauchy-Schwarz, selon laquelle

$$\operatorname{Re}\langle x - y, p_C(y) - p_C(x) \rangle \leq |\langle x - y, p_C(y) - p_C(x) \rangle| \leq \|x - y\| \cdot \|p_C(y) - p_C(x)\|.$$

II.3. Exercices

1. Espaces de Banach

Exercice II.3.1. — Soient f continue et périodique de période 1 et α irrationnel. Montrer que l'on a

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=1}^N f(n\alpha) = \int_0^1 f(t) dt$$

(commencer par un polynôme trigonométrique).

Exercice II.3.2. — a) La forme linéaire $f \mapsto \int_{-\infty}^{\infty} f(t) dt = I(f)$ sur $\mathcal{C}_c(\mathbf{R})$ s'étend-elle en une forme linéaire continue sur $L^2(\mathbf{R})$?

b) Soit e_n la fonction valant $\frac{1}{n}$ sur $[0, n]$, et 0 ailleurs. Montrer que $e_n \in L^2(\mathbf{R})$, et que la suite $(e_n)_{n \geq 1}$ tend vers 0 dans L^2 . En déduire que le sous-espace $\{f \in \mathcal{C}_c(\mathbf{R}) : I(f) = 0\}$ est dense dans $L^2(\mathbf{R})$.

c) Relation entre a) et b) ?

Exercice II.3.3. — Soit I un intervalle de \mathbf{R} , et soit $(f_n)_{n \in \mathbf{N}}$ une suite de fonctions continues sur I tendant simplement vers une fonction f . Si $j \in \mathbf{N}$, soit

$$F_{n,j} = \cap_{p \geq n} \{x \in I, |f_{n+p}(x) - f_n(x)| \leq 2^{-j}\}.$$

(i) Montrer que $F_{n,j}$ est fermé et que $\cup_{n \in \mathbf{N}} F_{n,j} = I$.

(ii) Soient $U_{n,j}$ l'intérieur de $F_{n,j}$ et $U_j = \cup_{n \in \mathbf{N}} U_{n,j}$. Montrer que U_j est dense dans I .

(iii) Montrer que si $x \in U_j$, il existe V_x ouvert contenant x tel que $|f(x) - f(y)| \leq 2^{2-j}$, si $y \in V_x$. En déduire que f est continue en au moins un point de I .

Exercice II.3.4. — Soit $a = (a_n)_{n \in \mathbf{N}} \in \ell^\infty$ une suite bornée de nombres complexes.

(i) Montrer que, si $x = (x_n)_{n \in \mathbf{N}} \in \ell^2$, alors $(a_n x_n)_{n \in \mathbf{N}} \in \ell^2$, et que l'application linéaire $f : \ell^2 \rightarrow \ell^2$ ainsi définie est continue.

(ii) Montrer que, si f est surjective, il existe $C > 0$ tel que $|a_n| \geq C$, pour tout $n \in \mathbf{N}$. (On pourra commencer par montrer que f est injective.)

Exercice II.3.5. — (i) Montrer que, si $a = (a_n)_{n \in \mathbf{N}} \in \ell^\infty$, et si $b = (b_n)_{n \in \mathbf{N}} \in \ell^1$, alors la série $\sum_{n \in \mathbf{N}} a_n b_n$ converge. On note $\Lambda_a(b)$ la somme de la série.

(ii) Montrer que l'application $\Lambda_a : \ell^1 \rightarrow \mathbf{C}$ ainsi définie appartient à $(\ell^1)^*$, et que $\|\Lambda_a\| = \|a\|_\infty$.

(iii) Montrer que $a \mapsto \Lambda_a$ est une isométrie de ℓ^∞ sur $(\ell^1)^*$. (On pourra s'inspirer de la seconde démonstration du th. II.2.11.)

Exercice II.3.6. — Montrer qu'un espace de Banach possédant une famille génératrice dénombrable est de dimension finie. (Si $(e_n)_{n \in \mathbf{N}}$ est une telle famille, on pourra considérer le sous-espace F_n engendré par les e_i , pour $i \leq n$.)

2. Espaces de Hilbert

Exercice II.3.7. — Quelle est la valeur maximale de $\int_{-1}^1 x f(x) dx$ pour $f \in L^2([-1, 1])$ soumis aux conditions $\int_{-1}^1 f(x) dx = 0$ et $\int_{-1}^1 f(x)^2 = 1$.

Exercice II.3.8. — Soit $(a_n)_{n \in \mathbf{N}}$ une suite de nombres complexes telle que, quelle que soit $(b_n)_{n \in \mathbf{N}} \in \ell^2$, la série $\sum_{n=0}^{+\infty} a_n b_n$ converge. Montrer que $\sum_{n=0}^{+\infty} |a_n|^2 < +\infty$. (On pourra considérer la suite de formes linéaires $\Lambda_k : \ell^2 \rightarrow \mathbf{C}$, pour $k \in \mathbf{N}$, définie par $\Lambda_k((b_n)_{n \in \mathbf{N}}) = \sum_{n=0}^k a_n b_n$. Les nostalgiques des classes préparatoires pourront considérer la suite de terme général $b_n = \frac{\bar{a}_n}{\sum_{i=0}^n |a_i|^2}$.)

Exercice II.3.9. — Soit E un espace de Hilbert de dimension infinie.

(i) Montrer que la boule unité de E n'est pas compacte.

(ii) Construire un sous-ensemble fermé F de E tel que 0 n'ait pas de projection sur F (c'est-à-dire tel qu'il n'existe pas d'élément de F de norme minimale).

Exercice II.3.10. — Soit E un espace de Hilbert et $(e_n)_{n \in \mathbf{N}}$ une famille orthonormale. Soit $(a_n)_{n \in \mathbf{N}}$ une suite de réels positifs. Montrer que $C = B(0, 1) \cap \{\sum_{n=0}^{+\infty} x_n e_n \mid x_n \in [-a_n, a_n]\}$ est compact si et seulement si $\sum_{n=0}^{+\infty} a_n^2 < +\infty$.

Exercice II.3.11. — (Polynômes de Legendre) On note P_n le polynôme $\frac{d^n}{dx^n}(1-x^2)^n$. Montrer que les P_n forment une famille orthogonale dans $L^2([-1, 1])$, et que les $P_n/\|P_n\|_2$ forment une base hilbertienne de $L^2([-1, 1])$.

Exercice II.3.12. — Soit $L^2([0, 1])$ le complété de $\mathcal{C}([0, 1])$ pour le produit scalaire $\langle f, g \rangle = \int_0^1 \overline{f(t)} g(t) dt$.

(i) Soient X_1, \dots, X_n des variables. Montrer que le déterminant de la matrice des $(\frac{1}{X_i + X_j})_{1 \leq i, j \leq n}$ est

$$\Delta_n(X_1, \dots, X_n) = \frac{\prod_{i < j} (X_i - X_j)^2}{\prod_{i, j} (X_i + X_j)}.$$

(ii) Soit $1 \leq a_1 < a_2 < \dots$ une suite d'entiers strictement croissante. Si $n \in \mathbf{N}$, notons $\text{Vect}(x^{a_1}, \dots, x^{a_n})$ le sous-espace de $\mathcal{C}([0, 1])$ engendré par les x^{a_i} , pour $1 \leq i \leq n$. Montrer que, si $k \in \mathbf{N}$, alors

$$d(x^k, \text{Vect}(x^{a_1}, \dots, x^{a_n}))^2 = \frac{\Delta_{n+1}(k + \frac{1}{2}, a_1 + \frac{1}{2}, \dots, a_n + \frac{1}{2})}{\Delta_n(a_1 + \frac{1}{2}, \dots, a_n + \frac{1}{2})}.$$

(iii) Montrer que $\text{Vect}(x^{a_i}, i \in \mathbf{N})$ est dense dans $L^2([0, 1])$, si et seulement si $\sum_{n=1}^{+\infty} \frac{1}{a_i} = +\infty$.

Exercice II.3.13. — Soit E un espace de Hilbert, et soit $A : E \rightarrow E$, linéaire, vérifiant $\langle Ax, y \rangle = \langle x, Ay \rangle$, pour tous $x, y \in E$. Montrer que A est continue. (On utilisera le résultat de l'ex. II.1.21.)

Exercice II.3.14. — Soit $(E, \|\cdot\|)$ un espace de Hilbert.

(i) Soient a, x, y des points de E vérifiant

$$\|x - a\|, \|y - a\| \leq r_2 \quad \text{et} \quad \left\| \frac{x+y}{2} - a \right\| \geq r_1.$$

Montrer que $\|x - y\| \leq 4(r_2^2 - r_1^2)$.

(ii) Soit $(C_n)_{n \in \mathbf{N}}$ une suite décroissante de convexes fermés non vides. Vérifier que $C = \bigcap_{n=0}^{+\infty} C_n$ est un convexe fermé. Montrer que C est non vide si et seulement si $\sup_{n \in \mathbf{N}} d(0, C_n) < +\infty$. On montrera en particulier que sous cette condition, si $x \in E$, alors $P_{C_n}(x)$ tend vers $P_C(x)$.

(iii) Soit $\varphi : E \rightarrow \mathbf{R}$ une fonction convexe (c'est-à-dire telle que $\varphi(tx + (1-t)y) \leq t\varphi(x) + (1-t)\varphi(y)$ quels que soient $t \in [0, 1]$ et $x, y \in E$) telle que $\lim_{\|x\| \rightarrow \infty} \varphi(x) = +\infty$. Montrer que φ est bornée inférieurement sur E et atteint son minimum.

Exercice II.3.15. — (i) Soit $\mathcal{C}_c([1, +\infty[)$ l'espace des $\phi : [1, +\infty[\rightarrow \mathbf{C}$, continues, à support compact. Montrer que $(f, g) \mapsto \langle f, g \rangle = \int_1^{+\infty} \overline{f(t)}g(t) \frac{dt}{t^2}$ est un produit scalaire sur $\mathcal{C}_c([1, +\infty[)$, pour lequel $\mathcal{C}_c([1, +\infty[)$ n'est pas complet.

(ii) On note E le complété de $\mathcal{C}_c([1, +\infty[)$ pour ce produit scalaire. Si n est un entier ≥ 2 , soit $\phi_n : [1, +\infty[\rightarrow \mathbf{C}$ la fonction définie par $\phi_n(t) = [\frac{t}{n}] - \frac{[t]}{n}$. Montrer que $\phi_n \in E$.

(iii) Montrer que l'adhérence dans E de l'espace engendré par les ϕ_n , pour $n \geq 2$, contient l'espace des fonctions constantes sur $[1, +\infty[$. (Indication : consulter l'ex. VII.5.2.)

3. Séries de Fourier

Exercice II.3.16. — Soit f la fonction périodique de période 1 telle que l'on ait $f(x) = x$ si $x \in]-\frac{1}{2}, \frac{1}{2}]$. Calculer les coefficients de Fourier de f ; en déduire la valeur de $\sum_{n=1}^{+\infty} \frac{1}{n^2}$.

Si $f \in \mathcal{C}(\mathbf{R}/\mathbf{Z})$, et si $N \in \mathbf{N}$, soit $\Lambda_N(f) = \sum_{k=-N}^N c_k(f)$. Plus généralement, si $x \in \mathbf{R}/\mathbf{Z}$, soit $\Lambda_{N,x}(f) = \sum_{k=-N}^N c_k(f) e^{2i\pi kx}$. Notre but est d'étudier la convergence des sommes partielles symétriques $\Lambda_{N,x}(f)$ de la série de Fourier de f vers $f(x)$.

Comme question préliminaire, on montrera que Λ_N est une forme linéaire continue sur $\mathcal{C}(\mathbf{R}/\mathbf{Z})$, et on établira la formule $\Lambda_N(f) = \int_0^1 S_N(t) f(t) dt$, où $S_N(t) = \frac{\sin(2N+1)\pi t}{\sin \pi t}$.

Exercice II.3.17. — (i) Montrer que $\|\Lambda_N\| = \|S_N\|_1$.

(ii) Montrer que $\|\Lambda_N\|$ tend vers $+\infty$; en déduire que l'ensemble des $f \in \mathcal{C}(\mathbf{R}/\mathbf{Z})$ tels que $\sup_{N \in \mathbf{N}} |\Lambda_N(f)| = +\infty$ est un G_δ dense dans $\mathcal{C}(\mathbf{R}/\mathbf{Z})$.

(iii) Montrer que, si $r \in \mathbf{Q}$, l'ensemble des $f \in \mathcal{C}(\mathbf{R}/\mathbf{Z})$ tels que $\sup_{N \in \mathbf{N}} |\Lambda_{N,r}(f)| = +\infty$ est un G_δ dense dans $\mathcal{C}(\mathbf{R}/\mathbf{Z})$. En déduire qu'il existe un G_δ dense X de $\mathcal{C}(\mathbf{R}/\mathbf{Z})$ tel que $\sup_{N \in \mathbf{N}} |\Lambda_{N,r}(f)| = +\infty$ quels que soient $f \in X$ et $r \in \mathbf{Q}$.

(iv) Montrer que, si $f \in X$, et si $M \in \mathbf{N}$, alors l'ensemble des $x \in \mathbf{R}/\mathbf{Z}$ tels que $\sup_{N \in \mathbf{N}} |\Lambda_{N,x}(f)| > M$ est un ouvert contenant \mathbf{Q}/\mathbf{Z} . En déduire que, si $f \in X$, il existe un G_δ dense Y_f de \mathbf{R}/\mathbf{Z} tel que $\sup_{N \in \mathbf{N}} |\Lambda_{N,x}(f)| = +\infty$, quel que soit $x \in Y_f$.

⁽¹⁸⁾Autrement dit, il existe un sous-ensemble non dénombrable dense X de $\mathcal{C}(\mathbf{R}/\mathbf{Z})$ tel que, si $f \in X$, alors la série de Fourier de f diverge en tout point d'un sous-ensemble non dénombrable dense de \mathbf{R}/\mathbf{Z} . Malgré ce résultat peu encourageant, Carleson (prix Abel 2006) a démontré en 1965 que, si $f \in \mathcal{C}(\mathbf{R}/\mathbf{Z})$ (et même si $f \in L^2(\mathbf{R}/\mathbf{Z})$), alors $\Lambda_{N,x}(f) \rightarrow f(x)$ pour presque tout x (au sens du chapitre suivant). Par contre, Kolmogorov (1926) a montré qu'il existe des éléments de $L^1(\mathbf{R}/\mathbf{Z})$ dont la transformée de Fourier diverge en tout point.

Exercice II.3.18. — Soit $L^1(\mathbf{R}/\mathbf{Z})$ le complété de $\mathcal{C}(\mathbf{R}/\mathbf{Z})$ pour la norme $\|f\|_1 = \int_0^1 |f(t)| dt$.

- (i) Montrer que, si $k \in \mathbf{Z}$, $f \mapsto c_k(f)$ s'étend par continuité à $L^1(\mathbf{R}/\mathbf{Z})$.
- (ii) Montrer que, si $f \in L^1(\mathbf{R}/\mathbf{Z})$, alors $c_k(f) \rightarrow 0$ quand $|k| \rightarrow +\infty$.
- (iii) On suppose f hölderienne d'exposant $\alpha > 0$ (i.e il existe $C > 0$ tel que l'on ait $|f(x) - f(y)| \leq C|x - y|^\alpha$ quels que soient $x, y \in [0, 1]$). Montrer que, si $x \in \mathbf{R}$, alors la suite de terme général $\sum_{k=-N}^N c_k(f) e^{2i\pi kx}$ tend vers $f(x)$.

CHAPITRE III

INTÉGRATION

III.1. Intégrale de Lebesgue

L'intégrale de Lebesgue (1902-1904) est une extension (ou plutôt une complétion) de l'intégrale de Riemann d'une extrême souplesse. Les espaces de fonctions intégrables deviennent complets, ce qui simplifie grandement les problèmes d'existence ou de convergence d'intégrales, et on dispose grâce au théorème de convergence dominée de Lebesgue, (th. III.1.27) d'un outil extrêmement puissant pour intervertir limites et intégrales. Tous les énoncés classiques de l'intégrale de Riemann (continuité et dérivabilité d'une intégrale dépendant d'un paramètre, théorème de Fubini) s'étendent avec des démonstrations souvent simplifiées. Il est toutefois toujours aussi difficile de calculer explicitement les intégrales dont on a montré l'existence, mais c'est un autre problème...

1. Dallages et fonctions en escalier

Soit $\mathbf{Z}[\frac{1}{2}]$ l'anneau des *nombre dyadiques* (i.e. des nombres rationnels dont le dénominateur est une puissance de 2). Une *dalle* de \mathbf{R}^m est un sous-ensemble de \mathbf{R}^m de la forme⁽¹⁾ $\prod_{j=1}^m [r_j, s_j[$, où $r_j < s_j$, pour $1 \leq j \leq m$, sont des nombres dyadiques. Un *dallage* est une réunion *finie* de dalles. Une *dalle élémentaire* (de taille 2^{-r}) est un ensemble de la forme $D_{r,\mathbf{k}}$, où $r \in \mathbf{N}$, $\mathbf{k} = (k_1, \dots, k_m) \in \mathbf{Z}^m$, et $D_{r,\mathbf{k}} = \prod_{j=1}^m [\frac{k_j}{2^r}, \frac{k_j+1}{2^r}[$.

Lemme III.1.1. — (i) Si D_1 et D_2 sont des dalles élémentaires (pas forcément de même taille), alors de deux choses l'une : soit D_1 et D_2 sont disjointes, soit l'une est incluse dans l'autre⁽²⁾.

(ii) Si $\mathbf{k} \in \mathbf{Z}^m$, alors $D_{r,\mathbf{k}}$ est la réunion disjointe des $D_{r+1,2\mathbf{k}+\mathbf{a}}$, pour $\mathbf{a} \in \{0, 1\}^m$.

(iii) Toute dalle (et tout dallage) est réunion disjointe d'un nombre fini de dalles élémentaires de même taille.

Démonstration. — Exercice.

⁽¹⁾Le lecteur est invité à supposer $m = 1$ ou $m = 2$, et à faire des dessins; cela rend les énoncés qui suivent parfaitement évidents.

⁽²⁾Autrement dit, les dalles élémentaires se comportent comme des billes de mercure.

On définit la *mesure* $\lambda(D_{r,\mathbf{k}})$ d'une dalle élémentaire $D_{r,\mathbf{k}}$ par la formule $\lambda(D_{r,\mathbf{k}}) = 2^{-mr}$. Si D est un dallage quelconque, réunion disjointes des D_{r,\mathbf{k}_i} , pour $i \in I$ ensemble fini, on définit $\lambda(D)$ par $\lambda(D) = \sum_{i \in I} \lambda(D_{r,\mathbf{k}_i})$. On vérifie que cela ne dépend pas du choix de la taille des dalles élémentaires choisies pour recouvrir D , en utilisant le (ii) du lemme précédent.

Si $r \in \mathbf{N}$, et si $\mathbf{k} \in \mathbf{Z}^m$, on note $e_{r,\mathbf{k}}$ la fonction caractéristique de $D_{r,\mathbf{k}}$. Si $r \in \mathbf{N}$, on note $\text{Esc}_r(\mathbf{R}^m)$, l'espace vectoriel des combinaisons linéaires des $e_{r,\mathbf{k}}$, pour $\mathbf{k} \in \mathbf{Z}^m$; c'est l'ensemble des *fonctions en escalier sur \mathbf{R}^m , constantes sur les dalles élémentaires de taille 2^{-r}* . Les $e_{r,\mathbf{k}}$, pour $\mathbf{k} \in \mathbf{Z}^m$, ayant des supports disjoints, forment en fait une base de $\text{Esc}_r(\mathbf{R}^m)$.

Comme $e_{r,\mathbf{k}} = \sum_{\mathbf{a} \in \{0,1\}^m} e_{r+1,2\mathbf{k}+\mathbf{a}}$, on a $\text{Esc}_r(\mathbf{R}^m) \subset \text{Esc}_{r+1}(\mathbf{R}^m)$, si $r \in \mathbf{N}$. La réunion $\text{Esc}(\mathbf{R}^m)$ des $\text{Esc}_r(\mathbf{R}^m)$, pour $r \in \mathbf{N}$, est donc un espace vectoriel; c'est l'espace vectoriel des *fonctions en escalier sur \mathbf{R}^m* . (On notera qu'une combinaison linéaire étant une somme finie, une fonction en escalier est à support compact.)

Si $X \subset \mathbf{R}^m$, on note $\text{Esc}(X)$ l'ensemble des fonctions en escalier sur X ; il peut se voir comme l'ensemble des fonctions en escalier sur \mathbf{R}^m qui sont nulles en dehors de X . Plus généralement, si $X \subset \mathbf{R}^m$, et si $U \subset \mathbf{C}$, on note $\text{Esc}(X, U)$ l'ensemble des fonctions, en escalier sur X , prenant leurs valeurs dans U . On a donc $\text{Esc}(X) = \text{Esc}(X, \mathbf{C})$.

Si $\phi \in \text{Esc}(\mathbf{R}^m)$, il existe $r \in \mathbf{N}$ tel que l'on puisse écrire ϕ sous la forme $\phi = \sum_{i \in I} \alpha_i e_{r,\mathbf{k}_i}$, et on définit son intégrale $\int \phi$ par $\int \phi = \sum_{i \in I} \alpha_i \lambda(D_{r,\mathbf{k}_i}) = 2^{-rm} \sum_{i \in I} \alpha_i$. On vérifie, en utilisant la formule $e_{r,\mathbf{k}} = \sum_{\mathbf{a} \in \{0,1\}^m} e_{r+1,2\mathbf{k}+\mathbf{a}}$, que cela ne dépend pas du choix de r , et que l'on a défini de la sorte une forme linéaire sur $\text{Esc}(\mathbf{R}^m)$. Suivant le contexte, l'intégrale de ϕ sera notée indifféremment

$$\int \phi = \int_{\mathbf{R}^m} \phi = \int_{\mathbf{R}^m} \phi d\lambda = \int_{\mathbf{R}^m} \phi(x) dx = \int_{\mathbf{R}^m} \phi(x) dx_1 \cdots dx_m.$$

Remarque III.1.2. — Si $\phi \in \mathcal{C}_c(\mathbf{R}^m)$ est une fonction continue à support compact, et si on note ϕ_r la fonction en escalier $\sum_{\mathbf{k} \in \mathbf{Z}^m} \phi(\frac{\mathbf{k}}{2^r}) e_{r,\mathbf{k}}$ prenant la valeur $\phi(\frac{k_1}{2^r}, \dots, \frac{k_m}{2^r})$ sur la dalle élémentaire $\prod_{j=1}^m [\frac{k_j}{2^r}, \frac{k_j+1}{2^r}[$, quel que soit $\mathbf{k} \in \mathbf{Z}^m$, alors $\phi_r \rightarrow \phi$ en norme $\|\cdot\|_\infty$ (i.e. uniformément⁽³⁾). Ceci permet de montrer que la suite $(\int \phi_r)_{r \in \mathbf{N}}$ converge, et on définit *l'intégrale de Riemann* $\int \phi$ comme la limite de cette suite. L'intégrale de Lebesgue va être définie de la même manière, mais en relâchant⁽⁴⁾ la condition de convergence uniforme, remplacée par la condition de *convergence simple presque partout*.

⁽³⁾C'est une conséquence de l'uniforme continuité d'une fonction continue sur un compact.

⁽⁴⁾Comme la convergence uniforme implique la convergence simple presque partout définie plus loin, les intégrales de Riemann et de Lebesgue d'une fonction continue à support compact coïncident, et l'intégrale de Lebesgue est bien une extension de l'intégrale de Riemann.

2. Ensembles de mesure nulle

On définit la *mesure extérieure* $\lambda^+(A)$ d'un ensemble A comme la borne inférieure des $\sum_{n \in \mathbf{N}} \lambda(D_n)$, où $(D_n)_{n \in \mathbf{N}}$ décrit l'ensemble des suites de dallages de \mathbf{R}^m telles que $A \subset \cup_{n \in \mathbf{N}} D_n$. En décomposant les dallages D_n en dalles élémentaires disjointes de taille décroissante avec n , et en éliminant les dalles élémentaires de D_n incluses dans un des D_i , pour $i \leq n-1$, on peut se ramener au cas où les D_n sont des dalles élémentaires disjointes⁽⁵⁾.

Proposition III.1.3. — (i) Si $B \subset A$, alors $\lambda^+(B) \leq \lambda^+(A)$.

(ii) Si $A_n \subset \mathbf{R}^m$, pour $n \in \mathbf{N}$, alors $\lambda^+(\cup_{n \in \mathbf{N}} A_n) \leq \sum_{n \in \mathbf{N}} \lambda^+(A_n)$.

Démonstration. — Le (i) est une évidence. Le (ii) est évident si $\sum_{n \in \mathbf{N}} \lambda^+(A_n) = +\infty$. Dans le cas contraire, soit $A = \cup_{n \in \mathbf{N}} A_n$, et soit $\varepsilon > 0$. Pour tout $n \in \mathbf{N}$, on peut trouver une suite $(D_{n,k})_{k \in \mathbf{N}}$ de dallages de \mathbf{R}^m tels que $A_n \subset \cup_{k \in \mathbf{N}} D_{n,k}$ et $\sum_{k \in \mathbf{N}} \lambda(D_{n,k}) \leq \lambda^+(A_n) + 2^{-1-n}\varepsilon$. Mais alors $A \subset \cup_{k,n \in \mathbf{N}} D_{n,k}$ et

$$\sum_{k,n \in \mathbf{N}} \lambda(D_{n,k}) \leq \sum_{n \in \mathbf{N}} (\lambda^+(A_n) + 2^{-1-n}\varepsilon) = \varepsilon + \sum_{n \in \mathbf{N}} \lambda^+(A_n).$$

On en déduit que $\lambda^+(A) \leq \varepsilon + \sum_{n \in \mathbf{N}} \lambda^+(A_n)$ quel que soit $\varepsilon > 0$, ce qui permet de conclure.

On dit que A est *de mesure nulle* si $\lambda^+(A) = 0$. En revenant à la définition, on voit que A est de mesure nulle si et seulement si, quel que soit $\varepsilon > 0$, il existe une suite $(D_n)_{n \in \mathbf{N}}$ de dallages de \mathbf{R}^m tels que $A \subset \cup_{n \in \mathbf{N}} D_n$ et $\sum_{n \in \mathbf{N}} \lambda(D_n) < \varepsilon$.

Proposition III.1.4. — (i) Tout sous-ensemble d'un ensemble de mesure nulle est de mesure nulle.

(ii) Une réunion dénombrable d'ensembles de mesure nulle est de mesure nulle.

Démonstration. — C'est une conséquence immédiate de la prop. III.1.3.

Exercice III.1.5. — Montrer que, si A est de mesure nulle dans \mathbf{R}^n , alors $A \times \mathbf{R}^m$ est de mesure nulle dans \mathbf{R}^{n+m} .

Exercice III.1.6. — (i) Montrer que la diagonale dans \mathbf{R}^2 est de mesure nulle.

(ii) Montrer, plus généralement, que le graphe dans \mathbf{R}^2 d'une application continue $f : \mathbf{R} \rightarrow \mathbf{R}$ est de mesure nulle, et qu'un hyperplan est de mesure nulle dans \mathbf{R}^m .

(iii) L'image dans \mathbf{R}^2 de $[0, 1]$ par une application continue est-elle nécessairement de mesure nulle ?

⁽⁵⁾Le résultat n'est pas sans rappeler la manière dont une image apparaît sur un ordinateur.

On dit qu'une propriété est *vraie presque partout*, ou bien est *vraie pour presque tout* $x \in \mathbf{R}^m$, ou encore *est vraie p.p.*, si l'ensemble des points ne la vérifiant pas est de mesure nulle. Par exemple, l'ensemble des rationnels étant dénombrable, presque tout réel est irrationnel. De même, presque tout ⁽⁶⁾ nombre complexe est transcendant. (Un nombre complexe est *algébrique* s'il est racine d'un polynôme unitaire à coefficients dans \mathbf{Q} . Un nombre complexe non algébrique est un *nombre transcendant*.)

Remarque III.1.7. — (i) Il ne faudrait pas croire qu'un sous-ensemble de \mathbf{R} de mesure nulle est forcément dénombrable. Par exemple, fixons une bijection $n \mapsto r_n$ de \mathbf{N} sur \mathbf{Q} , et soient $U_k = \cup_{n \in \mathbf{N}}]r_n - 2^{-n-k}, r_n + 2^{-n-k}[$, si $k \in \mathbf{N}$, et $A = \cap_{k \in \mathbf{N}} U_k$. Alors A est de mesure nulle puisqu'il est inclus dans U_k , pour tout k , et que $\lambda^+(U_k) \leq \sum_{n=0}^{+\infty} 2^{1-n-k} = 2^{2-k}$ peut être rendu arbitrairement petit. D'un autre côté, U_k est un ouvert dense pour tout k , et donc A est dense et non dénombrable d'après le lemme de Baire (cf. n° 9.2 du chap. « Vocabulaire mathématique »). En particulier, A contient bien d'autres éléments que les rationnels, ce qui n'est pas totalement transparent sur sa construction.

(ii) Un ensemble de mesure nulle peut avoir des propriétés assez surprenantes. Par exemple, A. Besicovitch (1919) a construit des ensembles de mesure nulle de \mathbf{R}^m , pour $m \geq 2$, contenant un segment de longueur 1 dans toutes les directions⁽⁷⁾.

Exercice III.1.8. — Montrer qu'une fonction continue, qui est nulle p.p., est identiquement nulle.

Théorème III.1.9. — (Borel-Cantelli) *Si $(A_n)_{n \in \mathbf{N}}$ est une suite de sous-ensembles de \mathbf{R}^m telle que $\sum_{n \in \mathbf{N}} \lambda^+(A_n) < +\infty$, alors presque tout $x \in \mathbf{R}^m$ n'appartient qu'à un nombre fini de A_n .*

Démonstration. — Il s'agit de prouver que l'ensemble A des $x \in \mathbf{R}$ appartenant à une infinité de A_n est de mesure nulle. Or $x \in A$ si et seulement si, quel que soit $n \in \mathbf{N}$, il

⁽⁶⁾Au vu de ces résultats, il est raisonnable de penser qu'un nombre n'ayant pas de bonnes raisons d'être algébrique est transcendant. Démontrer qu'un nombre donné est transcendant ou même irrationnel est, en général, très difficile. Par exemple, il a fallu attendre 1979 pour que R. Apéry démontre que $\zeta(3)$ est irrationnel, et il n'y a aucun entier n impair ≥ 5 , pour lequel on sache prouver que $\zeta(n)$ est irrationnel. Le seul résultat dans cette direction est un résultat de T. Rivoal (2000) qui a démontré que $\zeta(n)$ est irrationnel pour une infinité de n impairs, et qu'au moins un des 9 nombres $\zeta(5), \dots, \zeta(21)$ est irrationnel (amélioré depuis par W. Zudilin : au moins un des 4 nombres $\zeta(5), \zeta(7), \zeta(9), \zeta(11)$ est irrationnel).

⁽⁷⁾Ces ensembles sont sources de contre-exemples empoisonnants en analyse (par exemple pour la convergence des séries de Fourier en dimension ≥ 2). Les analystes seraient plutôt contents si on arrivait à démontrer (problème de Kakeya) qu'un ensemble de Besicovitch n'est pas trop petit, et plus précisément qu'un ensemble de Besicovitch de \mathbf{R}^m est de dimension de Minkowski m : la *dimension de Minkowski* d'un ensemble X , si elle existe, est la limite de $\frac{\log N(X, k)}{\log k}$ quand $k \rightarrow +\infty$, où $N(X, k)$ est le nombre minimum de boules de rayon $\frac{1}{k}$ nécessaires pour recouvrir complètement X . Le meilleur résultat connu pour le moment est dû à N. Katz et T. Tao (2001) : la dimension de Minkowski d'un ensemble de Besicovitch de \mathbf{R}^m est au moins $(2 - \sqrt{2})(m - 4) + 3$, si $m > 4$.

existe $p \geq n$ tel que $x \in A_p$; autrement dit, on a $A = \bigcap_{n \in \mathbf{N}} \left(\bigcup_{p \geq n} A_p \right)$. On en déduit que $\lambda^+(A) \leq \sum_{p \geq n} \lambda^+(A_p)$, quel que soit $n \in \mathbf{N}$, et comme la série $\sum_{n \in \mathbf{N}} \lambda^+(A_n)$ est supposée convergente, son reste tend vers 0, ce qui permet de conclure.

Exercice III.1.10. — (i) Montrer que, si $\varepsilon > 0$ et $C > 0$, alors pour presque tout⁽⁸⁾ nombre réel x , l'ensemble des couples d'entiers (p, q) , tels que $|x - \frac{p}{q}| \leq Cq^{-2-\varepsilon}$, est fini.

(ii) Montrer que l'ensemble des nombres de Liouville est de mesure nulle, non dénombrable, et dense dans \mathbf{R} . (Un réel x est *de Liouville* si ce n'est pas un nombre rationnel et si, quel que soit $n \in \mathbf{N}$, il existe un couple d'entiers (p, q) , $q \geq 2$, tels que $|x - \frac{p}{q}| \leq q^{-n}$.)

3. Fonctions mesurables, ensembles mesurables

Une fonction $f : \mathbf{R}^m \rightarrow \mathbf{C}$ est dite *mesurable* si elle est limite p.p. d'une suite de fonctions en escalier. Autrement dit, f est mesurable s'il existe $A \subset \mathbf{R}^m$ de mesure nulle, et une suite $(f_n)_{n \in \mathbf{N}}$ d'éléments de $\text{Esc}(\mathbf{R}^m)$ tels que $f_n(x) \rightarrow f(x)$, quel que soit $x \notin A$.

Exercice III.1.11. — (i) Montrer que les fonctions mesurables forment une algèbre (i.e. que λf , $f + g$ et fg sont mesurables si $\lambda \in \mathbf{C}$, f, g mesurables).

(ii) Montrer que, si f est mesurable sur \mathbf{R}^n et g est mesurable sur \mathbf{R}^m , alors $h(x, y) = f(x)g(y)$ est mesurable sur \mathbf{R}^{n+m} .

(iii) Montrer que, si f et g sont mesurables, alors $\inf(f, g)$ et $\sup(f, g)$ sont mesurables.

(iv) Montrer qu'une fonction continue est mesurable.

Proposition III.1.12. — Une limite simple p.p. de fonctions mesurables est mesurable.

Démonstration. — Cette proposition est moins évidente qu'il n'y paraît (cf. ex. II.1.11). La démonstration sera faite au n° 4 du § III.4.

Exercice III.1.13. — (i) Montrer qu'une fonction est mesurable si et seulement si elle est limite simple p.p. d'une suite de fonctions continues.

(ii) En déduire qu'une limite simple p.p. de limites simples p.p. de fonctions continues est limite simple p.p. d'une suite de fonctions continues. Comparer avec l'exercice II.1.11.

⁽⁸⁾K. Roth a obtenu la médaille Fields en 1958 pour avoir démontré que les nombres algébriques ont cette propriété (c'est évident pour les rationnels ou les nombres algébriques de degré 2 (si α est un nombre algébrique, le *degré* de α est le minimum des degrés des polynômes $P \in \mathbf{Q}[X]$ non nuls, avec $P(\alpha) = 0$), mais le cas général représente un tour de force). On ne sait pas démontrer que π vérifie cette propriété. Dans le même genre d'idées, si x est un nombre réel non rationnel, on peut prendre son développement en fractions continues (il s'agit de la suite d'entiers $(a_n)_{n \in \mathbf{N}}$ définie par l'algorithme $x_0 = x$, $a_n = [x_n]$, $x_{n+1} = \frac{1}{x_n - a_n}$; les nombres rationnels $\frac{u_n}{v_n}$, obtenus en écrivant x en termes de a_0, \dots, a_{n-1} et x_n , et en remplaçant x_n par a_n dans l'expression, sont les meilleures approximations de x par des nombres rationnels). On montre facilement que, pour presque tout nombre réel x , la suite a_n n'est pas bornée, ce qui équivaut à la nullité de la borne inférieure de l'ensemble des $q|qx - p|$, pour $p, q \in \mathbf{Z}$, $q \geq 1$. Si x est algébrique de degré 2, la suite a_n devient périodique à partir d'un certain rang et donc est bornée. On est persuadé que si x est algébrique de degré ≥ 3 , alors la suite a_n n'est pas bornée, mais on ne sait le démontrer dans aucun cas.

On note $\text{Mes}(\mathbf{R}^m)$ le \mathbf{C} -espace vectoriel des fonctions mesurables sur \mathbf{R}^m . Plus généralement, si $D \subset \mathbf{R}^m$ et $F \subset \mathbf{C}$ (ou $F \subset \overline{\mathbf{R}}_+$), on note $\text{Mes}(D, F)$ l'ensemble des fonctions mesurables sur \mathbf{R}^m , qui sont nulles en dehors de D et qui prennent leurs valeurs dans F .

Un sous-ensemble X de \mathbf{R}^m est *mesurable* si sa fonction caractéristique $\mathbf{1}_X$ est une fonction mesurable⁽⁹⁾. Si X est mesurable, et si $(f_k)_{k \in \mathbf{N}}$ est une suite de fonctions en escalier tendant simplement vers $\mathbf{1}_X$ en dehors de A , où A est de mesure nulle, on a aussi $\mathbf{1}_{X_k} \rightarrow \mathbf{1}_X$ en dehors de A , si X_k est le dallage des $x \in \mathbf{R}^m$ vérifiant $|f_k(x) - 1| \leq \frac{1}{2}$. Autrement dit, X est mesurable si et seulement si il existe une suite de dallages $(X_k)_{k \in \mathbf{N}}$ telle que $\mathbf{1}_{X_k}$ tende vers $\mathbf{1}_X$ p.p.

Exercice III.1.14. — (i) Montrer que si X est mesurable, alors $\mathbf{R}^m - X$ est mesurable.

(ii) Montrer qu'une réunion dénombrable et une intersection dénombrable d'ensembles mesurables sont mesurables.

(iii) Montrer que, si X est mesurable dans \mathbf{R}^n et Y est mesurable dans \mathbf{R}^m , alors $X \times Y$ est mesurable dans \mathbf{R}^{n+m} .

(iv) Montrer qu'un ouvert est mesurable, qu'un fermé est mesurable.

(v) Montrer que, si $S \subset [0, 1]$ est un système de représentants de \mathbf{R}/\mathbf{Q} (i.e. tout élément de \mathbf{R} peut s'écrire de manière unique sous la forme $s + q$, avec $s \in S$ et $q \in \mathbf{Q}$), alors S n'est pas mesurable.

4. Définition de l'intégrale de Lebesgue

Une fois qu'on a défini ce qu'était une fonction mesurable, on peut se demander comment la mesurer. Une mesure naturelle pour beaucoup de questions est fournie par l'intégrale de Lebesgue (et la mesure de Lebesgue). La présentation que nous avons choisie est purement axiomatique⁽¹⁰⁾; elle n'est pas sans rappeler le point de vue concernant les

⁽⁹⁾S. Banach et A. Tarski (1924) ont construit un découpage d'une boule de rayon 1 dans \mathbf{R}^3 en un nombre fini de morceaux (5 morceaux suffisent), tel que si on réarrange ces morceaux (i.e. si on les bouge par des isométries de \mathbf{R}^3), on obtient deux boules de rayon 1 (*paradoxe de Banach-Tarski*). Ces morceaux ne sont pas mesurables, et la construction de Banach et Tarski utilise l'axiome du choix.

D'un autre côté, R. Solovay (1966) a démontré que, si on s'interdit l'axiome du choix non dénombrable, tout en gardant l'axiome du choix dénombrable, on peut, sans introduire de contradiction supplémentaire aux mathématiques, supposer que tout ensemble est mesurable (en analyse, on peut difficilement se passer de l'axiome du choix dénombrable, qui est de toute façon assez raisonnable : on peut parfaitement imaginer que face au problème de choisir un nombre dénombrable d'éléments, on devienne de plus en plus performant, et donc arriver à choisir tous ces éléments en un temps fini ; avec un nombre non dénombrable d'éléments, ceci est voué à l'échec, si on doit les choisir un par un). La leçon à retenir de ce résultat est, qu'en pratique, toutes les fonctions rencontrées en analyse sont mesurables et qu'il est inutile de passer son temps à vérifier qu'elles le sont ; le problème est différent en théorie des probabilités où un même événement peut être mesurable ou non mesurable suivant les conditions.

⁽¹⁰⁾Ceci a pour avantage de faire glisser sous le tapis certains points assez délicats sans diminuer la facilité d'utilisation de la théorie. Toutefois, dans le but d'atténuer le sentiment d'inconfort que ressent toujours un peu un esprit mathématicien à l'idée d'utiliser un résultat dont il n'a pas vu (et souvent oublié) une démonstration, nous avons inclus, au § III.4, une construction de l'intégrale satisfaisant les axiomes.

nombres réels adopté en classes préparatoires (où \mathbf{R} est présenté comme un corps totalement ordonné dans lequel toute partie non vide admet une borne supérieure, le rôle joué par la propriété de la borne supérieure dans cette présentation de \mathbf{R} étant tenu ici par le théorème de convergence monotone).

Théorème III.1.15. — *Il existe une unique application $f \mapsto \int f$ de $\text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$ dans $\overline{\mathbf{R}}_+$ vérifiant les propriétés (i)–(v) ci-dessous.*

- (i) Si f est en escalier, alors $\int f$ est la quantité précédemment définie.
- (ii) (linéarité) $\int (af + bg) = a \int f + b \int g$, si $a, b \in \mathbf{R}_+$, et $f, g \in \text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$.
- (iii) $\int f = 0$ si et seulement si $f = 0$ p.p.
- (iv) Si $f \leq g$ p.p., alors $\int f \leq \int g$.
- (v) (théorème de convergence monotone) Si (f_n) est une suite croissante d'éléments de $\text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$, alors $\lim_{n \rightarrow +\infty} \int f_n = \int \lim_{n \rightarrow +\infty} f_n$.

Remarque III.1.16. — (i) Les propriétés fondamentales sont la linéarité de l'intégration (propriété (ii)) et le théorème de convergence monotone (propriété (v)).

(ii) La propriété (iv) découle des (ii) et (iii) : en effet, on a $\int g + \int -\inf(0, g - f) = \int f + \int \sup(g - f, 0)$, où toutes les fonctions sont positives et $-\inf(0, g - f) = 0$ p.p. par hypothèse.

(iii) La propriété (i) est une normalisation qui définit une mesure bien particulière sur \mathbf{R}^m , à savoir la mesure de Lebesgue. Comme nous le verrons (th. III.3.7), cette propriété implique que la mesure de Lebesgue est invariante par translation. Réciproquement, si $f \mapsto \int f$ est linéaire et invariante par translation [ce qui signifie que, pour tous $f \in \text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$ et $a \in \mathbf{R}^m$, on a $\int T_a(f) = \int f$, où $(T_a(f))(t) = f(t + a)$], et si $\int e_{0,0} = 1$, alors $\int e_{r,k}$ ne dépend pas de k , grâce à l'invariance par translation, et donc vaut 2^{-rm} , grâce à la normalisation $\int e_{0,0} = 1$ (cela suit, par une récurrence sur r , de l'indépendance de $\int e_{r,k}$ par rapport à k et de la formule $e_{r,k} = \sum_{a \in \{0,1\}^m} e_{r+1,k+a}$) ; la linéarité implique alors que $\int f$ vérifie la propriété (i). Autrement dit, on aurait pu remplacer⁽¹¹⁾ la propriété (i) par l'invariance par translation et la normalisation $\int e_{0,0} = 1$.

Si $X \subset \mathbf{R}^m$ est mesurable, sa mesure de Lebesgue est définie par $\lambda(X) = \int \mathbf{1}_X \in \overline{\mathbf{R}}_+$.

Proposition III.1.17. — *Soit $f \in \text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$.*

- (i) $\int f = 0$ si et seulement si $\lambda(\{t, f(t) \neq 0\}) = 0$.
- (ii) Si $\int f < +\infty$, alors $\lambda(\{t, f(t) = +\infty\}) = 0$.

⁽¹¹⁾Ce procédé s'étend à tout groupe localement compact, ce qui inclut \mathbf{R}^m , les groupes finis, les groupes compacts, \mathbf{Z} , le groupe additif \mathbf{Q}_p , les groupes multiplicatifs \mathbf{R}^* , \mathbf{C}^* ou \mathbf{Q}_p^* , ou plus généralement, si $d \geq 1$, les groupes $\mathbf{GL}_d(\mathbf{R})$, $\mathbf{GL}_d(\mathbf{C})$, $\mathbf{GL}_d(\mathbf{Q}_p)$ et leurs sous-groupes fermés... Un tel groupe possède, à multiplication près par une constante > 0 , une unique mesure invariante à droite (resp. à gauche), ce qui signifie que $\int T_a^d(f) = \int f$ (resp. $\int T_a^g(f) = \int f$), où $(T_a^d(f))(x) = f(xa^{-1})$ et $(T_a^g(f))(x) = f(ax)$; cette mesure est « la » mesure de Haar à droite (resp. à gauche).

Démonstration. — Soit $A = \{t, f(t) \neq 0\}$. On a $\inf(f, N) \leq N \mathbf{1}_A$, et donc $\int \inf(f, N) \leq N \lambda(A)$, pour tout $N \in \mathbf{N}$. Comme $\inf(f, N)$ tend en croissant vers f , le théorème de convergence monotone montre que $\int f = 0$ si $\lambda(A) = 0$. Réciproquement, si $\int f = 0$, alors Nf tend en croissant vers $+\infty \mathbf{1}_A$, et comme $\int Nf = 0$ pour tout N , on a $+\infty \lambda(A) = 0$ et donc $\lambda(A) = 0$. Ceci démontre le (i).

Pour démontrer le (ii), il suffit de remarquer que si $A = \{t, f(t) = +\infty\}$, alors $f \geq M \mathbf{1}_A$ quel que soit $M \in \mathbf{R}_+$, et donc $M \lambda(A) \leq \int f$ quel que soit $M \in \mathbf{R}_+$.

Corollaire III.1.18. — X est de mesure nulle si et seulement si $\lambda(X) = 0$.

Démonstration. — Si $X \subset \mathbf{R}^m$ est mesurable, alors

$$X \text{ est de mesure nulle} \iff \mathbf{1}_X = 0 \text{ p.p.} \iff \int \mathbf{1}_X = 0 \iff \lambda(X) = 0.$$

La première équivalence est par définition, la seconde suit du (iii) du th. III.1.15, et la dernière résulte du (i) de la prop. III.1.17.

Le cor. III.1.18 est un cas particulier du résultat suivant, démontré au n° 2 du § III.4.

Théorème III.1.19. — Si X est mesurable, alors $\lambda(X) = \lambda^+(X)$.

Exercice III.1.20. — Soient $X_n \subset \mathbf{R}^m$ mesurables, pour $n \in \mathbf{N}$, et soient $B = \cup_{n \in \mathbf{N}} X_n$ et $C = \cap_{n \in \mathbf{N}} X_n$.

(i) Montrer que si la suite X_n est croissante, alors $\lambda(B) = \lim_{n \rightarrow +\infty} \lambda(X_n)$.

(ii) Montrer que si la suite X_n est décroissante et si $\lambda(X_0) < +\infty$, alors $\lambda(C) = \lim_{n \rightarrow +\infty} \lambda(X_n)$. Le résultat est-il toujours valable sans l'hypothèse $\lambda(X_0) < +\infty$?

Exercice III.1.21. — Soit $X \subset \mathbf{R}^m$ mesurable de mesure finie. Montrer que, quel que soit $\alpha \in [0, 1]$, il existe $B \subset X$ mesurable tel que $\lambda(B) = \alpha \lambda(X)$. (On pourra s'intéresser à $f(t) = \lambda(X \cap [-t, t]^m)$.)

Exercice III.1.22. — (i) Montrer, en revenant à la définition, que $\lambda^+([0, 1]) = 1$ (il n'est pas si évident que $\lambda^+([0, 1]) \neq 0 \dots$).

(ii) Un *pavé* de \mathbf{R}^m est un ensemble de la forme $\prod_{j=1}^m I_j$, où I_j est un intervalle de n'importe quel type (ouvert ou fermé à chacune des extrémités). Montrer que, si $P = \prod_{j=1}^m I_j$ est un pavé, alors $\lambda^+(P) = \prod_{j=1}^m \lambda(I_j)$, où $\lambda(I_j)$ est la longueur de l'intervalle I_j [i.e. si $a \leq b$ sont deux réels, alors $\lambda([a, b]) = \lambda([a, b[) = \lambda(]a, b]) = \lambda(]a, b[) = b - a]$.

Définition III.1.23. — (Intégrale de Lebesgue)

(i) On dit que $f \in \text{Mes}(\mathbf{R}^m, \mathbf{C})$ est *sommable*, si $\int |f| < +\infty$. On note $\mathcal{L}^1(\mathbf{R}^m) \subset \text{Mes}(\mathbf{R}^m)$ l'ensemble des fonctions sommables.

(ii) Si $f \in \mathcal{L}^1(\mathbf{R}^m)$, les fonctions⁽¹²⁾ $\text{Re}^+(f)$, $\text{Re}^+(-f)$, $\text{Re}^+(if)$ et $\text{Re}^+(-if)$ sont sommables puisque majorées par $|f|$, et on pose

$$\int f = \int \text{Re}^+(f) - \int \text{Re}^+(-f) + i \int \text{Re}^+(-if) - i \int \text{Re}^+(if).$$

⁽¹²⁾Si $z \in \mathbf{C}$, on note $\text{Re}^+(z)$ le nombre réel $\sup(0, \text{Re}(z))$; on a alors

$$z = \text{Re}^+(z) - \text{Re}^+(-z) + i \text{Re}^+(-iz) - i \text{Re}^+(iz).$$

Remarque III.1.24. — (i) L'intégrale $f \mapsto \int f$ est \mathbf{C} -linéaire : cela résulte formellement du (ii) du th. III.1.15.

(ii) Si f est sommable, si $\int f = \rho e^{i\theta}$, et si h est définie par $f = |f|e^{ih}$ p.p., alors

$$\int |f| - \left| \int f \right| = \operatorname{Re} \left(\int |f| - \left| \int f \right| \right) = \operatorname{Re} \left(\int |f| - \int e^{-i\theta} f \right) = \operatorname{Re} \left(\int |f|(1 - e^{ih-i\theta}) \right) \geq 0,$$

puisque la fonction $\operatorname{Re}(|f|(1 - e^{ih-i\theta}))$ est positive. *En résumé*, $\left| \int f \right| \leq \int |f|$.

(iii) Si $X \subset \mathbf{R}^m$ est mesurable, et si $\phi \in \mathcal{L}^1(\mathbf{R}^m)$, alors $\mathbf{1}_X \phi \in \mathcal{L}^1(\mathbf{R}^m)$ car $|\mathbf{1}_X \phi| \leq |\phi|$. On note $\int_X \phi$, si $\phi \in \mathcal{L}^1(\mathbf{R}^m)$, l'intégrale $\int_{\mathbf{R}^m} \mathbf{1}_X \phi$.

5. Les théorèmes de convergence monotone et de convergence dominée

Soit X un sous-ensemble mesurable de \mathbf{R}^m . On dit que $\phi : X \rightarrow \mathbf{C}$ est mesurable si la fonction obtenue en prolongeant ϕ par 0 à \mathbf{R}^m est mesurable. L'ensemble $\operatorname{Mes}(X)$ des fonctions mesurables sur X est donc naturellement un sous-espace de $\operatorname{Mes}(\mathbf{R}^m)$, et on définit $\mathcal{L}^1(X)$ comme l'intersection de $\operatorname{Mes}(X)$ avec $\mathcal{L}^1(\mathbf{R}^m)$.

L'application $\phi \rightarrow \mathbf{1}_X \phi$ est une projection de $\mathcal{L}^1(\mathbf{R}^m)$ sur son sous-espace $\mathcal{L}^1(X)$.

Théorème III.1.25. — (de convergence monotone)

(i) Si $(f_n)_{n \in \mathbf{N}}$ est une suite croissante d'éléments de $\operatorname{Mes}(X, \overline{\mathbf{R}}_+)$, alors la limite est mesurable et $\lim_{n \rightarrow +\infty} \int f_n = \int \lim_{n \rightarrow +\infty} f_n$.

(ii) Si $(u_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\operatorname{Mes}(X, \overline{\mathbf{R}}_+)$, alors la somme de la série est mesurable et $\int \sum_{n \in \mathbf{N}} u_n = \sum_{n \in \mathbf{N}} \int u_n$.

Démonstration. — Le (i) est immédiat (compte-tenu du théorème de convergence uniforme sur \mathbf{R}^m , cf. (v) du th. III.1.15). Le (ii) se déduit du (i) en considérant les sommes partielles.

Proposition III.1.26. — (Lemme de Fatou) Si $f_n \in \operatorname{Mes}(X, \overline{\mathbf{R}}_+)$, alors

$$\int (\liminf f_n) \leq \liminf \int f_n.$$

Démonstration. — Soit $g_n = \inf_{k \geq n} f_k$. Alors g_n est mesurable comme limite simple de fonctions mesurables, et $g_n \rightarrow \liminf f_n$ en croissant, par définition de la limite inférieure. D'après le théorème de convergence monotone, on a donc $\int g_n \rightarrow \int \liminf f_n$. Par ailleurs, $\int g_n \leq \int f_n$ quel que soit n . On a donc

$$\int \liminf f_n = \lim \int g_n = \liminf \int g_n \leq \liminf \int f_n,$$

ce qu'il fallait démontrer.

Si $f \in \mathcal{L}^1(X)$, on définit sa (semi)-norme $\|f\|_1$ par $\|f\|_1 = \int |f|$. Il résulte du (iii) du th. III.1.15 que l'on a $\|f\|_1 = 0$ si et seulement si $f = 0$ p.p. On dit que f_k tend vers f dans $\mathcal{L}^1(X)$ (ou que f_k tend vers f en moyenne), si $\|f_k - f\|_1 \rightarrow 0$. Comme $\|\cdot\|_1$ est une

semi-norme et pas une norme, la limite d'une suite n'est pas unique; de fait si $f_k \rightarrow f$ dans $\mathcal{L}^1(X)$, alors $f_k \rightarrow g$ dans $\mathcal{L}^1(X)$ pour tout g tel que $g - f = 0$ p.p.

Il résulte du (ii) de la rem. III.1.24 que $|\int f| \leq \|f\|_1$, et donc que $f \mapsto \int f$ est linéaire continue (et même 1-lipschitzienne) sur $\mathcal{L}^1(X)$.

Théorème III.1.27. — (de convergence dominée)

Si $(f_n)_{n \in \mathbf{N}}$ est une suite d'éléments $\mathcal{L}^1(X)$ vérifiant :

- il existe $g \in \mathcal{L}^1(X)$ telle que, quel que soit $n \in \mathbf{N}$, on a $|f_n| \leq g$ p.p. (domination),
- f_n converge simplement presque partout,

alors la limite presque partout f de la suite $(f_n)_{n \in \mathbf{N}}$ appartient à $\mathcal{L}^1(X)$, et f_n tend vers f dans $\mathcal{L}^1(X)$; en particulier $\lim_{n \rightarrow +\infty} \int f_n = \int f$.

Démonstration. — f est mesurable comme limite simple p.p. de fonctions mesurables et sommable car $|f| \leq g$ p.p. D'autre part, quitte à modifier g et les f_n sur un ensemble de mesure nulle, on peut supposer que l'on a $|f_n| \leq g$ partout. On a alors $|f_n - f| \leq 2g$, et on peut appliquer le lemme de Fatou à $h_n = 2g - |f_n - f|$. Comme $h_n \rightarrow 2g$, on obtient

$$\int 2g \leq \liminf \int 2g - |f_n - f| = \int 2g - \limsup \int |f_n - f|,$$

et comme $\int 2g$ est fini, on en tire $\limsup \int |f_n - f| \leq 0$, et donc $\int |f_n - f| \rightarrow 0$, puisque $\int |f_n - f| \geq 0$, quel que soit $n \in \mathbf{N}$. Ceci permet de conclure.

Exercice III.1.28. — Si $n \geq 1$, soient $f_n = n\mathbf{1}_{[0,1/n]}$ et $g_n = \frac{1}{n}\mathbf{1}_{[0,n]}$. Montrer que $f_n \rightarrow 0$ p.p et $g_n \rightarrow 0$ p.p., et que $\int f_n = \int g_n = 1$. Peut-on en déduire que $1 = 0$?

Exercice III.1.29. — (i) Montrer que $\Gamma(s) = \int_0^{+\infty} e^{-t} t^{s-1} dt < +\infty$, si $s > 0$.

(ii) Montrer que $\lim_{n \rightarrow \infty} \int_0^n (1 - \frac{t}{n})^n t^{s-1} dt \rightarrow \Gamma(s)$ quand $n \rightarrow \infty$.

(iii) En déduire la formule $\Gamma(s) = \lim_{n \rightarrow \infty} \frac{n! n^s}{s(s+1)\dots(s+n)}$ de Gauss.

Exercice III.1.30. — Soit $f : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ décroissante et sommable. Montrer que $\lim_{t \rightarrow +\infty} tf(t) = 0$.

Exercice III.1.31. — Soient $f \in \mathcal{L}^1(\mathbf{R}^m)$ et A_n , pour $n \in \mathbf{N}$, des sous-ensembles mesurables de \mathbf{R}^m .

(i) Montrer que si A_n est croissante et $\cup_{n \in \mathbf{N}} A_n = \mathbf{R}^m$, alors $\int_{\mathbf{R}^d} f = \lim_{n \rightarrow \infty} \int_{A_n} f$.

(ii) Montrer que, si $\lambda(A_n) \rightarrow 0$, alors $\int_{A_n} f \rightarrow 0$.

6. Premières applications

Le résultat qui suit est très utile pour calculer explicitement des intégrales en dimension 1; le théorème de Fubini dont il sera question plus loin permet de ramener le calcul d'intégrales en dimension quelconque à une suite d'intégrations en une variable (on n'est heureusement pas forcé de revenir à la définition!).

Théorème III.1.32. — (Théorème fondamental de l'analyse)

Si $f : [a, b] \rightarrow \mathbf{C}$ est continue, et si $F : [a, b] \rightarrow \mathbf{C}$ est dérivable de dérivée f , alors $\int_{[a,b]} f = F(b) - F(a)$.

Démonstration. — Par linéarité, en considérant séparément $\text{Im}(F)$ et $\text{Re}(F)$, on peut se ramener au cas où F est à valeurs réelles. Si $n \in \mathbf{N}$, et si $i \in \{0, 1, \dots, n\}$, soit $c_{n,i} = a + i \frac{b-a}{n}$, et soit $f_n : [a, b[\rightarrow \mathbf{R}$, la fonction valant $\frac{n}{b-a} (F(c_{n,i+1}) - F(c_{n,i}))$ sur $[c_{n,i}, c_{n,i+1}[$, pour tout $i \in \{0, \dots, n-1\}$. Alors $\int f_n = \sum_{i=0}^{n-1} (F(c_{n,i+1}) - F(c_{n,i})) = F(b) - F(a)$, pour tout n . Le théorème des accroissements finis montre qu'il existe $u_{n,i} \in [c_{n,i}, c_{n,i+1}[$ tel que $f_n(t) = f(u_{n,i})$, si $t \in [c_{n,i}, c_{n,i+1}[$, et donc que $f_n(t) = f(u_n(t))$, avec $|t - u_n(t)| \leq \frac{b-a}{n}$. On en déduit que $|f_n| \leq \|f\|_\infty$ pour tout n et, f étant continue, que f_n tend simplement vers f sur $[a, b[$. On peut donc appliquer le théorème de convergence dominée pour intervertir limite et intégrale, ce qui nous donne $\int_{[a,b[} f = F(b) - F(a)$. Le résultat s'en déduit en remarquant que $\{b\}$ étant de mesure nulle, on a $\int_{[a,b]} f = \int_{[a,b[} f$.

Exercice III.1.33. — Soit $f : [a, b] \rightarrow \mathbf{C}$ dérivable. Montrer que f' est mesurable et que, si f' est bornée, alors $\int_a^b f' = f(b) - f(a)$.

Théorème III.1.34. — (convergence dominée pour les séries)

Soient I un ensemble dénombrable et $(a_{n,i})_{n \in \mathbf{N}, i \in I}$ des nombres complexes vérifiant :

- il existe $(b_i)_{i \in I}$, avec $\sum_{i \in I} |b_i| < +\infty$, telle que $|a_{n,i}| \leq b_i$ pour tous $n \in \mathbf{N}$ et $i \in I$,
- $\lim_{n \rightarrow +\infty} a_{n,i}$ existe pour tout $i \in I$.

Alors $\lim_{n \rightarrow +\infty} \sum_{i \in I} a_{n,i} = \sum_{i \in I} \lim_{n \rightarrow +\infty} a_{n,i}$.

Démonstration. — On peut supposer que $I = \mathbf{N}$, et on transforme les séries en intégrales de fonctions « localement constantes » en associant à une suite $(u_n)_{n \in \mathbf{N}}$, la fonction valant u_n sur $[n, n+1[$ et 0 sur \mathbf{R}_+^* . L'énoncé se déduit alors du th. III.1.27 (en fait, on peut le démontrer directement, en se fatiguant un peu).

III.2. Quelques espaces fonctionnels

1. L'espace $L^1(X)$

Dans tout ce qui suit, X est un sous-ensemble mesurable de \mathbf{R}^m .

Théorème III.2.1. — (Fubini sur $\mathbf{N} \times X$)

Soit $(u_n)_{n \in \mathbf{N}}$ une suite d'éléments de $\mathcal{L}^1(X)$ telle que $\sum_{n \in \mathbf{N}} \int |u_n| < +\infty$.

- (i) La série $\sum_{n \in \mathbf{N}} u_n(t)$ converge (absolument) p.p.
- (ii) Si $g = \sum_{n \in \mathbf{N}} u_n$ p.p., alors $g \in \mathcal{L}^1(X)$, et la série de terme général u_n converge vers g dans $\mathcal{L}^1(X)$; en particulier, $\int_X g = \sum_{n \in \mathbf{N}} \int_X u_n$.

Démonstration. — Soit $h(t) = \sum_{n \in \mathbf{N}} |u_n(t)|$. D'après le théorème de convergence monotone, on a $\int_X h = \sum_{n \in \mathbf{N}} \int |u_n|$, et l'hypothèse implique que $\int_X h < +\infty$. La prop. III.1.17 permet d'en conclure que $A_1 = \{t \in X, \sum_{n \in \mathbf{N}} |u_n(t)| = +\infty\}$ est de mesure nulle, et comme A_1 est précisément l'ensemble des $t \in X$ tels que la série $\sum_{n \in \mathbf{N}} u_n(t)$ ne converge pas absolument, on en déduit le premier point.

Soit A_2 l'ensemble des points tels que $g(t) \neq \sum_{n \in \mathbf{N}} u_n(t)$. Alors A_2 est de mesure nulle par hypothèse, et donc $A = A_1 \cup A_2$ est aussi de mesure nulle comme réunion de deux ensembles de mesure nulle. Soit S la fonction définie par $S(t) = \sum_{n \in \mathbf{N}} u_n(t)$, si $t \notin A$, et $S(t) = 0$ si $t \in A$. Si $N \in \mathbf{N}$, soit $S_N = \sum_{n \leq N} u_n$ la somme partielle de la série. On a $S = g$ p.p., et pour démontrer le second point, il suffit de prouver que S est sommable et $\|S - S_N\|_1 \rightarrow 0$. Or $|S_N(t)| \leq h(t)$ quels que soient $t \notin A$ et $N \in \mathbf{N}$, et donc $|S(t)| \leq h(t)$, quel que soit $t \notin A$. On en déduit le fait que S est sommable. De plus, $|S_N - S| \leq 2h$ en dehors de A et $|S_N(t) - S(t)| \rightarrow 0$, si $t \notin A$. Comme A est de mesure nulle, on est dans les conditions d'application du théorème de convergence dominée; on en déduit que $\int |S_N - S| \rightarrow 0$, ce qui permet de conclure.

Exercice III.2.2. — Soient $f \in L^1(\mathbf{R})$ et $T \in \mathbf{R}_+^*$. Montrer que $f_T(x) = \sum_{n \in \mathbf{Z}} f(x + nT)$ converge presque partout et que f_T est sommable sur $[0, T]$.

Exercice III.2.3. — Soit $f : \mathbf{R} \rightarrow \mathbf{R}_+$ définie par $f(x) = |x|^{-1/2}$, si $0 < |x| < 1$, et $f(x) = 0$, sinon.

(i) Montrer que f est sommable.

(ii) Soit $n \mapsto r_n$ une bijection de \mathbf{N} sur \mathbf{Q} . Montrer que $\sum_{n=0}^{+\infty} (\frac{-1}{2})^n f(x - r_n)$ converge absolument p.p., que la somme $F(x)$ est sommable, et calculer $\int_{\mathbf{R}} F(x) dx$. À quoi ressemble le graphe de F ?

Corollaire III.2.4. — Si $(f_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\mathcal{L}^1(X)$ tendant vers f dans $\mathcal{L}^1(X)$, on peut extraire de la suite $(f_n)_{n \in \mathbf{N}}$ une sous-suite convergeant p.p. vers f .

Démonstration. — Extrayons de $(f_n)_{n \in \mathbf{N}}$ une sous-suite $(g_n)_{n \in \mathbf{N}}$ telle que $\|f - g_n\|_1 \leq 2^{-n}$, pour tout $n \in \mathbf{N}$. Soit $u_n = g_n - g_{n-1}$, si $n \geq 1$, et $u_0 = g_0$. Alors $\|u_n\|_1 \leq \|g_n - f\|_1 + \|g_{n-1} - f\|_1 \leq 2^{1-n}$, et donc $\sum_{n \in \mathbf{N}} \|u_n\|_1 < +\infty$. D'après le th. III.2.1, la série $\sum_{n \in \mathbf{N}} u_n(x)$ converge presque partout, la limite presque partout g appartient à $\mathcal{L}^1(X)$, et $g_n = \sum_{i=0}^n u_i$ tend vers g dans $\mathcal{L}^1(X)$. Comme $g_n \rightarrow f$ dans $\mathcal{L}^1(X)$, on en déduit que $\|f - g\|_1 = 0$, ce qui implique que $f = g$ p.p., et que g_n tend vers f p.p., ce qui permet de conclure.

Exercice III.2.5. — Si $2^k \leq n < 2^{k+1}$, soit f_n la fonction caractéristique de $[\frac{n-2^k}{2^k}, \frac{n+1-2^k}{2^k}]$. Montrer que $f_n \rightarrow 0$ dans $\mathcal{L}^1([0, 1])$, mais que $f_n(x)$ ne tend vers 0 pour aucun $x \in [0, 1]$. Ce résultat n'est-il pas en contradiction avec le corollaire précédent?

L'espace $\mathcal{L}^1(X)$ muni de la semi-norme $\|\cdot\|_1$ n'est pas séparé puisque deux fonctions différant d'une fonction nulle p.p. sont à distance nulle. On note $L^1(X)$ le séparé de $\mathcal{L}^1(X)$. Comme $\|f\|_1 = 0$ si et seulement si $f = 0$ p.p., $L^1(X)$ est le quotient de $\mathcal{L}^1(X)$ par le sous-espace $\text{Npp}(X)$ des fonctions nulles presque partout; on peut donc penser à $L^1(X)$

comme étant l'espace $\mathcal{L}^1(X)$ des fonctions sommables, en *considérant comme égales deux fonctions qui sont égales presque partout* ⁽¹³⁾.

Remarque III.2.6. — L'intégrale $f \mapsto \int f$ est bien définie sur $L^1(X)$ car $\int(f - g) = 0$ si $f = g$ p.p. De plus, $f \mapsto \int f$, qui est linéaire puisqu'elle l'est sur $\mathcal{L}^1(X)$, est continue car $|\int f| \leq \int |f| = \|f\|_1$.

Théorème III.2.7. — *L'espace $L^1(X)$ est un espace de Banach.*

Démonstration. — On s'est débrouillé pour que $\|\cdot\|_1$ soit une norme sur $L^1(X)$; il suffit donc de prouver que $L^1(X)$ est complet, et pour cela, il suffit de vérifier que toute série normalement convergente est convergente, ce qui est précisément le contenu du th. III.2.1.

2. L'espace $L^2(X)$

Si X est un sous-ensemble mesurable de \mathbf{R}^m , on note $\mathcal{L}^2(X)$ l'ensemble des $f : X \rightarrow \mathbf{C}$ mesurables et *de carré sommable* (i.e. telles que $\int_X |f(t)|^2 dt < +\infty$). Il est immédiat que $\mathcal{L}^2(X)$ est stable par multiplication par un scalaire, et un peu moins qu'il est stable par addition, mais cela résulte de l'inégalité $|a + b|^2 \leq 2|a|^2 + 2|b|^2$, si $a, b \in \mathbf{C}$ dont on déduit que $\int_X |f + g|^2 \leq (2 \int_X |f|^2 + 2 \int_X |g|^2)$, si $f, g \in \text{Mes}(X)$. Autrement dit, $\mathcal{L}^2(X)$ est un espace vectoriel.

Maintenant, comme $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$, si $a, b \in \mathbf{C}$, on en déduit que, si $f, g \in \mathcal{L}^2(X)$, alors $\overline{f}g \in \mathcal{L}^1(X)$, ce qui permet de définir $\langle f, g \rangle \in \mathbf{C}$ par $\langle f, g \rangle = \int_X \overline{f}g$. L'application $\langle \cdot, \cdot \rangle$ vérifie toutes les propriétés d'un produit scalaire, sauf celle d'être définie. En effet, on a

$$\langle f, f \rangle = 0 \iff \int_X |f(t)|^2 = 0 \iff f \in \text{Npp}(X).$$

Autrement dit, l'application $f \mapsto \|f\|_2 = \langle f, f \rangle^{1/2}$ définit une semi-norme hilbertienne sur $\mathcal{L}^2(X)$. On note $L^2(X)$ l'espace séparé associé; d'après ce qui précède, c'est le quotient de $\mathcal{L}^2(X)$ par $\text{Npp}(X)$. Ceci permet, comme pour $L^1(X)$, de penser à $L^2(X)$ comme étant l'espace $\mathcal{L}^2(X)$ des fonctions de carré sommable, en considérant comme égales deux fonctions qui sont égales presque partout. Comme $\langle f, g \rangle = 0$, si f ou g est nulle p.p., la forme sesquilinéaire $\langle \cdot, \cdot \rangle$ passe au quotient, et induit un produit scalaire sur $L^2(X)$, étant donné qu'on a fait ce qu'il fallait pour la rendre définie.

Théorème III.2.8. — (Fischer-Riesz, 1907) *L'espace $L^2(X)$, muni de la norme $\|\cdot\|_2$ définie par $\|f\|_2 = (\int_X |f|^2)^{1/2}$, est un espace de Hilbert.*

⁽¹³⁾ Cette représentation mentale de $L^1(X)$ est probablement la plus facile d'utilisation; il faut quand même faire attention qu'un élément de $L^1(X)$ a beau être défini presque partout, il n'a de valeur précise en aucun point puisqu'on peut modifier arbitrairement sa valeur sur un ensemble de mesure nulle. Autrement dit, on peut parler de $f(x)$, où x est pensé comme une variable (par exemple pour les changements de variable dans les intégrales, ou pour définir le produit d'une fonction de $L^1(X)$ et d'une fonction bornée), mais pas de $f(x_0)$.

Démonstration. — D'après la discussion précédant le théorème, il suffit de prouver que $L^2(X)$ est complet, et pour cela, il suffit de prouver qu'une série normalement convergente a une limite. La démonstration, très analogue à celle du th. III.2.7, fait l'objet de l'exercice ci-dessous. La convergence dans $L^2(X)$ est dite *en moyenne quadratique*.

Exercice III.2.9. — Soit $(u_n)_{n \in \mathbf{N}}$ une suite d'éléments de $\mathcal{L}^2(X)$ telle que $\sum_{n \in \mathbf{N}} \|u_n\|_2 < +\infty$, et soit $h : X \rightarrow \overline{\mathbf{R}}_+$ définie par $h(t) = (\sum_{n \in \mathbf{N}} |u_n(t)|)^2$.

(i) Montrer que $\int_X h < +\infty$. En déduire qu'il existe $A \subset X$ de mesure nulle tel que, si $t \in X - A$, la série $\sum_{n \in \mathbf{N}} u_n(t)$ converge absolument.

(ii) On note $S(t)$ la somme de la série $\sum_{n \in \mathbf{N}} u_n(t)$, si $t \notin A$, et on prolonge S par 0 à A , et, si $N \in \mathbf{N}$, on note $S_N(t)$ la somme partielle $\sum_{n \leq N} u_n(t)$. Montrer que $|S - S_N|^2$ est majoré par $4h$ quel que soit $N \in \mathbf{N}$.

(iii) Montrer que $S \in \mathcal{L}^2(X)$ et que $S_N \rightarrow S$ dans $L^2(X)$.

(iv) Montrer que, si $f_n \rightarrow f$ dans $\mathcal{L}^2(X)$, on peut extraire de la suite $(f_n)_{n \in \mathbf{N}}$ une sous-suite convergeant p.p. vers f .

3. Convergence dans L^1 et L^2

Remarque III.2.10. — (i) Les espaces $L^1(X)$ et $L^2(X)$ n'ont a priori rien à voir. Toutefois, comme ils sont tous deux obtenus en prenant le quotient d'un sous-espace de $\text{Mes}(X)$ par $\text{Npp}(X)$, nous commettrons l'abus de notations de désigner par $L^1(X) + L^2(X)$ (resp. $L^1(X) \cap L^2(X)$) le quotient de $\mathcal{L}^1(X) + \mathcal{L}^2(X)$ (resp. $\mathcal{L}^1(X) \cap \mathcal{L}^2(X)$) par $\text{Npp}(X)$. Autrement dit, nous verrons un élément de $L^1(X) \cap L^2(X)$ comme une fonction qui est à la fois sommable et de carré sommable, à addition près d'une fonction nulle presque partout.

(ii) Il n'y a d'inclusion dans aucun sens entre $L^1(\mathbf{R}^m)$ et $L^2(\mathbf{R}^m)$. Par contre, si X est de mesure finie, alors $L^2(X) \subset L^1(X)$. En effet, l'inégalité de Cauchy-Schwarz montre que, si f est de carré sommable sur X , alors

$$\int_X |f| \leq \left(\int_X 1 \right)^{1/2} \left(\int_X |f|^2 \right)^{1/2} < +\infty.$$

En fait, l'inégalité ci-dessus montre que l'inclusion ι de $L^2(X)$ dans $L^1(X)$ est continue, et que l'on a $\|\iota\| \leq \lambda(X)^{1/2}$.

Soit X un ouvert de \mathbf{R}^m . On rappelle que $\mathcal{C}_c(X)$, $\mathcal{C}_c^k(X)$ et $\mathcal{C}_c^\infty(X)$ désignent respectivement l'espace des fonctions continues (resp. de classe \mathcal{C}^k , resp. de classe \mathcal{C}^∞) sur X , dont le support est compact. Comme une fonction nulle p.p. est identiquement nulle, la projection naturelle de $\mathcal{L}^1(X)$ sur $L^1(X)$ induit une injection de chacun des espaces ci-dessus dans $L^1(X)$, ce qui permet de les considérer comme des sous-espaces de $L^1(X)$. Pour la même raison, $\text{Esc}(X)$ est, de manière naturelle, un sous-espace de $L^1(X)$.

Théorème III.2.11. — Si X est un ouvert de \mathbf{R}^m , et si E est un des espaces $\text{Esc}(X)$, $\mathcal{C}_c(X)$, $\mathcal{C}_c^k(X)$, avec $k \in \mathbf{N}$, ou $\mathcal{C}_c^\infty(X)$, alors E est dense dans $L^1(X)$ et $L^2(X)$. De plus,

si $\phi \in L^1(X) \cap L^2(X)$, il existe une suite d'éléments de E convergeant vers ϕ , à la fois dans $L^1(X)$ et dans $L^2(X)$.

Démonstration. — Si D_n est la réunion des dalles élémentaires de taille 2^{-n} incluses dans $X \cap ([-2^n, 2^n]^m)$, alors X est la réunion croissante des dallages D_n , pour $n \in \mathbf{N}$. Soit F un des espaces $L^1(X)$, $L^2(X)$ ou $L^1(X) \cap L^2(X)$, et soit $\phi \in F$ [on munit $L^1(X) \cap L^2(X)$ de la norme $\sup(\|\cdot\|_1, \|\cdot\|_2)$]. Si $n \in \mathbf{N}$, soit ϕ_n la fonction valant $\phi(x)$, si $x \in D_n$ et $|\phi_n(x)| \leq n$, et valant 0 si $x \notin D_n$ ou si $|\phi(x)| > n$. Alors $\phi_n(x) \rightarrow \phi(x)$ quel que soit $x \in X$, et comme $|\phi_n| \leq |\phi|$ et $|\phi_n^2| \leq |\phi^2|$, cela implique que ϕ_n converge vers ϕ dans F , d'après le théorème de convergence dominée. On peut donc, si $j \in \mathbf{N}$, trouver n_j tel que $\|\phi - \phi_{n_j}\|_F \leq 2^{-j}$. Maintenant, comme ϕ_{n_j} est une fonction mesurable bornée à support borné, il existe $f_j \in \text{Esc}(D_{n_j})$ tel que $\|f_j - \phi_{n_j}\|_F \leq 2^{-j}$. On a donc $\|f_j - \phi\|_F \leq 2^{1-j}$. On en déduit la densité de $\text{Esc}(X)$ dans F , ce qui prouve le théorème pour $E = \text{Esc}(X)$. Le reste s'en déduit en utilisant l'existence de fonctions \mathcal{C}^∞ sympathiques (cf. exercice ci-dessous).

Exercice III.2.12. — On admet que la fonction φ_0 définie par $\varphi_0(x) = 0$ si $x \leq 0$ et $\varphi_0(x) = e^{-1/x}$ si $x > 0$ est une fonction \mathcal{C}^∞ sur \mathbf{R} .

(i) A partir de φ_0 , construire successivement des fonctions \mathcal{C}^∞ sur \mathbf{R} :

- $\varphi_1 : \mathbf{R} \rightarrow [0, 1]$, nulle en dehors de $[0, 1]$, avec $\int_0^1 \varphi_1 = 1$;
- $\varphi_2 : \mathbf{R} \rightarrow [0, 1]$, valant 0 si $x \leq 0$ et 1 si $x \geq 1$;
- $\varphi_\varepsilon : \mathbf{R} \rightarrow [0, 1]$, pour $\varepsilon \in]0, \frac{1}{2}[$, nulle en dehors de $[0, 1]$ et valant 1 sur $[\varepsilon, 1 - \varepsilon]$.

(ii) Terminer la démonstration du théorème III.2.11.

Corollaire III.2.13. — Si $f \in L^1(X)$ vérifie $\int_X \varphi f = 0$, pour tout $\varphi \in \mathcal{C}_c^\infty(X)$, alors $f = 0$.

Démonstration. — Il résulte de la démonstration du th. III.2.11 que, si $g \in L^1(X)$ est bornée, on peut trouver une suite $(g_n)_{n \in \mathbf{N}}$ d'éléments de $\mathcal{C}_c^\infty(X)$ tendant vers g dans $L^1(X)$ et vérifiant de plus $\|g_n\|_\infty \leq \|g\|_\infty$, pour tout n . En outre, quitte à extraire une sous-suite, on peut s'arranger (cor. III.2.4) pour que $g_n \rightarrow g$ p.p.

Ce qui précède s'applique en particulier, si $Y \subset X$ est un ouvert de mesure finie, à la fonction g_Y définie par $g_Y(x) = 0$, si $f(x) = 0$ ou $x \notin Y$, et $g_Y(x) = \bar{f}(x)/|f(x)|$, si $f(x) \neq 0$ et $x \in Y$. La suite $g_n f$ tend alors vers $\mathbf{1}_Y |f|$ p.p. et est majorée par $\|g\|_\infty |f| = |f|$ pour tout n . On est donc sous les conditions d'application du th. de convergence dominée, ce qui permet de montrer que $\int g_n f \rightarrow \int_Y |f|$. L'hypothèse $\int_X \varphi f = 0$ pour tout $\varphi \in \mathcal{C}_c^\infty(X)$ entraîne donc $\int_Y |f| = 0$, pour tout Y , et donc aussi $\int |f| = 0$ par le théorème de convergence monotone. Ceci permet de conclure.

Terminons ce n° par ce petit résultat qui sera utile plus tard.

Lemme III.2.14. — Soit X un ouvert de \mathbf{R}^m .

- (i) Soient $\phi \in L^1(X)$, et $g \in \mathcal{C}(X)$. On suppose qu'il existe une suite $(\phi_j)_{j \in \mathbf{N}}$ d'éléments de $L^1(X)$ tendant vers ϕ dans $L^1(X)$, telle que $\phi_j(x) \rightarrow g(x)$ p.p. Alors $g \in L^1(X)$ et $g = \phi$ p.p.
- (ii) On a le même résultat en remplaçant $L^1(X)$ par $L^2(X)$.

Démonstration. — La démonstration est exactement la même dans les deux cas. Quitte à extraire une sous-suite de la suite $(\phi_j)_{j \in \mathbf{N}}$, on peut supposer (cf. cor. III.2.4 et (iv) de l'ex. III.2.9) que $\phi_j(x) \rightarrow \phi(x)$ p.p., ce qui permet de conclure.

4. Espaces L^p

Les espaces $L^1(X)$ et $L^2(X)$ vivent dans une famille $L^p(X)$, pour $p \in [1, +\infty]$, d'espaces de Banach introduits par F. Riesz en 1910, obtenus comme séparés de sous-espaces $\mathcal{L}^p(X)$ de $\text{Mes}(X)$ (de fait, dans tous les cas, on passe de $\mathcal{L}^p(X)$ à $L^p(X)$ en quotientant par $\text{Npp}(X)$). Les espaces $\mathcal{L}^p(X)$ sont définis comme suit.

- Si $1 \leq p < +\infty$, on définit $\mathcal{L}^p(X)$ comme le sous-espace de $\text{Mes}(X)$ des f tels que $|f|^p$ soit sommable. L'inégalité de Minkowski $(\int |f+g|^p)^{1/p} \leq (\int |f|^p)^{1/p} + (\int |g|^p)^{1/p}$ permet de montrer que $\mathcal{L}^p(X)$ est un espace vectoriel et que $f \mapsto \|f\|_p = (\int |f|^p)^{1/p}$ est une semi-norme sur $\mathcal{L}^p(X)$.
- Si $p = +\infty$, on définit l'espace $\mathcal{L}^\infty(X)$ comme le sous-espace de $\text{Mes}(X)$ des f qui sont *essentiellement bornées*, c'est-à-dire qu'il existe $A \subset X$ de mesure nulle et $M \in \mathbf{R}_+$ tels que $|f(t)| \leq M$, quel que soit $t \in X - A$. On note $\|f\|_\infty$ la *borne supérieure essentielle* de $|f|$, c'est-à-dire la borne inférieure de l'ensemble des $M \in \mathbf{R}_+$, tels qu'il existe $A \subset X$ de mesure nulle tel que $|f(t)| \leq M$, quel que soit $t \in X - A$. Alors $\|\cdot\|_\infty$ est une semi-norme sur $\mathcal{L}^\infty(X)$.

Si X est un ouvert de \mathbf{R}^m , et si $p < +\infty$, les fonctions en escalier sont denses dans $L^p(X)$. Ce n'est plus le cas si $p = +\infty$, l'adhérence des fonctions en escalier étant l'ensemble des fonctions bornées tendant vers 0 à l'infini.

L'espace $L^2(X)$ étant un espace de Hilbert, il est son propre dual d'après le théorème de Riesz. Dans le cas général, on note $q \in [1, +\infty]$ l'exposant conjugué de p , défini par $1/p + 1/q = 1$. L'inégalité de Hölder $\int |fg| \leq \|f\|_p \|g\|_q$ montre que si $g \in L^q(X)$, alors $f \mapsto \int fg$ définit une forme linéaire Λ_g continue sur $L^p(X)$, et que l'on a $\|\Lambda_g\| \leq \|g\|_q$ (le résultat est trivial si $p = 1$ ou $p = +\infty$). Si $p \neq +\infty$, on peut montrer qu'en fait $g \mapsto \Lambda_g$ est une isométrie⁽¹⁴⁾ de $L^q(X)$ sur le dual de l'espace de Banach $L^p(X)$. Par contre, le dual de $L^\infty(\mathbf{R}^m)$ est nettement plus gros que $L^1(\mathbf{R}^m)$.

- Exercice III.2.15.* — (i) Montrer que $L^1(X)$ et $L^2(X)$ sont séparables, si $X \subset \mathbf{R}^m$ est mesurable.
- (ii) Montrer que $L^\infty(\mathbf{R}^m)$ n'est pas séparable.

⁽¹⁴⁾ On remarquera que, si X est de mesure finie, les $L^p(X)$ forment une famille décroissante d'espaces de Banach (on a $L^p(X) \subset L^{p'}(X)$ si $p \geq p'$), alors que leurs duals $L^q(X)$ forment une famille croissante. Autrement dit, plus l'espace fonctionnel est petit, plus son dual est gros, ce qui est un peu étrange quand on pense au cas des espaces vectoriels de dimension finie, mais conduit à la construction des distributions pour laquelle L. Schwartz a obtenu la médaille Fields (1950).

III.3. Intégrales multiples

1. Le théorème de Fubini

Une somme finie $\sum_{j,k} a_{j,k}$ peut se calculer en sommant d'abord sur j puis sur k ou en sommant d'abord sur k puis sur j . Le théorème de Fubini ci-dessous dit qu'il en est de même pour des intégrales (à condition que tout soit sommable).

Lemme III.3.1. — (Fubini pour les fonctions en escalier) *L'application $f \mapsto \int_{\mathbf{R}^n} f$ (resp. $f \mapsto \int_{\mathbf{R}^m} f$) est une application linéaire de $\text{Esc}(\mathbf{R}^{n+m})$ dans $\text{Esc}(\mathbf{R}^m)$ (resp. $\text{Esc}(\mathbf{R}^n)$), et on a*

$$\begin{aligned} \int_{\mathbf{R}^m} \left| \int_{\mathbf{R}^n} f \right| &\leq \int_{\mathbf{R}^{n+m}} |f| \quad \text{et} \quad \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^m} f \right| \leq \int_{\mathbf{R}^{n+m}} |f| \\ \int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} f \right) &= \int_{\mathbf{R}^{n+m}} f = \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^m} f \right). \end{aligned}$$

Démonstration. — La linéarité est une conséquence de la linéarité de l'intégrale. Maintenant, soit $r \in \mathbf{N}$, et soient $\mathbf{j} = (j_1, \dots, j_n) \in \mathbf{Z}^n$ et $\mathbf{k} = (k_1, \dots, k_m) \in \mathbf{Z}^m$. On note (\mathbf{j}, \mathbf{k}) l'élément $(j_1, \dots, j_n, k_1, \dots, k_m)$ de \mathbf{Z}^{n+m} . Un calcul immédiat montre alors que

$$\begin{aligned} \int_{\mathbf{R}^n} e_{r,(\mathbf{j},\mathbf{k})} &= 2^{-rn} e_{r,\mathbf{k}}, & \int_{\mathbf{R}^m} e_{r,(\mathbf{j},\mathbf{k})} &= 2^{-rm} e_{r,\mathbf{j}}, \\ \int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} e_{r,(\mathbf{j},\mathbf{k})} \right) &= 2^{-rn} \int_{\mathbf{R}^m} e_{r,\mathbf{k}} = 2^{-r(n+m)} = \int_{\mathbf{R}^{n+m}} e_{r,(\mathbf{j},\mathbf{k})} = \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^m} e_{r,(\mathbf{j},\mathbf{k})} \right). \end{aligned}$$

Le résultat s'en déduit en décomposant f sous la forme $f = \sum_{(\mathbf{j},\mathbf{k})} a_{(\mathbf{j},\mathbf{k})} e_{r,(\mathbf{j},\mathbf{k})}$, pour r assez grand (la somme étant une somme finie).

Proposition III.3.2. — (Fubini dans L^1) *Il existe une unique application linéaire continue $f \mapsto \int_{\mathbf{R}^n} f$ (resp. $f \mapsto \int_{\mathbf{R}^m} f$) de $L^1(\mathbf{R}^{n+m})$ dans $L^1(\mathbf{R}^m)$ (resp. $L^1(\mathbf{R}^n)$) coïncidant avec l'application du même nom sur $\text{Esc}(\mathbf{R}^{n+m})$, et on a*

$$\int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} f \right) = \int_{\mathbf{R}^{n+m}} f = \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^m} f \right).$$

Démonstration. — L'application linéaire $\int_{\mathbf{R}^n} : \text{Esc}(\mathbf{R}^{n+m}) \rightarrow L^1(\mathbf{R}^m)$ est, d'après le lemme III.3.1, uniformément continue (en fait 1-lipschitzienne) si on munit tous les espaces de la norme $\|\cdot\|_1$. Comme $L^1(\mathbf{R}^m)$ est complet, et comme $\text{Esc}(\mathbf{R}^{n+m})$ est dense dans $L^1(\mathbf{R}^{n+m})$, l'application $\int_{\mathbf{R}^n} : \text{Esc}(\mathbf{R}^{n+m}) \rightarrow L^1(\mathbf{R}^m)$ s'étend (de manière unique) par continuité à $L^1(\mathbf{R}^{n+m})$. De même, $\int_{\mathbf{R}^m} : \text{Esc}(\mathbf{R}^{n+m}) \rightarrow L^1(\mathbf{R}^n)$ s'étend par continuité à $L^1(\mathbf{R}^{n+m})$. Enfin, les trois applications

$$f \mapsto \int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} f \right), \quad f \mapsto \int_{\mathbf{R}^{n+m}} f, \quad f \mapsto \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^m} f \right)$$

sont continues sur $L^1(\mathbf{R}^{n+m})$ et coïncident sur $\text{Esc}(\mathbf{R}^{n+m})$. Comme cet espace est dense dans $L^1(\mathbf{R}^{n+m})$, elles coïncident sur $L^1(\mathbf{R}^{n+m})$ tout entier. Ceci permet de conclure.

On peut rendre la prop. III.3.2 plus concrète (et plus facilement utilisable pour le calcul d'intégrales multiples) sous la forme du (i) du théorème suivant.

Théorème III.3.3. — (Fubini)

(i) Si $f \in \mathcal{L}^1(\mathbf{R}^{n+m})$, alors

- $f(\cdot, y) \in \mathcal{L}^1(\mathbf{R}^n)$, pour presque tout $y \in \mathbf{R}^m$, et $y \mapsto \int_{\mathbf{R}^n} f(x, y) dx \in \mathcal{L}^1(\mathbf{R}^m)$,
- $f(x, \cdot) \in \mathcal{L}^1(\mathbf{R}^m)$, pour presque tout $x \in \mathbf{R}^n$, et $x \mapsto \int_{\mathbf{R}^m} f(x, y) dy \in \mathcal{L}^1(\mathbf{R}^n)$,

$$\int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} f(x, y) dx \right) dy = \int_{\mathbf{R}^{n+m}} f(x, y) dx dy = \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^m} f(x, y) dy \right) dx.$$

(ii) Si $f : \mathbf{R}^{n+m} \rightarrow \overline{\mathbf{R}}_+$ est mesurable, alors les fonctions

$$y \mapsto \int_{\mathbf{R}^n} f(x, y) dx \quad \text{et} \quad x \mapsto \int_{\mathbf{R}^m} f(x, y) dy,$$

à valeurs dans $\overline{\mathbf{R}}_+$, sont mesurables, et

$$\int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} f(x, y) dx \right) dy = \int_{\mathbf{R}^{n+m}} f(x, y) dx dy = \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^m} f(x, y) dy \right) dx.$$

Remarque III.3.4. — (i) Si $X \subset \mathbf{R}^n$ et $Y \subset \mathbf{R}^m$ sont mesurables, alors $X \times Y$ est mesurable dans \mathbf{R}^{n+m} , et on a un énoncé analogue à celui ci-dessus, en remplaçant \mathbf{R}^n par X , \mathbf{R}^m par Y et \mathbf{R}^{n+m} par $X \times Y$. Il se déduit de celui sur \mathbf{R}^{n+m} en écrivant une intégrale sur $X \times Y$ sous la forme $\int_{\mathbf{R}^{n+m}} \mathbf{1}_{X \times Y} \phi$.

(ii) Dans la pratique, on commence par utiliser le (ii) pour vérifier que $|f|$ est sommable, avant d'utiliser le (i) pour calculer des intégrales, intervertir les variables...

(iii) Le (i) s'utilise aussi comme un théorème de semi-existence : il affirme que si $f(x, y)$ est sommable, alors $\int f(x, y) dx$ converge pour presque tout y , et $\int f(x, y) dy$ converge pour presque tout x ; par contre on ne peut en déduire la convergence de ces intégrales pour aucun x ou y particulier.

Démonstration. — Soit $(f_k)_{k \in \mathbf{N}}$ une suite d'éléments de $\text{Esc}(\mathbf{R}^{n+m})$, tendant vers f dans $\mathcal{L}^1(\mathbf{R}^{n+m})$, et telle que $\sum_{k \in \mathbf{N}} \|f_{k+1} - f_k\|_1 < +\infty$. Il résulte du lemme III.3.1 que la série $\sum_{k \in \mathbf{N}} \int_{\mathbf{R}^n} (f_{k+1} - f_k)$ est à termes dans $\text{Esc}(\mathbf{R}^m)$, converge normalement dans $\mathcal{L}^1(\mathbf{R}^m)$, et que, si on note $g \in \mathcal{L}^1(\mathbf{R}^m)$, la limite, alors $\int_{\mathbf{R}^n} g = \int_{\mathbf{R}^{n+m}} f$. Le problème est donc de montrer que, pour y en dehors d'un ensemble B de mesure nulle, la fonction $x \mapsto f(x, y)$ appartient à $\mathcal{L}^1(\mathbf{R}^n)$, et $g(y) = \int_{\mathbf{R}^n} f(x, y) dx$. Nous aurons besoin du lemme suivant.

Lemme III.3.5. — Si $A \subset \mathbf{R}^{n+m}$ est de mesure nulle, alors les ensembles A_1 et A_2 définis par

$$A_1 = \{x \in \mathbf{R}^n, A \cap (\{x\} \times \mathbf{R}^m) \text{ n'est pas de mesure nulle},\}$$

$$A_2 = \{y \in \mathbf{R}^m, A \cap (\mathbf{R}^n \times \{y\}) \text{ n'est pas de mesure nulle},\}$$

sont de mesure nulle dans \mathbf{R}^n et \mathbf{R}^m respectivement.

Démonstration. — Par symétrie (entre n et m), il suffit de le prouver pour A_1 . Soit $\varepsilon > 0$. Comme A est de mesure nulle, on peut trouver une suite $(D_k)_{k \in \mathbf{N}}$ de dalles de \mathbf{R}^n , et une suite $(E_k)_{k \in \mathbf{N}}$ de dalles de \mathbf{R}^m , telles que $A \subset \cup_{k \in \mathbf{N}} D_k \times E_k$, et $\sum_{k \in \mathbf{N}} \lambda(D_k) \lambda(E_k) \leq \varepsilon$. Si $r \in \mathbf{N}$ et $j \in \mathbf{N}$, soit $B_{\varepsilon, r, j}$ l'ensemble des $x \in \mathbf{R}^n$ tels que $\sum_{k \leq j, x \in D_k} \lambda(E_k) > 2^{-r}$; c'est un dallage de \mathbf{R}^n , et on a $2^{-r} \lambda(B_{\varepsilon, r, j}) \leq \sum_{k \leq j} \lambda(D_k) \lambda(E_k) \leq \varepsilon$. De plus, la suite des $(B_{\varepsilon, r, j})_{j \in \mathbf{N}}$ est une suite croissante de dallages finis, et donc $\lambda^+(\cup_{j \in \mathbf{N}} B_{\varepsilon, r, j}) = \lim_{j \rightarrow +\infty} \lambda(B_{\varepsilon, r, j}) \leq 2^r \varepsilon$. Or l'ensemble B_r des $x \in \mathbf{R}^n$ tel que $\lambda^+(A \cap (\{x\} \times \mathbf{R}^m)) >$

2^{-r} est inclus dans $\cup_{j \in \mathbf{N}} B_{\varepsilon, r, j}$, quel que soit $\varepsilon > 0$; c'est donc un ensemble de mesure nulle puisque de mesure extérieure $\leq 2^r \varepsilon$ quel que soit $\varepsilon > 0$. Finalement, $A_1 = \cup_{r \in \mathbf{N}} B_r$ est de mesure nulle, en tant que réunion dénombrable d'ensembles de mesures nulles.

Revenons à la démonstration du théorème de Fubini. Comme $\sum_{k \in \mathbf{N}} \|f_{k+1} - f_k\|_1 < +\infty$, il résulte du th. III.2.1, que la série $\sum_{k \in \mathbf{N}} f_{k+1}(x, y) - f_k(x, y)$ converge absolument vers $f(x, y)$ pour tout (x, y) en dehors d'un ensemble de mesure nulle A . Par ailleurs, d'après le lemme III.3.5, il existe un sous-ensemble de mesure nulle B_1 de \mathbf{R}^m tel que, si $y \notin B_1$, alors $A \cap (\mathbf{R}^n \times \{y\})$ est de mesure nulle dans \mathbf{R}^n . Si $y \notin B_1$, on a donc $f_k(x, y) \rightarrow f(x, y)$ pour x en dehors d'un ensemble de mesure nulle.

De même, comme $\sum_{k \in \mathbf{N}} \int_{\mathbf{R}^n} (f_{k+1} - f_k) dx$ converge normalement dans $\mathcal{L}^1(\mathbf{R}^m)$ vers g , il existe $B_2 \subset \mathbf{R}^m$ tel que, si $y \notin B_2$, alors $\int_{\mathbf{R}^n} f_k(x, y) dx \rightarrow g(y)$.

Maintenant, en appliquant Fubini pour les fonctions en escalier à $u_k = |f_{k+1} - f_k|$, puis deux fois de suite le théorème de convergence monotone, on obtient :

$$\begin{aligned} \sum_{k \in \mathbf{N}} \int_{\mathbf{R}^{n+m}} |u_k(x, y)| dx dy &= \sum_{k \in \mathbf{N}} \int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} |u_k(x, y)| dx \right) dy \\ &= \int_{\mathbf{R}^m} \left(\sum_{k \in \mathbf{N}} \int_{\mathbf{R}^n} |u_k(x, y)| dx \right) dy \\ &= \int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^n} \sum_{k \in \mathbf{N}} |u_k(x, y)| dx \right) dy. \end{aligned}$$

Comme on a fait l'hypothèse que $\sum_{k \in \mathbf{N}} \int_{\mathbf{R}^{n+m}} |u_k(x, y)| dx dy < +\infty$, l'ensemble B_3 des $y \in \mathbf{R}^m$ tels que $\int_{\mathbf{R}^n} \sum_{k \in \mathbf{N}} |u_k(x, y)| dx = +\infty$ est de mesure nulle. Si $y \notin B_3$, la fonction $h_y(x) = \sum_{k \in \mathbf{N}} |u_k(x, y)|$ est donc sommable, et comme $f_k = \sum_{j=0}^{k-1} u_j$, on a $|f_k(x, y)| \leq h_y(x)$ quel que soit $k \in \mathbf{N}$.

Si $B = B_1 \cup B_2 \cup B_3$, et si $y \notin B$, on est dans les conditions d'application du théorème de convergence dominée, et donc

$$g(y) = \lim_{k \rightarrow +\infty} \int_{\mathbf{R}^n} f_k(x, y) dx = \int_{\mathbf{R}^n} \left(\lim_{k \rightarrow +\infty} f_k(x, y) \right) dx = \int_{\mathbf{R}^n} f(x, y) dx.$$

Ceci permet de terminer la démonstration du (i).

Le (ii) est une conséquence du (i) sauf si $\int_{\mathbf{R}^{n+m}} f(x, y) dx dy = +\infty$. Mais, dans ce cas, il suffit de prendre une suite $(f_k)_{k \in \mathbf{N}}$ d'éléments de \mathcal{L}^1 tendant vers f en croissant, de constater que les trois membres de l'identité dépendent de manière croissante de f , et d'utiliser le théorème de convergence monotone pour montrer que $\lim_{k \rightarrow +\infty} \int_{\mathbf{R}^{n+m}} f_k(x, y) dx dy = +\infty$, et en déduire que les trois membres sont égaux à $+\infty$. Ceci permet de conclure.

Exercice III.3.6. — Soit $f : \mathbf{R}_+ \times \mathbf{R}_+ \rightarrow \mathbf{R}$ définie par $f(x, y) = e^{x-y}$ si $x \leq y$, et $f(x, y) = -e^{y-x}$, si $y < x$. Calculer

$$\int_0^{+\infty} \left(\int_0^{+\infty} f(x, y) dx \right) dy \quad \text{et} \quad \int_0^{+\infty} \left(\int_0^{+\infty} f(x, y) dy \right) dx.$$

Peut-on en déduire que $-1 = 1$?

2. La formule du changement de variable

Théorème III.3.7. — (invariance par translation de la mesure de Lebesgue)

Si $v \in \mathbf{R}^m$, et si $\phi : \mathbf{R}^m \rightarrow \mathbf{C}$ est mesurable, alors $x \mapsto \phi(x+v)$ est mesurable, et on a

$$\int |\phi(x+v)| dx = \int |\phi(x)| dx \quad \text{et, si } \phi \text{ est sommable, } \int \phi(x+v) dx = \int \phi(x) dx.$$

Démonstration. — Notons T_v l'application $\phi \mapsto T_v(\phi)$, définie par $(T_v(\phi))(x) = \phi(x+v)$. Si D est une dalle, et si les coordonnées de v sont des nombres dyadiques, alors $D-v = \{x-v, x \in D\}$ est encore une dalle, et on a $T_v(\mathbf{1}_D) = \mathbf{1}_{D-v}$; on en déduit, en découpant D et $D-v$ en dalles élémentaires de même taille, que $\int T_v(\mathbf{1}_D) = \int \mathbf{1}_D$. Dans le cas général, on prend une suite $(v_n)_{n \in \mathbf{N}}$ d'éléments de \mathbf{R}^m à coordonnées dyadiques tendant vers le vecteur v de la translation : la suite des $\mathbf{1}_{D-v_n}$ tend simplement vers $\mathbf{1}_{D-v}$ en dehors des faces, mais celles-ci sont de mesure nulle ((ii) de l'ex. III.1.6, par exemple), ce qui montre que $\mathbf{1}_{D-v}$ est mesurable en tant que limite simple p.p. de fonctions mesurables, et permet d'utiliser le théorème de convergence dominée pour en déduire que $\int T_v(\mathbf{1}_D) = \int \mathbf{1}_D$.

Par linéarité, on en déduit que $T_v(\phi)$ est mesurable et $\int T_v(\phi) = \int \phi$, si $\phi \in \text{Esc}(\mathbf{R}^m)$. Le théorème s'en déduit en utilisant la densité de $\text{Esc}(\mathbf{R}^m)$ dans $L^1(\mathbf{R}^m)$ (et le théorème de convergence monotone pour traiter le cas $\int |\phi| = +\infty$).

Soit Ω un ouvert de \mathbf{R}^m , et soit $\varphi : \Omega \rightarrow \varphi(\Omega)$ un *difféomorphisme* de Ω sur un ouvert $\varphi(\Omega)$ de \mathbf{R}^m , c'est-à-dire une application bijective de classe \mathcal{C}^1 dont l'inverse est aussi de classe \mathcal{C}^1 . L'application φ s'écrit, en coordonnées

$$\varphi(x) = (\varphi_1(x_1, \dots, x_m), \dots, \varphi_m(x_1, \dots, x_m)),$$

et la condition « φ est de classe \mathcal{C}^1 » se traduit par le fait que les dérivées partielles $\frac{\partial \varphi_i}{\partial x_j}$ sont continues sur Ω . On définit la *matrice jacobienne* $\text{Jac}_\varphi(x)$ de φ au point x par $\text{Jac}_\varphi(x) = (\frac{\partial \varphi_i}{\partial x_j}(x))_{1 \leq i, j \leq m}$. On définit le *jacobien* $J_\varphi(x)$ de φ au point x comme le déterminant de $\text{Jac}_\varphi(x)$; le fait que φ soit un difféomorphisme implique que $\text{Jac}_\varphi(x)$ est inversible, et donc que $J_\varphi(x) \neq 0$, pour tout $x \in \Omega$.

Théorème III.3.8. — Si f est une fonction mesurable sur $\varphi(\Omega)$, alors f est sommable si et seulement si la fonction $x \mapsto f(\varphi(x)) \cdot J_\varphi(x)$ est sommable sur Ω , et on a

$$\int_{\varphi(\Omega)} f(y) dy = \int_{\Omega} f(\varphi(x)) \cdot |J_\varphi(x)| dx.$$

Remarque III.3.9. — En dimension 1, la matrice jacobienne de φ n'est autre que la dérivée φ' de φ , et on tombe sur la formule $\int_{\varphi([a,b])} f(y) dy = \int_{[a,b]} f(\varphi(x)) |\varphi'(x)| dx$. Comme φ' ne s'annule pas sur $]a, b[$, il y a deux cas : soit $\varphi' > 0$ sur $]a, b[$ et $\varphi([a, b]) =]\varphi(a), \varphi(b)[$, soit $\varphi' < 0$ sur $]a, b[$ et $\varphi([a, b]) =]\varphi(b), \varphi(a)[$. Dans le premier cas, la formule devient $\int_{\varphi(a)}^{\varphi(b)} f(y) dy = \int_a^b f(\varphi(x)) \varphi'(x) dx$, et dans le second, elle devient $-\int_{\varphi(a)}^{\varphi(b)} f(y) dy = -\int_a^b f(\varphi(x)) \varphi'(x) dx$; on retombe donc bien, dans les deux cas, sur la formule classique.

Démonstration. — Commençons par regarder ce qui se passe dans le cas $\Omega = \mathbf{R}^m$ et φ affine (i.e. de la forme $x \mapsto A \cdot x + b$, avec $A \in \mathbf{GL}_m(\mathbf{R})$, et $b \in \mathbf{R}^m$). Dans ce cas, la matrice jacobienne de φ est constante égale à A ; on a donc $J_\varphi(x) = |\det A|$ pour tout x .

Par ailleurs, si $r \in \mathbf{N}$ et si $\mathbf{k} \in \mathbf{Z}^m$, alors $\varphi(D_{r,\mathbf{k}})$ est un translaté de $\varphi(D_{r,\mathbf{0}})$ et donc a même volume. On en déduit l'existence de $C(A) \in \mathbf{R}_+^*$ tel que $\lambda(\varphi(D_{r,\mathbf{k}})) = 2^{-r}C(A)$, quels que soient $r \in \mathbf{N}$ et $\mathbf{k} \in \mathbf{Z}^m$. Par linéarité on a $\int_{\mathbf{R}^m} \phi(x) dx = C(A) \int_{\mathbf{R}^m} \phi(A \cdot x + b) dx$, quel que soit $\phi \in \text{Esc}(\mathbf{R}^m)$. Par continuité, cette égalité s'étend à $L^1(\mathbf{R}^m)$. Pour conclure dans le cas affine, il reste à vérifier que $C(A) = |\det A|$, ce qui constitue l'interprétation géométrique du déterminant⁽¹⁵⁾, et fait l'objet de l'exercice III.3.10 dans la note ci-dessous.

On démontre le cas général en utilisant le fait que, plus on regarde de près autour d'un point x , plus φ ressemble à l'application affine $h \mapsto \varphi(x) + \text{Jac}_\varphi(x) \cdot h$, et donc que quand r tend vers $+\infty$, l'image de $x + D_{r,\mathbf{0}}$ ressemble de plus en plus au parallélépipède $\varphi(x) + \text{Jac}_\varphi(x) \cdot D_{r,\mathbf{0}}$, dont le volume est $|J_\varphi(x)|\lambda(D_{r,\mathbf{0}})$.

De manière plus précise, on démontre, en utilisant la continuité uniforme de $x \mapsto \text{Jac}_\varphi(x)$ (et de $x \mapsto \text{Jac}_{\varphi^{-1}}(x)$), que si $K \subset \Omega$ est compact, il existe une suite de fonctions $\varepsilon_{K,r} : K \rightarrow \mathbf{R}_+$, pour r assez grand, tendant uniformément vers 0 sur K , telle que, quel que soit $x \in K$, on ait

$$\varphi(x) + (1 - \varepsilon_{K,r}(x))\text{Jac}_\varphi(x) \cdot D_{r,\mathbf{0}} \subset \varphi(x + D_{r,\mathbf{0}}) \subset \varphi(x) + (1 + \varepsilon_{K,r}(x))\text{Jac}_\varphi(x) \cdot D_{r,\mathbf{0}}.$$

On en déduit l'existence de $\varepsilon'_{K,r}$, tendant uniformément vers 0 sur K , telle que, quel que soit $x \in K$, on ait $\lambda(\varphi(x + D_{r,\mathbf{0}})) = (1 + \varepsilon'_{K,r}(x))|J_\varphi(x)|\lambda(x + D_{r,\mathbf{0}})$. Maintenant, on peut écrire Ω comme une réunion croissante de dallages D_n dont l'adhérence est contenue dans Ω , et il suffit de prouver que la formule du théorème est valable pour une fonction en escalier f à support dans un des D_n , car ces fonctions forment un sous-espace dense dans $\mathcal{L}^1(\Omega)$. Par linéarité, on est ramené à prouver que $\lambda(\varphi(D)) = \int_D |J_\varphi(x)| dx$, si D est une dalle élémentaire dont l'adhérence K est incluse dans Ω . Si r est assez grand, D est la réunion disjointe des $D_{r,\mathbf{k}}$ contenues dans D , et comme $D_{r,\mathbf{k}} = \frac{\mathbf{k}}{2^r} + D_{r,\mathbf{0}}$, on tire de la discussion ci-dessus l'identité

$$\lambda(\varphi(D)) = \sum_{D_{r,\mathbf{k}} \subset D} \lambda(\varphi(\frac{\mathbf{k}}{2^r} + D_{r,\mathbf{0}})) = \sum_{D_{r,\mathbf{k}} \subset D} (1 + \varepsilon'_{K,r}(\frac{\mathbf{k}}{2^r}))|J_\varphi(\frac{\mathbf{k}}{2^r})|\lambda(D_{r,\mathbf{k}}) = \int_D \phi_r(x) dx,$$

où ϕ_r est la fonction en escalier sur D valant $(1 + \varepsilon'_{K,r}(\frac{\mathbf{k}}{2^r}))|J_\varphi(\frac{\mathbf{k}}{2^r})|$, sur $D_{r,\mathbf{k}}$. Comme $\varepsilon'_{K,r}$ tend uniformément vers 0 sur K , ϕ_r tend uniformément vers $|J_\varphi(x)|$ sur D . On en déduit le résultat.

⁽¹⁵⁾Le volume du parallélépipède supporté par des vecteurs v_1, \dots, v_m de \mathbf{R}^m est égal à la valeur absolue du déterminant de ces vecteurs.

- Exercice III.3.10.* — (i) Prouver que $C(AB) = C(A)C(B)$, quels que soient $A, B \in \mathbf{GL}_m(\mathbf{R})$.
(ii) Montrer que $C(A) = |\det A|$ si A est une matrice diagonale ou une matrice de permutation (i.e. si $x \mapsto A \cdot x$ induit une permutation des vecteurs de la base canonique de \mathbf{R}^m).
(iii) Montrer que toute matrice unipotente supérieure ou inférieure (i.e. triangulaire avec des 1 sur la diagonale) peut s'écrire sous la forme $DUD^{-1}U^{-1}$, avec D diagonale, et U unipotente supérieure ou inférieure. En déduire que $C(A) = 1$ si U est unipotente inférieure ou supérieure.
(iv) En utilisant la méthode du pivot, montrer que $\mathbf{GL}_m(\mathbf{R})$ est engendré par les matrices diagonales, les unipotentes inférieures et supérieures et les matrices de permutation.
(v) En déduire que $C(A) = |\det A|$ quel que soit $A \in \mathbf{GL}_m(\mathbf{R})$.

3. L'intégrale de la gaussienne

Nous allons utiliser les résultats des deux précédents n° pour établir les formules

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi} \quad \text{et} \quad \int_{-\infty}^{+\infty} e^{-\pi x^2} dx = 1.$$

On passe de la première à la seconde par le changement de variable $x = \sqrt{\pi} u$; il suffit donc de démontrer la première formule. Pour cela, notons $I = \int_{-\infty}^{+\infty} e^{-x^2} dx$ l'intégrale à calculer, et posons $I_0 = \int_{\mathbf{R}^2} e^{-(x^2+y^2)} dx dy$.

- $e^{-(x^2+y^2)} = e^{-x^2} e^{-y^2}$ et, d'après le théorème de Fubini pour les fonctions positives,

$$I_0 = \int_{\mathbf{R}} \left(\int_{\mathbf{R}} e^{-x^2} e^{-y^2} dx \right) dy = \int_{\mathbf{R}} I e^{-y^2} dy = I^2.$$

• Soit Δ la demi-droite $] -\infty, 0] \times \{0\}$, et soit $\Omega' = \mathbf{R}^2 - \Delta$. Comme Δ est de mesure nulle, on a aussi $I_0 = \int_{\Omega'} e^{-(x^2+y^2)} dx dy$.

• Soit $\Omega = \mathbf{R}_+^* \times]-\pi, \pi[$. Alors φ défini par $\varphi(r, \theta) = (r \cos \theta, r \sin \theta)$ est un difféomorphisme de Ω sur Ω' , et comme $dx = \cos \theta dr - r \sin \theta d\theta$ et $dy = \sin \theta dr + r \cos \theta d\theta$, la matrice jacobienne de φ est $\begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$, et son jacobien est $J_\varphi(r, \theta) = r \cos^2 \theta + r \sin^2 \theta = r$. La formule du changement de variable appliquée à $f(x, y) = e^{-(x^2+y^2)}$ nous donne :

$$I_0 = \int_{\Omega} e^{-(r^2 \cos^2 \theta + r^2 \sin^2 \theta)} |J_\varphi(r, \theta)| dr d\theta = \int_{\Omega} e^{-r^2} r dr d\theta.$$

- On utilise de nouveau le théorème de Fubini pour les fonctions positives :

$$I_0 = \int_{\mathbf{R}_+^* \times]-\pi, \pi[} e^{-r^2} r dr d\theta = \int_0^{+\infty} \left(\int_{-\pi}^{\pi} e^{-r^2} r d\theta \right) dr = 2\pi \int_0^{+\infty} e^{-r^2} r dr.$$

- Enfin, le changement de variable $r^2 = u$, et donc $r dr = \frac{1}{2} du$ nous donne :

$$I_0 = \pi \int_0^{+\infty} e^{-u} du = \pi [-e^{-u}]_0^{+\infty} = \pi.$$

Comme $I_0 = I^2$, cela permet de conclure.

4. Exercices

Exercice III.3.11. — (Convolution de deux fonctions sommables) Soient $f, g \in \mathcal{L}^1(\mathbf{R}^m)$.

- Montrer que $\int \int |f(x-y)g(y)| = \|f\|_1 \|g\|_1$.
- En déduire que, pour presque tout x la fonction $y \mapsto f(x-y)g(y)$ est sommable et que $f * g$ définie p.p. par $f * g(x) = \int f(x-y)g(y) dy$ est elle-aussi sommable.
- Montrer que, si $f_1 = f_2$ p.p. et $g_1 = g_2$ p.p., alors $f_1 * g_1 = f_2 * g_2$ p.p. (L'application $(f, g) \mapsto f * g$ passe donc au quotient et définit une application encore notée $(f, g) \mapsto f * g$ de $L^1(\mathbf{R}^m) \times L^1(\mathbf{R}^m)$ dans $L^1(\mathbf{R}^m)$.)
- Montrer que $(f, g) \mapsto f * g$ induit une application bilinéaire continue de $L^1(\mathbf{R}^m) \times L^1(\mathbf{R}^m)$ dans $L^1(\mathbf{R}^m)$, et que l'on a $f * g = g * f$ et $f * (g * h) = (f * g) * h$, si $f, g, h \in L^1(\mathbf{R}^m)$.

Exercice III.3.12. — Soit ϕ une fonction \mathcal{C}^∞ sur \mathbf{R}^m , à support dans $[-1, 1]^m$, à valeurs dans \mathbf{R}_+ , et vérifiant $\int_{\mathbf{R}^m} \phi = 1$. Si $\varepsilon > 0$, soit ϕ_ε définie par $\phi_\varepsilon(x) = \varepsilon^{-m} \phi(\frac{x}{\varepsilon})$.

- (i) Montrer que $\int_{\mathbf{R}^m} \phi_\varepsilon = 1$, et que ϕ_ε est à support dans $[-\varepsilon, \varepsilon]^m$.
- (ii) Montrer que, si f est une fonction en escalier, alors $f * \phi_\varepsilon \rightarrow f$ p.p., quand $\varepsilon \rightarrow 0$.
- (iii) Montrer que, si f est une fonction en escalier, alors $f * \phi_\varepsilon \rightarrow f$ dans $L^1(\mathbf{R}^m)$, quand $\varepsilon \rightarrow 0$.
- (iv) Montrer que, si $f \in L^1(\mathbf{R}^m)$, alors $f * \phi_\varepsilon \rightarrow f$ dans $L^1(\mathbf{R}^m)$, quand $\varepsilon \rightarrow 0$.
- (v) Montrer que, si $f \in L^1(\mathbf{R}^m) + L^2(\mathbf{R}^m)$ vérifie $\int_{\mathbf{R}^m} \phi f = 0$, pour tout $\phi \in \mathcal{C}_c^\infty(\mathbf{R}^m)$, alors $f = 0$ p.p.

Exercice III.3.13. — Soient $f, g \in L^2(\mathbf{R}^n)$.

- (i) Montrer que $\|f * g\|_\infty \leq \|f\|_2 \|g\|_2$.
- (ii) Montrer que $f * g$ est continue et tend vers 0 à l'infini. (Commencer par des fonctions en escalier.)
- (iii) Montrer que si A et B sont deux sous-ensembles de \mathbf{R}^n de mesure strictement positive, alors $A + B$ contient un ouvert.
- (iv) Un fermé d'intérieur vide est-il forcément de mesure nulle ?

Exercice III.3.14. — (i) Établir la formule $\int_0^1 \int_0^1 \frac{dx dy}{1-x^2 y^2} = \frac{3}{4} \zeta(2)$.

(ii) Soient $\Omega_1 = \{(u, v), u > 0, v > 0, u + v < \frac{\pi}{2}\}$ et $\Omega_2 = \{(x, y), 0 < x, y < 1\}$. Montrer que φ défini par $\varphi(u, v) = (\frac{\sin u}{\cos v}, \frac{\sin v}{\cos u})$ induit un difféomorphisme de Ω_1 sur Ω_2 .

(iii) En déduire la formule $\zeta(2) = \frac{\pi^2}{6}$.

Exercice III.3.15. — Soit $N \in \mathbf{N} - \{0\}$, et soit $\varphi : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ l'application déduite de $z \mapsto z^N$ de \mathbf{C} dans \mathbf{C} , en identifiant \mathbf{R}^2 à \mathbf{C} . On rappelle par anticipation (cf. (i) de la rem. V.1.13) que le jacobien de φ en $z_0 \in \mathbf{C}$ est $|N z_0^{N-1}|^2$. Montrer que, si f est sommable sur \mathbf{R}^2 , alors

$$\int_{\mathbf{R}^2} f(u, v) du dv = N \int_{\mathbf{R}^2} f(\varphi(x, y)) (x^2 + y^2)^{N-1} dx dy.$$

Exercice III.3.16. — (Volume⁽¹⁶⁾ de la boule unité de \mathbf{R}^m)

Soit $m \geq 1$. On munit \mathbf{R}^m de la norme euclidienne standard $\| \cdot \|$. Si $\rho \in \mathbf{R}_+$, soit $B(\rho)$ la boule unité fermée de centre 0 et de rayon ρ . On note C_m le volume de $B(1)$.

- (i) Montrer que $\lambda(B(\rho)) = C_m \rho^m$, si $\rho \in \mathbf{R}_+$.
- (ii) Soit $\phi_r = \sum_{k=1}^{2^r} (\frac{k}{2^r})^{1-m} \mathbf{1}_{B(\frac{k}{2^r}) - B(\frac{k-1}{2^r})}$. Montrer que $\int_{\mathbf{R}^m} \phi_r \rightarrow \int_{B(1)} \|x\|^{1-m} dx$.
- (iii) En déduire que $\int_{B(1)} \|x\|^{1-m} dx = m C_m$.
- (iv) Montrer que, si $\phi \in \text{Esc}(\mathbf{R}_+)$, alors $m C_m \int_{\mathbf{R}_+} \phi(t) dt = \int_{\mathbf{R}^m} \|x\|^{1-m} \phi(\|x\|) dx$. En déduire que $\phi \in L^1(\mathbf{R}_+)$ si et seulement si $\|x\|^{1-m} \phi(\|x\|) \in L^1(\mathbf{R}^m)$, et que l'on a $m C_m \int_{\mathbf{R}_+} \phi(t) dt = \int_{\mathbf{R}^m} \|x\|^{1-m} \phi(\|x\|) dx$, quel que soit $\phi \in L^1(\mathbf{R}_+)$.
- (v) Soit $s \in \mathbf{R}$. Montrer que $\int_{\|x\| \leq 1} \|x\|^s < +\infty$ si et seulement si $s > -m$ et que $\int_{\|x\| \geq 1} \|x\|^s < +\infty$ si et seulement si $s < -m$.
- (vi) En appliquant ce qui précède à la fonction $t^{m-1} e^{-\pi t^2}$, et en utilisant la définition de $\pi = C_2$, en déduire la valeur de $\int_{\mathbf{R}} e^{-\pi x^2} dx$, et montrer que $C_m = \frac{\pi^{m/2}}{\Gamma(1 + \frac{m}{2})}$.

III.4. Construction de l'intégrale de Lebesgue

Ce § est consacré à la démonstration des th. III.1.15 et III.1.19 et de la prop. III.1.12 sur lesquels reposent toute la théorie. Commençons par remarquer que l'unicité n'a pas été utilisée pour déduire le

⁽¹⁶⁾Pour $n = 3$, le résultat était connu du mathématicien indien Bhaskaracarya, vers 1150.

théorème de convergence dominée du th. III.1.15. On en déduit que, si on peut définir l'intégrale, alors le résultat suivant doit être vrai.

Proposition III.4.1. — Soient D un dallage et $M > 0$. Soit $h \in \text{Mes}(D, [0, M])$ et soit $(h_n)_{n \in \mathbf{N}}$ une suite d'éléments de $\text{Esc}(D, [0, M])$ tendant vers h p.p. Alors $\int h_n$ a une limite qui ne dépend que de h .

De plus, si ce résultat est vrai, les théorèmes de convergence dominée et de convergence monotone montrent que l'on doit définir l'intégrale de Lebesgue⁽¹⁷⁾ de la manière suivante.

(L1) Si $f \in \text{Mes}(D, [0, M])$, alors $\int f \in \mathbf{R}_+$ est la limite, quand $n \rightarrow +\infty$, de $\int f_n$, où $(f_n)_{n \in \mathbf{N}}$ est n'importe quelle suite d'éléments de $\text{Esc}(D, [0, M])$ qui converge vers f p.p.

(L2) Si $f \in \text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$, alors $\int f \in \overline{\mathbf{R}}_+$ est la limite, quand $N \rightarrow +\infty$, de la suite croissante de terme général $\int_{D_N} \inf(f, N)$, où D_N est la dalle de sommets $(\pm N, \dots, \pm N)$.

On en déduit l'unicité d'une application $f \mapsto \int f$ satisfaisant aux conclusions du th. III.1.15. On est donc ramené à démontrer la prop. III.4.1, et à vérifier les th. III.1.15 et III.1.19 pour l'intégrale définie par les propriétés (L1) et (L2). La démonstration se fait en plusieurs étapes, la plus délicate étant la démonstration de la prop. III.4.1.

1. Le théorème de convergence dominée pour les fonctions en escalier bornées

Ce n° est consacré à la démonstration de la prop. III.4.1 sous la forme renforcée ci-dessous.

Dans tout ce qui suit, D est un dallage de \mathbf{R}^m , et $M \in \mathbf{R}_+$.

Proposition III.4.2. — Soit $h \in \text{Mes}(D, [0, M])$ et soit $(h_n)_{n \in \mathbf{N}}$ une suite d'éléments de $\text{Esc}(D, [0, M])$ tendant vers h p.p. Alors $\int h_n$ a une limite notée $\int h$ qui ne dépend que de h . De plus, $|h - h_n| \in \text{Mes}(D, [0, M])$, et $\int |h - h_n| \rightarrow 0$.

Lemme III.4.3. — Si A est de mesure extérieure finie, alors, quel que soit $\varepsilon > 0$, il existe un ouvert U contenant A et tel que $\lambda^+(U) \leq \lambda^+(A) + \varepsilon$.

Démonstration. — Si $D = \prod_{j=1}^m [a_j, b_j]$ est une dalle, on peut, quel que soit $\eta > 0$, trouver un pavé ouvert $P = \prod_{j=1}^m]a'_j, b_j[$ contenant D , tel que $\lambda^+(P) \leq \lambda(D) + \eta$. Soit alors $(D_n)_{n \in \mathbf{N}}$ une suite de dalles élémentaires telle que $A \subset \cup_{n \in \mathbf{N}} D_n$, et $\lambda^+(A) \leq \sum_{n \in \mathbf{N}} \lambda(D_n) + \frac{\varepsilon}{2}$. D'après la discussion précédente, il existe P_n , pavé ouvert contenant D_n , tel que $\lambda^+(P_n) \leq \lambda(D_n) + \frac{\varepsilon}{2^{n+3}}$, et $U = \cup_{n \in \mathbf{N}} P_n$ répond aux exigences du lemme.

⁽¹⁷⁾La présentation choisie dans ce texte fait ressembler beaucoup l'intégrale de Lebesgue à celle de Riemann : on part des fonctions en escalier, on définit l'intégrale d'une fonction mesurable par passage à la limite, et enfin on définit la notion d'ensemble mesurable et de mesure d'un ensemble. L'approche originelle de Lebesgue était inverse. Son point de départ était le suivant : pour calculer l'aire d'une surface sous le graphe d'une fonction d'un intervalle $[a, b]$ dans \mathbf{R}_+ , on peut soit découper verticalement (ce que fait Riemann), soit horizontalement (ce que fait Lebesgue). Comme le dit Lebesgue, pour calculer la quantité d'argent dans un tas contenant des pièces de différentes valeurs, l'intégrale de Riemann consiste à prendre chaque pièce à son tour et à ajouter sa valeur au total, alors que l'intégrale de Lebesgue consiste à commencer par trier les pièces et compter combien il y en a de chaque sorte. Évidemment, en découpant horizontalement, on tombe sur des ensembles nettement plus compliqués que verticalement comme un petit dessin le prouvera aisément. L'approche originelle de Lebesgue a l'avantage de s'étendre telle quelle à une théorie de la mesure valable dans un cadre très général (et indispensable en théorie des probabilités). Celle suivie dans ce texte permet d'éviter certains aspects un peu rébarbatifs de théorie des ensembles. Elle a l'inconvénient d'être limitée à des espaces ressemblant (au moins localement) à \mathbf{R}^n .

Lemme III.4.4. — Si $(X_n)_{n \in \mathbf{N}}$ est une suite décroissante de dallages telle que $\cap_{n \in \mathbf{N}} X_n$ est de mesure nulle, alors $\lim_{n \rightarrow +\infty} \lambda(X_n) = 0$.

Démonstration. — Soit \mathcal{L} la réunion des hyperplans de la forme $x_i = r$, pour $1 \leq i \leq m$ et $r \in \mathbf{Z}[\frac{1}{2}]$. Comme $\{1, \dots, m\} \times \mathbf{Z}[\frac{1}{2}]$ est dénombrable, \mathcal{L} est de mesure nulle. Maintenant, si $n \in \mathbf{N}$, et si \bar{X}_n désigne l'adhérence de X_n , on a $\bar{X}_n \subset X_n \cup \mathcal{L}$, et donc

$$\cap_{n \in \mathbf{N}} \bar{X}_n \subset \cap_{n \in \mathbf{N}} (X_n \cup \mathcal{L}) \subset (\cap_{n \in \mathbf{N}} X_n) \cup \mathcal{L};$$

on en déduit que $\cap_{n \in \mathbf{N}} \bar{X}_n$ est de mesure nulle.

Soit alors $\varepsilon > 0$. Comme $\cap_{n \in \mathbf{N}} \bar{X}_n$ est de mesure nulle, il existe un ouvert U_ε contenant $\cap_{n \in \mathbf{N}} \bar{X}_n$, avec $\lambda^+(U_\varepsilon) < \varepsilon$. Soit F_ε le complémentaire de U_ε dans \bar{X}_0 . On a $F_\varepsilon \cap (\cap_{n \in \mathbf{N}} \bar{X}_n) = \emptyset$, ce qui implique qu'il existe $n_0 \in \mathbf{N}$ tel que $F_\varepsilon \cap (\cap_{n \leq n_0} \bar{X}_n) = \emptyset$ car \bar{X}_0 est compact et F_ε et les \bar{X}_n sont fermés dans \bar{X}_0 . Comme la suite $(\bar{X}_n)_{n \in \mathbf{N}}$ est décroissante, on en déduit le fait que $F_\varepsilon \cap \bar{X}_{n_0} = \emptyset$, et donc que $X_n \subset \bar{X}_n \subset U_\varepsilon$, quel que soit $n \geq n_0$. On a donc prouvé que, quel que soit $\varepsilon > 0$, il existe n_0 , tel que $\lambda(X_n) \leq \lambda^+(U_\varepsilon) < \varepsilon$, si $n \geq n_0$. Ceci permet de conclure.

Lemme III.4.5. — Si $(h_n)_{n \in \mathbf{N}}$ est une suite décroissante d'éléments de $\text{Esc}(D, [0, M])$, tendant vers 0 presque partout, alors $\lim_{n \rightarrow +\infty} \int h_n = 0$.

Démonstration. — Si $\varepsilon > 0$, et si $n \in \mathbf{N}$, soit $X_{n,\varepsilon} = \{x \in D, h_n(x) \geq \varepsilon\}$. Alors $(X_{n,\varepsilon})_{n \in \mathbf{N}}$ est une suite décroissante de dallages, et $\cap_{n \in \mathbf{N}} X_{n,\varepsilon}$ est de mesure nulle puisque h_n tend vers 0 presque partout. D'après le lemme III.4.4, ceci implique que $\lim_{n \rightarrow +\infty} \lambda(X_{n,\varepsilon}) = 0$; et donc qu'il existe $n \in \mathbf{N}$ tel que $\lambda(X_{n+p,\varepsilon}) \leq \varepsilon$, si $p \in \mathbf{N}$. Comme $h_{n+p}(x) \leq M$, si $x \in X_{n+p,\varepsilon}$, et $h_{n+p}(x) \leq \varepsilon$, si $x \notin X_{n+p,\varepsilon}$, on a $\int h_{n+p} \leq \varepsilon(\lambda(D) + M)$, quel que soit $p \in \mathbf{N}$. Ceci permet de conclure.

Lemme III.4.6. — Si $(h_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\text{Esc}(D, [0, M])$, tendant vers 0 presque partout, telle que $\sum_{n \in \mathbf{N}} \int |h_{n+1} - h_n| < +\infty$, alors $\lim_{n \rightarrow +\infty} \int h_n = 0$.

Démonstration. — Supposons le contraire. Il existe alors $C > 0$ et une infinité de $n \in \mathbf{N}$ tels que $\int h_n \geq C$. Quitte à extraire une sous-suite de la suite h_n , on peut donc supposer que $\int h_n \geq C$ quel que soit $n \in \mathbf{N}$ (cela ne change pas la condition $\sum_{n \in \mathbf{N}} \int |h_{n+1} - h_n| < +\infty$ car, si $\varphi : \mathbf{N} \rightarrow \mathbf{N}$ est strictement croissante, on a $|h_{\varphi(n+1)} - h_{\varphi(n)}| \leq \sum_{k=\varphi(n)}^{\varphi(n+1)-1} |h_{k+1} - h_k|$). Comme la série $\sum_{n \in \mathbf{N}} \int |h_{n+1} - h_n|$ converge, on peut, quitte à rempacer n par $n + n_0$, supposer de plus que $\sum_{k \in \mathbf{N}} \int |h_{k+1} - h_k| \leq \frac{C}{2}$. Soit alors $g_n = \inf_{k \leq n} h_k$. Par construction, g_n est une suite décroissante d'éléments de $\text{Esc}(D, [0, M])$, qui tend vers 0 presque partout car $g_n \leq h_n$. D'après le lemme III.4.5, cela implique $\lim_{n \rightarrow +\infty} \int g_n = 0$. Par ailleurs, on a $g_n(x) \geq h_0(x) - \sum_{k=0}^{n-1} |h_{k+1}(x) - h_k(x)|$ (avec égalité si et seulement si la suite $(h_n(x))_{n \in \mathbf{N}}$ est décroissante). On en déduit que $\int g_n \geq \int h_0 - \sum_{k=0}^{n-1} \int |h_{k+1} - h_k| \geq C - \frac{C}{2}$, quel que soit $n \in \mathbf{N}$. D'où une contradiction qui permet de conclure.

Passons à la démonstration de la prop. III.4.2. Si $n \leq p$, soient

$$f_{n,p} = \inf_{n \leq k \leq p} h_k \quad \text{et} \quad g_{n,p} = \sup_{n \leq k \leq p} h_k.$$

Alors, quand n est fixé, $f_{n,p}$ (resp. $g_{n,p}$) est une suite décroissante (resp. croissante) d'éléments de $\text{Esc}(D, [0, M])$, alors que quand p est fixé $f_{n,p}$ (resp. $g_{n,p}$) est croissante (resp. décroissante). En particulier, si n est fixé, la suite $\int f_{n,p}$ (resp. $\int g_{n,p}$) est une suite décroissante (resp. croissante) d'éléments de $[0, M\lambda(D)]$; elle admet donc une limite et est de Cauchy. On peut donc trouver $\varphi_0(n) \geq n$ tel que, quels que soient $p_1, p_2 \geq \varphi_0(n)$, on ait

$$\left| \int f_{n,p_1} - \int f_{n,p_2} \right| \leq 2^{-n} \quad \text{et} \quad \left| \int g_{n,p_1} - \int g_{n,p_2} \right| \leq 2^{-n}.$$

On note a_n la limite, quand p tend vers $+\infty$ de la suite croissante $\int g_{n,p} - f_{n,p}$, et on choisit $\varphi : \mathbf{N} \rightarrow \mathbf{N}$, avec $\varphi(n) \geq \varphi_0(n)$, et $\int u_n \geq \frac{1}{2}a_n$, où $u_n = g_{n,\varphi(n)} - f_{n,\varphi(n)}$. Par construction, $u_n \in \text{Esc}(\mathbf{D}, [0, M])$, et $u_n \rightarrow 0$ p.p. car h_n a une limite p.p.

Si $n \in \mathbf{N}$, et si $p \geq \varphi(n)$,

$$\int |g_{n+1,\varphi(n+1)} - g_{n,\varphi(n)}| \leq \int |g_{n+1,\varphi(n+1)} - g_{n+1,p}| + \int |g_{n+1,p} - g_{n,p}| + \int |g_{n,p} - g_{n,\varphi(n)}|.$$

Maintenant, par hypothèse, on a

$$\int |g_{n+1,\varphi(n+1)} - g_{n+1,p}| + \int |g_{n,p} - g_{n,\varphi(n)}| \leq 2^{-n-1} + 2^{-n} \leq 2^{1-n},$$

et comme la suite $(g_{n,p})_{n \leq p}$ est décroissante, on a $|g_{n+1,p} - g_{n,p}| = g_{n,p} - g_{n+1,p}$. On en déduit, quel que soit $p \geq \max_{n \leq N} \varphi(n)$, la majoration

$$\sum_{n=0}^{N-1} \int |g_{n+1,\varphi(n+1)} - g_{n,\varphi(n)}| \leq \sum_{n=0}^{N-1} (2^{1-n} + \int g_{n,p} - g_{n+1,p}) \leq 4 + \int g_{0,p} - g_{N,p} \leq 4 + M\lambda(\mathbf{D}),$$

la dernière inégalité venant de ce que $g_{0,p} - g_{N,p} \in \text{Esc}(\mathbf{D}, [0, M])$. On montre de même que

$$\sum_{n=0}^{N-1} \int |f_{n+1,\varphi(n+1)} - f_{n,\varphi(n)}| \leq 4 + M\lambda(\mathbf{D}).$$

On en déduit que

$$\sum_{n=0}^{+\infty} \int |u_n - u_{n+1}| \leq \sum_{n=0}^{+\infty} \left(\int |f_{n+1,\varphi(n+1)} - f_{n,\varphi(n)}| + \int |g_{n+1,\varphi(n+1)} - g_{n,\varphi(n)}| \right) \leq 8 + 2M\lambda(\mathbf{D}) < +\infty,$$

ce qui permet d'utiliser le lemme III.4.6 pour montrer que $\int u_n \rightarrow 0$ et donc que $a_n \rightarrow 0$. Or on a

$$a_n = \sup_{p \geq n} \left(\left(\int \sup_{n \leq k \leq p} h_k \right) - \left(\int \inf_{n \leq k \leq p} h_k \right) \right) \geq \left(\sup_{k \geq n} \int h_k \right) - \left(\inf_{k \geq n} \int h_k \right).$$

On en déduit que les limites inférieure et supérieure de la suite $(\int h_n)_{n \in \mathbf{N}}$ sont égales et donc que $(\int h_n)_{n \in \mathbf{N}}$ a une limite.

Maintenant, si on part de suites $(h_n)_{n \in \mathbf{N}}$ et $(h'_n)_{n \in \mathbf{N}}$ d'éléments de $\text{Esc}(\mathbf{D}, [0, M])$ convergeant vers h p.p., on peut fabriquer une troisième suite $(h''_n)_{n \in \mathbf{N}}$ d'éléments de $\text{Esc}(\mathbf{D}, [0, M])$ convergeant vers h p.p., en posant $h''_{2n} = h_n$ et $h''_{2n+1} = h'_n$, si $n \in \mathbf{N}$. L'existence de la limite de $\int h''_n$ quand n tend vers $+\infty$ implique l'égalité de $\lim_{n \rightarrow +\infty} \int h_n$ et $\lim_{n \rightarrow +\infty} \int h'_n$, ce qui prouve que la limite ne dépend que de h et pas de la suite $(h_n)_{n \in \mathbf{N}}$.

Finalement, on peut appliquer ce qui précède à $f_n = \inf_{k \geq n} h_k = \lim_{p \rightarrow +\infty} f_{n,p}$ et $g_n = \sup_{k \geq n} h_k = \lim_{p \rightarrow +\infty} g_{n,p}$ qui sont des éléments de $\text{Mes}(\mathbf{D}, [0, M])$ par construction. D'après ce qui précède, on a $\int g_n - f_n = \lim_{p \rightarrow +\infty} \int g_{n,p} - f_{n,p} = a_n$, et $a_n \rightarrow 0$. Comme par ailleurs, $f_n \leq h_n \leq g_n$, et $f_n \leq h \leq g_n$ p.p., ce qui implique $|h - h_n| \leq g_n - f_n$ p.p., on a $\int |h - h_n| \leq a_n$, et donc $\int |h - h_n| \rightarrow 0$.

Ceci termine la démonstration de la proposition.

2. Mesure et mesure extérieure des ensembles mesurables

Ce n° est consacré à la démonstration du th. III.1.19, selon lequel $\lambda(A) = \lambda^+(A)$ pour tout ensemble mesurable A . Si A est mesurable, et si $A_N = A \cap [-N, N]^m$, alors $\lambda(A) = \lim_{N \rightarrow +\infty} \lambda(A_N)$ par définition. Par ailleurs, on déduit des (i) et (ii) de la prop. III.1.3, que $\lambda^+(A) = \sup_{N \in \mathbf{N}} \lambda^+(A_N)$. Il suffit donc de prouver que $\lambda(A_N) = \lambda^+(A_N)$ pour tout N pour prouver que $\lambda(A) = \lambda^+(A)$. Autrement dit, on peut supposer que A est borné.

Dans le reste de ce n^o, on fixe un dallage D de \mathbf{R}^m , et tous les ensembles considérés sont inclus dans D . On dit que $A \subset D$ est *dallable* s'il existe une suite $(D_n)_{n \in \mathbf{N}}$ de dalles élémentaires disjointes telles que $A = \coprod_{n \in \mathbf{N}} D_n$; une telle décomposition de A est une *décomposition en dalles élémentaires*.

Lemme III.4.7. — (i) Si $A \subset D$ est dallable, alors A est mesurable et $\lambda^+(A) = \lambda(A) = \sum_{n \in \mathbf{N}} \lambda(D_n)$, pour toute décomposition $A = \coprod_{n \in \mathbf{N}} D_n$ de A en dalles élémentaires.

(ii) Si $A \subset D$, alors $\lambda^+(A) = \inf \lambda(B)$, où B décrit l'ensemble des ensembles dallables, avec $A \subset B \subset D$.

(iii) Si A et B sont mesurables, alors $A \cup B$ et $A \cap B$ sont mesurables et $\lambda(A \cap B) + \lambda(A \cup B) = \lambda(A) + \lambda(B)$.

Démonstration. — Si $n \in \mathbf{N}$, soit $f_n = \sum_{i \leq n} \mathbf{1}_{D_i}$. Alors $(f_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\text{Esc}(D, [0, 1])$ tendant vers $\mathbf{1}_A$ en tout point, et donc $\int f_n = \sum_{i \leq n} \lambda(D_i)$ tend vers $\int \mathbf{1}_D = \lambda(D)$, d'après la prop. III.4.2. On en déduit que A est mesurable et que $\lambda(A) = \sum_{n \in \mathbf{N}} \lambda(D_n)$. Le (ii) s'en déduit en revenant à la définition de $\lambda^+(A)$, et l'égalité $\lambda^+(A) = \lambda(A)$ du (i) est alors une conséquence immédiate du (ii). Finalement, le (iii) suit de ce que $\mathbf{1}_{A \cup B} = \sup(\mathbf{1}_A, \mathbf{1}_B)$ et $\mathbf{1}_{A \cap B} = \inf(\mathbf{1}_A, \mathbf{1}_B)$ sont mesurables, et de ce que $\mathbf{1}_{A \cup B} + \mathbf{1}_{A \cap B} = \mathbf{1}_A + \mathbf{1}_B$.

Lemme III.4.8. — (i) Si $(A_n)_{n \in \mathbf{N}}$ est une suite croissante de sous-ensembles dallables de D , alors $A = \cup_{n \in \mathbf{N}} A_n$ est dallable, et $\lambda(A) = \sup_{n \in \mathbf{N}} \lambda(A_n)$.

(ii) Si $(A_n)_{n \in \mathbf{N}}$ est une suite croissante de sous-ensembles de D , et si $A = \cup_{n \in \mathbf{N}} A_n$, alors $\lambda^+(A) = \sup_{n \in \mathbf{N}} \lambda^+(A_n)$.

Démonstration. — (i) Écrivons chaque A_n comme une réunion disjointe dénombrable de dalles élémentaires $D_{n,i}$, pour $i \in I_n$. Si $x \in A$, soit D_x la plus grande dalle élémentaire contenant x parmi les $D_{n,i}$, pour $n \in \mathbf{N}$ et $i \in I_n$ (l'existence de D_x vient de ce que les dalles élémentaires se comportent comme des billes de mercure). Soit $J \subset \cup_{n \in \mathbf{N}} (\{n\} \times I_n)$ l'ensemble des (n, i) tels qu'il existe $x \in A$ avec $D_x = D_{n,i}$. Alors A est la réunion disjointe des $D_{n,i}$, pour $(n, i) \in J$, et donc A est dallable. (On remarquera que l'on n'a pas utilisé le fait que la suite $(A_n)_{n \in \mathbf{N}}$ est croissante dans cette partie de la démonstration.)

Maintenant, l'inégalité $\lambda(A) \geq \sup_{n \in \mathbf{N}} \lambda(A_n)$ est immédiate puisque A contient chacun des A_n . Réciproquement, si J' est un sous-ensemble fini de J , alors $\coprod_{(n,i) \in J'} D_{n,i}$ est inclus dans un des A_m , et on a $\sum_{(n,i) \in J'} \lambda(D_{n,i}) \leq \sup_{n \in \mathbf{N}} \lambda(A_n)$, quel que soit $J' \subset J$ fini. En passant à la limite, on obtient l'inégalité $\lambda(A) \leq \sup_{n \in \mathbf{N}} \lambda(A_n)$ qui termine la démonstration du (i).

(ii) L'inégalité $\lambda^+(A) \geq \sup_{n \in \mathbf{N}} \lambda^+(A_n)$ est immédiate. Montrons l'inégalité dans l'autre sens. Si $\varepsilon > 0$, on peut, quel que soit $n \in \mathbf{N}$, trouver D_n dallable tel que $\lambda(D_n) \leq \lambda^+(A_n) + \frac{\varepsilon}{2^{n+1}}$. Soit $D'_n = \cup_{i \leq n} D_i$. Alors D'_n est un ensemble dallable (comme réunion finie d'ensembles dallables), contenant A_n , et la suite D'_n est croissante par construction. De plus, on a

$$\lambda(D'_{n+1}) - \lambda^+(A_{n+1}) = \lambda(D'_n) + \lambda(D_{n+1}) - \lambda(D'_n \cap D_{n+1}) - \lambda^+(A_{n+1}) \leq \lambda(D'_n) - \lambda^+(A_n) + \frac{\varepsilon}{2^{n+1}},$$

car $D'_n \cap D_{n+1}$ contient A_n . On en déduit, par récurrence, que $\lambda(D'_{n+1}) - \lambda^+(A_{n+1}) \leq (1 - \frac{1}{2^{n+1}})\varepsilon$. Maintenant, d'après le (i), $D = \cup_{n \in \mathbf{N}} D'_n$, qui contient A , est dallable, et $\lambda(D) = \sup_{n \in \mathbf{N}} \lambda(D'_n) \leq \varepsilon + \sup_{n \in \mathbf{N}} \lambda^+(A_n)$. On a donc $\lambda^+(A) \leq \varepsilon + \sup_{n \in \mathbf{N}} \lambda^+(A_n)$, quel que soit $\varepsilon > 0$. Ceci permet de conclure.

Revenons à la démonstration de la proposition III.1.19. Si B est un ensemble dallable contenant A , on a $\lambda(A) \leq \lambda(B)$ puisque $\mathbf{1}_A \leq \mathbf{1}_B$. En prenant la borne inférieure sur tous les B dallables contenant A , on obtient $\lambda(A) \leq \lambda^+(A)$, d'après le (ii) du lemme III.4.7.

Pour prouver l'autre inégalité, considérons une suite $(X_k)_{k \in \mathbf{N}}$ de dallages contenus dans D , tels que $\mathbf{1}_{X_k} \rightarrow \mathbf{1}_A$ p.p. Alors, par définition, $\lambda(A) = \lim_{k \rightarrow +\infty} \lambda(X_k)$. Soit $A' = \cup_{k \in \mathbf{N}} \cap_{\ell \geq k} X_\ell$. Par construction, on a $\mathbf{1}_{A'} = \liminf \mathbf{1}_{X_k}$, et donc $\mathbf{1}_{A'} = \mathbf{1}_A$ p.p. On en déduit les égalités $\lambda(A) = \lambda(A')$ et $\lambda^+(A) = \lambda^+(A')$. Maintenant, $\lambda^+(\cap_{\ell \geq k} X_\ell) \leq \inf_{\ell \geq k} \lambda(X_\ell)$, puisque X_ℓ est un dallage contenant $\cap_{\ell \geq k} X_\ell$, si $\ell \geq k$; comme la

suite $\cap_{\ell \geq k} X_\ell$ est croissante, on a $\lambda^+(A') = \sup_{k \in \mathbf{N}} \lambda^+(\cap_{\ell \geq k} X_\ell) \leq \sup_{k \in \mathbf{N}} \inf_{\ell \geq k} \lambda(X_\ell) = \liminf \lambda(X_k) = \lambda(A)$. Ceci permet de conclure.

3. Le théorème de convergence monotone pour les fonctions bornées à support compact

Lemme III.4.9. — Si $f \in \text{Mes}(\mathbf{R}^m, \mathbf{R}_+)$, et si $a \in \mathbf{R}_+$, il existe Y mesurable tel que

$$\{x, f(x) > a\} \subset Y \subset \{x, f(x) \geq a\}.$$

Démonstration. — Soit $(f_k)_{k \in \mathbf{N}}$ une suite d'éléments de $\text{Esc}(\mathbf{R}^m, \mathbf{R}_+)$ tendant vers f en dehors de B , avec B de mesure nulle. Soit $X_k = \{x \in \mathbf{R}^m, f_k(x) > a\}$. Alors X_k est un dallage, quel que soit $k \in \mathbf{N}$, et il existe X tel que, si $x \notin B$, alors $\mathbf{1}_{X_k}(x) \rightarrow \mathbf{1}_X(x)$, et donc $\mathbf{1}_{X_k} \rightarrow \mathbf{1}_X$ p.p., ce qui prouve que X est mesurable. Or X contient $\{x, f(x) > a\} - B$ et est contenu dans $\{x, f(x) \geq a\} \cup B$, et il suffit de poser $Y = (X - (X \cap B)) \cup (\{x, f(x) \geq a\} \cap B)$ pour obtenir un ensemble mesurable (Y diffère de X par un ensemble de mesure nulle) ayant les propriétés requises.

Lemme III.4.10. — Si $(h_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\text{Mes}(\mathbf{R}^m, \mathbf{R}_+)$ tendant simplement vers h p.p., et si $\int h_n \rightarrow 0$, alors $h = 0$ p.p.

Démonstration. — h est la limite p.p. de toute suite extraite de la suite $(h_n)_{n \in \mathbf{N}}$, ce qui permet, quitte à extraire une sous-suite, de supposer que l'on a $\int h_n \leq 2^{-n}$ quel que soit $n \in \mathbf{N}$. Pour montrer que $h = 0$ p.p., il suffit de montrer que, quel que soit $j \in \mathbf{N}$, l'ensemble X_j des x tels que $h(x) > 2^{-j}$ est de mesure nulle. Si $n \in \mathbf{N}$, soit $X_{j,n}$ un ensemble mesurable vérifiant $\{x, h_n(x) > 2^{-j}\} \subset X_{j,n} \subset \{x, h_n(x) \geq 2^{-j}\}$ (le lemme III.4.9 assure l'existence d'un tel $X_{j,n}$). Comme $2^{-j} \mathbf{1}_{X_{j,n}} \leq h_n$, on a $\lambda(X_{j,n}) \leq 2^j \int h_n \leq 2^{j-n}$. Comme $\sum 2^{j-n} < +\infty$, l'ensemble des x tels que $h_n(x) > 2^{-j}$ pour une infinité de $n \in \mathbf{N}$ est de mesure nulle d'après le théorème de Borel-Cantelli, et comme X_j est inclus dans cet ensemble (à l'ensemble près des x tels que $h_n(x) \not\rightarrow h(x)$, qui est de mesure nulle), cela permet de conclure.

Lemme III.4.11. — Si $(g_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\text{Esc}(\mathbf{R}^m)$, telle que $\sum_{n \in \mathbf{N}} \int |g_n| < +\infty$, alors la série $\sum_{n \in \mathbf{N}} g_n(x)$ converge absolument p.p.

Démonstration. — Soit $C = \sum_{n \in \mathbf{N}} \int |g_n|$. Si $M \in \mathbf{R}_+$, soit $X_{M,n}$ l'ensemble des $x \in \mathbf{R}^m$ tel que $\sum_{k \leq n} |g_k(x)| \geq M$. Alors $X_{M,n}$ est une suite croissante de dallages, et on a

$$M\lambda(X_{M,n}) \leq \int_{X_{M,n}} \sum_{k \leq n} |g_k| \leq \int \sum_{k \leq n} |g_k| = \sum_{k \leq n} \int |g_k| \leq C,$$

quel que soit $n \in \mathbf{N}$. On en déduit que $\lambda^+(\cup_{n \in \mathbf{N}} X_{M,n}) \leq M^{-1}C$, et comme l'ensemble A des $x \in \mathbf{R}^m$ tels que $\sum_{n \in \mathbf{N}} |g_n(x)| = +\infty$ est l'intersection des $X_{M,n}$, pour $M \in \mathbf{R}_+$, on a $\lambda^+(A) \leq M^{-1}C$ quel que soit $M \in \mathbf{R}_+$. Ceci implique que A est de mesure nulle, et permet de conclure.

Lemme III.4.12. — Si $(h_k)_{k \in \mathbf{N}}$ est une suite croissante d'éléments de $\text{Mes}(D, [0, M])$, alors la limite h de la suite $(h_k)_{k \in \mathbf{N}}$ est mesurable, et $\int h = \lim \int h_k = \sup \int h_k$.

Démonstration. — La suite $\int h_k$ est croissante et majorée par $M\lambda(D)$; elle admet donc une limite ℓ finie, et, quitte à extraire une sous-suite, on peut supposer que $\ell - \int h_k \leq 2^{-k}$, quel que soit $k \in \mathbf{N}$.

Maintenant, comme h_k est supposée mesurable, il existe une suite de fonctions en escalier $f_{k,\ell}$ tendant vers h_k p.p. Comme h_k est à support dans D et à valeurs dans $[0, M]$, on peut, quitte à remplacer $f_{k,\ell}$ par la fonction valant 0 si $x \notin D$, ou si $\text{Re}(f_{k,\ell}(x)) \leq 0$, valant $\text{Re}(f_{k,\ell}(x))$ si $\text{Re}(f_{k,\ell}(x)) \in [0, M]$ et $x \in D$, et valant M si $\text{Re}(f_{k,\ell}(x)) \geq M$ et $x \in D$, supposer que $f_{k,\ell} \in \text{Esc}(D, [0, M])$. D'après la prop. III.4.2,

$\lim_{\ell \rightarrow +\infty} \int |h_k - f_{k,\ell}| = 0$; il existe donc $\ell(k)$ tel que, si on pose $f_k = f_{k,\ell(k)}$, alors $\int |f_k - h_k| \leq 2^{-k}$. On a donc

$$\int |f_{k+1} - f_k| \leq \int |f_{k+1} - h_{k+1}| + \int |h_{k+1} - h_k| + \int |h_k - f_k| \leq 3 \cdot 2^{-k},$$

et comme $\sum_{k \in \mathbf{N}} 3 \cdot 2^{-k} < +\infty$, le lemme III.4.11 montre que $f_k(x)$ a une limite simple $f(x)$ p.p. La fonction f est alors mesurable comme limite simple p.p. de fonctions en escalier, et $\int f_k \rightarrow \int f$ d'après la prop. III.4.2. Comme $\int f_k$ et $\int h_k$ ont même limite, il suffit, pour terminer la démonstration, de prouver que $f = h$ p.p. Or $f - h$ est la limite p.p. de $f_k - h_k$ et $\int |f_k - h_k| \rightarrow 0$ par construction, ce qui permet d'utiliser le lemme III.4.10 pour conclure.

4. Limites simples p.p. de fonctions mesurables

Le but de ce n° est de démontrer la prop. III.1.12 selon laquelle une limite simple p.p. de fonctions mesurables est mesurable. Le lemme ci-dessous nous permet de nous ramener au cas où toutes les fonctions considérées sont positives.

Lemme III.4.13. — (i) $f : \mathbf{R}^m \rightarrow \mathbf{C}$ est mesurable si et seulement si les fonctions $\operatorname{Re}^+(f)$, $\operatorname{Re}^+(-f)$, $\operatorname{Re}^+(if)$ et $\operatorname{Re}^+(-if)$ sont mesurables.

(ii) Une fonction positive est mesurable si et seulement si elle est limite simple p.p. de fonctions en escalier positives.

Démonstration. — Cela suit de ce que $\operatorname{Re}^+(g)$ est en escalier, si g est en escalier et de ce que Re^+ est continue.

Lemme III.4.14. — Soit f une fonction positive sur \mathbf{R}^m , et si $j \in \mathbf{N}$, soit D_j le dallage de sommets $(\pm 2^j, \dots, \pm 2^j)$.

(i) Si $\mathbf{1}_{D_j} f$ est mesurable pour tout $j \in \mathbf{N}$, alors f est mesurable.

(ii) Si $\mathbf{1}_{D_j} \inf(f, N)$ est mesurable pour tous $j, N \in \mathbf{N}$, alors f est mesurable.

Démonstration. — (i) Soit $(f_{j,k})_{k \in \mathbf{N}}$, si $j \in \mathbf{N}$, une suite de fonctions en escalier tendant vers $\mathbf{1}_{D_j} f$ p.p. Soit g_k la fonction en escalier, nulle en dehors de D_k et égale à $f_{j,k}$ sur $D_j - D_{j-1}$, si $j \leq k$. On a donc $g_k(x) = f_{j,k}(x)$ si $x \in D_j - D_{j-1}$ et $k \geq j$, ce qui permet de prouver que g_k tend vers f p.p. On en déduit le (i).

(ii) Pour démontrer le (ii), compte tenu du (i), on peut supposer que f est à support dans D_j , et donc que $\mathbf{1}_{D_j} \inf(f, N) = \inf(f, N)$. Soit $(f_{N,k})_{k \in \mathbf{N}}$, si $N \in \mathbf{N}$, une suite de fonctions en escalier tendant vers $\inf(f, N)$ p.p. Soit g_k la fonction en escalier valant $f_{1,k}$ si $f_{1,k} < 1$, $f_{2,k}$ si $f_{1,k} = 1$ et $f_{2,k} < 2$, $f_{3,k}$ si $f_{1,k} = 1$, $f_{2,k} = 2$ et $f_{3,k} < 3$, ..., et valant k si $f_{i,k} = i$ quel que soit $i \leq k$. Soit A_N l'ensemble des points tels que $f_{N,k}$ ne tende pas vers $\inf(f, N)$, et soit $A = \bigcup_{N \in \mathbf{N}} A_N$. Alors A est un ensemble de mesure nulle comme réunion dénombrable d'ensembles de mesure nulle, et si $x \notin A$, on a $g_k(x) = f_{i,k}(x)$, si k est assez grand et $f(x) \in]i-1, i[$, et $g_k(x) \in \{f_{i,k}(x), f_{i-1,k}(x)\}$, si k est assez grand et $f(x) = i$. On en déduit que $g_k(x)$ tend vers $f(x)$ si $x \notin A$, ce qui permet de conclure.

Passons à la démonstration de la prop. III.1.12. Soit donc (ϕ_k) une suite de fonctions mesurables positives ayant une limite ϕ p.p. Pour montrer que ϕ est mesurable, il suffit, d'après le lemme III.4.14 de montrer que $\mathbf{1}_{D_j}(\sup(\phi, N))$ est mesurable, quels que soient $j, N \in \mathbf{N}$. Comme $\mathbf{1}_{D_j}(\sup(\phi, N))$ est la limite p.p. de $\mathbf{1}_{D_j}(\sup(\phi_k, N))$, on peut supposer que ϕ et les ϕ_k sont à support dans D_j et à valeurs dans $[0, N]$.

Or on a démontré (lemme III.4.12) qu'une suite croissante d'éléments h_n de $\operatorname{Mes}(D_j, [0, N])$ est mesurable; il en est de même pour une suite décroissante comme on le voit en remplaçant h_n par $N - h_n$. Comme $\phi_k \rightarrow \phi$ p.p., ϕ est aussi la limite inférieure de ϕ_k et donc est mesurable en tant que limite p.p.

de la suite croissante $g_k = \inf_{\ell \geq k} \phi_\ell$, où chaque g_k est mesurable comme limite de la suite décroissante $g_{k,n} = \inf_{k \leq \ell \leq n} \phi_\ell$. Ceci permet de conclure.

5. Le théorème de convergence monotone et ses conséquences

Théorème III.4.15. — Si $(f_n)_{n \in \mathbf{N}}$ est une suite croissante de fonctions mesurables positives sur \mathbf{R}^m , alors $\lim_{n \rightarrow +\infty} \int f_n = \int \lim_{n \rightarrow +\infty} f_n$.

Démonstration. — Notons f la limite de la suite f_n ; c'est un élément de $\text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$. Si $n, N \in \mathbf{N}$, soit $a_{n,N} = \int_{D_N} \inf(f_n, N)$. Par définition, on a $\int f_n = \lim_{N \rightarrow +\infty} a_{n,N}$, et donc $\int f_n = \sup_{N \in \mathbf{N}} a_{n,N}$ puisque la suite est croissante. Par ailleurs, comme $f_n \rightarrow f$ p.p., cela implique que $\inf(f_n, N) \rightarrow \inf(f, N)$ p.p. sur D_N , et comme la suite $\inf(f_n, N)$ est croissante, il résulte du lemme III.4.12 que $\int_{D_N} \inf(f, N) = \sup_{n \in \mathbf{N}} a_{n,N}$. On a donc

$$\int f = \sup_{N \in \mathbf{N}} \left(\sup_{n \in \mathbf{N}} a_{n,N} \right) = \sup_{(n,N) \in \mathbf{N} \times \mathbf{N}} a_{n,N} = \sup_{n \in \mathbf{N}} \left(\sup_{N \in \mathbf{N}} a_{n,N} \right) = \sup_{n \in \mathbf{N}} \int f_n,$$

et comme la suite $(\int f_n)_{n \in \mathbf{N}}$ est croissante, on a aussi $\sup_{n \in \mathbf{N}} \int f_n = \lim_{n \rightarrow +\infty} \int f_n$, ce qui permet de conclure.

On peut maintenant prouver que l'intégrale que l'on a construite satisfait les propriétés (i)-(v) du th. III.1.15, ce qui terminera la preuve de l'existence de l'intégrale de Lebesgue.

- On vient de prouver la propriété (v).
- La (i) est incluse dans la construction.
- Si $f, g \in \text{Mes}(\mathbf{R}^m, \overline{\mathbf{R}}_+)$, alors

$$\inf(f + g, N) \leq \inf(f, N) + \inf(g, N) \leq \inf(f + g, 2N),$$

et donc

$$\mathbf{1}_{D_N} \inf(f + g, N) \leq \mathbf{1}_{D_N} \inf(f, N) + \mathbf{1}_{D_N} \inf(g, N) \leq \mathbf{1}_{D_{2N}} \inf(f + g, 2N).$$

On en déduit, en faisant tendre N vers $+\infty$, et en utilisant le théorème de convergence monotone, les inégalités $\int(f + g) \leq \int f + \int g \leq \int(f + g)$, ce qui prouve que $\int(f + g) = \int f + \int g$. Par récurrence, on en déduit que $\int n f = n \int f$, si $n \in \mathbf{N}$, puis $\int a f = a \int f$, si $a \in \mathbf{Q}_+$, et finalement, en utilisant la croissance de $a \mapsto \int a f$, que $\int a f = a \int f$, si $a \in \mathbf{R}_+$. On en déduit la linéarité de l'intégration (propriété (ii)).

• On sait que, si le théorème de convergence monotone (propriété (v) démontrée ci-dessus) est vérifié, alors la propriété (iii) est un cas particulier du th. III.1.19 (cf. cor. III.1.18), démontré au n° 2.

- On a remarqué (cf. (ii) de la rem. III.1.16) que la (iv) pouvait se déduire des (ii) et (iii).

Ceci permet de conclure.

CHAPITRE IV

TRANSFORMÉE DE FOURIER

La représentation d'une fonction périodique comme somme d'une série de Fourier est un outil très efficace pour la résolution de certaines équations aux dérivées partielles. La formule d'inversion de Fourier, qui permet d'écrire une fonction raisonnable sur \mathbf{R}^m comme somme continue de caractères linéaires unitaires⁽¹⁾ continus, rend le même genre de services pour des équations aux dérivées partielles sur \mathbf{R}^m . Cette « boîte à outils » de Fourier s'adapte à tout groupe commutatif localement compact ; nous l'avons déjà rencontrée dans le cadre des groupes finis (n° 5 du § I.2) ; nous la rencontrerons de nouveau pour \mathbf{Q}_p et pour le groupe des adèles de \mathbf{Q} (cf. n° 2 du § D.2).

IV.1. Intégrales dépendant d'un paramètre

De nombreuses fonctions sont définies comme intégrales de fonctions plus simples⁽²⁾, et le théorème de convergence dominée permet, bien souvent, de démontrer leur continuité et leur dérivabilité.

Théorème IV.1.1. — (Continuité d'une intégrale dépendant d'un paramètre) *Soit X un espace métrique, et soit $x_0 \in X$. Soit $f : X \times \mathbf{R}^m \rightarrow \mathbf{C}$ vérifiant :*

- *la fonction $t \mapsto f(x, t)$ est mesurable, quel que soit $x \in X$;*
- *pour presque tout $t \in \mathbf{R}^m$, la fonction $x \mapsto f(x, t)$ est continue en x_0 ;*
- *il existe $h \in \mathcal{L}^1(\mathbf{R}^m)$ tel que, quel que soit $x \in X$, on ait $|f(x, t)| \leq h(t)$ p.p.*

Alors, si $x \in X$, l'intégrale $\int_{\mathbf{R}^m} f(x, t) dt$ est bien définie et la fonction $F : X \rightarrow \mathbf{C}$ définie par $F(x) = \int_{\mathbf{R}^m} f(x, t) dt$ est continue en x_0 .

Démonstration. — La fonction $t \mapsto f(x, t)$ appartient à \mathcal{L}^1 , quel que soit $x \in X$, puisqu'on l'a supposée mesurable et majorée en module par un élément de $\mathcal{L}^1(\mathbf{R}^m)$. La fonction

⁽¹⁾Un caractère linéaire de \mathbf{R}^m est une fonction $\chi : \mathbf{R}^m \rightarrow \mathbf{C}^*$ vérifiant $\chi(x+y) = \chi(x)\chi(y)$, quels que soient $x, y \in \mathbf{R}^m$; un tel caractère est unitaire si $|\chi(x)| = 1$, quel que soit $x \in \mathbf{R}^m$.

⁽²⁾C'est par exemple le cas de la fonction Γ d'Euler définie par $\Gamma(s) = \int_{\mathbf{R}_+} e^{-ts} \frac{dt}{t}$, ou de la transformée de Fourier \hat{f} d'une fonction sommable f définie (en une variable) par $\hat{f}(x) = \int_{\mathbf{R}} e^{-2i\pi xt} f(t) dt$.

F est donc bien définie. Pour montrer que F est continue en x_0 , il suffit (car X est un espace métrique) de prouver que pour toute suite $(y_n)_{n \in \mathbf{N}}$ convergeant vers x_0 , la suite $(F(y_n))_{n \in \mathbf{N}}$ tend vers $F(x_0)$, c'est-à-dire

$$\lim_{n \rightarrow +\infty} \int_{\mathbf{R}^m} f(y_n, t) dt = \int_{\mathbf{R}^m} f(x_0, t) dt.$$

Posons $g_n(t) = f(y_n, t)$, et $g(t) = f(x_0, t)$. On a alors

- $\lim_{n \rightarrow +\infty} g_n(t) = g(t)$ p.p., car $x \mapsto f(x, t)$ est continue en x_0 , pour presque tout $t \in \mathbf{R}^m$;

- $|g_n(t)| \leq h(t)$, si t n'appartient pas à $\cup_{n \in \mathbf{N}} \{t, |f(y_n, t)| > h(t)\}$, qui est un ensemble de mesure nulle comme réunion dénombrable d'ensembles de mesure nulle.

On est donc dans les conditions d'application du théorème de convergence dominée, et $\lim_{n \rightarrow +\infty} \int_{\mathbf{R}^m} g_n(t) dt = \int_{\mathbf{R}^m} g(t) dt$, ce qui permet de conclure.

Théorème IV.1.2. — (de dérivation sous le signe somme) Soient I un intervalle de \mathbf{R} et $f : I \times \mathbf{R}^m \rightarrow \mathbf{C}$ vérifiant :

- $t \mapsto f(x, t)$ est sommable, quel que soit $x \in I$;
- il existe un ensemble de mesure nulle $A \subset \mathbf{R}^m$ et $h : \mathbf{R}^m \rightarrow \mathbf{R}_+$ sommable, tels que $\frac{\partial f}{\partial x}(x, t)$ existe en tout point de $\mathbf{R}^m - A$ et $|\frac{\partial f}{\partial x}(x, t)| \leq h(t)$, quels que soient ⁽³⁾ $x \in I$ et $t \notin A$.

Alors la fonction F définie sur I par $F(x) = \int_{\mathbf{R}^m} f(x, t) dt$ est dérivable et, quel que soit $x \in I$, on a

$$F'(x) = \int_{\mathbf{R}^m} \frac{\partial f}{\partial x}(x, t) dt.$$

Démonstration. — Quitte à remplacer f par la fonction valant $f(x, t)$, si $t \notin A$, et valant 0, si $t \in A$, ce qui ne change pas la valeur des intégrales, on peut supposer que $A = \emptyset$.

Fixons $x \in I$. Soit $(x_n)_{n \in \mathbf{N}}$ une suite d'éléments de $I - \{x\}$ tendant vers x quand n tend vers $+\infty$. Soit $g(t) = \frac{\partial f}{\partial x}(x, t)$, et si $n \in \mathbf{N}$, soit $g_n(t) = \frac{f(x_n, t) - f(x, t)}{x_n - x}$. Alors g_n tend vers g simplement, et d'après le théorème des accroissements finis, on a

$$|g_n(t)| \leq \sup_{0 \leq \theta \leq 1} \left| \frac{\partial f}{\partial x}(x + \theta(x_n - x), t) \right| \leq h(t),$$

quel que soit $t \in \mathbf{R}^m$. On est donc dans les conditions d'application du théorème de convergence dominée, et $\lim_{n \rightarrow +\infty} \int_{\mathbf{R}^m} g_n = \int_{\mathbf{R}^m} g$. Autrement dit, on a

$$\lim_{n \rightarrow +\infty} \frac{F(x_n) - F(x)}{x_n - x} = \int_{\mathbf{R}^m} \frac{\partial f}{\partial x}(x, t) dt,$$

⁽³⁾La dérivabilité étant une propriété locale, pour démontrer la dérivabilité sur un intervalle I, il suffit de la démontrer sur une suite d'intervalles dont la réunion est I. En d'autres termes, on n'a pas vraiment besoin d'une majoration sur I tout entier, mais sur une suite d'intervalles dont la réunion est I. Cette remarque s'applique aussi au cor. IV.1.3 pour lequel on peut aussi commencer par diminuer Ω .

pour toute suite $(x_n)_{n \in \mathbf{N}}$ d'éléments de $I - \{x\}$ tendant vers x quand n tend vers $+\infty$. On en déduit le résultat.

Si $\ell = (\ell_1, \dots, \ell_n) \in \mathbf{N}^n$, on pose $|\ell| = \ell_1 + \dots + \ell_n$, et $\partial^\ell = (\frac{\partial}{\partial x_1})^{\ell_1} \dots (\frac{\partial}{\partial x_n})^{\ell_n}$.

Corollaire IV.1.3. — Soient Ω un ouvert de \mathbf{R}^n , $k \in \mathbf{N}$ et $f : \Omega \times \mathbf{R}^m \rightarrow \mathbf{C}$ vérifiant :

- $t \mapsto f(x, t)$ est sommable, quel que soit $x \in \Omega$;
- il existe un ensemble de mesure nulle $A \subset \mathbf{R}^m$ et $h : \mathbf{R}^m \rightarrow \mathbf{R}_+$ sommable, tels que, si $t \in \mathbf{R}^m - A$, la fonction $x \mapsto f(x, t)$ est de classe \mathcal{C}^k sur Ω , et $|\partial^\ell f(x, t)| \leq h(t)$, quels que soient $x \in \Omega$, $\ell \in \mathbf{N}^n$, avec $|\ell| \leq k$, et $t \notin A$.

Alors la fonction F définie sur Ω par $F(x) = \int_{\mathbf{R}^m} f(x, t) dt$ est de classe \mathcal{C}^k et, quels que soient $\ell \in \mathbf{N}^n$, avec $|\ell| \leq k$, et $x \in \Omega$, on a

$$\partial^\ell F(x) = \int_{\mathbf{R}^m} \partial^\ell f(x, t) dt.$$

Démonstration. — Cela se déduit du théorème de dérivation sous le signe somme par une récurrence immédiate.

Exercice IV.1.4. — Soit $I =]0, 1[$, et soit $f : I \times \mathbf{R} \rightarrow \mathbf{R}$ définie par $f(x, t) = 0$ si $t \leq 0$ ou si $t \geq x$, et $f(x, t) = 1$ si $0 < t < x$. Calculer explicitement $F(x) = \int_{\mathbf{R}} f(x, t) dt$ et $F'(x)$. Expliquer pourquoi le théorème de dérivation sous le signe somme ne permet pas d'en déduire $1 = 0$. A quel endroit cela intervient-il dans la démonstration du dit théorème ?

Exercice IV.1.5. — (Fonction Γ d'Euler)

- (i) Montrer que l'intégrale $\Gamma(s) = \int_0^{+\infty} e^{-t} t^s \frac{dt}{t}$ est bien définie si $s \in \mathbf{R}_+^*$.
- (ii) Montrer que Γ est de classe \mathcal{C}^∞ sur \mathbf{R}_+^* et tend vers $+\infty$ en $s = 0$.
- (iii) Montrer que $\Gamma(s+1) = s\Gamma(s)$ si $s > 0$; en déduire que $\Gamma(n+1) = n!$, si $n \in \mathbf{N}$.
- (iv) *Formule de Stirling* : montrer que $\Gamma(s+1) \sim (\frac{s}{e})^s \sqrt{2\pi s}$ au voisinage de $+\infty$. On fera le changement de variable $t = s + u\sqrt{s}$, et on montrera que

$$-u\sqrt{s} + s \log(1 + \frac{u}{\sqrt{s}}) \leq \begin{cases} -\frac{u^2}{2} & \text{si } -\sqrt{s} < u \leq 0, \\ -u + \log(1 + u) & \text{si } u \geq 0 \text{ et } s \geq 1. \end{cases}$$

- (v) En déduire la *formule de Gauss* : $\frac{1}{\Gamma(s)} = \lim_{n \rightarrow +\infty} \frac{s(s+1)\dots(s+n)}{n!n^s}$, si $s \in \mathbf{R}_+^*$.

Exercice IV.1.6. — (i) Montrer que l'intégrale $\int_0^{+\infty} \frac{\sin t}{t} dt$ est semi-convergente (c'est-à-dire que $\frac{\sin t}{t}$ est sommable sur $[0, T]$, pour tout T , et que $\int_0^T \frac{\sin t}{t} dt$ a une limite quand $T \rightarrow +\infty$).

(ii) On définit $F(\lambda)$ pour $\lambda \in \mathbf{R}_+$, par $F(\lambda) = \int_0^{+\infty} e^{-\lambda t} \frac{\sin t}{t} dt$. Montrer que F est de classe \mathcal{C}^1 sur \mathbf{R}_+^* et calculer $F'(\lambda)$.

(iii) Montrer que F tend vers 0 en $+\infty$; en déduire $F(\lambda)$, pour $\lambda > 0$.

(iv) Montrer que F est continue en 0 ; en déduire la valeur de $\int_0^{+\infty} \frac{\sin t}{t} dt$.

Exercice IV.1.7. — Soient $f \in \mathcal{L}^1(\mathbf{R}^m)$ et $g \in \mathcal{C}_c^k(\mathbf{R}^m)$, où $k \in \mathbf{N} \cup \{\infty\}$. Montrer que $x \mapsto f * g(x) = \int f(x-y)g(y) dy$ est de classe \mathcal{C}^k sur \mathbf{R}^m .

IV.2. Transformée de Fourier dans L^1

1. Caractères linéaires de \mathbf{R} et \mathbf{R}^m

Si $x = (x_1, \dots, x_m)$ et $t = (t_1, \dots, t_m)$ sont deux éléments de \mathbf{R}^m , on note $x \cdot t$ leur produit scalaire $\sum_{i=1}^m x_i t_i$; on a $x \cdot t = t \cdot x$. Rappelons que, si $A \in \mathbf{M}_m(\mathbf{R})$, on note ${}^t A$ la matrice transposée de A . On a alors ${}^t A x \cdot t = x \cdot A t$ quels que soient $x, t \in \mathbf{R}^m$.

Proposition IV.2.1. — (i) Les caractères linéaires continus de \mathbf{R} sont les $t \mapsto e^{\lambda t}$, pour $\lambda \in \mathbf{C}$, et les caractères linéaires unitaires continus sont les $t \mapsto e^{2i\pi x t}$, pour $x \in \mathbf{R}$.

(ii) Les caractères linéaires unitaires continus de \mathbf{R}^m sont les $t \mapsto e^{2i\pi x \cdot t}$, pour $x \in \mathbf{R}^m$.

Démonstration. — Soit $\chi : \mathbf{R} \rightarrow \mathbf{C}^*$ un caractère linéaire continu. On a en particulier $\chi(0) = 1$, et la continuité implique l'existence de $j \in \mathbf{N}$ tel que $|\chi(t) - 1| \leq \frac{1}{2}$, si $|t| \leq 2^{-j}$. Notons $\log : \mathbf{C}^* \rightarrow \{x + iy, -\pi < y \leq \pi\}$ le logarithme complexe. Comme $B(1, \frac{1}{2})$ est incluse dans le secteur angulaire $|\arg(z)| \leq \frac{\pi}{4}$, l'application $g = \log \circ \chi : B(1, \frac{1}{2}) \rightarrow \mathbf{C}$ est à valeur dans la bande $\{x + iy, |y| \leq \frac{\pi}{4}\}$. Maintenant, $\log(z_1 z_2) = \log z_1 + \log z_2 \in 2i\pi$, pour tous $z_1, z_2 \in \mathbf{C}^*$, et comme $\chi(t_1 + t_2) = \chi(t_1)\chi(t_2)$, on a $g(t_1 + t_2) = g(t_1) + g(t_2)$ si t_1, t_2 et $t_1 + t_2$ appartiennent à $B(1, \frac{1}{2})$. On en déduit, par récurrence sur n , que $g(2^{-j-n}) = 2^{-n}g(2^{-j})$, pour tout $n \in \mathbf{N}$, et, par récurrence sur k , que $g(k2^{-j-n}) = kg(2^{-j-n})$, si $k \in \{-2^n, \dots, 2^n\}$. On a donc $g(r2^{-j}) = rg(2^{-j})$, si r est un nombre dyadique dans l'intervalle $[-1, 1]$, et comme g est continue, on en déduit que $g(t) = \lambda t$, avec $\lambda = 2^j g(2^{-j})$, pour tout $t \in [-2^{-j}, 2^{-j}]$. Par définition de g , cela implique que $\chi(t) = e^{\lambda t}$ pour tout $t \in [-2^{-j}, 2^{-j}]$. Enfin, si $t \in \mathbf{R}$, il existe $n \in \mathbf{N}$ tel que $t/n \in [-2^{-j}, 2^{-j}]$, et comme $\chi(t) = \chi(t/n)^n$, on a $\chi(t) = e^{\lambda t}$ pour tout $t \in \mathbf{R}$.

Maintenant, si χ est unitaire, on doit avoir $\lambda t \in i\mathbf{R}$, pour tout $t \in \mathbf{R}$, et il existe donc $x \in \mathbf{R}$ tel que $\lambda = 2i\pi x$. On en déduit le (i).

Si $\chi : \mathbf{R}^m \rightarrow \mathbf{C}^*$ est un caractère linéaire unitaire continu, alors sa restriction à $\mathbf{R}e_j$ définit un caractère linéaire unitaire continu de \mathbf{R} , pour tout $j \in \{1, \dots, m\}$. Il existe donc $x_j \in \mathbf{R}$ tel que $\chi(t_j e_j) = e^{2i\pi x_j t_j}$. Or $t = \sum_{j=1}^m t_j e_j$, et donc $\chi(t) = \prod_{j=1}^m \chi(t_j e_j) = e^{2i\pi x \cdot t}$, avec $x = (x_1, \dots, x_m)$. Ceci permet de conclure.

2. Définition et premières propriétés

Si $f \in \mathcal{L}^1(\mathbf{R}^m)$, on a $|e^{-2i\pi x \cdot t} f(t)| = |f(t)|$ quels que soient $x, t \in \mathbf{R}^m$; l'intégrale $\int_{\mathbf{R}^m} e^{-2i\pi x \cdot t} f(t) dt$ est donc bien définie pour toute valeur de $x \in \mathbf{R}^m$. On appelle *transformée de Fourier* de f la fonction \hat{f} définie, pour $x \in \mathbf{R}^m$, par

$$\hat{f}(x) = \int_{\mathbf{R}^m} e^{-2i\pi x \cdot t} f(t) dt.$$

Elle ne dépend que de la classe de f dans $L^1(\mathbf{R}^m)$, ce qui permet de définir la transformée de Fourier d'un élément de $L^1(\mathbf{R}^m)$ par la même formule. On note souvent, pour des

raisons typographiques, $\mathcal{F}f$ au lieu de \hat{f} , la transformée de Fourier de f , et on définit $\overline{\mathcal{F}f}$ par $\overline{\mathcal{F}f}(x) = \hat{f}(-x)$.

Exemple IV.2.2. — La transformée de Fourier de $\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}[}$ est $\alpha(x) = \frac{\sin \pi x}{\pi x}$, comme le montre un calcul immédiat.

Soit $f \in L^1(\mathbf{R}^m)$. Des changements de variable immédiats montrent que :

- $\mathcal{F}(f(at))(x) = |a|^{-m} \hat{f}(a^{-1}x)$, si $a \in \mathbf{R}^*$; la transformée de Fourier transforme une dilatation en dilatation de rapport inverse;
- $\mathcal{F}(f(t+b))(x) = e^{2i\pi b \cdot x} \hat{f}(x)$ et $\mathcal{F}(e^{2i\pi c \cdot t} f(t))(x) = \hat{f}(x-c)$, si $b, c \in \mathbf{R}^m$; la transformée de Fourier échange les translations et les multiplications par un caractère.

Exercice IV.2.3. — Montrer plus généralement que, si $f \in L^1(\mathbf{R}^m)$, si $A \in \mathbf{GL}_m(\mathbf{R})$, si $b, c \in \mathbf{R}^m$, et si $g(t) = e^{2i\pi c \cdot t} f(At+b)$, alors $\hat{g}(x) = |\det A|^{-1} e^{2i\pi {}^t A^{-1}(x-c) \cdot b} \hat{f}({}^t A^{-1}(x-c))$.

Exercice IV.2.4. — Soit f une fonction continue bornée et sommable sur \mathbf{R} , et soit $x_0 \in \mathbf{R}$. Montrer que, quand λ tend vers 0^+ , la fonction

$$F(\lambda) = \int_{-\infty}^{+\infty} e^{2i\pi x_0 y - \lambda |y|} \hat{f}(y) dy$$

tend vers une limite que l'on calculera. En déduire que si \hat{f} est identiquement nulle, alors f est identiquement nulle.

Exercice IV.2.5. — Soit $f \in L^1(\mathbf{R})$. Montrer que $\lim_{T \rightarrow +\infty} T^{-1} \int_{-T}^T |\hat{f}(x)|^2 dx = 0$.

3. Le théorème de Riemann-Lebesgue

On rappelle que $\mathcal{C}_0(\mathbf{R}^m)$ désigne l'espace des fonctions continues sur \mathbf{R}^m , tendant vers 0 à l'infini.

Proposition IV.2.6. — (i) Si $r \in \mathbf{N}$, et si $\mathbf{k} \in \mathbf{Z}^m$, alors

$$\hat{e}_{r,\mathbf{k}}(x) = 2^{-rm} \prod_{j=1}^m \left(e^{-2^{1-r} i\pi(k_j + \frac{1}{2})x_j} \alpha(2^{-r}x_j) \right).$$

(ii) Si f est une fonction en escalier, alors $\hat{f} \in \mathcal{C}_0(\mathbf{R}^m)$.

Démonstration. — Le (i) est une conséquence de la formule

$$e_{r,\mathbf{k}}(t) = \prod_{j=1}^m \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}[}(2^r t_j - k_j - \frac{1}{2}),$$

de la formule de l'exemple IV.2.2, et des formules pour les dilatations-translations. Le (ii) est une conséquence du (i), de ce que les $e_{r,\mathbf{k}}$ forment une famille génératrice de $\text{Esc}(\mathbf{R}^m)$, et de ce que α est une fonction continue sur \mathbf{R} , tendant vers 0 à l'infini.

Théorème IV.2.7. — L'application $f \mapsto \hat{f}$ est une application linéaire 1-lipschitzienne de $L^1(\mathbf{R}^m)$ dans $\mathcal{C}_0(\mathbf{R}^m)$. Autrement dit, si $f \in L^1(\mathbf{R}^m)$, alors \hat{f} est une fonction continue sur \mathbf{R}^m , tendant vers 0 à l'infini, et on a $\|\hat{f}\|_\infty \leq \|f\|_1$.

Démonstration. — La linéarité de $f \mapsto \hat{f}$ suit de la linéarité de l'intégration, et l'inégalité $\|\hat{f}\|_\infty \leq \|f\|_1$ suit de la majoration

$$\left| \int_{\mathbf{R}^m} e^{-2i\pi x \cdot t} f(t) dt \right| \leq \int_{\mathbf{R}^m} |e^{-2i\pi x \cdot t} f(t)| dt = \int_{\mathbf{R}^m} |f(t)| dt = \|f\|_1.$$

Maintenant, si f est une fonction en escalier, la prop. IV.2.6, montre que $\hat{f} \in \mathcal{C}_0(\mathbf{R}^m)$. Comme les fonctions en escalier sont denses dans $L^1(\mathbf{R}^m)$, comme $f \mapsto \hat{f}$ est linéaire continue de $L^1(\mathbf{R}^m)$ dans l'espace $\mathcal{B}(\mathbf{R}^m)$ des fonctions bornées sur \mathbf{R}^m (muni de la norme $\|\cdot\|_\infty$) dans lequel $\mathcal{C}_0(\mathbf{R}^m)$ est fermé, on en déduit que $\hat{f} \in \mathcal{C}_0(\mathbf{R}^m)$ quel que soit $f \in L^1(\mathbf{R}^m)$. Ceci permet de conclure.

4. Transformée de Fourier et dérivation

Une des propriétés fondamentales de la transformée de Fourier est d'échanger la régularité et la décroissance à l'infini (i.e. plus une fonction est régulière, plus sa transformée de Fourier est petite à l'infini, et réciproquement, plus une fonction est petite à l'infini et plus sa transformée de Fourier est régulière), ainsi que dérivations et multiplications par des polynômes, ce qui, combiné avec la formule d'inversion de Fourier (prop. IV.3.24), facilite grandement l'étude de certaines équations aux dérivées partielles⁽⁴⁾. On a en particulier le résultat suivant.

Théorème IV.2.8. — (i) Si $f \in \mathcal{C}^k(\mathbf{R}^m)$ a toutes ses dérivées partielles d'ordre $\leq k$ sommables, alors $(1 + \|x\|^2)^{k/2} \hat{f}(x)$ tend vers 0 à l'infini, et $\mathcal{F}(\partial^\ell f) = (2i\pi x)^\ell \hat{f}$ si $\ell \in \mathbf{N}^m$ vérifie⁽⁵⁾ $|\ell| \leq k$.

(ii) Si $(1 + \|t\|^2)^{k/2} f(t)$ est sommable, alors \hat{f} est de classe \mathcal{C}^k , et on a $\partial^\ell \hat{f}(x) = (-2i\pi)^\ell \mathcal{F}(t^\ell f)$, si $|\ell| \leq k$.

Démonstration. — Si $f \in \mathcal{C}_c^k(\mathbf{R}^m)$, la formule $\mathcal{F}(\partial^\ell f) = (2i\pi x)^\ell \hat{f}$ s'obtient en intégrant par partie (on intègre $\partial^\ell f$ et on dérive $e^{-2i\pi x \cdot t}$). Par exemple, si $f \in \mathcal{C}_c^1(\mathbf{R}^2)$, on déduit

⁽⁴⁾ Soit $P = \sum_{\ell} a_{\ell} X^{\ell} \in \mathbf{C}[X_1, \dots, X_m]$, et soit $P(\partial)$ l'opérateur différentiel $\sum_{\ell} a_{\ell} \partial^{\ell}$. Si on cherche à résoudre l'équation aux dérivées partielles $P(\partial)u = \phi$, où ϕ est donnée, et supposée suffisamment sympathique, on peut appliquer formellement la transformée de Fourier aux deux membres, pour obtenir $P(2i\pi x)\hat{u}(x) = \hat{\phi}(x)$. En appliquant la formule d'inversion de Fourier, cela nous donne $u = \mathcal{F}\left(\frac{\hat{\phi}(x)}{P(2i\pi x)}\right)$. Ce qui précède est un jeu d'écriture purement formel, mais donne des résultats utilisables dans de nombreux cas provenant de problèmes physiques. Cette méthode de résolution d'équations aux dérivées partielles acquiert une efficacité maximale dans le cadre de la théorie des distributions.

⁽⁵⁾ On rappelle que $(2i\pi x)^{\ell} = \prod_{j=1}^m (2i\pi x_j)^{\ell_j}$, si $x = (x_1, \dots, x_m)$ et $\ell = (\ell_1, \dots, \ell_m)$.

du théorème de Fubini, que

$$\widehat{\partial_1 f}(x_1, x_2) = \int_{\mathbf{R}^2} e^{-2i\pi(x_1 t_1 + x_2 t_2)} \frac{\partial f}{\partial t_1}(t_1, t_2) dt_1 dt_2 = \int_{\mathbf{R}} e^{-2i\pi x_2 t_2} \left(\int_{\mathbf{R}} e^{-2i\pi x_1 t_1} \frac{\partial f}{\partial t_1}(t_1, t_2) dt_1 \right) dt_2.$$

Une intégration par partie dans l'intégrale entre parenthèses nous donne

$$\int_{\mathbf{R}} e^{-2i\pi x_1 t_1} \frac{\partial f}{\partial t_1}(t_1, t_2) dt_1 = [e^{-2i\pi x_1 t_1} f(t_1, t_2)]_{t_1=-\infty}^{+\infty} + 2i\pi x_1 \int_{\mathbf{R}} e^{-2i\pi x_1 t_1} f(t_1, t_2) dt_1,$$

et comme f est à support compact, le premier terme du second membre est nul. On réinjecte alors le second terme dans l'intégrale double, et on réutilise le théorème de Fubini, pour obtenir

$$\begin{aligned} \widehat{\partial_1 f}(x_1, x_2) &= \int_{\mathbf{R}} e^{-2i\pi x_2 t_2} \left(2i\pi x_1 \int_{\mathbf{R}} e^{-2i\pi x_1 t_1} f(t_1, t_2) dt_1 \right) dt_2 \\ &= 2i\pi x_1 \int_{\mathbf{R}^2} e^{-2i\pi(x_1 t_1 + x_2 t_2)} f(t_1, t_2) dt_1 dt_2 = 2i\pi x_1 \hat{f}(x_1, x_2). \end{aligned}$$

Le cas général se traite, par récurrence, de la même manière.

Pour traiter le cas f général, choisissons $\phi \in \mathcal{C}_c^k(\mathbf{R}^m)$ valant 1 sur $[-1, 1]^m$, et définissons f_j par $f_j(x) = f(x)\phi(2^{-j}x)$, si $j \in \mathbf{N}$. Alors $f_j \in \mathcal{C}_c^k(\mathbf{R}^m)$ et un petit calcul utilisant la formule de Leibnitz pour la dérivée d'un produit montre que $\partial^\ell f_j$ tend simplement vers $\partial^\ell f$ et que $\partial^\ell f_j$ est majorée par une somme de dérivées \mathbf{k} -ièmes de f , avec $|\mathbf{k}| \leq |\ell|$, ce qui implique que $\partial^\ell f_j$ tend vers $\partial^\ell f$ dans $L^1(\mathbf{R}^m)$. Ceci permet de déduire l'identité $\mathcal{F}(\partial^\ell f) = (2i\pi x)^\ell \hat{f}$ par passage à la limite (en utilisant la continuité de $\phi \mapsto \hat{\phi}(x)$ qui découle de la continuité de $\phi \mapsto \hat{\phi}$ de $L^1(\mathbf{R}^m)$ dans $\mathcal{C}_0(\mathbf{R}^m)$). On en déduit le (i) car $\mathcal{F}(\partial^\ell f)$ tendant vers 0 à l'infini quel que soit $\ell \in \mathbf{N}^m$ vérifiant $|\ell| \leq k$, il en est de même de $|x|^\ell \hat{f}$, et donc aussi de $|(1 + \|x\|^2)^{k/2} \hat{f}(x)|$ car $(1 + \|x\|^2)^{k/2} \leq (1 + \sum_{j=1}^m |x_j|)^k$.

Le (ii) est, quant à lui, une simple application du théorème de dérivation sous le signe somme, une fois que l'on a remarqué que $\partial^\ell(e^{-2i\pi x \cdot t}) = (-2i\pi t)^\ell e^{-2i\pi x \cdot t}$.

IV.3. Formules d'inversion

1. Séries de Fourier

Une fonction $f : \mathbf{R} \rightarrow \mathbf{C}$ est *périodique de période 1*, si $f(x+1) = f(x)$, pour tout $x \in \mathbf{R}$. On a alors $f(x+n) = f(x)$ pour tous $x \in \mathbf{R}$ et $n \in \mathbf{Z}$, et donc f est périodique de période \mathbf{Z} .

On peut aussi (et c'est souvent nettement plus agréable) voir une fonction périodique de période 1 comme une fonction de $\mathbf{T} = \mathbf{R}/\mathbf{Z}$ dans \mathbf{C} . Le passage d'un point de vue à l'autre se fait de la manière suivante, en notant $\pi : \mathbf{R} \rightarrow \mathbf{T}$ l'application envoyant $x \in \mathbf{R}$ sur sa classe modulo \mathbf{Z} : si f est une fonction sur \mathbf{T} , alors $f \circ \pi : \mathbf{R} \rightarrow \mathbf{C}$ est une fonction périodique de période 1, et réciproquement, si $g : \mathbf{R} \rightarrow \mathbf{C}$ est périodique de période 1, alors il existe $f : \mathbf{T} \rightarrow \mathbf{C}$ unique, telle que $g = f \circ \pi$.

\mathbf{T} est muni de la topologie quotient, ce qui signifie que $f : \mathbf{T} \rightarrow \mathbf{C}$ est continue si et seulement si $f \circ \pi : \mathbf{R} \rightarrow \mathbf{C}$ est continue. L'espace $\mathcal{C}(\mathbf{T})$ des fonctions continues sur \mathbf{T} s'identifie donc à l'espace des fonctions continues sur \mathbf{R} , périodiques de période 1.

\mathbf{T} est un groupe (quotient du groupe commutatif $(\mathbf{R}, +)$ par son sous-groupe \mathbf{Z}), et, par construction, $\pi : \mathbf{R} \rightarrow \mathbf{T}$ est un morphisme de groupes dont le noyau est \mathbf{Z} . Si $n \in \mathbf{Z}$, alors $t \mapsto e^{2i\pi nt}$ est un caractère continu de \mathbf{R} dont le noyau contient \mathbf{Z} , et donc est un caractère continu de \mathbf{T} en vertu de notre identification entre les fonctions périodiques de période 1 et les fonctions sur \mathbf{T} . On note χ_n le caractère $t \mapsto e^{2i\pi nt}$ de \mathbf{T} , si $n \in \mathbf{N}$.

Proposition IV.3.1. — *Les caractères linéaires continus de \mathbf{T} sont les χ_n , pour $n \in \mathbf{Z}$.*

Démonstration. — Si $\chi : \mathbf{T} \rightarrow \mathbf{C}^*$ est un caractère linéaire continu, alors $\psi = \chi \circ \pi$ est un caractère continu de \mathbf{R} , périodique de période 1. D'après la prop. IV.2.1, il existe $\lambda \in \mathbf{C}$ tel que $\psi(t) = e^{\lambda t}$, pour tout $t \in \mathbf{R}$. La périodicité de ψ équivaut alors à $\psi(1) = \psi(0) = 1$, ce qui montre que λ doit être de la forme $2i\pi n$, avec $n \in \mathbf{Z}$. Ceci permet de conclure.

Si $a \in \mathbf{R}$, tout élément t de \mathbf{R} peut s'écrire de manière unique sous la forme $t = x + n$, avec $x \in [a, a + 1[$ et $n \in \mathbf{Z}$. Autrement dit, l'intervalle $[a, a + 1[$ est un système de représentants de \mathbf{R} modulo \mathbf{Z} , pour tout $a \in \mathbf{R}$. On en déduit que l'application ι_a , qui à une fonction f sur \mathbf{T} associe sa restriction (plus précisément, la restriction de $f \circ \pi$) à $[a, a + 1[$, est un isomorphisme de l'espace des fonctions sur \mathbf{T} sur celui des fonctions sur $[a, a + 1[$. On va utiliser ces isomorphismes pour définir un certain nombre d'espaces de fonctions sur \mathbf{T} .

On vérifie facilement, en utilisant l'invariance de l'intégrale de Lebesgue par translation, que les définitions suivantes, pour $f : \mathbf{T} \rightarrow \mathbf{C}$, ne dépendent pas du choix de $a \in \mathbf{R}$:

f est dite nulle p.p., si $\iota_a(f)$ est nulle p.p.,

f est dite sommable, si $\iota_a(f)$ est sommable ; si f est sommable, on définit $\int_{\mathbf{T}} f$ par $\int_{\mathbf{T}} f = \int_a^{a+1} f(t) dt$, et $\|f\|_1$ par $\|f\|_1 = \int_{\mathbf{T}} |f|$.

f est dite de carré sommable, si $\iota_a(f)$ est de carré sommable ; si f est de carré sommable, on pose $\|f\|_2 = (\int_{\mathbf{T}} |f|^2)^{1/2}$, et si f, g sont de carrés sommables, on définit leur produit scalaire $\langle f, g \rangle$ par la formule $\langle f, g \rangle = \int_{\mathbf{T}} \bar{f} g$.

On note, $\mathcal{L}^1(\mathbf{T})$ (resp. $\mathcal{L}^2(\mathbf{T})$) l'espace des fonctions sommables (resp. de carré sommable), et $L^1(\mathbf{T})$ (resp. $L^2(\mathbf{T})$) son quotient par l'espace des fonctions nulles p.p. Par définition, ces espaces sont isométriques (via ι_a) aux espaces $\mathcal{L}^1([a, a + 1[)$, $\mathcal{L}^2([a, a + 1[)$, $L^1([a, a + 1[)$ et $L^2([a, a + 1[)$ respectivement. On en déduit que $\text{Esc}(\mathbf{T})$ est dense dans $L^1(\mathbf{T})$ et $L^2(\mathbf{T})$: en effet, $\text{Esc}([0, 1])$ est dense dans $L^1([0, 1])$ et $L^2([0, 1])$ (cf. th. III.2.11). Comme $[a, a + 1[$ est de volume fini, on a $L^2(\mathbf{T}) \subset L^1(\mathbf{T})$.

On définit $\text{Esc}(\mathbf{T})$ comme l'image inverse de $\text{Esc}([0, 1])$ par ι_0 ; si $r \in \mathbf{N}$, et si $k \in \{0, \dots, 2^r - 1\}$, on note encore $e_{r,k}$ la fonction sur \mathbf{T} dont l'image par ι_0 est $e_{r,k}$. Les $e_{r,k}$, pour $r \in \mathbf{N}$ et $k \in \{0, \dots, 2^r - 1\}$ forment une famille génératrice de $\text{Esc}(\mathbf{T})$.

Soit $\text{Trig}(\mathbf{T})$ l'espace des polynômes trigonométriques (i.e. des combinaisons linéaires des χ_n , pour $n \in \mathbf{Z}$).

Théorème IV.3.2. — (i) $\text{Trig}(\mathbf{T})$ est dense dans $L^2(\mathbf{T})$.

(ii) Les χ_n , pour $n \in \mathbf{Z}$, forment une base hilbertienne de $L^2(\mathbf{T})$.

Démonstration. — On a

$$\langle \chi_n, \chi_m \rangle = \int_a^{a+1} e^{2i\pi(m-n)t} dt = \begin{cases} 1 & \text{si } m = n, \\ \left[\frac{1}{2i\pi(m-n)} e^{2i\pi(m-n)t} \right]_a^{a+1} = 0 & \text{si } m \neq n. \end{cases}$$

Les χ_n forment donc une famille orthonormale, et le (ii) est une conséquence du (i), puisque les χ_n engendrent $\text{Trig}(\mathbf{T})$. La manière standard (cf. ex. II.2.2) pour démontrer la densité de $\text{Trig}(\mathbf{T})$ dans $L^2(\mathbf{T})$ est de passer par sa densité dans $\mathcal{C}(\mathbf{T})$ (cas particulier du théorème de Stone-Weierstrass). Nous proposons ci-dessous une autre approche, via les fonctions en escalier.

Soit F l'adhérence de $\text{Trig}(\mathbf{T})$ dans $L^2(\mathbf{T})$. Le lemme IV.3.3 ci-dessous montre que $\phi_0 \in F$, où $\phi_0 \in L^2(\mathbf{T})$ est définie par $\phi_0(t) = t$, si $t \in [-\frac{1}{2}, \frac{1}{2}]$. Nous allons en déduire que F contient $\text{Esc}(\mathbf{T})$, ce qui permettra de conclure, cet espace étant dense dans $L^2(\mathbf{T})$.

Soit $T_a : L^2(\mathbf{T}) \rightarrow L^2(\mathbf{T})$ définie par $(T_a(\phi))(t) = \phi(t+a)$. Alors T_a est une isométrie grâce à l'invariance par translation de l'intégration ; en particulier, T_a est continue. Comme $\text{Trig}(\mathbf{T})$ est stable par T_a , il en est de même de F (si $\phi \in F$, et si $(f_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\text{Trig}(\mathbf{T})$ tendant vers ϕ dans $L^2(\mathbf{T})$, alors $(T_a(f_n))_{n \in \mathbf{N}}$ est une suite d'éléments de $\text{Trig}(\mathbf{T})$ tendant vers $T_a(\phi)$ dans $L^2(\mathbf{T})$). Maintenant, si $r \in \mathbf{N}$, et si $k \in \{0, \dots, 2^r - 1\}$, on a $e_{r,k} = 2^{-r} - T_{-\frac{1}{2} - \frac{k}{2^r}}(\phi_0) + T_{-\frac{1}{2} - \frac{k+1}{2^r}}(\phi_0)$, comme le montre un petit calcul ; on déduit donc de l'appartenance de ϕ_0 et des constantes à F , celle de $e_{r,k}$, pour tous $r \in \mathbf{N}$ et $k \in \{0, \dots, 2^r - 1\}$. Les $e_{r,k}$ formant une famille génératrice de $\text{Esc}(\mathbf{T})$, cela implique $\text{Esc}(\mathbf{T}) \subset F$, ce que l'on cherchait à démontrer.

Lemme IV.3.3. — (i) Si $t \in]-\frac{1}{2}, \frac{1}{2}[$, et si $z = e^{2i\pi t}$, alors $\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{2i\pi n} (z^n - \bar{z}^n) = t$.

(ii) La série $\sum_{n \neq 0} \frac{(-1)^{n-1}}{2i\pi n} e^{2i\pi n t}$ tend vers ϕ_0 dans $L^2(\mathbf{T})$.

Démonstration. — Si $|a| = 1$ et $a \neq -1$, on a

$$\sum_{n=1}^N \frac{(-1)^{n-1}}{2i\pi n} a^n = \frac{a}{2i\pi} \int_0^1 \sum_{n=1}^N (-ua)^{n-1} du = \frac{a}{2i\pi} \int_0^1 \frac{1 - (-ua)^N}{1 + ua} du.$$

Comme $|a| = 1$, la suite de fonctions $\frac{1 - (-ua)^N}{1 + ua}$, pour $N \in \mathbf{N}$, tend simplement vers $\frac{1}{1 + ua}$, sur $[0, 1[$, et est majorée en module par $\frac{2}{|1 + ua|}$, qui est sommable, puisque $a \neq -1$. On peut donc utiliser le théorème de convergence dominée pour intervertir limite et intégrale. On

en déduit que la série qui nous intéresse converge vers :

$$\begin{aligned} \frac{z}{2i\pi} \int_0^1 \frac{du}{1+uz} - \frac{\bar{z}}{2i\pi} \int_0^1 \frac{du}{1+u\bar{z}} &= \frac{1}{\pi} \int_0^1 \frac{\sin 2\pi t}{(u + \cos 2\pi t)^2 + \sin^2 2\pi t} du \\ &= \frac{1}{\pi} \left[-\operatorname{arctg} \frac{\sin 2\pi t}{u + \cos 2\pi t} \right]_0^1 = \frac{1}{\pi} (\operatorname{arctg} (\operatorname{tg} 2\pi t) - \operatorname{arctg} (\operatorname{tg} \pi t)) = t. \end{aligned}$$

Ceci démontre le (i). Maintenant, $\sum_{n \neq 0} \frac{1}{4\pi^2 n^2} < +\infty$, et comme les χ_n forment une famille orthonormale, la série $\sum_{n \neq 0} \frac{(-1)^{n-1}}{2i\pi n} \chi_n$ est, d'après le lemme II.2.4, sommable dans $L^2(\mathbf{T})$. On note f sa somme. On peut alors extraire (cf. ex. III.2.9) une sous-suite de la suite de ses sommes partielles tendant vers f p.p. Or le (i) montre que toute sous-suite de ses sommes partielles tend simplement vers ϕ_0 en dehors de $\frac{1}{2}$. On en déduit que $f = \phi_0$ p.p., et donc que la série tend vers ϕ_0 dans $L^2(\mathbf{T})$. Ceci permet de conclure.

Si $f \in L^1(\mathbf{T})$, on note $c_n(f) = \langle \chi_n, f \rangle = \int_0^1 e^{-2i\pi nt} f(t) dt$ son n -ième coefficient de Fourier.

Corollaire IV.3.4. — Si $f \in L^2(\mathbf{T})$, alors $\sum_{n \in \mathbf{Z}} c_n(f) \chi_n$ est sommable, de somme f , dans $L^2(\mathbf{T})$, et $\|f\|_2 = (\sum_{n \in \mathbf{Z}} |c_n(f)|^2)^{1/2}$.

Démonstration. — C'est une simple application du th. II.2.6.

Exercice IV.3.5. — Calculer de deux manières $\|\phi_0\|_2$. En déduire la formule $\zeta(2) = \frac{\pi^2}{6}$.

Proposition IV.3.6. — Si $f \in \mathcal{C}(\mathbf{T})$, et si $\sum_{n \in \mathbf{Z}} |c_n(f)| < +\infty$, alors $\sum_{n \in \mathbf{Z}} c_n(f) \chi_n$ tend uniformément vers f . En particulier, $f(t) = \sum_{n \in \mathbf{Z}} c_n(f) e^{2i\pi nt}$, pour tout t .

Démonstration. — La série $\sum_{n \in \mathbf{Z}} c_n(f) \chi_n$ converge normalement dans $\mathcal{C}(\mathbf{T})$ (muni de $\|\cdot\|_\infty$), et la somme g est donc une fonction continue. De plus, comme $\|h\|_2 \leq \|h\|_\infty$, la série $\sum_{n \in \mathbf{Z}} c_n(f) \chi_n$ converge vers g aussi dans $L^2(\mathbf{T})$. Par ailleurs, il résulte du cor. IV.3.4 que la série converge aussi vers f dans $L^2(\mathbf{T})$, et donc que $f = g$ dans $L^2(\mathbf{T})$. Ceci se traduit par $\int_0^1 |f(t) - g(t)|^2 dt = 0$ et, f et g étant continues, cela implique que $f - g$ est identiquement nulle. Ceci permet de conclure.

Remarque IV.3.7. — La condition $\sum_{n \in \mathbf{Z}} |c_n(f)| < +\infty$ est en particulier vérifiée si f est de classe \mathcal{C}^1 . En effet, si $n \neq 0$, une intégration par partie nous donne

$$c_n(f) = \int_0^1 f(t) e^{-2i\pi nt} dt = \frac{1}{2i\pi n} \int_0^1 f'(t) e^{-2i\pi nt} dt = \frac{1}{2i\pi n} c_n(f'),$$

et comme $(\sum_{n \neq 0} |c_n(f')|^2)^{1/2} \leq \|f'\|_2$, on a

$$\sum_{n \neq 0} \left| \frac{1}{2i\pi n} c_n(f') \right| \leq \left(\sum_{n \neq 0} \frac{1}{4\pi^2 n^2} \right)^{1/2} \left(\sum_{n \neq 0} |c_n(f')|^2 \right)^{1/2} < +\infty.$$

Remarque IV.3.8. — Comme $L^2([a, a+1[)$ est isométrique à $L^2(\mathbf{T})$, les χ_n , pour $n \in \mathbf{N}$, forment aussi une base hilbertienne de $L^2([a, a+1[)$ et donc aussi de $L^2([a, a+1])$ ou $L^2([a, a+1])$ puisque $[a, a+1[,]a, a+1[$ et $[a, a+1]$ ne diffèrent que par des ensembles de mesure nulle. Il est très facile d'en déduire une base hilbertienne de $L^2(I)$, pour tout intervalle I de longueur finie.

2. Séries de Fourier multidimensionnelles

L'étude des séries de Fourier en dimension m se ramène de manière assez formelle⁽⁶⁾ à celle des séries de Fourier en dimension 1.

2.1. Le cas du réseau \mathbf{Z}^m

Une fonction $f : \mathbf{R}^m \rightarrow \mathbf{C}$ est *périodique de période \mathbf{Z}^m* , si $f(x + \omega) = f(x)$, pour tous $x \in \mathbf{R}^m$ et $\omega \in \mathbf{Z}^m$. Pour que ceci soit le cas, il suffit que l'on ait $f(x + e_j) = f(x)$, pour tout $x \in \mathbf{R}^m$, et tout $j \in \{1, \dots, m\}$, où e_1, \dots, e_m est la base canonique de \mathbf{R}^m .

Comme en dimension 1, on voit une fonction périodique de période \mathbf{Z}^m comme une fonction du groupe $\mathbf{T}^m = (\mathbf{R}/\mathbf{Z})^m = \mathbf{R}^m/\mathbf{Z}^m$ dans \mathbf{C} , et l'espace $\mathcal{C}(\mathbf{T}^m)$ des fonctions continues sur \mathbf{T}^m s'identifie à l'espace des fonctions continues sur \mathbf{R}^m , périodiques de période \mathbf{Z}^m .

Si $\mathbf{n} = (n_1, \dots, n_m) \in \mathbf{Z}^m$, on note $\chi_{\mathbf{n}}$ le caractère de \mathbf{T}^m défini par $\chi_{\mathbf{n}}(t) = e^{2i\pi \mathbf{n} \cdot t}$.

Proposition IV.3.9. — *Les caractères linéaires continus de \mathbf{T}^m sont les $\chi_{\mathbf{n}}$, pour $\mathbf{n} \in \mathbf{Z}^m$.*

Démonstration. — Si $\chi : \mathbf{T}^m \rightarrow \mathbf{C}^*$ est un caractère linéaire continu, la restriction de χ à $\mathbf{R}e_j$ est un caractère linéaire continu, périodique de période \mathbf{Z} , et donc de la forme $t_j \mapsto e^{2i\pi n_j t_j}$, d'après la prop. IV.3.1. Comme $t = \sum_{j=1}^m t_j e_j$, on a $\chi(t) = \prod_{j=1}^m e^{2i\pi n_j t_j} = e^{2i\pi \mathbf{n} \cdot t}$, où $\mathbf{n} = (n_1, \dots, n_m)$. Ceci permet de conclure.

Comme en dimension 1, si $a = (a_1, \dots, a_m) \in \mathbf{R}^m$, et si $X_a = \prod_{j=1}^m [a_j, a_j + 1[$, l'application ι_a , qui à une fonction f sur \mathbf{T}^m associe la restriction de f à X_a , est un isomorphisme de l'espace des fonctions sur \mathbf{T}^m sur celui des fonctions sur X_a . Ceci permet de définir, comme en dimension 1 :

- des espaces $L^1(\mathbf{T}^m) \cong L^1(X_a)$ et $L^2(\mathbf{T}^m) \cong L^2(X_a)$,
- une intégrale $f \mapsto \int_{\mathbf{T}^m} f = \int_{X_a} f$ sur $L^1(\mathbf{T}^m)$,
- une norme $\|f\|_1 = \int_{\mathbf{T}^m} |f|$ sur $L^1(\mathbf{T}^m)$,
- un produit scalaire $(f, g) \mapsto \langle f, g \rangle = \int_{\mathbf{T}^m} \bar{f} g$ sur $L^2(\mathbf{T}^m)$, et la norme $\|f\|_2 = \langle f, f \rangle^{1/2}$ qui va avec.

L'invariance de l'intégrale de Lebesgue par translation implique que ce qu'on obtient ne dépend pas du choix de $a \in \mathbf{R}^m$ (cf. lemme IV.3.16 pour un énoncé plus général).

On définit $\text{Esc}(\mathbf{T}^m)$ comme l'image inverse de $\text{Esc}([0, 1]^m)$ par ι_0 ; si $r \in \mathbf{N}$, et si $\mathbf{k} \in \{0, \dots, 2^r - 1\}^m$, on note encore $e_{r, \mathbf{k}}$ la fonction sur \mathbf{T}^m dont l'image par ι_0 est $e_{r, \mathbf{k}}$. Les $e_{r, \mathbf{k}}$, pour $r \in \mathbf{N}$ et $\mathbf{k} \in \{0, \dots, 2^r - 1\}^m$ forment une famille génératrice de $\text{Esc}(\mathbf{T}^m)$.

⁽⁶⁾On peut s'amuser à formaliser le lemme IV.3.11 et son utilisation à grands coups de produits tensoriels et de produits tensoriels complétés. On a vu (ex. I.3.1) que l'espace des fonctions sur $I \times J$ est le produit tensoriel des espaces de fonctions sur I et J , si I et J sont finis. Si I et J ne sont pas finis, la situation est plus compliquée, mais certains sous-espaces de fonctions sur $I \times J$ sont encore des produits tensoriels d'espaces de fonctions sur I et J . C'est par exemple le cas des polynômes trigonométriques sur $\mathbf{R}^m/\mathbf{Z}^m$ qui sont le produit tensoriel de m copies des polynômes trigonométriques sur \mathbf{R}/\mathbf{Z} . L'espace $L^2(\mathbf{R}^m/\mathbf{Z}^m)$ est, quant-à-lui, obtenu en complétant le produit tensoriel de m copies de $L^2(\mathbf{R}/\mathbf{Z})$. Ce procédé de réduction à la dimension est 1 est extrêmement efficace pour beaucoup de questions.

Soit $\text{Trig}(\mathbf{T}^m)$ l'espace des polynômes trigonométriques sur \mathbf{T}^m (i.e. des combinaisons linéaires des $\chi_{\mathbf{n}}$, pour $\mathbf{n} \in \mathbf{Z}^m$).

Si $f \in L^1(\mathbf{T}^m)$, on définit ses *coefficients de Fourier* $c_{\mathbf{n}}(f)$, pour $\mathbf{n} \in \mathbf{Z}^m$, par la formule $c_{\mathbf{n}}(f) = \langle \chi_{\mathbf{n}}, f \rangle = \int_{[0,1]^m} e^{-2i\pi \mathbf{n} \cdot t} f(t) dt$.

Théorème IV.3.10. — (i) $\text{Trig}(\mathbf{T}^m)$ est dense dans $L^2(\mathbf{T}^m)$.

(ii) Les $\chi_{\mathbf{n}}$, pour $\mathbf{n} \in \mathbf{Z}^m$ forment une base hilbertienne de $L^2(\mathbf{T}^m)$.

(iii) Si $f \in L^2(\mathbf{T}^m)$, alors $f = \sum_{\mathbf{n} \in \mathbf{Z}^m} c_{\mathbf{n}}(f) \chi_{\mathbf{n}}$ dans $L^2(\mathbf{T}^m)$.

(iv) Si $f \in \mathcal{C}(\mathbf{T}^m)$, et si $\sum_{\mathbf{n} \in \mathbf{Z}^m} |c_{\mathbf{n}}(f)| < +\infty$, alors $f = \sum_{\mathbf{n} \in \mathbf{Z}^m} c_{\mathbf{n}}(f) \chi_{\mathbf{n}}$ dans $\mathcal{C}(\mathbf{T}^m)$, et en particulier, $f(t) = \sum_{\mathbf{n} \in \mathbf{Z}^m} c_{\mathbf{n}}(f) e^{2i\pi \mathbf{n} \cdot t}$ pour tout t .

Démonstration. — Nous allons déduire cet énoncé de l'énoncé analogue en dimension 1. Si ϕ_1, \dots, ϕ_m sont des fonctions de \mathbf{R} dans \mathbf{C} , on note $\phi_1 \otimes \dots \otimes \phi_m$ ou, de manière plus compacte, $\otimes_i \phi_i$, la fonction de \mathbf{R}^m dans \mathbf{C} définie par

$$(\otimes_i \phi_i)(t) = \phi_1 \otimes \dots \otimes \phi_m(t) = \prod_{i=1}^m \phi_i(t_i), \quad \text{si } t = (t_1, \dots, t_m).$$

Par exemple,

$$\begin{aligned} \chi_{\mathbf{n}} &= \otimes_i \chi_{n_i}, & \text{si } \mathbf{n} = (n_1, \dots, n_m) \in \mathbf{Z}^m, \\ e_{r, \mathbf{k}} &= \otimes_i e_{r, k_i}, & \text{si } r \in \mathbf{N}, \text{ et si } \mathbf{k} = (k_1, \dots, k_m) \in \mathbf{Z}^m. \end{aligned}$$

Lemme IV.3.11. — (i) Si $\phi_i = \psi_i$ p.p., pour tout i , alors $\otimes_i \phi_i = \otimes_i \psi_i$ p.p.

(ii) Si $\phi_i \in \mathcal{L}^1(\mathbf{T})$ pour tout i , alors $\otimes_i \phi_i \in \mathcal{L}^1(\mathbf{T}^m)$, et on a $\|\otimes_i \phi_i\|_1 = \prod_i \|\phi_i\|_1$.

(iii) Si $\phi_i \in \mathcal{L}^2(\mathbf{T})$ pour tout i , alors $\otimes_i \phi_i \in \mathcal{L}^2(\mathbf{T}^m)$, et on a $\|\otimes_i \phi_i\|_2 = \prod_i \|\phi_i\|_2$.

Démonstration. — Soit $A_i \subset \mathbf{R}$ (resp. $A \subset \mathbf{R}^m$) l'ensemble des x tels que $\phi_i(x) \neq \psi_i(x)$ (resp. $\otimes_i \phi_i(x) \neq \otimes_i \psi_i(x)$). Alors A est inclus dans la réunion des $A_i \times \prod_{j \neq i} \mathbf{R}$, qui sont tous de mesure nulle dans \mathbf{R}^m , puisque A_i est de mesure nulle dans \mathbf{R} par hypothèse. Cela démontre le (i).

Les (ii) et (iii) sont des conséquences immédiates du théorème de Fubini.

Revenons à la démonstration du th. IV.3.10. Le (i) du lemme précédent montre que l'application $(\phi_1, \dots, \phi_m) \mapsto \otimes_i \phi_i$ passe au quotient (des deux côtés à la fois) modulo les fonctions nulles p.p. Comme cette application est linéaire en chacune des ϕ_i , les (ii) et (iii) montre que l'on obtient ainsi des applications multilinéaires continues $L^1(\mathbf{T})^m \rightarrow L^1(\mathbf{T}^m)$ et $L^2(\mathbf{T})^m \rightarrow L^2(\mathbf{T}^m)$.

Pour montrer que $\text{Trig}(\mathbf{T}^m)$ est dense dans $L^2(\mathbf{T}^m)$, il suffit de montrer que son adhérence contient $\text{Esc}(\mathbf{T}^m)$, et, par linéarité, il suffit de vérifier qu'elle contient les $e_{r, \mathbf{k}}$, pour $r \geq 1$ et $\mathbf{k} \in \{0, \dots, 2^r - 1\}^m$. Pour cela, on écrit $e_{r, \mathbf{k}}$ sous la forme $e_{r, \mathbf{k}} = \otimes_i e_{r, k_i}$, et on choisit, pour chaque i , une suite $P_{i, n}$ d'éléments de $\text{Trig}(\mathbf{T})$ tendant vers e_{r, k_i} dans $L^2(\mathbf{T})$. Il résulte de la continuité de $(\phi_1, \dots, \phi_m) \mapsto \otimes_i \phi_i$ que $P_n = \otimes_i P_{i, n}$ tend vers $e_{r, \mathbf{k}}$ dans $L^2(\mathbf{T}^m)$, et comme $P_n \in \text{Trig}(\mathbf{T}^m)$, on en déduit le (i).

On déduit du théorème de Fubini que $\langle \chi_{\mathbf{n}}, \chi_{\mathbf{\ell}} \rangle = \prod_{i=1}^m \langle \chi_{n_i}, \chi_{\ell_i} \rangle$, si $\mathbf{n} = (n_1, \dots, n_m)$ et $\mathbf{\ell} = (\ell_1, \dots, \ell_m)$, ce qui permet de déduire l'orthonormalité de la famille des $\chi_{\mathbf{n}}$, pour $\mathbf{n} \in \mathbf{Z}^m$, de celle des χ_n , pour $n \in \mathbf{Z}$; le (ii) est donc une conséquence du (i).

Le (iii) est alors une application du th. II.2.6, et le (iv) se déduit du (iii) comme dans la démonstration de la prop. IV.3.6.

Ceci permet de conclure.

Exercice IV.3.12. — Montrer que, si $k > \frac{m}{2}$, et si f est périodique de période \mathbf{Z}^m et de classe \mathcal{C}^k , alors $f(t) = \sum_{\mathbf{n} \in \mathbf{Z}^m} c_{\mathbf{n}}(f) e^{2i\pi \mathbf{n} \cdot t}$, pour tout t .

2.2. Le cas d'un réseau quelconque

Un réseau Λ de \mathbf{R}^m est un sous-groupe de \mathbf{R}^m de la forme $\mathbf{Z}v_1 + \cdots + \mathbf{Z}v_m$, où (v_1, \dots, v_m) est une base de \mathbf{R}^m . On dit alors que (v_1, \dots, v_m) est une base de Λ (sous-entendu sur \mathbf{Z}). L'exemple le plus simple est le réseau \mathbf{Z}^m dont une base est la base canonique de \mathbf{R}^m .

Le réseau dual Λ^* de Λ , est l'ensemble des $x \in \mathbf{R}^m$ tels que $x \cdot \omega \in \mathbf{Z}$, quel que soit $\omega \in \Lambda$. Le réseau dual de \mathbf{Z}^m est \mathbf{Z}^m de manière évidente. Le cas général est décrit par le lemme suivant.

Lemme IV.3.13. — Soit (v_1, \dots, v_m) une base de Λ , soit A la matrice dont la j -ième colonne est le vecteur v_j . Alors $\Lambda^* = {}^t A^{-1} \mathbf{Z}^m$.

Démonstration. — Le fait que (v_1, \dots, v_m) soit une base de Λ se traduit par $\Lambda = A \mathbf{Z}^m = \{A\mathbf{n}, \mathbf{n} \in \mathbf{Z}^m\}$. Or $x \cdot A\mathbf{n} = {}^t A x \cdot \mathbf{n}$. Comme le réseau dual de \mathbf{Z}^m est \mathbf{Z}^m , on en déduit que $x \in \Lambda^*$, si et seulement si ${}^t A x \in \mathbf{Z}^m$; autrement dit, on a $\Lambda^* = {}^t A^{-1} \mathbf{Z}^m$.

Une fonction $f : \mathbf{R}^m \rightarrow \mathbf{C}$ est *périodique de période Λ* , si $f(x + \omega) = f(x)$, pour tous $x \in \mathbf{R}^m$ et $\omega \in \Lambda$. Pour que ceci soit le cas, il suffit que l'on ait $f(x + v_j) = f(x)$, pour tout $x \in \mathbf{R}^m$, et tout $j \in \{1, \dots, m\}$, si (v_1, \dots, v_m) est une base Λ .

Comme d'habitude, on voit une fonction périodique de période Λ comme une fonction du groupe $\mathbf{T}(\Lambda) = \mathbf{R}^m / \Lambda$ dans \mathbf{C} , et l'espace $\mathcal{C}(\mathbf{T}(\Lambda))$ des fonctions continues sur $\mathbf{T}(\Lambda)$ s'identifie à l'espace des fonctions continues sur \mathbf{R}^m , périodiques de période Λ .

Si $\omega \in \Lambda^*$, alors $t \mapsto e^{2i\pi \omega \cdot t}$ est un caractère de \mathbf{R}^m dont le noyau contient Λ ; c'est donc un caractère de $\mathbf{T}(\Lambda)$; nous le noterons χ_ω .

Lemme IV.3.14. — Les caractères linéaires continus de $\mathbf{T}(\Lambda)$ sont les χ_ω , pour $\omega \in \Lambda^*$.

Démonstration. — Un caractère de $\mathbf{T}(\Lambda)$ est de la forme $e^{2i\pi x \cdot t}$, avec $x \in \mathbf{R}^m$, trivial sur Λ , ce qui équivaut à $x \cdot \omega \in \mathbf{Z}$, pour tout $\omega \in \Lambda$. On en déduit le résultat.

Si Λ est un réseau de \mathbf{R}^m , un *domaine fondamental* de \mathbf{R}^m modulo Λ , est un ensemble mesurable $X \subset \mathbf{R}^m$ tel que tout élément x de \mathbf{R}^m puisse s'écrire de manière unique sous la forme $\omega + a$, avec $a \in X$ et $\omega \in \Lambda$. Par exemple, si (v_1, \dots, v_m) est une base Λ , alors $\{t_1 v_1 + \cdots + t_m v_m, 0 \leq t_i < 1, \text{ pour tout } i\}$ est un domaine fondamental modulo Λ .

Soit X un domaine fondamental de \mathbf{R}^m modulo Λ . On note $\text{Vol}(\Lambda)$ la mesure de Lebesgue de X . Le lemme IV.3.16 ci-dessous montre que ceci ne dépend pas du choix de X ; on a donc $\text{Vol}(\Lambda) = |\det(v_1, \dots, v_m)|$, si (v_1, \dots, v_m) est une base de Λ ; en particulier, $\text{Vol}(\mathbf{Z}^m) = 1$.

L'application ι_X , qui à une fonction f sur $\mathbf{T}(\Lambda)$ associe la restriction de f à X , est un isomorphisme de l'espace des fonctions sur $\mathbf{T}(\Lambda)$ sur celui des fonctions sur X . Ceci permet de définir, comme d'habitude,

- des espaces $L^1(\mathbf{T}(\Lambda)) \cong L^1(X)$ et $L^2(\mathbf{T}(\Lambda)) \cong L^2(X)$,
- une intégrale $f \mapsto \int_{\mathbf{T}(\Lambda)} f = \int_X f$ sur $L^1(\mathbf{T}(\Lambda))$,
- une norme $\|f\|_1 = \frac{1}{\text{Vol}(\Lambda)} \int_{\mathbf{T}(\Lambda)} |f|$ sur $L^1(\mathbf{T}(\Lambda))$,
- un produit scalaire $(f, g) \mapsto \langle f, g \rangle = \frac{1}{\text{Vol}(\Lambda)} \int_{\mathbf{T}(\Lambda)} \bar{f} g$ sur $L^2(\mathbf{T}(\Lambda))$, et la norme $\|f\|_2 = \langle f, f \rangle^{1/2}$ qui va avec.

Le lemme IV.3.16 ci-dessous montre que ce qu'on obtient ne dépend pas du choix de X .

Si $f \in L^1(\mathbf{T}(\Lambda))$, on définit ses *coefficients de Fourier* $c_\omega(f)$, pour $\omega \in \Lambda^*$, par la formule $c_\omega(f) = \langle \chi_\omega, f \rangle = \frac{1}{\text{Vol}(\Lambda)} \int_X e^{-2i\pi \omega \cdot t} f(t) dt$, où X est un domaine fondamental modulo Λ .

Théorème IV.3.15. — (i) Les χ_ω , pour $\omega \in \Lambda^*$, forment une base hilbertienne de $L^2(\mathbf{T}(\Lambda))$.

(ii) Si $f \in L^2(\mathbf{T}(\Lambda))$, alors $f = \sum_{\omega \in \Lambda^*} c_\omega(f) \chi_\omega$ dans $L^2(\mathbf{T}(\Lambda))$.

(iii) Si $f \in \mathcal{C}(\mathbf{T}(\Lambda))$, et si $\sum_{\omega \in \Lambda^*} |c_\omega(f)| < +\infty$, alors $f = \sum_{\omega \in \Lambda^*} c_\omega(f) \chi_\omega$ dans $\mathcal{C}(\mathbf{T}(\Lambda))$; en particulier, $f(t) = \sum_{\omega \in \Lambda^*} c_\omega(f) e^{2i\pi \omega \cdot t}$ pour tout t .

Démonstration. — Soit (v_1, \dots, v_d) une base de Λ , et soit A la matrice dont la j -ième colonne est v_j . Alors $A \cdot \mathbf{Z}^m = \Lambda$, ce qui fait que $\phi \mapsto \phi \circ A$ transforme une fonction périodique de période Λ en une fonction périodique de période \mathbf{Z}^m . De plus, le facteur $\frac{1}{\text{Vol}(\Lambda)} = \frac{1}{|\det A|}$ dans la définition du produit scalaire sur $L^2(\mathbf{T}(\Lambda))$ fait que $\phi \mapsto \phi \circ A$ est une isométrie de $L^2(\mathbf{T}(\Lambda))$ sur $L^2(\mathbf{T}^m)$. Ceci permet, en remarquant que $\chi_\omega(At) = \chi_{\mathbf{n}}(t)$ si $\omega = {}^t A^{-1} \mathbf{n} \in \Lambda^*$ (cf. lemme IV.3.13), de déduire le théorème ci-dessus du résultat pour \mathbf{Z}^m .

Lemme IV.3.16. — Si $f : \mathbf{T}(\Lambda) \rightarrow \mathbf{R}_+$ est mesurable, alors $\int_X f$ ne dépend pas du choix du domaine fondamental X modulo Λ .

Démonstration. — Le fait pour X d'être un domaine fondamental modulo Λ peut se réécrire sous la forme $\sum_{\omega \in \Lambda} \mathbf{1}_{\omega+X} = 1$, où $\omega + X = \{\omega + a, a \in X\}$.

Soient X_1, X_2 des domaines fondamentaux modulo Λ . En utilisant successivement :

- l'identité $1 = \sum_{\omega \in \Lambda} \mathbf{1}_{\omega+X_2}$,
- le théorème de convergence monotone pour échanger somme et intégrale,
- le changement de variable $x = \omega + t$ et l'invariance de f par ce changement de variable,
- de nouveau le théorème de convergence monotone,
- l'identité $\sum_{\omega \in \Lambda} \mathbf{1}_{-\omega+X_1} = 1$,

on obtient

$$\begin{aligned} \int_{X_1} f &= \int \mathbf{1}_{X_1} f = \int \left(\sum_{\omega \in \Lambda} \mathbf{1}_{\omega+X_2} \right) \mathbf{1}_{X_1} f = \sum_{\omega \in \Lambda} \int \mathbf{1}_{\omega+X_2} \mathbf{1}_{X_1} f \\ &= \sum_{\omega \in \Lambda} \int \mathbf{1}_{X_2} \mathbf{1}_{-\omega+X_1} f = \int \left(\sum_{\omega \in \Lambda} \mathbf{1}_{-\omega+X_1} \right) \mathbf{1}_{X_2} f = \int \mathbf{1}_{X_2} f = \int_{X_2} f, \end{aligned}$$

ce qui permet de conclure.

3. La formule de Poisson

On note $\mathcal{S}(\mathbf{R}^m)$ l'espace de Schwartz des fonctions de classe \mathcal{C}^∞ sur \mathbf{R}^m , qui sont à décroissance rapide à l'infini ainsi que toutes leur dérivées. (g est à décroissance rapide à l'infini, si $(1+\|t\|^2)^N g(t)$ est bornée sur \mathbf{R}^m , quel que soit $N \in \mathbf{N}$.) On déduit du th. IV.2.8 le résultat suivant.

Corollaire IV.3.17. — L'image de $\mathcal{S}(\mathbf{R}^m)$ par la transformée de Fourier est incluse dans $\mathcal{S}(\mathbf{R}^m)$.

Théorème IV.3.18. — (Formule de Poisson) Si $f \in \mathcal{S}(\mathbf{R})$ ou, plus généralement, si f est de classe \mathcal{C}^1 , et f et f' sont des $O(|t|^{-2})$ au voisinage de $\pm\infty$, alors

$$\sum_{n \in \mathbf{Z}} f(n) = \sum_{n \in \mathbf{Z}} \hat{f}(n).$$

Démonstration. — Soit $F(t) = \sum_{n \in \mathbf{Z}} f(n+t)$. La série converge pour tout t grâce à l'hypothèse $f = O(|t|^{-2})$, et la fonction F est périodique de période 1. De plus, sur $[0, 1]$, la série des $f'(n+t)$ converge normalement, grâce à l'hypothèse $f' = O(|t|^{-2})$, ce qui

implique que $F \in \mathcal{C}^1(\mathbf{T})$. On en déduit, en utilisant la rem. IV.3.7, que F est somme de sa série de Fourier en tout point. Or, si $k \in \mathbf{Z}$, on a (l'interversion de l'intégrale et de la série ci-dessous est justifiée par la convergence uniforme sur $[0, 1]$) :

$$\begin{aligned} c_k(F) &= \int_0^1 e^{-2i\pi kt} F(t) dt = \int_0^1 e^{-2i\pi kt} \left(\sum_{n \in \mathbf{Z}} f(n+t) \right) dt = \sum_{n \in \mathbf{Z}} \int_0^1 e^{-2i\pi kt} f(n+t) dt \\ &= \sum_{n \in \mathbf{Z}} \int_n^{n+1} e^{-2i\pi kt} f(t) dt = \int_{-\infty}^{+\infty} e^{-2i\pi kt} f(t) dt = \hat{f}(k). \end{aligned}$$

On en déduit la formule du théorème en comparant la série donnant $F(0)$ avec la série de Fourier de F en 0.

On démontre de même, en dimension quelconque, le résultat suivant.

Théorème IV.3.19. — (Formule de Poisson dans \mathbf{R}^m) Si $f \in \mathcal{S}(\mathbf{R}^m)$, et si Λ est un réseau de \mathbf{R}^m , alors

$$\sum_{\omega \in \Lambda} f(\omega) = \frac{1}{\text{Vol}(\Lambda)} \sum_{\omega \in \Lambda^*} \hat{f}(\omega).$$

Exercice IV.3.20. — Comparer ce que donnent la formule du th. IV.3.18 et celle du th. IV.3.19 pour évaluer $\sum_{n \in \mathbf{Z}} f(\lambda n)$, avec $\lambda \in \mathbf{R}^*$.

4. La formule d'inversion de Fourier dans \mathcal{S}

Théorème IV.3.21. — Si $\varphi \in \mathcal{S}(\mathbf{R}^m)$, alors $\overline{\mathcal{F}} \mathcal{F} \varphi = \varphi$ et $\mathcal{F} \overline{\mathcal{F}} \varphi = \varphi$.

Démonstration. — Comme on passe de \mathcal{F} à $\overline{\mathcal{F}}$ en changeant x en $-x$, il suffit de démontrer une des deux formules ; nous démontrerons la première. Soient $u \in \mathbf{R}^m$ et $r \in \mathbf{N}$. Si on applique la formule de Poisson à la fonction $f(t) = \varphi(u + 2^r t)$ et au réseau $\Lambda = \mathbf{Z}^m$ (et donc $\Lambda^* = \mathbf{Z}^m$), on obtient la formule suivante (où l'on a utilisé l'identité $\hat{f}(x) = 2^{-rm} e^{2i\pi u \cdot 2^{-r} x} \hat{\varphi}(2^{-r} x)$, conséquence des formules pour les dilations-translations) :

$$\sum_{\mathbf{k} \in \mathbf{Z}^m} \varphi(u + 2^r \mathbf{k}) = 2^{-rm} \sum_{\mathbf{k} \in \mathbf{Z}^m} 2^{-rm} e^{2i\pi u \cdot 2^{-r} \mathbf{k}} \hat{\varphi}(2^{-r} \mathbf{k}).$$

Nous allons montrer que, quand $r \rightarrow +\infty$, le membre de gauche tend vers $\varphi(u)$ et le membre de droite vers $\overline{\mathcal{F}} \hat{\varphi}(u)$, ce qui permettra de conclure.

• On commence par remarquer, que comme φ tend vers 0 à l'infini, on a $\varphi(u + 2^r \mathbf{k}) \rightarrow 0$ si $\mathbf{k} \neq 0$, et $\varphi(u + 2^r \mathbf{k}) = \varphi(u)$ si $\mathbf{k} = 0$; pour prouver que le membre de gauche tend vers $\varphi(u)$, il n'y a donc qu'à justifier l'interversion de la somme et de la limite. La décroissance rapide de φ à l'infini implique en particulier l'existence de C tel que $|\varphi(t)| \leq C(1 + \|t\|^2)^{-(m+1)/2}$ pour tout $t \in \mathbf{R}^m$. Or $\sum_{\mathbf{k} \in \mathbf{Z}^m} (1 + a\|\mathbf{k}\|^2)^{-(m+1)/2} < +\infty$ pour tout $a > 0$, et, si $2^r \geq 2\|u\|$, on a $\|u + 2^r \mathbf{k}\| \geq a\|\mathbf{k}\|$, avec $a = \|u\|$, pour tout $\mathbf{k} \in \mathbf{Z}^m$. La série de terme général $C(1 + a\|\mathbf{k}\|^2)^{-(m+1)/2}$ est donc un majorant sommable, pour tout r

assez grand, de la série de terme général $\varphi(u + 2^r \mathbf{k})$. On conclut en utilisant le théorème de convergence dominée pour les séries.

• Passons à l'étude de la série dans le membre de droite. Le procédé habituel transformant une somme en intégrale d'une fonction en escalier montre que cette série est égale à $\int \psi_r$, où ψ_r est la fonction définie par

$$\psi_r = \sum_{\mathbf{k} \in \mathbf{Z}^m} \psi(2^{-r} \mathbf{k}) e_{r, \mathbf{k}}, \quad \text{et } \psi(x) = e^{2i\pi u \cdot x} \hat{\varphi}(x).$$

Maintenant, si $t = (t_1, \dots, t_m)$, alors $\psi_r(t) = \psi(t^{(r)})$, où $t_i^{(r)} = 2^{-r} [2^r t_i] \in [t_i - 2^{-r}, t_i]$. En particulier, $t^{(r)} \rightarrow t$, et ψ étant continue, on a $\psi_r(t) \rightarrow \psi(t)$, pour tout $t \in \mathbf{R}^m$. De plus, si $C(t)$ désigne le cube $\prod_{i=1}^m [t_i - 1, t_i]$, on a $t^{(r)} \in C(t)$, pour tout $r \in \mathbf{N}$. La fonction ψ_r est donc majorée, pour tout $r \in \mathbf{N}$, par g , où $g(t) = \sup_{u \in C(t)} |\psi(t)| = \sup_{u \in C(t)} |\hat{\varphi}(t)|$. Enfin, comme $\hat{\varphi}$ est à décroissance rapide à l'infini, il en est de même de g qui est, de ce fait, sommable, ce qui permet d'utiliser le théorème de convergence dominée pour en déduire que $\int \psi_r \rightarrow \int \psi$. Comme $\int \psi = \overline{\mathcal{F}} \hat{\varphi}(u)$, cela permet de conclure.

Corollaire IV.3.22. — \mathcal{F} et $\overline{\mathcal{F}}$ sont des isomorphismes de $\mathcal{S}(\mathbf{R}^m)$ dans $\mathcal{S}(\mathbf{R}^m)$ inverses⁽⁷⁾ l'un de l'autre.

5. Formules d'inversion dans L^1

Proposition IV.3.23. — Si $f, g \in L^1(\mathbf{R}^m)$, alors

$$\int_{\mathbf{R}^m} g \mathcal{F} f = \int_{\mathbf{R}^m} f \mathcal{F} g \quad \text{et} \quad \int_{\mathbf{R}^m} g \overline{\mathcal{F}} f = \int_{\mathbf{R}^m} f \overline{\mathcal{F}} g.$$

Démonstration. — Remarquons que les deux membres sont bien définis car \hat{f} et \hat{g} sont bornées puisqu'éléments de $\mathcal{C}_0(\mathbf{R}^m)$. Soit $h(x, t) = g(x) f(t) e^{-2i\pi x \cdot t}$. Alors h est sommable sur $\mathbf{R}^m \times \mathbf{R}^m$, car d'après le théorème de Fubini pour les fonctions positives, on a

$$\begin{aligned} \int_{\mathbf{R}^m \times \mathbf{R}^m} |h(x, t)| dx dt &= \int_{\mathbf{R}^m \times \mathbf{R}^m} |f(t)| |g(x)| dx dt = \int_{\mathbf{R}^m} \left(\int_{\mathbf{R}^m} |f(t)| |g(x)| dx \right) dt \\ &= \int_{\mathbf{R}^m} \|g\|_1 |f(t)| dt = \|g\|_1 \|f\|_1 < +\infty. \end{aligned}$$

Une application immédiate du théorème de Fubini montre alors que les deux membres sont égaux à $\int_{\mathbf{R}^m \times \mathbf{R}^m} h(x, t) dx dt$, ce qui permet de conclure.

Proposition IV.3.24. — Si $h \in L^1(\mathbf{R}^m)$ a une transformée de Fourier sommable, alors $\overline{\mathcal{F}} \mathcal{F} h = h$ p.p.

⁽⁷⁾Ce résultat, combiné avec la formule de la prop. IV.3.23, est à la base de la définition de la transformée de Fourier d'une distribution (cf. cours de F. Golse en seconde année).

Démonstration. — Soit $\varphi \in \mathcal{C}_c^\infty(\mathbf{R}^m)$. Alors $\mathcal{F}\overline{\mathcal{F}}\varphi = \varphi$ d'après le th. IV.3.21. Par ailleurs, en appliquant la prop. IV.3.23 successivement à $f = \mathcal{F}h$ et $g = \varphi$, puis à $f = h$ et $g = \overline{\mathcal{F}}\varphi$ (ce qui est licite car h et $\mathcal{F}h$ sont dans $L^1(\mathbf{R}^m)$ par hypothèse, et φ et $\overline{\mathcal{F}}\varphi$ sont dans $\mathcal{S}(\mathbf{R}^m)$ qui est inclus dans $L^1(\mathbf{R}^m)$), on obtient :

$$\int_{\mathbf{R}^m} \overline{\mathcal{F}}\mathcal{F}h \varphi = \int_{\mathbf{R}^m} \mathcal{F}h \overline{\mathcal{F}}\varphi = \int_{\mathbf{R}^m} h \mathcal{F}\overline{\mathcal{F}}\varphi = \int_{\mathbf{R}^m} h \varphi.$$

La fonction $\overline{\mathcal{F}}\mathcal{F}h$ n'est pas a priori dans $L^1(\mathbf{R}^m)$ mais, comme elle est continue, sa restriction à tout ouvert borné X est dans $L^1(X)$. Or $\int_X (\overline{\mathcal{F}}\mathcal{F}h - h) \varphi = 0$ pour tout $\varphi \in \mathcal{C}_c^\infty(X)$, d'après ce qui précède, et le cor. III.2.13 permet d'en déduire la nullité (presque partout) de $\overline{\mathcal{F}}\mathcal{F}h - h$, sur tout ouvert borné X . Ceci permet de conclure.

On trouvera une autre démonstration sous forme d'exercice (utilisant la convolution introduite dans les ex. III.3.11 et III.3.12) ci-dessous, et encore une autre en passant par la transformée de Fourier dans L^2 au n° 3 du § IV.4.

Exercice IV.3.25. — (Formule d'inversion)

- 1) Soit $h : \mathbf{R} \rightarrow \mathbf{R}$ définie par $h(t) = e^{-|t|}$, et, si $\varepsilon > 0$, soit $h_\varepsilon(t) = h(\varepsilon t)$. Soit $\delta(x) = \hat{h}(-x)$.
 - a) Calculer $\delta(x)$ et vérifier que $\int_{\mathbf{R}} \delta = 1$ et que $\hat{h}_\varepsilon(-x) = \delta_\varepsilon(x)$, avec $\delta_\varepsilon(x) = \frac{1}{\varepsilon} \delta(x/\varepsilon)$.
 - b) Soit $f \in L^1(\mathbf{R})$. Montrer que $\|f * \delta_\varepsilon - f\|_{L^1} \rightarrow_{\varepsilon \rightarrow 0} 0$. (Commencer par une fonction en escalier.)
 - c) En déduire qu'il existe une suite $\varepsilon_n \rightarrow_{n \rightarrow +\infty} 0$, telle que $f * \delta_{\varepsilon_n}(x) \rightarrow_{n \rightarrow +\infty} f(x)$ pour presque tout $x \in \mathbf{R}$.
 - d) Montrer que $(f * \delta_\varepsilon)(x) = \int_{\mathbf{R}} h(\varepsilon t) \hat{f}(t) e^{2i\pi tx} dt$.
- 2) On suppose de plus que $\hat{f} \in L^1(\mathbf{R})$. On pose $g(x) = \int_{\mathbf{R}} \hat{f}(t) e^{2i\pi tx} dt$. Vérifier que g est continue, bornée, que $\int_{\mathbf{R}} h(\varepsilon t) \hat{f}(t) e^{2i\pi tx} dt \rightarrow g(x)$, quand $\varepsilon \rightarrow 0$, quel que soit $x \in \mathbf{R}$. En déduire que $f(x) = g(x)$, pour presque tout $x \in \mathbf{R}$.

6. Exercices

Exercice IV.3.26. — Soit $f : \mathbf{R} \rightarrow \mathbf{C}$ définie par $f(t) = \frac{1}{(t+i)^3}$.

- (i) Montrer que \hat{f} est bien définie, est de classe \mathcal{C}^1 , et que $|t^N \hat{f}(t)| \rightarrow 0$ quand $|t| \rightarrow +\infty$, pour tout $N \in \mathbf{N}$.
- (ii) Soit $g : \mathbf{R} \rightarrow \mathbf{C}$ définie par $g(t) = e^{-2\pi t}$, si $t > 0$, et $g(t) = 0$, si $t \leq 0$. Calculer \hat{g} ; en déduire la transformée de Fourier de $h(t) = t^2 g(t)$, puis \hat{f} .
- (iii) Montrer que la série $\sum_{n \in \mathbf{Z}} \frac{1}{(n+i)^3}$ est absolument convergente, et calculer sa somme.

Exercice IV.3.27. — (i) Montrer que, si $\phi \in \mathcal{C}_c^\infty$, l'équation différentielle $u'' - u = \phi$ a une unique solution dans $\mathcal{S}(\mathbf{R})$. Cette solution est-elle toujours à support compact ?

(ii) À quelle condition portant sur $\hat{\phi}$, l'équation différentielle $u'' + u = \phi$ a-t-elle une solution dans $\mathcal{S}(\mathbf{R})$? Si solution il y a, est-elle toujours à support compact ?

Exercice IV.3.28. — Soit $f : \mathbf{R} \rightarrow \mathbf{R}$ définie par $f(t) = e^{-\pi t^2}$.

- (i) Montrer que \hat{f} est \mathcal{C}^∞ sur \mathbf{R} , et vérifie l'équation différentielle $\hat{f}'(x) = -2\pi x \hat{f}(x)$. En déduire, en utilisant la formule $\int_{\mathbf{R}} e^{-\pi t^2} dt = 1$, que $\hat{f}(x) = e^{-\pi x^2}$.
- (ii) Si $u \in \mathbf{R}_+^*$, calculer la transformée de Fourier de f_u définie par $f_u(t) = e^{-\pi u t^2}$.
- (iii) Si $u \in \mathbf{R}_+^*$, et si $F(u) = \sum_{n \in \mathbf{Z}} e^{-\pi n^2 u}$, montrer que $F(u) = \frac{1}{\sqrt{u}} F(\frac{1}{u})$.

Exercice IV.3.29. — Si $\lambda > 0$, soit $\phi_\lambda(t) = e^{-\pi\lambda|t|} \frac{\sin \pi t}{\pi t}$.

(i) Montrer que, si $x \in \mathbf{R}$, alors $\lambda \mapsto \hat{\phi}_\lambda(x)$ est dérivable sur \mathbf{R}_+^* , et calculer sa dérivée. En déduire $\hat{\phi}_\lambda(x)$.

(ii) Remarquer que $\frac{\sin \pi t}{\pi t}$ est la transformée de Fourier de $\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}[}$, et retrouver directement de le (i).

Exercice IV.3.30. — (i) Montrer que $(\frac{d}{dx})^n e^{-\pi x^2} = e^{-\pi x^2} H_n(x)$, où H_n est un polynôme de degré exactement n et dont on calculera le coefficient du terme de plus haut degré.

(ii) On définit une fonction Ψ_n par la formule $\Psi_n(x) = e^{-\frac{\pi x^2}{2}} H_n(x)$. Calculer le produit scalaire $\langle \Psi_n, \Psi_m \rangle_{L^2}$.

(iii) Soit $f \in L^2(\mathbf{R})$ orthogonale aux Ψ_n . Montrer que g définie par $g(x) = e^{-\frac{\pi x^2}{2}} f(x)$ est sommable et que \hat{g} est identiquement nulle. En déduire que les Ψ_n engendrent un sous-espace dense de $L^2(\mathbf{R})$.

(iv) En utilisant le fait que la transformée de Fourier de $e^{-\pi t^2}$ est $e^{-\pi x^2}$, calculer, par récurrence, la transformée de Fourier de Ψ_n . En déduire une construction de la transformée de Fourier dans $L^2(\mathbf{R})$, et démontrer la formule d'inversion de Fourier dans $L^2(\mathbf{R})$.

IV.4. Transformée de Fourier dans L^2

Si $\phi \in L^2(\mathbf{R}^m)$ n'est pas sommable, l'intégrale $\int_{\mathbf{R}^m} e^{-2i\pi xt} \phi(t) dt$ ne converge pour aucune valeur de x . Malgré ce petit problème, on peut définir la transformée de Fourier d'un élément de $L^2(\mathbf{R}^m)$, par un passage à la limite, et ce qu'on obtient fournit une théorie ayant des propriétés nettement plus agréables que dans $L^1(\mathbf{R}^m)$. Il y a des tas de manières d'arriver au résultat. Nous avons choisi de privilégier les fonctions en escalier jusqu'au bout. On trouvera d'autres approches dans les exercices.

1. Transformée de Fourier des fonctions en escalier

Commençons par quelques calculs en dimension 1.

Lemme IV.4.1. — (i) Si $\lambda > 0$, la transformée de Fourier de $e^{-\pi\lambda|t|} \frac{\sin \pi t}{\pi t}$ est

$$\frac{1}{\pi} \left(\operatorname{arctg}\left(\frac{2x+1}{\lambda}\right) - \operatorname{arctg}\left(\frac{2x-1}{\lambda}\right) \right).$$

(ii) La transformée de Fourier de $\frac{\sin^2 \pi t}{(\pi t)^2}$ est $\frac{1}{2}(|x+1| + |x-1| - 2|x|)$.

Démonstration. — (i) On cherche à calculer $\int_{\mathbf{R}} e^{-2i\pi xt} e^{-\pi\lambda|t|} \frac{\sin \pi t}{\pi t} dt$. Or $\frac{\sin \pi t}{\pi t}$ est la transformée de Fourier de $\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}[}$, alors que la transformée de Fourier de $e^{-2i\pi xt} e^{-\pi\lambda|t|}$ est

$$\begin{aligned} y \mapsto \int_{\mathbf{R}} e^{-2i\pi(x+y)t} e^{-\pi\lambda|t|} dt &= \int_0^{+\infty} e^{-\pi t(\lambda+2i(x+y))} + e^{-\pi t(\lambda-2i(x+y))} dt \\ &= \frac{1}{\pi} \left(\frac{1}{\lambda+2i(x+y)} + \frac{1}{\lambda-2i(x+y)} \right) = \frac{2\lambda}{\pi(\lambda^2+4(x+y)^2)}. \end{aligned}$$

Il résulte donc de la prop. IV.3.23 que $\int_{\mathbf{R}} e^{-2i\pi xt} e^{-\pi\lambda|t|} \frac{\sin \pi t}{\pi t} dt = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{2\lambda}{\pi(\lambda^2+4(x+y)^2)} dy$. On conclut en utilisant le fait que la primitive de $\frac{2\lambda}{\pi(\lambda^2+4(x+y)^2)}$ est $\frac{1}{\pi} \operatorname{arctg}\left(\frac{2(x+y)}{\lambda}\right)$.

(ii) Nous allons plutôt calculer la transformée de Fourier de $e^{-\pi\lambda|t|} \frac{\sin^2 \pi t}{(\pi t)^2}$, pour $\lambda > 0$, et en déduire le résultat en faisant tendre λ vers 0. Si $x \in \mathbf{R}$, et $\lambda \geq 0$, soit

$$G_x(\lambda) = \int_{\mathbf{R}} g_x(\lambda, t) dt, \quad \text{avec } g_x(\lambda, t) = e^{-2i\pi xt} e^{-\pi\lambda|t|} \frac{\sin^2 \pi t}{(\pi t)^2}.$$

Comme la fonction $\frac{\sin^2 \pi t}{(\pi t)^2}$ est un majorant sommable de $g_x(\lambda, t)$ pour tout $\lambda \geq 0$, et comme $\lambda \mapsto g_x(\lambda, t)$ est continu sur \mathbf{R}_+ , pour tout $t \geq 0$, le théorème de continuité d'une intégrale dépendant d'un paramètre montre que G_x est continue sur \mathbf{R}_+ . La quantité qui nous intéresse $G_x(0)$ est donc la limite, quand $\lambda \rightarrow 0$, de $G_x(\lambda)$.

On va utiliser la prop. IV.3.23, avec $f = \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}[}$ et $g = e^{-2i\pi xt} e^{-\pi\lambda|t|} \frac{\sin \pi t}{\pi t}$, puisque $G_x(\lambda) = \int_{\mathbf{R}} \hat{f} g$. En utilisant le (i), on obtient $G_x(\lambda) = \int_{\mathbf{R}} f \hat{g} = \frac{1}{\pi} \int_{-1/2}^{1/2} h_x(\lambda, y) dy$, avec

$$h_x(\lambda, y) = \left(\operatorname{arctg}\left(\frac{2(x+y)+1}{\lambda}\right) - \operatorname{arctg}\left(\frac{2(x+y)-1}{\lambda}\right) \right).$$

Maintenant, $|h_x(\lambda, y)|$ est majoré par π quel que soit $\lambda > 0$, et comme $\operatorname{arctg}(\frac{a}{\lambda})$ tend vers $\operatorname{sign}(a) \frac{\pi}{2}$, quand $\lambda \rightarrow 0^+$, on obtient, en utilisant le théorème de continuité d'une intégrale dépendant d'un paramètre,

$$G_x(0) = \frac{1}{2} \int_{-1/2}^{1/2} (\operatorname{sign}(2(x+y)+1) - \operatorname{sign}(2(x+y)-1)) dy.$$

On conclut en utilisant la formule $\int_{-1/2}^{1/2} \operatorname{sign}(2y+a) dy = |\frac{a+1}{2}| - |\frac{a-1}{2}|$.

Proposition IV.4.2. — Soit $m \geq 1$.

- (i) Si $r \in \mathbf{N}$, et si $\mathbf{k} \in \mathbf{Z}^m$, alors $\hat{e}_{r,\mathbf{k}} \in L^2(\mathbf{R}^m)$.
- (ii) Si $\mathbf{k}, \mathbf{k}' \in \mathbf{Z}^m$, alors

$$\langle \hat{e}_{r,\mathbf{k}}, \hat{e}_{r,\mathbf{k}'} \rangle = \begin{cases} 0 & \text{si } \mathbf{k} \neq \mathbf{k}' \\ 2^{-rm} & \text{si } \mathbf{k} = \mathbf{k}' \end{cases} = \langle e_{r,\mathbf{k}}, e_{r,\mathbf{k}'} \rangle.$$

(iii) L'application $\phi \mapsto \hat{\phi}$ induit une isométrie de $\operatorname{Esc}(\mathbf{R}^m)$, muni de la norme $\|\cdot\|_2$, sur un sous-espace de $L^2(\mathbf{R}^m)$.

Démonstration. — Le (i) est une conséquence du fait que $\alpha(t) = \frac{\sin \pi t}{\pi t}$ est de carré sommable dans \mathbf{R} , et de la formule de la prop. IV.2.6 exprimant $\hat{e}_{r,\mathbf{k}}$ en terme de α .

Un calcul immédiat montre que $\langle e_{r,\mathbf{k}}, e_{r,\mathbf{k}'} \rangle$ est nul si $\mathbf{k} \neq \mathbf{k}'$, et vaut 2^{-rm} si $\mathbf{k} = \mathbf{k}'$. Par ailleurs, en partant de la formule

$$\hat{e}_{r,\mathbf{k}}(x) = \prod_{j=1}^m (2^{-r} e^{-2^{1-r} i \pi (k_j + \frac{1}{2}) x_j} \alpha(2^{-r} x_j))$$

de la prop. IV.2.6, on obtient (après une application immédiate du théorème de Fubini),

$$\begin{aligned} \langle \hat{e}_{r,\mathbf{k}}, \hat{e}_{r,\mathbf{k}'} \rangle &= 2^{-2rm} \prod_{j=1}^m \left(\int_{\mathbf{R}} e^{-2^{1-r} i \pi (k'_j - k_j) x_j} \alpha(2^{-r} x_j)^2 dx_j \right) \\ &= 2^{-rm} \prod_{j=1}^m \left(\int_{\mathbf{R}} e^{-2i\pi (k'_j - k_j) u_j} \alpha(u_j)^2 du_j \right). \end{aligned}$$

On reconnaît dans la parenthèse la transformée de Fourier de $\alpha^2(t) = \frac{\sin^2 \pi t}{(\pi t)^2}$, ce qui permet d'utiliser le (ii) du lemme IV.4.1 pour montrer que $\int_{\mathbf{R}} e^{-2i\pi (k'_j - k_j) u_j} \alpha(u_j)^2 du_j$ est nul si $k_j \neq k'_j$, et vaut 1 si $k_j = k'_j$. On en déduit le (ii).

Finalement, le (iii) suit (en utilisant le fait que les $e_{r,\mathbf{k}}$, pour $\mathbf{k} \in \mathbf{Z}^m$, forment une base orthogonale de $\operatorname{Esc}_r(\mathbf{R}^m)$) du (ii) et de ce que $\operatorname{Esc}(\mathbf{R}^m)$ est la réunion des $\operatorname{Esc}_r(\mathbf{R}^m)$, pour $r \in \mathbf{N}$.

2. Définition de la transformée de Fourier dans L^2

Théorème IV.4.3. — La transformée de Fourier $\mathcal{F} : \text{Esc}(\mathbf{R}^m) \rightarrow L^2(\mathbf{R}^m)$ se prolonge par continuité en une isométrie de $L^2(\mathbf{R}^m)$ sur $L^2(\mathbf{R}^m)$ dont l'inverse est $\phi \mapsto \overline{\mathcal{F}}\phi$, avec $\overline{\mathcal{F}}\phi(x) = \mathcal{F}\phi(-x)$. Autrement dit, on a

- (i) $\langle \mathcal{F}\phi_1, \mathcal{F}\phi_2 \rangle = \langle \phi_1, \phi_2 \rangle$, si $\phi_1, \phi_2 \in L^2(\mathbf{R}^m)$ (\mathcal{F} est une isométrie),
- (ii) $\mathcal{F}(\mathcal{F}\phi)(x) = \phi(-x)$, si $\phi \in L^2(\mathbf{R}^m)$ (formule d'inversion de Fourier dans L^2).

De plus, les formules du n° IV.2.2, pour les dilatations, translations et multiplications par un caractère, sont encore valables dans $L^2(\mathbf{R}^m)$, et si $\phi \in L^1(\mathbf{R}^m) \cap L^2(\mathbf{R}^m)$, les deux définitions de $\mathcal{F}\phi$ coïncident.

Démonstration. — Il résulte du (iii) de la prop. IV.4.2 et de la densité de $\text{Esc}(\mathbf{R}^m)$ dans $L^2(\mathbf{R}^m)$, que \mathcal{F} peut se prolonger par continuité, de manière unique, en une isométrie de $L^2(\mathbf{R}^m)$ sur un sous-espace⁽⁸⁾ de $L^2(\mathbf{R}^m)$. De plus, les formules pour les dilatations, translations et multiplications par un caractère s'étendent à $L^2(\mathbf{R}^m)$ par continuité.

Maintenant, si $\phi \in L^1(\mathbf{R}^m) \cap L^2(\mathbf{R}^m)$, notons (provisoirement), $\mathcal{F}_1\phi$ (resp. $\mathcal{F}_2\phi$) la transformée de Fourier de ϕ dans $L^1(\mathbf{R}^m)$ (resp. $L^2(\mathbf{R}^m)$). Soit $(f_j)_{j \in \mathbf{N}}$ une suite d'éléments de $\text{Esc}(\mathbf{R}^m)$ convergeant vers ϕ à la fois dans $L^1(\mathbf{R}^m)$ et dans $L^2(\mathbf{R}^m)$ (cf. th. III.2.11). Alors $\mathcal{F}_1 f_j = \mathcal{F}_2 f_j$ tend uniformément vers la fonction continue $\mathcal{F}_1\phi$ et en norme $\|\cdot\|_2$ vers $\mathcal{F}_2\phi$. D'après le lemme III.2.14, cela implique $\mathcal{F}_1\phi \in L^2(\mathbf{R}^m)$ et $\mathcal{F}_1\phi = \mathcal{F}_2\phi$ p.p.

Pour conclure, il suffit donc de prouver l'on a $\mathcal{F} \circ \mathcal{F} = s$, où $s : L^2(\mathbf{R}^m) \rightarrow L^2(\mathbf{R}^m)$ est l'application déduite de celle sur $\mathcal{L}^2(\mathbf{R}^m)$ et envoyant ϕ sur $s(\phi)$, avec $s(\phi)(x) = \phi(-x)$. En effet ceci prouve à la fois la surjectivité et le fait que $\overline{\mathcal{F}}$ est l'inverse de \mathcal{F} . Par continuité des applications $\mathcal{F} \circ \mathcal{F}$ et s , il suffit de vérifier qu'elles coïncident sur $\text{Esc}(\mathbf{R}^m)$, qui est un sous-espace dense, et par linéarité, il suffit de le prouver pour une famille génératrice de $\text{Esc}(\mathbf{R}^m)$, par exemple la famille des $\mathbf{e}_{r,\mathbf{k}}$, pour $r \in \mathbf{N}$, et $\mathbf{k} \in \mathbf{Z}^m$. Finalement, comme les $\mathbf{e}_{r,\mathbf{k}}$ sont obtenues par translation et dilatation à partir de la fonction $\beta_m(t) = \prod_{j=1}^m \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}(t_j)$, il suffit de prouver que $\mathcal{F}(\mathcal{F}\beta_m)(t) = \beta_m(-t)$ dans $L^2(\mathbf{R}^m)$. Or on a $(\mathcal{F}\beta_m)(x) = \prod_{j=1}^m \frac{\sin \pi x_j}{\pi x_j}$, et on est ramené à calculer la transformée de Fourier de $\prod_{j=1}^m \frac{\sin \pi x_j}{\pi x_j}$. Les arguments étant les mêmes pour $m = 1$ que pour m quelconque, nous ne traiterons que le cas $m = 1$, ce qui a pour avantage de raccourcir un peu les expressions.

On doit donc calculer la transformée de Fourier de $\alpha(t) = \frac{\sin \pi t}{\pi t}$. Comme α n'est pas sommable, on ne peut pas utiliser l'expression $\hat{\alpha}(x) = \int_{\mathbf{R}} e^{-2i\pi xt} \alpha(t) dt$ pour faire le calcul. Pour contourner le problème, constatons que $\alpha_\lambda(t) = e^{-\pi\lambda|t|} \alpha(t)$ tend vers α dans $L^2(\mathbf{R}^m)$, quand λ tend vers 0, car α_λ est majoré, pour tout $\lambda \geq 0$, par α qui est de carré sommable, et $\alpha_\lambda(t)$ tend vers $\alpha(t)$ quand λ tend vers 0, pour tout $t \in \mathbf{R}$. Par continuité de \mathcal{F} , on en déduit que $\mathcal{F}\alpha_\lambda$ tend vers $\mathcal{F}\alpha$ dans L^2 quand λ tend vers 0. Or on a calculé $\mathcal{F}\alpha_\lambda$ (cf. IV.4.1), et il est apparent sur la formule que $\mathcal{F}\alpha_\lambda(x)$ tend vers $\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}(x)$, quand λ tend vers 0, pour tout $x \neq \pm \frac{1}{2}$. On en déduit que $\mathcal{F}\alpha = \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}$ p.p., ce qui permet de conclure.

Remarque IV.4.4. — On a été confronté, au cours de la démonstration, au problème du calcul de la transformée de Fourier d'une fonction ϕ , de carré sommable, mais non sommable. On s'en est sorti par un passage à la limite, en multipliant ϕ par $e^{-\pi\lambda(|t_1| + \dots + |t_m|)}$, et en faisant tendre λ vers 0. Une autre manière de procéder est de prendre une suite croissante $(X_N)_{N \in \mathbf{N}}$ d'ensembles bornés dont la réunion est \mathbf{R}^m , et d'écrire $\mathcal{F}\phi$ comme la limite de $\mathcal{F}(\mathbf{1}_{X_N}\phi)$ quand N tend vers $+\infty$ (en effet, $\mathbf{1}_{X_N}\phi$ tend vers ϕ dans $L^2(\mathbf{R}^m)$ d'après le théorème de convergence dominée, et $\mathbf{1}_{X_N}\phi$ est sommable puisque de carré sommable et à support borné).

Corollaire IV.4.5. — Si $f, g \in L^2(\mathbf{R}^m)$, alors $\int_{\mathbf{R}^m} g\hat{f} = \int_{\mathbf{R}^m} f\hat{g}$.

⁽⁸⁾ Comme on est en dimension infinie, l'injectivité n'implique pas, a priori, la surjectivité.

Démonstration. — Posons $\phi_1 = \overline{\mathcal{F}f}$, et $\phi_2 = g$. On a $\mathcal{F}\phi_1 = \overline{f}$, et donc

$$\int_{\mathbf{R}^m} g \hat{f} = \langle \phi_1, \phi_2 \rangle = \langle \mathcal{F}\phi_1, \mathcal{F}\phi_2 \rangle = \int_{\mathbf{R}^m} f \hat{g},$$

ce qu'il fallait démontrer.

3. Comparaison des transformées de Fourier dans L^1 et L^2

Proposition IV.4.6. — Si $f_1 \in L^1(\mathbf{R}^m)$ et $f_2 \in L^2(\mathbf{R}^m)$ ont même transformée de Fourier, alors $f_1 = f_2$.

Démonstration. — Si $\phi \in \mathcal{C}_c^\infty(\mathbf{R}^m)$, alors $\overline{\mathcal{F}\phi} \in L^1(\mathbf{R}^m) \cap L^2(\mathbf{R}^m)$ d'après le cor. IV.3.22 ; en particulier, $\mathcal{F}(\overline{\mathcal{F}\phi}) = \phi$. Comme f et g ont même transformée de Fourier, on déduit de la prop. IV.3.23 et du cor. IV.4.5, utilisés pour $g = \overline{\mathcal{F}\phi}$, que $\int_{\mathbf{R}^m} f_1 \phi = \int_{\mathbf{R}^m} f_2 \phi$. D'après l'exercice III.3.12, ceci implique $f_1 = f_2$.

Proposition IV.4.7. — (Formule d'inversion de Fourier dans L^1) Si $\phi \in L^1(\mathbf{R}^m)$ a une transformée de Fourier sommable, alors $\phi = \overline{\mathcal{F}\hat{\phi}}$.

Démonstration. — Par hypothèse, $\hat{\phi} \in L^1(\mathbf{R}^m)$. Par ailleurs, on sait que $\hat{\phi}$ est bornée puisque $\phi \in L^1(\mathbf{R}^m)$. Donc $|\hat{\phi}|^2 \leq \|\hat{\phi}\|_\infty \cdot |\hat{\phi}|$, et $\hat{\phi} \in L^2(\mathbf{R}^m)$, ce qui fait que $\overline{\mathcal{F}\hat{\phi}}$ est un élément de $L^2(\mathbf{R}^m)$ ayant même transformée de Fourier que ϕ (grâce à la formule d'inversion de Fourier dans $L^2(\mathbf{R}^m)$). La prop. IV.4.6 permet de conclure.

4. Dérivation

Théorème IV.4.8. — (i) Si $f \in H^k(\mathbf{R}^m)$, alors $(1 + \|x\|^2)^{k/2} \hat{f}(x)$ est de carré sommable, $\mathcal{F}(\partial^\ell f) = (2i\pi x)^\ell \hat{f}$ si $\ell \in \mathbf{N}^m$ vérifie $|\ell| \leq k$, et $\|(1 + \|x\|^2)^{k/2} \hat{f}\|_2 \leq (1 + \frac{m}{2\pi})^k \|f\|_{H^k}$.

(ii) Réciproquement, si $(1 + \|t\|^2)^{k/2} f(t)$ est de carré sommable, alors $\hat{f} \in H^k(\mathbf{R}^m)$, on a $\partial^\ell \hat{f}(x) = (-2i\pi)^\ell \mathcal{F}(t^\ell f)$, si $|\ell| \leq k$, et $\|\hat{f}\|_{H^k} \leq (2\pi)^k \|(1 + \|t\|^2)^{k/2} f\|_2$.

Remarque IV.4.9. — (i) Le (i) s'applique en particulier à une fonction de classe \mathcal{C}^k dont toutes les dérivées partielles d'ordre $\leq k$ sont de carré sommable.

(ii) On déduit du théorème une définition « par Fourier » de l'espace de Sobolev $H^k(\mathbf{R}^m)$. C'est l'ensemble des $f \in L^2(\mathbf{R}^m)$ tel que $(1 + \|x\|^2)^{k/2} \hat{f} \in L^2(\mathbf{R}^m)$. On peut étendre cette définition à $k \in \mathbf{R}_+$, et parler de l'espace de Sobolev $H^s(\mathbf{R}^m)$, pour $s \in \mathbf{R}_+$ (et même pour $s \in \mathbf{R}$ en supprimant la condition $f \in L^2(\mathbf{R}^m)$, mais on tombe alors sur un espace de distributions).

Démonstration. — On sait déjà que $\mathcal{F}(\partial^\ell f) = (2i\pi x)^\ell \hat{f}$, si $f \in \mathcal{C}_c^k(\mathbf{R}^m)$. Maintenant, $\mathcal{F} \circ \partial^\ell : H^k(\mathbf{R}^m) \rightarrow L^2(\mathbf{R}^m)$ est continue, comme composée de $\partial^\ell : H^k(\mathbf{R}^m) \rightarrow H^{k-|\ell|}(\mathbf{R}^m)$, qui est continue d'après le n° 4 du § II.2, de l'inclusion de $H^{k-|\ell|}(\mathbf{R}^m)$ dans $L^2(\mathbf{R}^m)$ qui est 1-lipschitzienne, et de $\mathcal{F} : L^2(\mathbf{R}^m) \rightarrow L^2(\mathbf{R}^m)$ qui est une isométrie. Comme $\mathcal{C}_c^k(\mathbf{R}^m)$ est dense dans $H^k(\mathbf{R}^m)$ par construction, cela permet d'en déduire que $\mathcal{F}(\partial^\ell f) = (2i\pi x)^\ell \hat{f}$, quel que soit $f \in H^k(\mathbf{R}^m)$. En utilisant le fait que $\mathcal{F} : L^2(\mathbf{R}^m) \rightarrow L^2(\mathbf{R}^m)$ est une isométrie, on obtient alors (en développant $(1 + \sum_{j=1}^m |x_j|^2)^k$ sous la forme $\sum_{|\ell| \leq k} a_\ell |x|^\ell$),

$$\|(1 + \|x\|^2)^{k/2} \hat{f}\|_2 \leq \|(1 + \sum_{j=1}^m |x_j|^2)^k \hat{f}\|_2 \leq \sum_{|\ell| \leq k} a_\ell \|x^\ell \hat{f}\|_2 = \sum_{|\ell| \leq k} a_\ell (2\pi)^{-|\ell|} \|\partial^\ell f\|_2.$$

On conclut la démonstration du (i) en majorant chaque $\|\partial^\ell f\|_2$ par $\|f\|_{H^k}$, et en remarquant que $\sum_{|\ell| \leq k} a_\ell (2\pi)^{-|\ell|} = (1 + \frac{m}{2\pi})^k$.

Passons à la démonstration du (ii). Soit E l'ensemble des f tels que $(1 + \|t\|^2)^{k/2}f$ soit de carré sommable. On munit E de la norme $\|\cdot\|_E$ définie par $\|f\|_E = \|(1 + \|t\|^2)^{k/2}f\|_2$. Par définition, l'application $f \mapsto (1 + \|t\|^2)^{k/2}f$ induit une isométrie de E sur $L^2(\mathbf{R}^m)$, ce qui fait que E est un espace de Hilbert.

Maintenant, si f est de la forme $(1 + \|t\|^2)^{-k/2}g$, où g est une fonction en escalier, alors $(1 + \|t\|^2)^{k/2}f$ est sommable, et il résulte du (ii) du théorème IV.2.8, que \hat{f} est de classe \mathcal{C}^k , et que $\partial^\ell \hat{f} = (-2i\pi)^{|\ell|} \mathcal{F}(t^\ell f)$, si $|\ell| \leq k$. En utilisant le fait que $\mathcal{F} : L^2(\mathbf{R}^m) \rightarrow L^2(\mathbf{R}^m)$ est une isométrie, on obtient alors

$$\|\hat{f}\|_{H^k} = \sup_{|\ell| \leq k} \|\partial^\ell \hat{f}\|_2 = \sup_{|\ell| \leq k} (2\pi)^{|\ell|} \|t^\ell f\|_2 \leq (2\pi)^k \|(1 + \|t\|^2)^{k/2}f\|_2 = (2\pi)^k \|f\|_E.$$

Autrement dit, $f \mapsto \hat{f}$ est une application $(2\pi)^k$ -lipschitzienne de $(1 + \|t\|^2)^{-k/2}\text{Esc}(\mathbf{R}^m)$ muni de la norme $\|\cdot\|_E$ dans $H^k(\mathbf{R}^m)$ muni de la norme $\|\cdot\|_{H^k}$. Comme $\text{Esc}(\mathbf{R}^m)$ est dense dans $L^2(\mathbf{R}^m)$, l'espace $(1 + \|t\|^2)^{-k/2}\text{Esc}(\mathbf{R}^m)$ est dense dans E , et comme $H^k(\mathbf{R}^m)$ est complet, on en déduit que $f \mapsto \hat{f}$ induit une application $(2\pi)^k$ -lipschitzienne de E dans H^k . Autrement dit, si $(1 + \|t\|^2)^{k/2}f$ est de carré sommable, alors $\hat{f} \in H^k(\mathbf{R}^m)$ et $\|\hat{f}\|_{H^k} \leq (2\pi)^k \|(1 + \|t\|^2)^{k/2}f\|_2$.

Il reste à prouver que l'on a $\partial^\ell \hat{f} = (-2i\pi)^{|\ell|} \mathcal{F}(t^\ell f)$, quel que soit $f \in E$, sachant que c'est vrai pour f dans un sous-espace dense. Or $\partial^\ell \circ \mathcal{F} : E \rightarrow L^2(\mathbf{R}^m)$ est continue, puisque l'on vient de prouver que $\mathcal{F} : E \rightarrow H^k(\mathbf{R}^m)$ est continue et que $\partial^\ell : H^k(\mathbf{R}^m) \rightarrow H^{k-|\ell|}(\mathbf{R}^m) \subset L^2(\mathbf{R}^m)$ est continue. De même, $f \mapsto t^\ell f$ est continue (et même 1-lipschitzienne) de E dans $L^2(\mathbf{R}^m)$, et donc $f \mapsto (-2i\pi)^{|\ell|} \mathcal{F}(t^\ell f)$ est continue de E dans $L^2(\mathbf{R}^m)$. On a deux applications continues coïncidant sur un sous-espace dense ; elles sont donc égales, ce qui permet de conclure.

CHAPITRE V

FONCTIONS HOLOMORPHES

Une fonction holomorphe sur un ouvert Ω est une fonction de Ω dans \mathbf{C} qui est \mathcal{C}^∞ au sens complexe et somme de sa série de Taylor autour de chaque point. Ces fonctions jouissent de propriétés de rigidité absolument remarquables et qui peuvent sembler miraculeuses si on se réfère à ce que l'on connaît des fonctions d'une variable réelle. Une des premières surprises que l'on rencontre est qu'une fonction est holomorphe si et seulement si elle est de classe \mathcal{C}^1 (au sens complexe)! Ceci est une des nombreuses conséquences de la formule intégrale de Cauchy (th. V.2.5). Cette formule et ses conséquences immédiates (rem. V.2.8, th. V.2.12, V.2.15 et V.2.18), couplées avec le principe du maximum (th. V.1.20), le théorème des zéros isolés (th. V.1.15) et la formule des résidus de Cauchy du chapitre suivant, permettent d'attaquer une variété de problèmes assez spectaculaire. Parmi ceux-ci, mentionnons le théorème fondamental de l'algèbre (ex. V.1.22, cor. V.2.10 ou ex. VI.2.21), le théorème des 4 carrés de Lagrange (ex. VII.6.9) ou encore le théorème des nombres premiers auquel l'annexe A est consacrée.

V.1. Fonctions holomorphes et fonctions analytiques complexes

1. Séries entières

Soit K un corps. Une *série entière* (ou *série formelle*) à coefficients dans K est une expression du type $\sum_{n=0}^{+\infty} a_n T^n$, où les a_n sont des éléments de K . L'ensemble $K[[T]]$ des séries entières à coefficients dans K contient l'anneau $K[T]$ des polynômes à coefficients dans K , et on munit $K[[T]]$ d'une structure d'anneau étendant celle de $K[T]$ en posant

$$\left(\sum_{n=0}^{+\infty} a_n T^n\right) + \left(\sum_{n=0}^{+\infty} b_n T^n\right) = \sum_{n=0}^{+\infty} (a_n + b_n) T^n, \quad \left(\sum_{n=0}^{+\infty} a_n T^n\right) \left(\sum_{n=0}^{+\infty} b_n T^n\right) = \sum_{n=0}^{+\infty} \left(\sum_{k=0}^n a_k b_{n-k}\right) T^n.$$

On définit la *valuation* $v_T(F)$ de $F = \sum_{n \in \mathbf{N}} a_n T^n$ comme étant le plus petit élément de l'ensemble des $n \in \mathbf{N}$ vérifiant $a_n \neq 0$, si $F \neq 0$, et comme étant $+\infty$, si $F = 0$. On a $v_T(FG) = v_T(F) + v_T(G)$, et $v_T(F + G) \geq \inf(v_T(F), v_T(G))$. En définissant $| \cdot |_T$, par $|F|_T = e^{-v_T(F)}$, cela se traduit par $|FG|_T = |F|_T |G|_T$ et $|F + G|_T \leq \sup(|F|_T, |G|_T)$.

Cette dernière inégalité (*ultramétrique*), plus forte que l'inégalité triangulaire, montre que $d(F, G) = |F - G|_T$ est une distance sur $K[[T]]$. On vérifie facilement, que $K[[T]]$ est complet pour cette distance, qu'une série $\sum_{n \in \mathbf{N}} F_n$ converge, si et seulement si $|F_n|_T \rightarrow 0$ (ce qui équivaut à $v_T(F_n) \rightarrow +\infty$), et que $K[T]$ est dense dans $K[[T]]$, ce qui prouve que $K[[T]]$ est le complété de $K[T]$ pour la (distance associée à la) valuation v_T . Ceci fournit une construction topologique⁽¹⁾ de $K[[T]]$ à partir de $K[T]$.

Si $F = \sum_{n=0}^{+\infty} a_n T^n \in K[[T]]$, on définit la dérivée F' de F par la formule $F' = \sum_{n=1}^{+\infty} n a_n T^{n-1}$. Et on définit par récurrence la dérivée k -ième $F^{(k)}$ comme la dérivée de $F^{(k-1)}$, en posant $F^{(0)} = F$. La dérivée seconde $F^{(2)}$ est souvent aussi notée F'' . Un calcul immédiat montre que $\frac{1}{k!} F^{(k)} = \sum_{n=0}^{+\infty} \binom{n+k}{k} T^n$, où $\binom{X}{k}$ est le *polynôme binomial*⁽²⁾ de degré k défini par $\binom{X}{0} = 1$ et $\binom{X}{k} = \frac{X(X-1)\cdots(X-k+1)}{k!}$, si $k \geq 1$.

Exercice V.1.1. — (i) Montrer que, si $F, G \in K[[T]]$, alors $(FG)' = F'G + FG'$.

(ii) Montrer que, si $(F_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $K[[T]]$ telle que $\sum_{n \in \mathbf{N}} F_n$ converge dans $K[[T]]$, et si F est la somme de la série, alors la série $\sum_{n \in \mathbf{N}} F'_n$ converge dans $K[[T]]$, et on a $F' = \sum_{n \in \mathbf{N}} F'_n$.

(ii) Montrer que, si $F = \sum_{n \in \mathbf{N}} a_n T^n \in K[[T]]$, et si $G \in TK[[T]]$, alors $\sum_{n \in \mathbf{N}} a_n G^n$ converge dans $K[[T]]$, et la limite $F \circ G$ vérifie $(F \circ G)' = (F' \circ G) G'$.

Exercice V.1.2. — Soit $(u_n)_{n \in \mathbf{N}}$ une suite d'éléments de \mathbf{C} vérifiant la relation de récurrence $u_{n+2} - 3u_{n+1} + 2u_n = 1$, pour tout $n \in \mathbf{N}$, et soit $F = \sum_{n \in \mathbf{N}} u_n T^n \in \mathbf{C}[[T]]$. Calculer $(1 - 3T + 2T^2)F(T)$ en fonction de u_0 et u_1 . En déduire une formule générale pour u_n .

On ne s'intéressera, par la suite, qu'au cas $K = \mathbf{C}$, en lien avec les fonctions holomorphes, mais les exemples qui suivent ont un sens sur un corps K de caractéristique 0 quelconque.

Exemple V.1.3. — • *La série exponentielle.* Si $\lambda \in \mathbf{C}$, on note $\exp(\lambda T)$ la série formelle $\sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} T^n$. C'est la solution formelle de l'équation différentielle $F' = \lambda F$ dont le terme constant est 1, et on a $\exp(\lambda T) \exp(\mu T) = \exp((\lambda + \mu)T)$.

• *Les séries puissances.* Si $\alpha \in \mathbf{N}$, alors $(1 + T)^\alpha = \sum_{n=0}^{+\infty} \binom{\alpha}{n} T^n$, d'après la formule du binôme, et l'annulation de $\binom{\alpha}{n}$, pour $n > \alpha$. Par analogie, on définit $(1 + T)^\alpha$, pour $\alpha \in \mathbf{C}$,

⁽¹⁾On obtient une construction algébrique en remarquant que $F \in K[[T]]$ est déterminée si on connaît ses n premiers coefficients, quel que soit $n \in \mathbf{N}$. Autrement dit F est déterminée par ses images F_n dans $K[T]/(T^n)$, pour $n \in \mathbf{N}$, et l'application $F \mapsto (F_n)_{n \in \mathbf{N}}$ permet d'identifier $K[[T]]$ à la *limite projective* $\varprojlim K[T]/(T^n)$ des $K[T]/(T^n)$, ensemble des suites $(F_n)_{n \in \mathbf{N}}$, où $F_n \in K[T]/(T^n)$, et $F_{n+1} \in K[T]/(T^{n+1})$ a pour image F_n modulo T^n , quel que soit $n \in \mathbf{N}$.

⁽²⁾Ce polynôme doit son nom au fait que $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ est un coefficient du binôme $(1 + X)^n$, si n est un entier $\geq k$; si $n \in \{0, \dots, k-1\}$, on a $\binom{n}{k} = 0$.

comme la série entière $\sum_{n=0}^{+\infty} \binom{\alpha}{n} T^n$. Si $\alpha, \beta \in \mathbf{N}$, on a

$$\sum_{n=0}^{+\infty} \left(\sum_{k=0}^n \binom{\alpha}{k} \binom{\beta}{n-k} \right) T^n = (1+T)^\alpha (1+T)^\beta = (1+T)^{\alpha+\beta} = \sum_{n=0}^{+\infty} \binom{\alpha+\beta}{n} T^n,$$

ce qui se traduit par le fait que les polynômes en 2 variables

$$\binom{X+Y}{n} \quad \text{et} \quad \sum_{k=0}^n \binom{X}{k} \binom{Y}{n-k}$$

prennent les mêmes valeurs sur $\mathbf{N} \times \mathbf{N}$. Par suite, ils sont égaux, ce qui permet de démontrer que $(1+T)^\alpha (1+T)^\beta = (1+T)^{\alpha+\beta}$ quels que soient $\alpha, \beta \in \mathbf{C}$.

- *La série du logarithme.* On définit $\log(1+T)$ comme la série $\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} T^n$.

Exercice V.1.4. — (i) Montrer que $\log(\exp(\lambda T)) = \lambda T$ dans $\mathbf{C}[[T]]$.

(ii) Montrer que $(1+T)^\alpha = \exp(\alpha \log(1+T))$ dans $\mathbf{C}[[T]]$. (On pourra appliquer l'opérateur $(1+T) \frac{d}{dT}$ aux deux membres.)

2. Rayon de convergence d'une série entière

Si $z_0 \in \mathbf{C}$ et si $r > 0$, on note $D(z_0, r)$ le *disque fermé* de centre z_0 et de rayon r (i.e. l'ensemble des $z \in \mathbf{C}$ vérifiant $|z - z_0| \leq r$), et $D(z_0, r^-)$ le *disque ouvert* de centre z_0 et de rayon r (i.e. l'ensemble des $z \in \mathbf{C}$ vérifiant $|z - z_0| < r$).

Rappelons que, si $F = \sum_{n=0}^{+\infty} a_n T^n \in \mathbf{C}[[T]]$, il existe un unique $\rho(F) \in \overline{\mathbf{R}}_+$, appelé *rayon de convergence* de F , tel que, si $|z| < \rho$, la série $\sum_{n=0}^{+\infty} a_n z^n$ soit normalement convergente, et si $|z| > \rho(F)$, la suite $a_n z^n$ ne soit pas bornée (et donc la série soit divergente). On a par ailleurs $\rho(F)^{-1} = \limsup |a_n|^{1/n}$.

Si $z \in D(0, \rho(F)^-)$, on note $F(z)$ la somme de la série $\sum_{n=0}^{+\infty} a_n z^n$. La vérification du résultat suivant est laissée au lecteur.

Lemme V.1.5. — Si $F, G \in \mathbf{C}[[T]]$, alors

$$\rho(F+G) \geq \inf(\rho(F), \rho(G)) \quad \text{et} \quad \rho(FG) \geq \inf(\rho(F), \rho(G)),$$

et on a

$$(F+G)(z) = F(z) + G(z) \quad \text{et} \quad (FG)(z) = F(z)G(z), \quad \text{si } |z| < \inf(\rho(F), \rho(G)).$$

Exemple V.1.6. — • Si $\lambda \in \mathbf{C}$, alors $\rho(\exp(\lambda T)) = +\infty$.

• $\rho(\log(1+T)) = 1$, et $\rho((1+T)^\alpha) = 1$, si $\alpha \notin \mathbf{N}$ (sinon on a affaire à un polynôme dont le rayon de convergence est infini). En effet, si $\alpha \in \mathbf{C} - \mathbf{N}$, alors $|\binom{\alpha}{n+1} / \binom{\alpha}{n}| = |\frac{\alpha-n}{n+1}| \rightarrow 1$, et donc $|\binom{\alpha}{n}|^{1/n} \rightarrow 1$ (version multiplicative de la moyenne de Césaro).

- La série⁽³⁾ $\sum_{n=0}^{+\infty} (-1)^n n! T^n$ a un rayon de convergence nul.

⁽³⁾C'est la série de Taylor en 0 de la fonction $f(x) = \int_0^{+\infty} \frac{e^{-t}}{1+tx} dt$, qui est \mathcal{C}^∞ sur \mathbf{R}_+ , ce qui faisait dire à Euler que $\sum_{n=0}^{+\infty} (-1)^n n! = \int_0^{+\infty} \frac{e^{-t}}{1+t} dt$.

Remarque V.1.7. — Comme $(1+T)^\alpha(1+T)^\beta = (1+T)^{\alpha+\beta}$ dans $\mathbf{C}[[T]]$, et comme les trois séries entières ci-dessus ont un rayon de convergence ≥ 1 , on a $(1+z)^\alpha(1+z)^\beta = (1+z)^{\alpha+\beta}$ quels que soient $z \in D(0, 1^-)$, et $\alpha, \beta \in \mathbf{C}$. En particulier, si $m \in \mathbf{N} - \{0\}$, alors $(1+z)^{1/m}$ est une racine m -ième de $1+z$, quel que soit $z \in D(0, 1^-)$.

Exercice V.1.8. — Soit $\alpha \in \mathbf{R} - \mathbf{N}$ ⁽⁴⁾.

(i) Exprimer $\binom{\alpha}{n}$ en termes de la fonction Γ .

(ii) En déduire, en utilisant la formule de Stirling (ex. IV.1.5), qu'il existe $C(\alpha) \in \mathbf{C}$ tel que $\binom{\alpha}{n} \sim C(\alpha)(-1)^n n^{-1-\alpha}$.

Proposition V.1.9. — (i) Si $F = \sum_{n=0}^{+\infty} a_n T^n \in \mathbf{C}[[T]]$, alors $\rho(F^{(k)}) \geq \rho(F)$, quel que soit $k \in \mathbf{N}$, et si $|z_0| + |z - z_0| < \rho(F)$, alors

$$F(z) = \sum_{k=0}^{+\infty} \left(\sum_{n=0}^{+\infty} \binom{n+k}{k} a_{n+k} z_0^n \right) (z - z_0)^k = \sum_{k=0}^{+\infty} \frac{F^{(k)}(z_0)}{k!} (z - z_0)^k.$$

(ii) Si $|z_0| < \rho(F)$, alors $\lim_{z \rightarrow z_0} \frac{F(z) - F(z_0)}{z - z_0} = F'(z_0)$.

Démonstration. — Si $|z_0| + |z - z_0| < \rho(F)$, on a

$$\begin{aligned} \sum_{k=0}^{+\infty} \left(\sum_{n=0}^{+\infty} \binom{n+k}{k} |a_{n+k} z_0^n| \right) |z - z_0|^k &= \sum_{m=0}^{+\infty} \sum_{n+k=m} \binom{m}{k} |a_m| |z_0^n| |z - z_0|^k \\ &= \sum_{m=0}^{+\infty} |a_m| (|z_0| + |z - z_0|)^m < +\infty, \end{aligned}$$

ce qui montre que la série double de la proposition est absolument convergente. En fixant k , on en déduit que $F^{(k)}$ converge si $|z_0| < \rho(F)$, et donc $\rho(F^{(k)}) \geq \rho(F)$. De plus, on peut réordonner les termes comme on veut, et commencer par sommer sur $n+k=m$, puis sommer sur $m \in \mathbf{N}$. La somme pour $n+k=m$ étant $\sum_{n+k=m} \binom{m}{n} a_m z_0^n (z - z_0)^k = a_m z^m$, cela démontre le (i).

Le (i) permet, en faisant le changement de variables $z = z_0 + h$, de supposer que $z_0 = 0$ pour démontrer le (ii). Or on a

$$\left| \frac{F(h) - F(0)}{h} - F'(0) \right| = \left| h \sum_{n=2}^{+\infty} a_n h^{n-2} \right| \leq |h| \left(\sum_{n=2}^{+\infty} |a_n| (\rho(F)/2)^{n-2} \right), \quad \text{si } |h| < \rho(F)/2,$$

et comme $\sum_{n=2}^{+\infty} |a_n| (\rho(F)/2)^{n-2} < +\infty$, on en déduit que $\frac{F(h) - F(0)}{h} - F'(0)$ tend vers 0 quand $h \rightarrow 0$, ce qui permet de conclure.

⁽⁴⁾Les résultats de cet exercice s'étendent à $\alpha \in \mathbf{C} - \mathbf{N}$, une fois que l'on dispose de la fonction Γ dans le plan complexe (cf. th. VII.2.1 et prop. VII.2.3).

Remarque V.1.10. — On peut reformuler le (ii) de la proposition précédente en disant que, si $F \in \mathbf{C}[[T]]$ est de rayon de convergence non nul, alors F est *dérivable au sens complexe* dans $D(0, \rho(F)^-)$, et que la dérivée au sens complexe de F est F' . Une récurrence immédiate permet donc de montrer que F est de classe \mathcal{C}^∞ au sens complexe sur $D(0, \rho(F)^-)$, et que la dérivée k -ième au sens complexe de F est la fonction associée à la série entière $F^{(k)}$.

3. Premières propriétés des fonctions holomorphes

3.1. Définition

Soit Ω un ouvert de \mathbf{C} , et soit $f : \Omega \rightarrow \mathbf{C}$. Si $z_0 \in \Omega$, si $r > 0$, et si $D(z_0, r^-) \subset \Omega$, on dit que f est *développable en série entière sur* $D(z_0, r^-)$, s'il existe $F \in \mathbf{C}[[T]]$ de rayon de convergence $\geq r$, telle que $f(z) = F(z - z_0)$, si $z \in D(z_0, r^-)$. Il résulte du (i) de la prop. V.1.9 que si f est développable en série entière sur $D(z_0, r^-)$, et si $D(z_1, r_1^-) \subset D(z_0, r^-)$, alors f est développable en série entière sur $D(z_1, r_1^-)$.

On dit que f est *développable en série entière autour de* z_0 s'il existe $r > 0$ tel que f soit développable en série entière sur $D(z_0, r^-)$, et on dit que f est *analytique sur* Ω ou encore que f est *holomorphe sur* Ω , si f est développable en série entière autour de tout point de Ω . Il résulte de la rem. V.1.10 qu'une fonction analytique sur Ω est de classe \mathcal{C}^∞ au sens complexe sur Ω .

Exemple V.1.11. — (i) Un polynôme est une fonction holomorphe sur \mathbf{C} .

(ii) $\exp(\lambda z)$ est holomorphe sur \mathbf{C} , quel que soit $\lambda \in \mathbf{C}$.

(iii) $\frac{1}{z}$ est holomorphe sur $\mathbf{C} - \{0\}$.

(iv) $\log(1 + z)$ et $(1 + z)^s$, si $s \in \mathbf{C}$, sont holomorphes sur $D(0, 1^-)$.

(v) Si f et g sont holomorphes sur Ω , alors $f + g$ et fg sont holomorphes sur Ω .

(vi) Si $f : \Omega_1 \rightarrow \Omega_2$ est holomorphe, et si $g : \Omega_2 \rightarrow \mathbf{C}$ est holomorphe, alors $g \circ f$ est holomorphe sur Ω_1 .

(vii) Si $f : \Omega_1 \rightarrow \Omega_2$ est holomorphe bijective, alors sa réciproque $f^{-1} : \Omega_2 \rightarrow \Omega_1$ est holomorphe.

(viii) Une fonction rationnelle est holomorphe sur l'ouvert de \mathbf{C} complémentaire de ses pôles.

Démonstration. — Le (i) est évident. Le (ii) et le (iv) suivent de l'exemple V.1.6. Le (iii) se démontre en constatant que $\frac{1}{z} = \sum_{n=0}^{+\infty} \frac{(z_0 - z)^n}{z_0^{n+1}}$, si $|z - z_0| < |z_0|$. Le (v) suit du lemme V.1.5. Le (vi) et (vii) pourraient se démontrer directement, mais nous attendrons d'en savoir plus pour en donner une démonstration élégante (cf. cor. V.2.7). Finalement, le (viii) se déduit des (i), (iii) et (v) et (vi).

Exercice V.1.12. — (i) Montrer que $\sin z$ est holomorphe sur \mathbf{C} , et que $\cotg \pi z$ est holomorphe sur $\mathbf{C} - \mathbf{Z}$.

(ii) Montrer que $z \mapsto |z|^2$ est \mathcal{C}^∞ au sens réel sur $\mathbf{C} \cong \mathbf{R}^2$, mais n'est pas holomorphe.

Remarque V.1.13. — (i) On peut aussi considérer f comme une fonction, à valeurs complexes, des variables $x = \operatorname{Re}(z)$ et $y = \operatorname{Im}(z)$. En notant P et Q les parties réelle et imaginaire de f , cela permet de voir f comme une fonction de classe \mathcal{C}^1 (et même de classe \mathcal{C}^∞) d'un ouvert de \mathbf{R}^2 dans \mathbf{R}^2 . La dérivabilité de f au sens complexe en z_0 s'exprime alors par le fait que la différentielle de f en z_0 est une similitude, ce qui se traduit par les *relations de Cauchy-Riemann*

$$\frac{\partial P}{\partial x}(z_0) = \frac{\partial Q}{\partial y}(z_0) = \operatorname{Re}(f'(z_0)) \quad \text{et} \quad \frac{\partial P}{\partial y}(z_0) = -\frac{\partial Q}{\partial x}(z_0) = -\operatorname{Im}(f'(z_0))$$

entre les dérivées partielles des parties réelle et imaginaire de f en z_0 . Le jacobien de f en z_0 est alors $|f'(z_0)|^2$; en particulier, sa nullité est équivalente à celle de $f'(z_0)$.

(ii) Comme les similitudes ont comme propriété de conserver les angles de vecteurs (mais pas les longueurs), une fonction holomorphe f hérite de cette propriété, et donc *conserve les angles*, ce qui signifie, par exemple, qu'elle transforme deux courbes se coupant à angle droit en z_0 en deux courbes se coupant à angle droit en $f(z_0)$, si $f'(z_0) \neq 0$.

(iii) Réciproquement, soit $f : \Omega \rightarrow \mathbf{C}$ une fonction, de classe \mathcal{C}^1 en tant que fonction de x et y . Soient $\frac{\partial}{\partial z}$ et $\frac{\partial}{\partial \bar{z}}$, les opérateurs différentiels définis par

$$\frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \quad \text{et} \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right).$$

Ces notations sont justifiées par les formules

$$\frac{\partial}{\partial z} z = 1, \quad \frac{\partial}{\partial z} \bar{z} = 0 \quad \text{et} \quad \frac{\partial}{\partial \bar{z}} z = 0, \quad \frac{\partial}{\partial \bar{z}} \bar{z} = 1.$$

On en déduit que l'on a $f(z_0 + h) = f(z_0) + \frac{\partial f}{\partial z}(z_0) \cdot h + \frac{\partial f}{\partial \bar{z}}(z_0) \cdot \bar{h} + o(|h|)$ au voisinage de $h = 0$. La différentielle de f est une similitude, si et seulement si $\frac{\partial f}{\partial \bar{z}}(z_0) = 0$, et donc f est dérivable au sens complexe en z_0 , si et seulement si $\frac{\partial f}{\partial \bar{z}}(z_0) = 0$. Ceci permet de voir une fonction de classe \mathcal{C}^1 au sens complexe, comme une fonction qui « ne dépend que de z et pas de \bar{z} »⁽⁵⁾, et permet d'expliquer un peu le miracle de la prop. V.2.6.

Exercice V.1.14. — (i) Soit $P \in \mathbf{C}[X]$ de degré ≥ 2 . Montrer que, si P ne s'annule pas dans \mathbf{C} , alors $1/P$ et ses dérivées partielles par rapport à x et y sont de carrés sommables sur $\mathbf{C} \cong \mathbf{R}^2$. En déduire que la transformée de Fourier de $1/P$ (vu comme fonction de x et y) est identiquement nulle, puis que \mathbf{C} est algébriquement clos.

(ii) Adapter l'argument ci-dessus pour n'utiliser que la transformée de Fourier dans $L^1(\mathbf{C})$.

3.2. Théorème des zéros isolés et unicité du prolongement analytique

Si f est analytique sur Ω , et si $z_0 \in \Omega$, on définit l'ordre du zéro $v_{z_0}(f) \in \mathbf{N} \cup \{+\infty\}$ de f en z_0 comme la valuation $v_T(F)$ de la série entière F telle que $f(z) = F(z - z_0)$. Si cet ordre est $+\infty$, c'est que la fonction f est identiquement nulle dans un voisinage de z_0 (i.e. il existe $r > 0$ tel que $f(z) = 0$ si $z \in D(z_0, r^-)$). Si cet ordre est fini, égal à k , on peut factoriser F sous la forme $F = T^k G$, avec $G(0) \neq 0$. Par continuité, $G(z - z_0)$ ne s'annule pas dans un voisinage de z_0 , ce qui prouve que z_0 est le seul zéro de f dans ce voisinage.

⁽⁵⁾Ceci ne doit pas être pris littéralement vu que \bar{z} et z ne sont pas indépendants, mais se comprend bien si on regarde un polynôme en $x = \frac{1}{2}(z + \bar{z})$ et $y = \frac{1}{2i}(z - \bar{z})$ comme un polynôme en z et \bar{z} : on voit que $P(x, y)$ définit une fonction holomorphe si et seulement si le polynôme $Q(z, \bar{z}) = P(\frac{1}{2}(z + \bar{z}), \frac{1}{2i}(z - \bar{z}))$ n'a pas de terme en \bar{z} .

Théorème V.1.15. — (des zéros isolés) Soit Ω un ouvert connexe de \mathbf{C} . Si f est une fonction holomorphe sur Ω qui n'est pas identiquement nulle, et si $z_0 \in \Omega$ est un zéro de f , alors il existe $r > 0$ tel que z_0 soit le seul zéro de f sur $D(z_0, r)$.

Démonstration. — D'après la discussion précédant le théorème, z_0 est un zéro non isolé si et seulement si $v_{z_0}(f) = +\infty$. L'ensemble des zéros non isolés est donc fermé car c'est l'intersection des fermés $\{z, f^{(n)}(z) = 0\}$, pour $n \in \mathbf{N}$, et ouvert car si $f^{(n)}(z_0) = 0$ pour tout n , alors f est identiquement nulle dans $D(z_0, r^-)$, si f est développable en série entière dans $D(z_0, r^-)$. Comme Ω est supposé connexe, on en déduit que l'ensemble des zéros non isolés est soit Ω (auquel cas f est identiquement nulle), soit l'ensemble vide. Ceci permet de conclure.

Exercice V.1.16. — Peut-on trouver une fonction \mathcal{C}^∞ , non identiquement nulle, à support compact dans \mathbf{R} , dont la transformée de Fourier soit à support compact ?

Corollaire V.1.17. — (Unicité du prolongement analytique) Soient $\Omega \subset \Omega'$ deux ouverts non vides de \mathbf{C} . Si Ω' est connexe, et si f et g sont deux fonctions analytiques sur Ω' ayant la même restriction à Ω , alors $f = g$ sur Ω' tout entier.

Démonstration. — Il suffit d'appliquer le théorème des zéros isolés à $f - g$.

Remarque V.1.18. — Soit f une fonction holomorphe sur un ouvert Ω de \mathbf{C} , et soit Ω' un ouvert connexe de \mathbf{C} contenant Ω . Il est, en général, impossible de trouver une fonction holomorphe g sur Ω' dont la restriction à Ω soit f . Dans le cas où c'est possible, le corollaire précédent montre que g est unique ; elle est appelée *le prolongement analytique de f à Ω'* .

On prendra garde au problème suivant. Si Ω_1 et Ω_2 sont deux ouverts connexes, contenant Ω , sur lesquels f admet des prolongements analytiques f_1 et f_2 , on ne peut pas en conclure, en général, que $f_1 = f_2$ sur $\Omega_1 \cap \Omega_2$ (le problème étant que $\Omega_1 \cap \Omega_2$ n'a aucune raison d'être connexe). Le logarithme fournira un exemple de ce phénomène. Évidemment, si f admet un prolongement analytique à \mathbf{C} tout entier, ce prolongement analytique est véritablement unique.

Exemple V.1.19. — La fonction analytique f définie sur $D(0, 1^-)$ par $f(z) = \sum_{n=0}^{+\infty} z^n$ admet un prolongement analytique à $\mathbf{C} - \{1\}$: en effet, elle coïncide⁽⁶⁾, sur $D(0, 1^-)$, avec

⁽⁶⁾On a donc envie d'écrire $1 + 2 + 4 + 8 + \dots = -1$, ce qu'Euler aurait fait sans aucun scrupule. Dans la même veine, on a (ex. VII.3.5)

$$1 + 1 + 1 + 1 + \dots = \zeta(0) = -\frac{1}{2}, \quad 1 + 2 + 3 + 4 + \dots = \zeta(-1) = -\frac{1}{12}, \quad 1 + 4 + 9 + 16 + \dots = \zeta(-2) = 0, \dots$$

Manipuler des séries divergentes est amusant mais demande un certain doigté. Par exemple, il ne faut pas confondre $\sum_{n=0}^{+\infty} 1 = 1 + \zeta(0) = \frac{1}{2}$ avec $\sum_{n=1}^{+\infty} 1 = \zeta(0) = -\frac{1}{2}$. Les physiciens ont acquis une dextérité certaine en la matière ; les mathématiciens ont plus de complexes depuis que Cauchy a déclaré, dans son *Cours d'analyse* à l'École polytechnique, paru en 1821 : « Je me suis vu forcé d'admettre plusieurs propositions qui paraîtront peut-être un peu dures au premier abord. Par exemple (...) qu'une série divergente n'a pas de somme. »

$\frac{1}{1-z}$ qui est holomorphe sur $\mathbf{C} - \{1\}$. Cet exemple n'est qu'à moitié convaincant car on est parti de la fonction $\frac{1}{1-z}$ qui est visiblement holomorphe sur $\mathbf{C} - \{1\}$, mais on verra des exemples moins banals plus tard (ex. V.2.20 et th. VII.3.4 et VII.4.4, par exemple).

3.3. Principe du maximum

Théorème V.1.20. — (Principe du maximum) *Si Ω est un ouvert connexe de \mathbf{C} , si $z_0 \in \Omega$, si f holomorphe sur Ω , et si $|f|$ admet un maximum local en z_0 , alors f est constante sur Ω .*

Démonstration. — Si f n'est pas constante sur Ω , la fonction $f - f(z_0)$ n'est pas identiquement nulle sur Ω , et Ω étant connexe, le développement en série entière de $f - f(z_0)$ autour de z_0 n'est pas identiquement nul. Il existe donc $k \in \mathbf{N} - \{0\}$ et $\alpha \in \mathbf{C}^*$ tels que $f(z) = f(z_0) + \alpha \cdot (z - z_0)^k + (z - z_0)^k \varepsilon_0(z)$, avec $\lim_{z \rightarrow z_0} \varepsilon_0(z) = 0$.

- Si $f(z_0) = 0$, alors z_0 est un zéro isolé de f , et $|f(z)| > |f(z_0)|$, quel que soit z dans un voisinage de z_0 ; en particulier, $|f|$ n'admet pas de maximum local en z_0 .

- Si $f(z_0) \neq 0$, soit β avec $\beta^k = \alpha^{-1} f(z_0)$. On a alors $f(z_0 + t\beta) = f(z_0)(1 + t^k + t^k \varepsilon_1(t))$, avec $\lim_{t \rightarrow 0} \varepsilon_1(t) = 0$. Ceci implique que $|f(z_0 + t\beta)| > |f(z_0)|$, si $t > 0$ est assez petit, et donc que $|f|$ n'admet pas de maximum local en z_0 .

Ceci permet de conclure.

Remarque V.1.21. — Le principe du maximum est souvent utilisé sous la forme : *une fonction holomorphe atteint son maximum au bord*. Autrement dit, si K est un compact et si f est holomorphe sur un ouvert contenant K , alors le maximum de $|f|$ sur K est atteint sur la frontière de K .

Exercice V.1.22. — Soit $P \in \mathbf{C}[X]$ non constant. Montrer que, si P ne s'annule pas sur \mathbf{C} , alors $|1/P|$ atteint son maximum en un point de \mathbf{C} . En déduire que \mathbf{C} est algébriquement clos.

Exercice V.1.23. — Soient $a < b \in \mathbf{R}$, et soit f une fonction holomorphe dans un voisinage ouvert de la bande verticale $B = \{z \in \mathbf{C}, a \leq \operatorname{Re}(z) \leq b\}$. On suppose que f est bornée sur les droites $\operatorname{Re}(z) = a$ et $\operatorname{Re}(z) = b$.

(i) Montrer que, s'il existe $C > 0$ et $c > 0$ tels que $|f(z)| \leq Ce^{c|\operatorname{Im}(z)|}$, quel que soit $z \in B$, alors f est bornée sur B . (On considérera la fonction $e^{\varepsilon z^2} f(z)$.)

(ii) Montrer de même que, s'il existe $C > 0$, $c > 0$, et $N \in \mathbf{N}$ tels que $|f(z)| \leq Ce^{c|\operatorname{Im}(z)|^N}$, quel que soit $z \in B$, alors f est bornée sur B .

(iii) Construire une fonction holomorphe sur \mathbf{C} , bornée sur chacune des droites $\operatorname{Re}(z) = a$ et $\operatorname{Re}(z) = b$, et non bornée sur B (penser à des exponentielles d'exponentielles).

Exercice V.1.24. — (Lemme de Schwarz) Soit $D = \{z \in \mathbf{C}, |z| < 1\}$, et soit f holomorphe de D dans D vérifiant $f(0) = 0$.

(i) Montrer que $|f(z)| \leq |z|$, quel que soit $z \in D$.

(ii) Montrer que, si $|f'(0)| = 1$, alors $f(z) = f'(0)z$ quel que soit $z \in D$.

(iii) Montrer que si f est une bijection holomorphe de D dans D , et si $f(0) = 0$, alors il existe $\theta \in [0, \pi]$ tel que $f(z) = e^{2i\theta}z$, quel que soit $z \in D$.

Exercice V.1.25. — Soit $\mathcal{H} = \{z \in \mathbf{C}, \operatorname{Im}(z) > 0\}$ le demi-plan de Poincaré.

(i) Vérifier que, si $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbf{R})$ (groupe des matrices 2×2 , à coefficients dans \mathbf{R} , de déterminant 1), alors $\gamma \cdot z = \frac{az+b}{cz+d} \in \mathcal{H}$, et que l'on a $\gamma_1 \gamma_2 \cdot z = \gamma_1 \cdot (\gamma_2 \cdot z)$.

(ii) En déduire que $\gamma \mapsto \varphi_\gamma$, où φ_γ est la transformation $z \mapsto \gamma \cdot z$, est un morphisme de groupes de $\mathbf{SL}_2(\mathbf{R})$ dans le groupe $\operatorname{Aut}(\mathcal{H})$ des bijections holomorphes de \mathcal{H} .

(iii) Montrer que, si $z \in \mathcal{H}$, il existe $\gamma \in \mathbf{SL}_2(\mathbf{R})$, avec $z = \gamma \cdot i$. En déduire que, si $\varphi \in \operatorname{Aut}(\mathcal{H})$, alors il existe $\gamma \in \mathbf{SL}_2(\mathbf{R})$ tel que i soit un point fixe de $\varphi_\gamma \circ \varphi$.

(iv) Soit $h(z) = \frac{i-z}{i+z}$. Montrer que h est une bijection holomorphe de \mathcal{H} dans D dont la réciproque h^{-1} est donnée par la formule $h^{-1}(z) = i \frac{1-z}{1+z}$.

(v) Montrer, en utilisant le lemme de Schwarz (ex. V.1.24), que si $\varphi \in \operatorname{Aut}(\mathcal{H})$ fixe i , alors il existe $\theta \in [0, \pi]$ tel que $\varphi(z) = \frac{z \cos \theta + \sin \theta}{-z \sin \theta + \cos \theta}$. (On considérera $\psi = h \circ \varphi \circ h^{-1}$.)

(vi) Montrer que $\gamma \mapsto \varphi_\gamma$ induit une surjection de $\mathbf{SL}_2(\mathbf{R})$ sur $\operatorname{Aut}(\mathcal{H})$; en déduire que $\operatorname{Aut}(\mathcal{H})$ est isomorphe à $\mathbf{SL}_2(\mathbf{R})/\{\pm 1\}$.

V.2. La formule intégrale de Cauchy et ses conséquences

1. Généralités sur les chemins

Un *chemin* γ (sous-entendu \mathcal{C}^1 par morceaux) dans un ouvert Ω de \mathbf{C} est une application $\gamma : [a, b] \rightarrow \Omega$, continue et \mathcal{C}^1 par morceaux, d'un intervalle compact $[a, b]$ de \mathbf{R} dans Ω . Le point $\gamma(a)$ est l'*origine* de γ et $\gamma(b)$ en est l'*extrémité*. On dit que γ est un *lacet* si son origine et son extrémité coïncident. Si $t \in [a, b]$ est un point où les dérivées à droite et à gauche de γ sont distinctes, on dit que $\gamma(t)$ est un *point anguleux* de γ . Par hypothèse, γ n'a qu'un nombre fini de points anguleux.

Si $\gamma(t) = x(t) + iy(t)$, on définit la *longueur* $\operatorname{lg}(\gamma)$ de γ par la formule

$$\operatorname{lg}(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt = \int_a^b |\gamma'(t)| dt.$$

La longueur est invariante par *reparamétrage* (un reparamétrage de γ est un chemin de la forme $\gamma_1 = \gamma \circ \varphi$, où $\varphi : [a_1, b_1] \rightarrow [a, b]$ est une bijection croissante de classe \mathcal{C}^1). En effet, on a

$$\begin{aligned} \operatorname{lg}(\gamma_1) &= \int_{a_1}^{b_1} \sqrt{(x \circ \varphi)'(t)^2 + (y \circ \varphi)'(t)^2} dt = \int_{a_1}^{b_1} \sqrt{(x' \circ \varphi)(t)^2 \varphi'(t)^2 + (y' \circ \varphi)(t)^2 \varphi'(t)^2} dt \\ &= \int_{a_1}^{b_1} \sqrt{(x' \circ \varphi)(t)^2 + (y' \circ \varphi)(t)^2} \varphi'(t) dt = \int_a^b \sqrt{x'(s)^2 + y'(s)^2} ds = \operatorname{lg}(\gamma). \end{aligned}$$

Si $\gamma : [a, b] \rightarrow \Omega$ est un chemin, on définit le *chemin opposé* $\gamma^{\text{opp}} : [a, b] \rightarrow \Omega$ de γ , par $\gamma^{\text{opp}}(t) = \gamma(a + b - t)$.

Si $\gamma_1 : [a_1, b_1] \rightarrow \Omega$ et $\gamma_2 : [a_2, b_2] \rightarrow \Omega$ sont deux chemins, on dit que γ_1 et γ_2 sont *composables* si l'origine de γ_2 coïncide avec l'extrémité de γ_1 , et on définit le chemin $\gamma_1 \cdot \gamma_2 : [a_1, b_2 - a_2 + b_1] \rightarrow \Omega$, *composé* de γ_1 et γ_2 , en posant $\gamma_1 \cdot \gamma_2(t) = \gamma_1(t)$, si $t \in [a_1, b_1]$ et $\gamma_1 \cdot \gamma_2(t) = \gamma_2(t - a_2 + b_1)$, si $t \in [b_1, b_2 - a_2 + b_1]$.

Exemple V.2.1. — (chemins standard)

- Si $z_0 \in \mathbf{C}$ et $r > 0$, on note $C(z_0, r) : [0, 1] \rightarrow \mathbf{C}$ le chemin $t \mapsto z_0 + re^{2i\pi t}$; c'est le cercle de centre z_0 et de rayon r parcouru dans le sens direct. Sa longueur est $2\pi r$.
- Si $a, b \in \mathbf{C}$, on note $[a, b] : [0, 1] \rightarrow \mathbf{C}$ le chemin $t \mapsto a + t(b - a)$; c'est le segment d'origine a et d'extrémité b . Sa longueur est $|b - a|$.

2. Intégration le long d'un chemin

On identifie \mathbf{C} à \mathbf{R}^2 , en écrivant z sous la forme $z = x + iy$. Une 1-forme sur un ouvert Ω de \mathbf{C} est une expression⁽⁷⁾ du type $P dx + Q dy$, où P, Q sont des fonctions continues sur Ω . Si f est une fonction de classe \mathcal{C}^1 sur Ω , la différentielle df de f peut être vue comme une 1-forme : on a $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$.

Si $\omega = P dx + Q dy$ est une 1-forme sur Ω , et si $\gamma : [a, b] \rightarrow \Omega$ est un chemin dans Ω , on définit l'intégrale $\int_\gamma \omega$ par la formule

$$\int_\gamma \omega = \int_a^b P(\gamma(t))d(x(\gamma(t))) + Q(\gamma(t))d(y(\gamma(t))) = \int_a^b P(\gamma(t))(x \circ \gamma)'(t) + Q(\gamma(t))(y \circ \gamma)'(t) dt.$$

Dans l'intégrale ci-dessus, $(x \circ \gamma)'(t)$ et $(y \circ \gamma)'(t)$ sont bien définis sauf aux points anguleux, mais comme ceux-ci sont en nombre fini, cela n'affecte pas l'intégrale.

Théorème V.2.2. — (i) $\int_\gamma \omega$ est invariante par reparamétrage : si $\varphi : [a_1, b_1] \rightarrow [a, b]$ une bijection croissante de classe \mathcal{C}^1 , et si $\gamma_1 = \gamma \circ \varphi$, alors $\int_{\gamma_1} \omega = \int_\gamma \omega$.

(ii) $\int_{\gamma_{\text{opp}}} \omega = - \int_\gamma \omega$.

(iii) Si γ_1 et γ_2 sont composables, alors $\int_{\gamma_1 \cdot \gamma_2} \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega$.

(iv) Si $\omega = df$, alors $\int_\gamma \omega = f(\gamma(b)) - f(\gamma(a))$.

⁽⁷⁾Plus généralement, si Ω est un ouvert de \mathbf{R}^n , et si $p \leq n$, une p -forme sur Ω est une expression du type

$$\omega = \sum_{i_1 < i_2 < \dots < i_p} f_{i_1, \dots, i_p} dx_{i_1} \wedge \dots \wedge dx_{i_p},$$

où les f_{i_1, \dots, i_p} sont des fonctions continues sur Ω . (Une 0-forme est juste, par convention, une fonction continue sur Ω .) On voit aussi souvent une p -forme ω sur Ω comme une fonction $x \mapsto \omega_x$, continue sur Ω , à valeurs dans les formes p -linéaires alternées : si u_1, \dots, u_p sont p vecteurs de \mathbf{R}^n , avec $u_i = (u_{i,1}, \dots, u_{i,n})$, et si $x \in \Omega$, alors

$$\omega_x(u_1, \dots, u_p) = \sum_{i_1 < i_2 < \dots < i_p} f_{i_1, \dots, i_p}(x) \sum_{\sigma \in S_p} \text{sign}(\sigma) u_{\sigma(1), i_1} \dots u_{\sigma(p), i_p}.$$

Ces objets jouent un rôle très important en géométrie différentielle. L'exemple le plus simple est la 1-forme df , si f est de classe \mathcal{C}^1 sur Ω . Sa valeur en un point x est la forme linéaire $u \mapsto \lim_{t \rightarrow 0} \frac{f(x+tu) - f(x)}{t} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) u_i$.

Démonstration. — Le (i) résulte du calcul suivant :

$$\begin{aligned}\int_{\gamma_1} \omega &= \int_{a_1}^{b_1} (P(\gamma \circ \varphi(t))(x \circ \gamma \circ \varphi)'(t) + Q(\gamma \circ \varphi(t))(y \circ \gamma \circ \varphi)'(t)) dt \\ &= \int_{\varphi(a_1)}^{\varphi(b_1)} (P(\gamma(s))(x \circ \gamma)'(s) + Q(\gamma(s))(y \circ \gamma)'(s)) ds = \int_{\gamma} \omega.\end{aligned}$$

Le (ii) suit, via le changement de variable $t \mapsto a+b-t$, de ce que les bornes d'intégration sont échangées.

Le (iii) est immédiat.

Le (iv) suit du calcul suivant :

$$\begin{aligned}\int_{\gamma} df &= \int_a^b \left(\frac{\partial f}{\partial x}(\gamma(t))(x \circ \gamma)'(t) + \frac{\partial f}{\partial y}(\gamma(t))(y \circ \gamma)'(t) \right) dt \\ &= \int_a^b (f \circ \gamma)'(t) dt = f(\gamma(b)) - f(\gamma(a)).\end{aligned}$$

On note dz la différentielle $dx + i dy$. Si f est une fonction continue sur Ω , la 1-forme $f(z)dz$ n'est donc autre que $f(x+iy) dx + i f(x+iy) dy$. Si $\gamma : [a, b] \rightarrow \Omega$ est un chemin \mathcal{C}^1 par morceaux, on a alors

$$\int_{\gamma} f(z) dz = \int_a^b f(\gamma(t))((x \circ \gamma)'(t) + i(y \circ \gamma)'(t)) dt = \int_a^b f(\gamma(t))\gamma'(t) dt.$$

La majoration du lemme suivant est immédiate, via la formule $\lg(\gamma) = \int_a^b |\gamma'(t)| dt$.

Lemme V.2.3. — $|\int_{\gamma} f(z) dz| \leq \lg(\gamma) \cdot \sup_{z \in \gamma([a,b])} |f(z)|$.

Exemple V.2.4. — (Intégration sur les chemins standard)

(i) Si $z_0 \in \mathbf{C}$ et $r > 0$, alors $\int_{C(z_0, r)} f(z) dz = 2i\pi r \int_0^1 f(z_0 + re^{2i\pi t})e^{2i\pi t} dt$.

(ii) Si $a, b \in \mathbf{C}$, alors $\int_{[a,b]} f(z) dz = (b-a) \int_0^1 f(a + t(b-a)) dt$.

3. Holomorphie des fonctions dérivables au sens complexe

Théorème V.2.5. — (Formule intégrale de Cauchy, 1825) Soit f de classe \mathcal{C}^1 au sens complexe sur Ω , ouvert de \mathbf{C} , et soient $z_0 \in \Omega$ et $r > 0$, tels que $D(z_0, r) \subset \Omega$. Alors, pour tout $z \in D(z_0, r^-)$, on a

$$f(z) = \frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{w-z} dw.$$

Ce théorème, dont nous ferons la démonstration plus loin (n° 3 du § VI.1), a des conséquences totalement surprenantes pour quelqu'un de familier avec la théorie des fonctions

d'une variable réelle⁽⁸⁾. La prop. V.2.6 ci-dessous en est un premier exemple, puisqu'il montre qu'une fonction de classe \mathcal{C}^1 au sens complexe est en fait holomorphe (et donc, a fortiori, de classe \mathcal{C}^∞). Le (i) de la rem. V.2.8 et le th. V.2.12 fournissent d'autres exemples de phénomènes un peu miraculeux.

Proposition V.2.6. — Si f vérifie la formule de Cauchy, et si on définit $a_n \in \mathbf{C}$, pour $n \in \mathbf{N}$, par

$$a_n = \frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{(w - z_0)^{n+1}} dw,$$

la série $F(T) = \sum_{n=0}^{+\infty} a_n T^n$ a un rayon de convergence $\geq r$, et on a $f(z) = F(z - z_0)$ quel que soit $z \in D(z_0, r^-)$.

Démonstration. — Soit $M = \sup_{w \in C(z_0, r)} |f(w)|$ (M est fini car f est continue sur le compact $C(z_0, r)$). On a alors

$$|a_n| \leq \frac{1}{2\pi} \lg(C(z_0, r)) \cdot \left(\sup_{w \in C(z_0, r)} \left| \frac{f(w)}{(w - z_0)^{n+1}} \right| \right) \leq M r^{-n}.$$

On en déduit que $\rho(F) \geq r$. De plus, si $|z - z_0| < r$, alors

$$\begin{aligned} \sum_{n=0}^{+\infty} a_n (z - z_0)^n &= \frac{1}{2i\pi} \sum_{n=0}^{+\infty} \int_{C(z_0, r)} \frac{(z - z_0)^n f(w)}{(w - z_0)^{n+1}} dw \\ &= \frac{1}{2i\pi} \int_{C(z_0, r)} \sum_{n=0}^{+\infty} \frac{(z - z_0)^n f(w)}{(w - z_0)^{n+1}} dw = \frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{w - z} dw. \end{aligned}$$

On conclut en utilisant la formule intégrale de Cauchy. (Pour justifier l'échange des signes \int et \sum ci-dessus, on peut, par exemple, faire appel au théorème de convergence dominée en majorant $|\sum_{n=0}^N \frac{(z - z_0)^n f(w)}{(w - z_0)^{n+1}}|$ sur $C(z_0, r)$ par $M \sum_{n=0}^{+\infty} \frac{|z - z_0|^n}{r^{n+1}} = \frac{M}{r - |z - z_0|}$.)

Corollaire V.2.7. — (i) Si f est holomorphe et ne s'annule pas sur Ω , alors $1/f$ est holomorphe sur Ω ; sa dérivée est $-\frac{f'}{f^2}$.

⁽⁸⁾• La fonction $\varphi_0(x) = e^{-x^{-2}}$, prolongée par continuité en 0, est \mathcal{C}^∞ sur \mathbf{R} , mais n'est pas développable en série entière autour de 0, puisque toutes ses dérivées sont nulles en 0 et $\varphi_0(x) > 0$ si $x \neq 0$.

• La fonction $\frac{1}{x^2+1}$ est analytique sur \mathbf{R} , mais n'est pas développable en série entière sur $](1+\varepsilon), 1+\varepsilon[$, si $\varepsilon > 0$, car le développement de Taylor $\sum_{n=0}^{+\infty} (-x^2)^n$ en 0 a un rayon de convergence égal à 1.

• Il existe des fonctions continues sur \mathbf{R} n'ayant de dérivée en aucun point (Weierstrass, 1875); en prenant la primitive d'une telle fonction cela montre qu'il existe des fonctions de classe \mathcal{C}^1 sur \mathbf{R} qui ne sont pas de classe \mathcal{C}^2 et donc a fortiori, pas analytiques sur \mathbf{R} .

D'un autre côté, la théorie des fonctions holomorphes a été développée avant la théorie des fonctions d'une variable réelle, et c'est les pathologies de cette dernière qui ont beaucoup surpris les mathématiciens du 19-ième siècle (et un peu ceux du 20-ième). Par exemple, H. Poincaré a eu quelques ennuis avec les fonctions \mathcal{C}^∞ qui ne sont pas somme de leur série de Taylor...

(ii) Si $f : \Omega_1 \rightarrow \Omega_2$ est holomorphe et si $g : \Omega_2 \rightarrow \mathbf{C}$ est holomorphe, alors $g \circ f : \Omega_1 \rightarrow \mathbf{C}$ est holomorphe ; sa dérivée est $(g' \circ f) f'$.

(iii) Si $f : \Omega_1 \rightarrow \Omega_2$ est holomorphe bijective, et si f' ne annule pas⁽⁹⁾ dans Ω , alors sa réciproque $f^{-1} : \Omega_2 \rightarrow \Omega_1$ est holomorphe ; sa dérivée est $\frac{1}{f' \circ f^{-1}}$.

Démonstration. — Il suffit de recopier la démonstration du cas réel pour montrer que $1/f$ (resp. $g \circ f$, resp. f^{-1}) est \mathcal{C}^1 au sens complexe sur Ω (resp. sur Ω_1 , resp. sur Ω_2) ; la prop. V.2.6 permet d'en déduire l'holomorphie de $1/f$ (resp. $g \circ f$, resp. f^{-1}). (On aurait aussi pu démontrer directement que $1/f$ (resp. $g \circ f$, resp. f^{-1}) est développable en série entière autour de chaque point, mais ça aurait été nettement plus fatigant.)

Remarque V.2.8. — (i) La formule $f(z) = F(z - z_0)$ montre que les a_n sont les coefficients du développement de Taylor de f en z_0 . Autrement dit, une fonction f , holomorphe sur un ouvert Ω , est somme de sa série de Taylor en z_0 sur tout disque ouvert $D(z_0, r^-)$ contenu dans Ω . De plus, on a

$$\frac{1}{n!} f^{(n)}(z_0) = \frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{(w - z_0)^{n+1}} dw.$$

On en déduit la majoration (appelée *inégalité de Cauchy*)

$$\left| \frac{1}{n!} f^{(n)}(z_0) \right| \leq r^{-n} \sup_{w \in C(z_0, r)} |f(w)| = r^{-n} \sup_{\theta \in [0, 2\pi]} |f(z_0 + re^{i\theta})|.$$

(ii) Plus généralement, si $K \subset \Omega$ est un compact, si U est un ouvert contenant K et dont l'adhérence \bar{U} est un compact contenu dans Ω , alors, quel que soit $n \in \mathbf{N}$, il existe $M = M(n, K, U)$, tel que

$$\sup_{z \in K} |f^{(n)}(z)| \leq M \sup_{z \in \bar{U}} |f(z)|.$$

En effet, si $r \leq d(K, \Omega - U)$, alors $C(z, r) \subset \bar{U} \subset \Omega$ quel que soit $z \in K$, et le (i) montre que $M = n!r^{-n}$ convient.

Corollaire V.2.9. — (Théorème de Liouville) Si f est une fonction holomorphe sur \mathbf{C} tout entier, et si f est bornée, alors f est constante.

Démonstration. — Par hypothèse, il existe $M > 0$ tel que $|f(z)| \leq M$, quel que soit $z \in \mathbf{C}$. On déduit de l'inégalité de Cauchy que, si $n \in \mathbf{N}$, alors $|\frac{1}{n!} f^{(n)}(0)| \leq r^{-n} M$, quel que soit $r > 0$. En faisant tendre r vers $+\infty$, on en déduit que $f^{(n)}(0) = 0$, si $n \geq 1$. Or f étant holomorphe sur \mathbf{C} tout entier, est somme de sa série de Taylor en 0 en tout point de \mathbf{C} et donc $f(z) = f(0)$, pour tout $z \in \mathbf{C}$. Ceci permet de conclure.

Corollaire V.2.10. — (Théorème de d'Alembert-Gauss) \mathbf{C} est algébriquement clos.

⁽⁹⁾Cette hypothèse est en fait inutile (cf. rem. V.3.3, deuxième point).

Démonstration. — Soit $P \in \mathbf{C}[T]$ ne s'annulant pas sur \mathbf{C} . Alors $f = 1/P$ est une fonction holomorphe. De plus f est bornée sur \mathbf{C} car elle a une limite quand $|z| \rightarrow +\infty$, et est continue. On déduit du théorème de Liouville que f est constante, et donc que P est de degré 0. En prenant la contraposée, cela montre que, si $\deg P \geq 1$, alors P a un zéro dans \mathbf{C} , ce qui permet de conclure.

Exercice V.2.11. — Montrer que $f(z) = \sum_{n=0}^{+\infty} z^{n!}$ est holomorphe sur $D(0, 1^-)$, mais n'a de prolongement analytique à aucun ouvert contenant strictement $D(0, 1^-)$.

4. Construction de fonctions holomorphes

4.1. Séries de fonctions holomorphes

Soit Ω un ouvert de \mathbf{C} . Rappelons qu'une suite de fonctions $(f_n)_{n \in \mathbf{N}}$ converge uniformément sur tout compact vers f , si pour tout compact $K \subset \Omega$, et tout $\varepsilon > 0$, il existe $N(K, \varepsilon)$, tel que $|f_n(z) - f(z)| < \varepsilon$, si $z \in K$ et $n \geq N(K, \varepsilon)$. Une série $\sum_{n \in \mathbf{N}} u_n$ converge uniformément sur tout compact, si la suite de ses somme partielles le fait ; elle converge normalement sur tout compact, si pour tout compact $K \subset \Omega$, on a $\sum_{n \in \mathbf{N}} \|u_n\|_K < +\infty$, où $\|u_n\|_K = \sup_{z \in K} |u_n(z)|$ est la norme de u_n pour la convergence uniforme sur K .

La convergence normale sur tout compact implique la convergence uniforme sur tout compact qui implique la convergence simple.

Théorème V.2.12. — Soit Ω un ouvert de \mathbf{C} .

(i) Si $(f_n)_{n \in \mathbf{N}}$ est une suite de fonctions holomorphes sur Ω convergeant uniformément sur tout compact de Ω , alors la limite f de la suite f_n est holomorphe sur Ω et, pour tout k , la suite $(f_n^{(k)})_{n \in \mathbf{N}}$ converge uniformément vers $f^{(k)}$ sur tout compact de Ω .

(ii) Si $(u_n)_{n \in \mathbf{N}}$ est une suite de fonctions holomorphes sur Ω telle que la série $\sum_{n \in \mathbf{N}} u_n$ converge uniformément (resp. normalement) sur tout compact de Ω , alors la somme f de la série est holomorphe sur Ω et, pour tout k , la série $\sum_{n \in \mathbf{N}} u_n^{(k)}$ converge uniformément (resp. normalement) vers $f^{(k)}$ sur tout compact de Ω .

Démonstration. — Le (i) se déduit du (ii), et pour montrer le (ii), il suffit de prouver que f est holomorphe, le reste s'en déduisant en utilisant le (ii) de la remarque V.2.8 ci-dessus.

Soit $z_0 \in \Omega$, et soit $r > 0$ tel que $D(z_0, r) \subset \Omega$. Alors, d'après la formule de Cauchy, on a, pour tout $z \in K$,

$$\begin{aligned} f(z) &= \sum_{n \in \mathbf{N}} u_n(z) = \frac{1}{2i\pi} \sum_{n \in \mathbf{N}} \int_{C(z_0, r)} \frac{u_n(w)}{w - z} dw \\ &= \frac{1}{2i\pi} \int_{C(z_0, r)} \sum_{n \in \mathbf{N}} \frac{u_n(w)}{w - z} dw = \frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{w - z} dw, \end{aligned}$$

l'interversion des signes \int et \sum ci-dessus étant justifiée par la convergence uniforme (resp. normale) de la série $\sum_{n \in \mathbf{N}} u_n(z)$ sur le compact $C(z_0, r)$. La prop. V.2.6 permet d'en déduire l'holomorphie de f dans $D(z_0, r^-)$, ce qui permet de conclure.

Exercice V.2.13. — Soient Ω un ouvert de \mathbf{C} et $(f_n)_{n \in \mathbf{N}}$ une suite de fonctions holomorphes sur Ω tendant simplement vers f . Montrer que, si pour tout compact K de Ω , il existe $M_K > 0$ tel que $|f_n(z)| \leq M_K$, pour tout $n \in \mathbf{N}$ et $z \in K$, alors f est holomorphe.

Exercice V.2.14. — (i) Montrer que si K est un compact de \mathbf{C} , il existe $n(K)$ telle que la série $\sum_{n=n(K)}^{+\infty} \left(\frac{1}{z+n} + \frac{1}{z-n} \right)$ soit uniformément convergente sur K . En déduire que la série $\frac{1}{z} + \sum_{n=1}^{+\infty} \left(\frac{1}{z+n} + \frac{1}{z-n} \right)$ définit une fonction F , holomorphe sur $\mathbf{C} - \mathbf{Z}$.

(ii) Montrer que $G(z) = F(z) - \pi \cotg \pi z$ est impaire et périodique de période 1.

(iii) Montrer que G se prolonge par continuité en une fonction holomorphe sur \mathbf{C} .

(iv) Montrer que G est bornée sur \mathbf{C} ; en déduire que $\frac{1}{z} + \sum_{n=1}^{+\infty} \left(\frac{1}{z+n} + \frac{1}{z-n} \right) = \pi \cotg \pi z$, si $z \in \mathbf{C} - \mathbf{Z}$.

(v) Soit $\sum_{n=0}^{+\infty} B_n \frac{t^n}{n!}$ le développement de Taylor de $\frac{t}{e^t - 1}$ en $t = 0$, et, si k est un entier ≥ 1 , soit $\zeta(2k) = \sum_{n=1}^{+\infty} \frac{1}{n^{2k}}$. Montrer que

$$\zeta(2k) = -\frac{1}{2} B_{2k} \frac{(2i\pi)^{2k}}{(2k)!}.$$

En déduire que $\pi^{-2k} \zeta(2k)$ est un nombre rationnel (Euler, 1734).

4.2. Produits infinis de fonctions holomorphes

Soit I un ensemble dénombrable, et soit $(u_i)_{i \in I}$ une suite de nombres complexes. On dit que le produit $\prod_{i \in I} u_i$ est *convergent* si $\sum_{i \in I} |u_i - 1| < +\infty$. Si $I_0 \subset I$ est l'ensemble des $i \in I$ tels que $u_i = 0$, la condition précédente équivaut à $\sum_{i \in I} |\log u_i| < +\infty$, où $\log : \mathbf{C}^* \rightarrow \mathbf{C}$ désigne la branche du logarithme dont la partie imaginaire appartient à $] -\pi, \pi]$ (en particulier $\log z = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} (z-1)^n$, si $|z-1| < 1$, et $\lim_{z \rightarrow 1} \frac{\log z}{z-1} = 1$). Si I est fini, le produit est convergent et sa valeur est le produit au sens usuel. Si I est infini, et si $n \mapsto i(n)$ est une bijection de \mathbf{N} sur I , alors $x_n = \prod_{k \leq n} u_{i(k)}$ est une suite convergente dont la limite ne dépend pas de la bijection de \mathbf{N} sur I choisie, et qui est, par définition, la valeur du produit infini. En effet :

- si $I_0 \neq \emptyset$, alors $x_n = 0$, pour tout n assez grand, et la valeur du produit est 0;
- si $I_0 = \emptyset$, on a $x_n = \exp \left(\sum_{k \leq n} \log u_{i(k)} \right)$, et comme $\sum_{i \in I} |\log u_i| < +\infty$, la série $\sum_{n \in \mathbf{N}} \log u_{i(n)}$ est convergente, et on a $\prod_{i \in I} x_i = \exp \left(\sum_{n \in \mathbf{N}} \log u_{i(n)} \right) = \exp \left(\sum_{i \in I} \log u_i \right)$.

En particulier, un produit convergent est nul si et seulement si un des termes du produit est nul.

Théorème V.2.15. — Soit Ω un ouvert connexe de \mathbf{C} , soit $(u_n)_{n \in \mathbf{N}}$ une suite de fonctions holomorphes sur \mathbf{C} telle que la série $\sum_{n \in \mathbf{N}} u_n$ converge normalement sur tout compact de Ω , et soit $f_n = \prod_{k=0}^n (1 + u_k)$, si $n \in \mathbf{N}$.

(i) La suite f_n converge uniformément sur tout compact (on dit que le produit infini $\prod_{n \in \mathbf{N}} (1 + u_n)$ converge uniformément sur tout compact), et la limite f de la suite f_n est une fonction holomorphe sur Ω .

(ii) Si aucune des u_n est identiquement -1 sur Ω , alors $f(z) = 0$ si et seulement si il existe $n \in \mathbf{N}$ tel que $1 + u_n(z) = 0$, et on a $v_z(f) = \sum_{n \in \mathbf{N}} v_z(1 + u_n)$, quel que soit $z \in \Omega$. De plus, la série $\sum_{n=0}^{+\infty} \frac{u'_n}{1+u_n}$ converge normalement sur tout compact de Ω sur lequel f ne

s'annule pas, et

$$\frac{f'(z)}{f(z)} = \sum_{n=0}^{+\infty} \frac{u'_n(z)}{1+u_n(z)}, \quad \text{si } f(z) \neq 0.$$

Démonstration. — Si K est un compact de Ω , soit $C(K) = \exp(\sum_{n \in \mathbf{N}} \|u_n\|_K)$; c'est une quantité finie, par hypothèse. Si $n \in \mathbf{N}$ et si $z \in K$, on a alors

$$\begin{aligned} \left| \prod_{k=0}^{n+1} (1+u_k(z)) - \prod_{k=0}^n (1+u_k(z)) \right| &= |u_{n+1}(z)| \cdot \left| \prod_{k=0}^n (1+u_k(z)) \right| \\ &\leq \|u_{n+1}\|_K \prod_{k=0}^n (1+\|u_k\|_K) \leq C(K) \|u_{n+1}\|_K. \end{aligned}$$

On en déduit, grâce à la convergence de $\sum_{n \in \mathbf{N}} \|u_n\|_K$, la convergence uniforme du produit sur K . La limite f est donc holomorphe sur Ω d'après le th. V.2.12.

Maintenant, si $z_0 \in \Omega$, on peut choisir $r > 0$ tel que $K = D(z_0, r) \subset \Omega$, et $N \in \mathbf{N}$ tel que $\|u_n\|_K \leq \frac{1}{2}$, si $n > N$. En particulier, $1+u_n$ ne s'annule pas sur K , et $|1+u_n(z)| \geq 1 - \|u_n\|_K \geq e^{-2\|u_n\|_K}$, quel que soit $z \in K = D(z_0, r)$. Soit $g_N = \prod_{n \geq N+1} (1+u_n)$. On déduit de ce qui précède la minoration $|g_N(z)| \geq \exp(-2 \sum_{n=N+1}^{+\infty} \|u_n\|_K)$, ce qui permet de prouver que g_N ne s'annule pas sur K . On en déduit que

$$v_{z_0}(f) = v_{z_0}\left(g_N \prod_{n=0}^N (1+u_n)\right) = \sum_{n=0}^N v_{z_0}(1+u_n) = \sum_{n=0}^{+\infty} v_{z_0}(1+u_n).$$

Maintenant, la convergence normale de la série des u_n implique, d'après le (ii) du th. V.2.12, celle de la série des u'_n . Comme $|1+u_n(z)| \geq \frac{1}{2}$, si $z \in K$ et $n \geq N$, on en déduit la convergence normale de la série des $\frac{u'_n}{1+u_n}$. De même, la convergence uniforme sur K de f_n vers f entraîne celle de f'_n vers f' ((i) du th. V.2.12). On en déduit que $\frac{f'_n(z)}{f_n(z)}$ tend vers $\frac{f'(z)}{f(z)}$, si $f(z) \neq 0$. Or $\frac{f'_n}{f_n} = \sum_{k=0}^n \frac{u'_k}{1+u_k}$ (cela suit, par récurrence de ce que $(v_1 v_2)' = v'_1 v_2 + v_1 v'_2$), et donc $\frac{f'(z)}{f(z)} = \sum_{n=0}^{+\infty} \frac{u'_n(z)}{1+u_n(z)}$, si $f(z) \neq 0$. Ceci permet de conclure.

Exercice V.2.16. — Reprendre la démonstration du théorème en utilisant la série des $\log u_n$ (attention au fait que $\log u_n$ n'est pas forcément holomorphe sur Ω tout entier).

Exercice V.2.17. — Montrer que $\prod_{n \geq 1} (1 - \frac{z^2}{n^2})$ est uniformément convergent sur tout compact de \mathbf{C} , et que sa valeur est $\frac{\sin \pi z}{\pi z}$. (On pourra utiliser les résultats de l'ex. V.2.14.)

4.3. Fonctions holomorphes définies par une intégrale

Théorème V.2.18. — Soit X un sous-ensemble mesurable de \mathbf{R}^m , et soit Ω un ouvert de \mathbf{C} . Soit $F : \Omega \times X \rightarrow \mathbf{C}$. On suppose que :

- $z \mapsto F(z, t)$ est holomorphe sur Ω , quel que soit $t \in X$;

• $t \mapsto F(z, t)$ est mesurable quel que soit $z \in \Omega$ et, quel que soit $z_0 \in \Omega$, il existe $r_{z_0} > 0$ et $g_{z_0} \in \mathcal{L}^1(X)$ tels que $D(z_0, r_{z_0}) \subset \Omega$ et $|F(z, t)| \leq g_{z_0}(t)$ quels que soient $z \in D(z_0, r_{z_0})$ et $t \in X$.

Alors la fonction $f : \Omega \rightarrow \mathbf{C}$ définie par $f(z) = \int_X F(z, t) dt$ est holomorphe sur Ω . De plus, si $k \in \mathbf{N}$, alors $(\frac{\partial}{\partial z})^k f(z) = \int_X (\frac{\partial}{\partial z})^k F(z, t) dt$.

Démonstration. — L'holomorphie étant une propriété locale, il suffit de prouver que, quel que soit $z_0 \in \Omega$, il existe $r > 0$ tel que f soit holomorphe sur $D(z_0, r^-)$. Fixons donc z_0 , et soit $r = r_{z_0}$. Comme $z \mapsto F(z, t)$ est holomorphe, quel que soit $t \in X$, on tire de la formule de Cauchy l'identité

$$f(z) = \frac{1}{2i\pi} \int_X \left(\int_{C(z_0, r)} \frac{F(w, t)}{w - z} dw \right) dt, \quad \text{si } z \in D(z_0, r^-).$$

Or on a $|\frac{F(w, t)}{w - z}| \leq \frac{g_{z_0}(t)}{r - |z - z_0|}$, si $(w, t) \in (C(z_0, r) \times X)$. Comme g_{z_0} est sommable sur X , la fonction $|\frac{F(w, t)}{w - z}|$ est sommable sur $C(z_0, r) \times X$, ce qui permet d'utiliser le théorème de Fubini pour permuter les deux intégrations. On obtient

$$f(z) = \frac{1}{2i\pi} \int_{C(z_0, r)} \left(\int_X \frac{F(w, t)}{w - z} dt \right) dw = \frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{w - z} dw, \quad \text{si } z \in D(z_0, r^-).$$

La prop. V.2.6 permet d'en déduire l'holomorphie de f sur $D(z_0, r^-)$.

Maintenant, on a $\frac{1}{k!} f^{(k)}(z_0) = \frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{(w - z_0)^{k+1}} dw$ d'après le (i) de la rem. V.2.8. En majorant $|\frac{F(w, t)}{(w - z_0)^{k+1}}|$ par $r^{-k-1} g_{z_0}(t)$ qui est sommable sur $C(z_0, r) \times X$, ce qui permet d'utiliser Fubini dans l'autre sens, on obtient

$$\frac{1}{k!} f^{(k)}(z_0) = \int_X \left(\frac{1}{2i\pi} \int_{C(z_0, r)} \frac{F(w, t)}{(w - z_0)^{k+1}} dw \right) dt = \int_X \frac{1}{k!} \left(\frac{\partial}{\partial z} \right)^k F(z_0, t) dt.$$

Ceci permet de conclure.

Corollaire V.2.19. — Soient Ω et D deux ouverts de \mathbf{C} . Si f est continue sur $\Omega \times D$, et si $z \mapsto f(z, w)$ est holomorphe sur Ω pour tout $w \in D$, alors $z \mapsto \int_\gamma f(z, w) dw$ est holomorphe sur Ω , quel que soit le chemin \mathcal{C}^1 par morceaux γ inclus dans D .

Démonstration. — On a $\int_\gamma f(z, w) dw = \int_0^1 f(z, \gamma(t)) \gamma'(t) dt$. Soit alors $z_0 \in \Omega$. On peut trouver $r > 0$ tel que $D(z_0, r) \subset \Omega$. La fonction continue $(z, t) \mapsto |f(z, \gamma(t))|$ est majorée sur le compact $D(z_0, r) \times [0, 1]$, et donc $f(z, \gamma(t)) \gamma'(t)$ admet un majorant du type $M |\gamma'(t)|$ sur $D(z_0, r) \times [0, 1]$. Comme $\int_0^1 |\gamma'(t)| dt < +\infty$, on est dans les conditions d'application du th. V.2.18, ce qui permet de conclure.

Exercice V.2.20. — (La fonction Γ dans le plan complexe).

(i) Montrer que $F(z) = \int_0^{+\infty} e^{-t} t^z \frac{dt}{t}$ converge si $\operatorname{Re}(z) > 0$, et que $F(z)$ est holomorphe sur le demi-plan $\operatorname{Re}(z) > 0$.

(ii) Montrer que, quel que soit $n \in \mathbf{N}$, on a $F(z) = \frac{F(z+n+1)}{z(z+1)\cdots(z+n)}$, si $\operatorname{Re}(z) > 0$. En déduire qu'il existe une unique fonction Γ , méromorphe sur \mathbf{C} , holomorphe en dehors de pôles simples aux entiers négatifs, telle que $\Gamma(z) = F(z)$ si $\operatorname{Re}(z) > 0$.

(iii) Montrer que $F(z)$ est bornée dans toute bande verticale $0 < a \leq \operatorname{Re}(z) \leq b < +\infty$; en déduire que Γ est à décroissance rapide à l'infini dans toute bande verticale de largeur finie, c'est-à-dire que, quels que soient $a < b \in \mathbf{R}$ et $N \in \mathbf{N}$, il existe $C(a, b, N)$ tel que l'on ait

$$|\Gamma(z)| \leq C(a, b, N) |\operatorname{Im}(z)|^{-N}, \quad \text{si } a \leq \operatorname{Re}(z) \leq b \text{ et } |\operatorname{Im}(z)| \geq 1.$$

V.3. Structure locale des fonctions holomorphes

1. Le théorème d'inversion locale holomorphe

Soient Ω_1, Ω_2 deux ouverts de \mathbf{C} . Une application $\varphi : \Omega_1 \rightarrow \Omega_2$ est *biholomorphe*, si elle est bijective et si φ et φ^{-1} sont holomorphe. (On verra plus tard (rem. V.3.3) qu'il suffit que φ soit injective et holomorphe pour qu'elle soit biholomorphe.)

Théorème V.3.1. — (d'inversion locale pour les fonctions holomorphes) *Soient Ω un ouvert de \mathbf{C} , f holomorphe sur Ω , et $z_0 \in \Omega$. Si $f'(z_0) \neq 0$, il existe un voisinage ouvert U de z_0 dans Ω et un voisinage ouvert V de $f(z_0)$ dans \mathbf{C} tels que f soit biholomorphe de U sur V .*

Démonstration. — Commençons par supposer que $z_0 = 0$, $f(z_0) = 0$ et $f'(z_0) = 1$; nous allons construire l'application réciproque g de f par une méthode du point fixe (cf. th.9.2) en partant du fait que g est point fixe de $\varphi \mapsto \varphi - f \circ \varphi + z$.

Il existe $\alpha \in \mathbf{C}$ tel que $f(z) = z + \alpha z^2 + O(|z|^3)$, ce qui fait que,

$$h(u, v) = f(u - v) - f(u) + v = \alpha(v^2 - 2uv) + O((|u| + |v|)^3).$$

Soit $r_0 < \frac{1}{2}\rho(f)$. On déduit du développement limité ci-dessus, l'existence de $C > 0$, tel que $|h(u, v)| < C|u| \cdot |v|$ si $|v| \leq |u| \leq r_0$. Comme $f(z) - z = O(|z|^2)$, on peut quitte à diminuer r_0 imposer de plus que $|f(z) - z| \leq \frac{1}{4}|z|$ si $|z| \leq r_0$, et on peut aussi imposer $r_0 \leq \frac{1}{2C}$.

Soit $r = \frac{2}{3}r_0$. On construit, par récurrence, deux suites de fonctions $(g_n)_{n \geq 1}$ et $(u_n)_{n \geq 1}$, en posant $g_1(z) = z$; $u_n(z) = f(g_n(z)) - z$; $g_{n+1}(z) = g_n(z) - u_n(z)$. Nous allons montrer que g_n et u_n sont holomorphes sur $D(0, r^-)$, et que

$$\left(\frac{1}{2} + \frac{1}{2^n}\right)|z| \leq |g_n(z)| \leq \left(\frac{3}{2} - \frac{1}{2^n}\right)|z| \quad \text{et} \quad |u_n(z)| \leq \frac{1}{2^{n+1}}|z|, \quad \text{quel que soit } z \in D(0, r^-).$$

Le résultat est évident pour $n = 1$ puisque $g_1(z) = z$ et $u_1(z) = f(z) - z$. Si la propriété est vérifiée pour n , alors l'encadrement de $|g_{n+1}(z)|$ est une conséquence immédiate de l'encadrement pour $|g_n(z)|$ et de la majoration pour $|u_n(z)|$. Par ailleurs, comme $|g_n(z)| \leq \frac{3}{2}|z| < r_0 < \rho(f)$, $f(g_n(z))$ est bien définie et holomorphe sur $D(0, r^-)$, et donc u_{n+1} est holomorphe sur $D(0, r^-)$. Maintenant, on a

$$u_{n+1}(z) = f(g_n(z) - u_n(z)) - z = h(g_n(z), u_n(z)) + f(g_n(z)) - u_n(z) - z = h(g_n(z), u_n(z)).$$

Comme $|u_n(z)| \leq \frac{1}{2^{n+1}}|z| \leq \left(\frac{1}{2} + \frac{1}{2^n}\right)|z| \leq |g_n(z)|$, on a

$$|u_{n+1}(z)| \leq |h(g_n(z), u_n(z))| \leq C|g_n(z)||u_n(z)| \leq \frac{3C|z|}{2} \cdot \frac{|z|}{2^{n+1}} \leq Cr_0 \frac{|z|}{2^{n+1}} \leq \frac{|z|}{2^{n+2}}.$$

Ceci montre que l'hypothèse de récurrence est vérifiée au rang $n + 1$.

On déduit de la majoration $|g_{n+1}(z) - g_n(z)| = |u_n(z)| \leq \frac{r}{2^{n+1}}$ la convergence uniforme de la suite g_n sur $D(0, r^-)$. La limite est donc une fonction holomorphe sur $D(0, r^-)$ qui vérifie $g(z) = g(z) - (f(g(z)) - z)$. Autrement dit, on a construit une fonction holomorphe g sur $D(0, r^-)$ telle que $f(g(z)) = z$ si $z \in D(0, r^-)$.

Dans le cas général, on peut appliquer ce qui précède à $F(z) = f'(z_0)^{-1}(f(z + z_0) - f(z_0))$ (et donc $f(z) = f'(z_0)(F(z - z_0) + f(z_0))$). On en déduit l'existence de $r_0 > 0$, et de G holomorphe sur $D(0, r_0^-)$, tels que $F(G(z)) = z$, si $z \in D(0, r_0^-)$. Si on définit g sur $D(f(z_0), r^-)$, où $r = |f'(z_0)|r_0$, par $g(z) = z_0 + G(f'(z_0)^{-1}(z - f(z_0)))$, un calcul immédiat montre que $f(g(z)) = z$, si $z \in D(f(z_0), r^-)$. Autrement, dit, on peut trouver un voisinage V_0 de $f(z_0)$ et g holomorphe sur V_0 tels que $f \circ g = \text{id}$ sur V_0 . Ceci implique en particulier que g est injective sur V_0 . De plus, on a $f'(g(z))g'(z) = 1$, si $z \in V_0$, et donc $g'(f(z_0)) \neq 0$. En appliquant ce qui précède à g au lieu de f , on en déduit l'existence d'un voisinage U de z_0 (que l'on peut supposer inclus dans Ω), et de $h : U \rightarrow V_0$ holomorphe, tels que $g \circ h = \text{id}$ sur U . Soit $V = h(U)$. Comme $g \circ h = \text{id}$ sur U , on a $g(V) = U$, et comme g est injective sur V_0 , on a $V = g^{-1}(U)$, ce qui prouve que V est un ouvert (dans V_0 et donc aussi dans \mathbf{C} puisque V_0 est ouvert dans \mathbf{C}) qui contient $z_0 = h(f(z_0))$. Finalement, on a $f = f \circ (g \circ h) = (f \circ g) \circ h = h$ sur U , et $f \circ g = \text{id}$ sur V (puisque $V \subset V_0$) et $g \circ f = g \circ h = \text{id}$ sur U . En résumé, f induit une bijection holomorphe de U sur V dont l'inverse g est aussi holomorphe. Ceci permet de conclure.

Théorème V.3.2. — Soient Ω un ouvert connexe de \mathbf{C} et f une fonction holomorphe non constante sur Ω . Si $z_0 \in \Omega$, et si $m = v_{z_0}(f - f(z_0))$, il existe un voisinage ouvert U de z_0 dans Ω , un réel $r > 0$ et $\varphi : U \rightarrow D(0, r^-)$ biholomorphe, avec $\varphi(z_0) = 0$, telle que $f(z) = f(z_0) + \varphi(z)^m$, quel que soit $z \in U$.

Démonstration. — Par définition de m , on a, dans un voisinage U_0 de z_0 ,

$$f(z) - f(z_0) = (z - z_0)^m(a_m + a_{m+1}(z - z_0) + \cdots),$$

avec $a_m \neq 0$. Soit $\alpha \in \mathbf{C}^*$ une racine m -ième de a_m , et soit

$$g = \frac{f(z) - f(z_0)}{a_m(z - z_0)^m} = 1 + b_1(z - z_0) + \cdots.$$

La fonction g est holomorphe dans U_0 , et comme elle vaut 1 en z_0 , il existe un voisinage ouvert U_1 de z_0 tel que $g(z) \in D(1, 1^-)$ si $z \in U_1$. Mais alors la fonction $h = g^{1/m}$ est holomorphe sur U_1 et vérifie $h^m = g$ (cf. rem. V.1.7). Ceci montre que, si on pose $\varphi(z) = \alpha(z - z_0)h(z)$, alors φ est holomorphe sur U_1 , et on a $f(z) = f(z_0) + \varphi(z)^m$ quel que soit $z \in U_1$. Finalement, comme $\varphi'(z_0) = \alpha \neq 0$, le théorème d'inversion locale holomorphe montre qu'il existe un voisinage ouvert U de z_0 dans U_1 et $r > 0$ tels que φ induise une bijection biholomorphe de U sur $D(0, r^-)$.

Remarque V.3.3. — Si $m \geq 1$, l'application $z \mapsto z^m$ induit une surjection de $D(0, r^-)$, sur $D(0, (r^m)^-)$, et tout point de $D(0, (r^m)^-) - \{0\}$ a exactement m antécédents. On en déduit (grâce au th. V.3.2), les résultats suivants.

- Si Ω est un ouvert connexe de \mathbf{C} , et si f est holomorphe non constante sur Ω , alors $f(\Omega)$ est un ouvert (théorème de l'application ouverte). En effet, en reprenant les

notations du théorème, on voit que si $z_0 \in \Omega$, alors il existe $r > 0$ tel que $f(\Omega)$ contienne $D(f(z_0), (r^m)^-)$.

• Si Ω est un ouvert de \mathbf{C} , et si f est holomorphe injective sur Ω , alors f est biholomorphe de Ω sur $f(\Omega)$ (inversion globale). En effet, si f est injective, on doit avoir $m = 1$ dans les notations du théorème⁽¹⁰⁾, et donc f' ne s'annule pas sur Ω . on déduit du théorème d'inversion locale que f^{-1} est holomorphe au voisinage de tout point de l'ouvert $f(\Omega)$.

Exercice V.3.4. — Montrer qu'il n'existe pas de fonction holomorphe $f : D(0, 1^-) \rightarrow \mathbf{C}$ qui soit bijective.

2. Logarithme et fonctions puissances

Soit $\alpha \in \mathbf{R}$. L'application exponentielle est injective sur la bande $\{z, \alpha < \operatorname{Im}(z) < \alpha + 2\pi\}$, et son image est \mathbf{C} privé de la demi-droite $\mathbf{R}_+ e^{i\alpha}$ d'argument α (ou $\alpha + 2\pi$, ce qui revient au même). D'après le théorème d'inversion globale ((ii) de la rem. V.3.3), son inverse

$$\log_\alpha : \mathbf{C} - \mathbf{R}_+ e^{i\alpha} \rightarrow \{z, \alpha < \operatorname{Im}(z) < \alpha + 2\pi\}.$$

est une fonction holomorphe.

Comme $\exp(\log_\alpha(z)) = z$ sur $\mathbf{C} - \mathbf{R}_+ e^{i\alpha}$, on obtient, en dérivant cette relation, que la dérivée de \log_α sur $\mathbf{C} - \mathbf{R}_+ e^{i\alpha}$ est $\frac{1}{z}$, et donc que \log_α vérifie les propriétés que l'on est en droit d'attendre d'une fonction logarithme. Le seul problème est qu'on ne peut pas définir le logarithme comme une fonction holomorphe sur \mathbf{C}^* tout entier, à cause de la discontinuité le long de la demi-droite $\mathbf{R}_+ e^{i\alpha}$ (les limites « à gauche » et « à droite » diffèrent de $2i\pi$, le long de cette demi-droite). Ce problème⁽¹¹⁾ est dans la nature des choses : le logarithme est une fonction holomorphe multivaluée sur \mathbf{C}^* , et chacune des fonctions \log_α ci-dessus correspond au choix d'une détermination du logarithme, holomorphe dans un ouvert aussi grand que possible. La *détermination principale du logarithme* est la fonction $\log_{-\pi}$; c'est donc la détermination du logarithme, définie sur \mathbf{C} privé de la demi-droite réelle négative \mathbf{R}_- , et qui vaut 0 en 1; ses valeurs ont une partie imaginaire dans l'intervalle $] -\pi, \pi[$. Dans la pratique, on ne met pas le α en indice, ce qui demande de *faire attention à ce que signifie exactement* $\log z$ dans la situation que l'on considère; les valeurs possibles de $\log z$ étant les $2i\pi k + \log_{-\pi} z$, pour $k \in \mathbf{Z}$.

Proposition V.3.5. — Soit \log la détermination principale du logarithme. Alors on a $\log(1+z) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} z^n$, si $|z| < 1$.

⁽¹⁰⁾Si $m \geq 2$, et si $0 < |u - f(z_0)| < r^m$, alors u a m antécédents par f dans U , à savoir les images réciproques par φ des m racines m -ièmes de $u - f(z_0)$.

⁽¹¹⁾Il a fortement perturbé nos ancêtres. Bernoulli et Leibnitz se sont battus (épistolairement), en 1712-1713, sur la valeur de $\log(-1)$, l'un soutenant que $\log(-1) = 0$ car $2\log(-1) = \log(-1)^2 = \log 1 = 0$, et l'autre que $\log(-1)$ devait être imaginaire car $\log(-1) = -2 - \frac{2^2}{2} - \frac{2^3}{3} \dots$ ne converge pas... Il a fallu attendre 1749 pour qu'Euler explique qu'un nombre complexe a une infinité de logarithmes.

Démonstration. — Soit $f(z) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} z^n$. C'est une fonction holomorphe sur $D(0, 1^-)$ dont la dérivée est $\sum_{n=1}^{+\infty} (-1)^{n-1} z^{n-1} = \frac{1}{1+z}$. Par ailleurs, $\log(1+z)$ est aussi holomorphe sur $D(0, 1^-)$, de dérivée $\frac{1}{1+z}$. La fonction $g(z) = f(z) - \log(1+z)$ est, d'après le (i) de la rem. V.2.8, somme de sa série de Taylor en 0 sur tout $D(0, 1^-)$, et comme sa dérivée est nulle, elle est constante sur $D(0, 1^-)$. On conclut en remarquant que $g(0) = 0$.

Exercice V.3.6. — Soit \log la détermination principale du logarithme.

- (i) Montrer que $\log z = \log |z| + i \arg z$, où $\arg z \in]-\pi, \pi[$ est l'argument principal de z .
- (ii) Soient $z_1, z_2 \in \mathbf{C} - \mathbf{R}_-$, tels que $z_1 z_2 \in \mathbf{C} - \mathbf{R}_-$. Déterminer, suivant les cas, la valeur de $\log(z_1 z_2) - \log z_1 - \log z_2$.

Si $s \in \mathbf{C}^*$, on définit z^s par la formule $z^s = \exp(s \log z)$. Comme la fonction $\log z$ est multivaluée, il en est de même de z^s , et il faut donc bien faire attention au choix que l'on a fait de $\log z$. Si a est une valeur possible de z^s , les autres sont les $e^{2i\pi k s} a$, pour $k \in \mathbf{Z}$. En particulier, on peut se débrouiller, par un choix convenable de la détermination de $\log z$, pour que z^s soit holomorphe dans $\mathbf{C} - \mathbf{R}_+ e^{i\alpha}$, mais, si $s \notin \mathbf{Z}$, la fonction z^s ne peut pas se prolonger en une fonction holomorphe sur \mathbf{C}^* . Même si on prend la détermination principale du logarithme, il n'est pas toujours vrai que $z_1^s z_2^s = (z_1 z_2)^s$. Par contre, la fonction $s \mapsto z^s$ est, elle, holomorphe sur \mathbf{C} tout entier, et on a $z^{s_1+s_2} = z^{s_1} z^{s_2}$ (si on ne s'amuse pas à changer la valeur de $\log z$). En particulier, si $m \in \mathbf{N} - \{0\}$, alors $z^{1/m}$ est une racine m -ième de z , quelle que soit la détermination du logarithme choisie.

Exercice V.3.7. — (i) Soit $\Omega = \mathbf{C} - [0, 1]$. Montrer qu'il existe deux fonctions f , holomorphes sur Ω , dont le carré est $z \mapsto z(z-1)$.

(ii) Soient a_1, \dots, a_{2n} , des éléments distincts de \mathbf{C} . Montrer qu'il existe une permutation σ de $1, \dots, 2n$ telle que les segments $[a_{\sigma(2i-1)}, a_{\sigma(2i)}]$, pour $1 \leq i \leq n$ soient disjoints. Montrer que si Ω est l'ouvert complémentaire de ces segments, il existe deux fonctions f , holomorphes sur Ω , dont le carré est $z \mapsto \prod_{i=1}^{2n} (z - a_i)$.

CHAPITRE VI

LA FORMULE DE CAUCHY ET CELLE DES RÉSIDUS (DE CAUCHY)

Le lecteur a déjà pu apprécier la puissance de la formule de Cauchy pour l'étude des fonctions holomorphes. La formule des résidus, qui en est une extension, est à la fois un outil pratique permettant de calculer sans douleur certaines intégrales, et un outil théorique extrêmement souple dont nous verrons certaines utilisations au chap. VII et dans l'annexe A. Par ailleurs, les idées qu'elle met en oeuvre sont le point de départ de constructions plus générales permettant d'associer des invariants aux variétés de dimension quelconque. Par exemple, on peut utiliser (les idées intervenant dans) la formule des résidus, pour démontrer qu'un ouvert de \mathbf{R}^2 n'est pas homéomorphe à un ouvert de \mathbf{R}^n , si $n \neq 2$. Démontrer, ce qui est vital si on veut pouvoir parler de dimension d'un point de vue topologique, qu'un ouvert de \mathbf{R}^n n'est pas homéomorphe à un ouvert de \mathbf{R}^m , si $n \neq m$, peut se faire en généralisant (grandement) ces idées.

VI.1. Homotopie de lacets et formule de Cauchy

1. Vocabulaire de topologie algébrique

Soient X et Y deux espaces topologiques et $f : X \rightarrow Y$ et $g : X \rightarrow Y$ deux applications continues. On dit que f et g sont *homotopes* si « on peut passer continûment de f à g ». De manière précise, f et g sont homotopes, s'il existe $u : [0, 1] \times X \rightarrow Y$ continue, telle que $u(0, x) = f(x)$ et $u(1, x) = g(x)$, quel que soit $x \in X$. Si on note $f_t : X \rightarrow Y$ l'application $f_t(x) = u(t, x)$, alors f_t est continue, quel que soit $t \in [0, 1]$, $t \mapsto f_t$ est une application continue de $[0, 1]$ dans l'espace des applications continues de X dans Y (muni de la topologie de la convergence simple), et $t \mapsto f_t$ connecte $f = f_0$ à $g = f_1$.

Un ouvert Ω de \mathbf{C} est *contractile* s'il est *homotope à un point*, c'est-à-dire, s'il existe $z_0 \in \Omega$ tel que $\text{id} : \Omega \rightarrow \Omega$ est homotope, dans Ω , à l'application constante $z \mapsto z_0$, quel que soit $z \in \Omega$. Autrement dit, Ω est contractile, s'il existe $z_0 \in \Omega$ et $u : [0, 1] \times \Omega \rightarrow \Omega$, continue, telle que $u(0, z) = z$ et $u(1, z) = z_0$, quel que soit $z \in \Omega$.

Un ouvert Ω est *étoilé* s'il existe $z_0 \in \Omega$ tel que, quel que soit $z \in \Omega$, le segment $[z_0, z]$ est inclus dans Ω . En particulier, un ouvert convexe est étoilé ; \mathbf{C} privé d'une demi-droite Y est

étoilé (on peut prendre pour z_0 n'importe quel élément de l'autre demi-droite de la droite contenant Y). Un ouvert étoilé est contractile : il suffit de poser $u(t, z) = z(1 - t) + tz_0$; la condition selon laquelle $[z_0, z] \subset \Omega$ fait que u est bien à valeurs dans Ω .

Si Ω est un ouvert de \mathbf{C} , une *homotopie de lacets dans* Ω de $\gamma_0 : (\mathbf{R}/\mathbf{Z}) \rightarrow \Omega$ sur $\gamma_1 : (\mathbf{R}/\mathbf{Z}) \rightarrow \Omega$ est une application continue $u : [0, 1] \times (\mathbf{R}/\mathbf{Z}) \rightarrow \Omega$, vérifiant $u(0, t) = \gamma_0(t)$ et $u(1, t) = \gamma_1(t)$; on peut aussi voir u comme une application continue $u : [0, 1] \times [0, 1] \rightarrow \Omega$ telle que $u(s, 0) = u(s, 1)$, quel que soit $s \in [0, 1]$. On dit qu'un lacet $\gamma : (\mathbf{R}/\mathbf{Z}) \rightarrow \Omega$ est *contractile dans* Ω , s'il existe une homotopie de lacets de γ sur un lacet constant. Si Ω est contractile⁽¹⁾, alors tout lacet est contractile dans Ω .

2. Un cas particulier de la formule de Stokes

Soient $A = (0, 0)$, $B = (1, 0)$, $C = (1, 1)$ et $D = (0, 1)$ les sommets du carré $[0, 1]^2$. On note $\partial([0, 1]^2)$ le bord orienté du carré $[0, 1]^2$; c'est la réunion des segments $[A, B]$, $[B, C]$, $[C, D]$ et $[D, A]$.

Proposition VI.1.1. — Si $\omega = P ds + Q dt$ est une 1-forme sur $[0, 1]^2$, avec P et Q de classe \mathcal{C}^1 , alors⁽²⁾

$$\int_{\partial([0, 1]^2)} \omega = \int_{[0, 1]^2} \left(-\frac{\partial P}{\partial t} + \frac{\partial Q}{\partial s} \right) ds dt.$$

Démonstration. — En revenant à la définition d'une intégrale curviligne, on obtient

$$\begin{aligned} \int_{[A, B]} \omega &= \int_0^1 P(s, 0) ds, & \int_{[C, D]} \omega &= \int_0^1 P(1 - v, 1) dv = - \int_0^1 P(s, 1) ds, \\ \int_{[B, C]} \omega &= \int_0^1 Q(1, t) dt, & \int_{[D, A]} \omega &= \int_0^1 Q(0, 1 - v) dv = - \int_0^1 Q(0, t) dt. \end{aligned}$$

⁽¹⁾ Plus généralement, un espace topologique connexe X est *simplement connexe*, si tout lacet dans X est contractile. D'après un célèbre théorème énoncé par B. Riemann (1851) et vraiment démontré par P. Koebe en 1914, (le *théorème de représentation conforme de Riemann*), un ouvert simplement connexe de \mathbf{C} , non égal à \mathbf{C} , est biholomorphe au disque unité $D = D(0, 1^-)$ ou, ce qui revient au même (cf. ex. V.1.25), au demi-plan de Poincaré. En particulier, cela permet de montrer que tout ouvert U simplement connexe de \mathbf{R}^2 est difféomorphe à \mathbf{R}^2 (i.e. il existe une bijection de classe \mathcal{C}^1 de U sur \mathbf{R}^2 dont la réciproque est aussi de classe \mathcal{C}^1). Comme une fonction holomorphe préserve les angles, on voit que l'on peut transformer (l'intérieur d') un cercle en (l'intérieur d') un triangle en conservant les angles...

⁽²⁾ La formule qui suit est un cas particulier de la formule de Stokes $\int_{\Omega} d\omega = \int_{\partial\Omega} \omega$, qui est une des formules mathématiques les plus esthétiques, et aussi une des plus utiles. Dans la formule de Stokes, ω est une p -forme sur un compact K de dimension $p + 1$, $d\omega$ est sa différentielle, Ω est l'intérieur de K et $\partial\Omega$ est la frontière orientée de Ω (et donc aussi celle de K). Donner un sens précis à ce qui précède demande un peu de travail (en particulier définir une orientation sur la frontière, ce qui n'est pas toujours possible). Si $p = 0$, la formule de Stokes est « le théorème fondamental de l'analyse » $\int_a^b df = f(b) - f(a)$, et si $p = 1$ et si Ω est de dimension 2, la formule de Stokes est aussi connue sous le nom de théorème de Green (on a alors $d(P dx + Q dy) = dP \wedge dx + dQ \wedge dy = \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x} \right) dx \wedge dy$, en tenant compte de ce que \wedge est bilinéaire alterné ; on retombe donc bien sur la formule de la proposition).

Ceci nous donne

$$\begin{aligned} \int_{\partial([0,1]^2)} \omega &= \int_0^1 (P(s, 0) - P(s, 1)) ds + \int_0^1 (Q(1, t) - Q(0, t)) dt \\ &= - \int_0^1 \left(\int_0^1 \frac{\partial P}{\partial t}(s, t) dt \right) ds + \int_0^1 \left(\int_0^1 \frac{\partial Q}{\partial s}(s, t) ds \right) dt \\ &= \int_{[0,1]^2} \left(- \frac{\partial P}{\partial t} + \frac{\partial Q}{\partial s} \right) ds dt. \end{aligned}$$

Si Ω est un ouvert de \mathbf{C} , si $u : [0, 1]^2 \rightarrow \Omega$ est une application de classe \mathcal{C}^1 , et si $f : \Omega \rightarrow \mathbf{C}$ est de classe \mathcal{C}^1 , on définit la 1-forme $u^*(f(z) dz)$ sur $[0, 1]^2$, par la formule :

$$u^*(f(z) dz) = f(u(s, t)) d(u(s, t)) = f(u(s, t)) \left(\frac{\partial u}{\partial s}(s, t) ds + \frac{\partial u}{\partial t}(s, t) dt \right).$$

Lemme VI.1.2. — Soit Ω un ouvert de \mathbf{C} . Si $f : \Omega \rightarrow \mathbf{C}$ est une fonction de classe \mathcal{C}^1 au sens complexe, si $u : [0, 1]^2 \rightarrow \Omega$ est une application de classe \mathcal{C}^2 , et si $u^*(f(z) dz) = P(s, t) ds + Q(s, t) dt$, alors $-\frac{\partial P}{\partial t} + \frac{\partial Q}{\partial s} = 0$.

Démonstration. — Comme f est holomorphe⁽³⁾, on a

$$f(u(s+h, t)) = f(u(s, t) + h \frac{\partial u}{\partial s}(s, t) + o(h)) = f(u(s, t)) + f'(u(s, t)) \left(\frac{\partial u}{\partial s}(s, t) h \right) + o(h).$$

On en déduit que (avec le même argument pour $\frac{\partial}{\partial t}$)

$$\frac{\partial(f \circ u)}{\partial s} = \frac{\partial u}{\partial s} \cdot f' \circ u \quad \text{et} \quad \frac{\partial(f \circ u)}{\partial t} = \frac{\partial u}{\partial t} \cdot f' \circ u.$$

Comme $P = \frac{\partial u}{\partial s} \cdot f' \circ u$ et $Q = \frac{\partial u}{\partial t} \cdot f' \circ u$, on obtient

$$\begin{aligned} -\frac{\partial P}{\partial t} + \frac{\partial Q}{\partial s} &= -\frac{\partial^2 u}{\partial t \partial s} \cdot f' \circ u - \frac{\partial u}{\partial s} \frac{\partial u}{\partial t} \cdot f'' \circ u + \frac{\partial^2 u}{\partial s \partial t} \cdot f' \circ u - \frac{\partial u}{\partial t} \frac{\partial u}{\partial s} \cdot f'' \circ u \\ &= \left(-\frac{\partial^2 u}{\partial t \partial s} + \frac{\partial^2 u}{\partial s \partial t} \right) f' \circ u = 0. \end{aligned}$$

Théorème VI.1.3. — Soient Ω un ouvert de \mathbf{C} et f une fonction de classe \mathcal{C}^1 au sens complexe sur Ω . Alors, si γ_0, γ_1 sont deux lacets homotopes dans Ω , on a

$$\int_{\gamma_0} f(z) dz = \int_{\gamma_1} f(z) dz.$$

⁽³⁾Une démonstration plus savante, mais plus naturelle, consisterait à remarquer que l'on a d'une part $(-\frac{\partial P}{\partial t} + \frac{\partial Q}{\partial s}) ds \wedge dt = d(u^*(f(z) dz))$, et d'autre part que $d(u^*(f(z) dz)) = u^*(d(f(z) dz))$, et $\frac{\partial f}{\partial \bar{z}} = 0$, puisque f est holomorphe, et donc $d(f(z) dz) = \frac{\partial f}{\partial \bar{z}} d\bar{z} \wedge dz = 0$.

Remarque VI.1.4. — Si γ est un lacet constant, on a $\int_{\gamma} \omega = 0$ quelle que soit la 1-forme ω . On en déduit, si $\omega = f(z) dz$, avec f holomorphe sur Ω , les résultats suivants.

- (i) Si γ est un lacet contractile dans Ω , alors $\int_{\gamma} f(z) dz = 0$.
- (ii) Si Ω est contractile (en particulier, si Ω est étoilé), alors $\int_{\gamma} f(z) dz = 0$, quel que soit le lacet γ dans Ω .

Démonstration. — Commençons par supposer qu'il existe, dans Ω , une homotopie de lacets, de classe \mathcal{C}^2 , de γ_0 sur γ_1 . Autrement dit, il existe $u : [0, 1]^2 \rightarrow \Omega$ de classe \mathcal{C}^2 , telle que $u(0, t) = \gamma_0(t)$, $u(1, t) = \gamma_1(t)$, et $u(s, 0) = u(s, 1)$ quel que soit $s \in [0, 1]$. Notons ω la 1-forme $f(z) dz$. Écrivons $u^*\omega$, sous la forme $P ds + Q dt$. D'après le lemme VI.1.2, on a $-\frac{\partial P}{\partial t} + \frac{\partial Q}{\partial s} = 0$. On déduit donc de la proposition VI.1.1, la nullité de $\int_{\partial([0,1]^2)} u^*\omega$. En notant, $A = (0, 0)$, $B = (1, 0)$, $C = (1, 1)$ et $D = (0, 1)$ les sommets du carré $[0, 1]^2$, on voit que

$$\int_{[B,C]} u^*\omega = \int_0^1 Q(1, t) dt = \int_0^1 f(u(1, t)) \frac{\partial u}{\partial t}(1, t) dt = \int_0^1 f(\gamma_1(t)) \gamma_1'(t) dt = \int_{\gamma_1} f(z) dz.$$

De même, $\int_{[D,A]} u^*\omega = -\int_{\gamma_0} f(z) dz$. Par ailleurs, on a $\int_{[A,B]} u^*\omega + \int_{[C,D]} u^*\omega = 0$, puisque $u(s, 0) = u(s, 1)$, quel que soit $s \in [0, 1]$. On obtient donc

$$0 = \int_{\partial([0,1]^2)} u^*\omega = \int_{[A,B]} u^*\omega + \int_{[B,C]} u^*\omega + \int_{[C,D]} u^*\omega + \int_{[D,A]} u^*\omega = \int_{\gamma_1} f(z) dz - \int_{\gamma_0} f(z) dz,$$

ce qui permet de conclure dans le cas particulier où γ_0 et γ_1 sont homotopes dans Ω par une homotopie de classe \mathcal{C}^2 .

Passons au cas général. Les chemins γ_0 et γ_1 sont \mathcal{C}^1 par morceaux, et on dispose de $u : [0, 1]^2 \rightarrow \Omega$, continue, telle que $u(0, t) = \gamma_0(t)$, $u(1, t) = \gamma_1(t)$, et $u(s, 0) = u(s, 1)$ quel que soit $s \in [0, 1]$. L'idée est d'approximer u par des fonctions u_ε de classe \mathcal{C}^2 , en régularisant comme dans l'ex. III.3.12, et de passer à la limite. La mise en œuvre de cette stratégie demande de prendre un peu de précautions pour s'assurer que les approximations que l'on considère sont bien des homotopies de lacets ne sortant pas de Ω .

On note $[x]$ la partie entière de $x \in \mathbf{R}$. Soit $\tilde{u} : [-1, 2]^2 \rightarrow \Omega$ définie par

$$\tilde{u}(s, t) = \begin{cases} u(0, t - [t]), & \text{si } s \leq 0, \\ u(s, t - [t]), & \text{si } 0 \leq s \leq 1, \\ u(1, t - [t]), & \text{si } s \geq 1. \end{cases}$$

Par construction, \tilde{u} est continue, coïncide avec u sur $[0, 1]^2$, est la restriction d'une fonction périodique (en t) de période 1, et son image coïncide avec celle de u . Comme $[-1, 2]^2$ est compact, son image K par \tilde{u} est compacte, et la distance $d(K, \mathbf{C} - \Omega)$ de K au fermé $\mathbf{C} - \Omega$ est > 0 . Il existe donc $\delta > 0$ tel que Ω contienne $K_\delta = \{z \in \mathbf{C}, d(z, K) \leq \delta\}$. Maintenant, si $\varepsilon > 0$, soit

$$\delta(\varepsilon) = \sup_{(s', t', s, t) \in [-1, 2]^4, |s' - s| \leq \varepsilon, |t' - t| \leq \varepsilon} |\tilde{u}(s', t') - \tilde{u}(s, t)|.$$

Comme $[-1, 2]^2$ est compact, \tilde{u} est uniformément continue sur $[-1, 2]^2$, ce qui se traduit par $\delta(\varepsilon) \rightarrow 0$ quand $\varepsilon \rightarrow 0$. On choisit $\varepsilon_0 \leq \frac{1}{2}$, tel que $\delta(\varepsilon) \leq \delta$, si $\varepsilon \leq \varepsilon_0$. Finalement, on choisit $\varphi : \mathbf{R} \rightarrow \mathbf{R}_+$, de classe

\mathcal{C}^2 , nulle en dehors de $[-1, 1]$, avec $\int_{\mathbf{R}} \varphi = 1$, et on définit φ_ε , par $\varphi_\varepsilon(x) = \varepsilon^{-1} \varphi(\varepsilon^{-1}x)$, ce qui fait de φ_ε une fonction positive de classe \mathcal{C}^2 sur \mathbf{R} , nulle en dehors de $[-\varepsilon, \varepsilon]$, avec $\int_{\mathbf{R}} \varphi_\varepsilon = 1$.

On note $\varphi_\varepsilon^{[2]} : \mathbf{R}^2 \rightarrow \mathbf{R}_+$ la fonction définie par $\varphi_\varepsilon^{[2]}(s, t) = \varphi_\varepsilon(s)\varphi_\varepsilon(t)$. Si $\varepsilon \leq \varepsilon_0$, soit u_ε la restriction de $\tilde{u} * \varphi_\varepsilon^{[2]}$ à $[-\frac{1}{2}, \frac{3}{2}]$. On a donc

$$u_\varepsilon(s, t) = \int_{s-\varepsilon}^{s+\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} \tilde{u}(x, y) \varphi_\varepsilon(s-x) \varphi_\varepsilon(t-y) dy dx.$$

Comme $|\tilde{u}(x, y) - \tilde{u}(s, t)| \leq \delta(\varepsilon)$, si $x \in [s-\varepsilon, s+\varepsilon]$ et $y \in [t-\varepsilon, t+\varepsilon]$, on obtient

$$\begin{aligned} |u_\varepsilon(s, t) - \tilde{u}(s, t)| &= \left| \int_{s-\varepsilon}^{s+\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} (\tilde{u}(x, y) - \tilde{u}(s, t)) \varphi_\varepsilon(s-x) \varphi_\varepsilon(t-y) dy dx \right| \\ &\leq \int_{s-\varepsilon}^{s+\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} |\tilde{u}(x, y) - \tilde{u}(s, t)| \varphi_\varepsilon(s-x) \varphi_\varepsilon(t-y) dy dx \\ &\leq \int_{s-\varepsilon}^{s+\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} \delta(\varepsilon) \varphi_\varepsilon(s-x) \varphi_\varepsilon(t-y) dy dx = \delta(\varepsilon), \end{aligned}$$

ce qui montre à la fois que u_ε est à valeurs dans $K_\delta \subset \Omega$, si $\varepsilon \leq \varepsilon_0$, et que u_ε tend uniformément vers \tilde{u} sur $[-\frac{1}{2}, \frac{3}{2}]^2$ quand ε tend vers 0.

Par ailleurs, comme $\varphi_\varepsilon^{[2]}$ est de classe \mathcal{C}^2 sur \mathbf{R}^2 et à support compact, il en est de même de $\tilde{u} * \varphi_\varepsilon^{[2]}$, et comme $\tilde{u}(s, t+1) = \tilde{u}(s, t)$, si $s \in [-1, 2]$, et $t \in [-2, 0]$, on a $u_\varepsilon(s, t+1) = u_\varepsilon(s, t)$, si $\varepsilon \leq \frac{1}{2}$, si $s \in [-\frac{1}{2}, \frac{3}{2}]$, et si $t \in [-\frac{1}{2}, \frac{1}{2}]$. Ceci s'applique en particulier à $t = 0$; on en déduit que $\gamma_{s,\varepsilon} = u_\varepsilon(s, \cdot)$, est un lacet de $[0, 1]$ dans Ω , pour tout $s \in [-\frac{1}{2}, \frac{3}{2}]$, et tout $\varepsilon \leq \varepsilon_0$. En résumé, on a prouvé que, si $\varepsilon \leq \varepsilon_0$, u_ε est une homotopie de lacets dans Ω , de classe \mathcal{C}^2 . On en déduit que $\int_{\gamma_{s,\varepsilon}} f(z) dz$ ne dépend pas de $s \in [-\frac{1}{2}, \frac{3}{2}]$. En particulier, on a

$$\int_{\gamma_{-1/2,\varepsilon}} f(z) dz = \int_{\gamma_{3/2,\varepsilon}} f(z) dz, \quad \text{quel que soit } \varepsilon \leq \varepsilon_0.$$

Or, par construction, $\tilde{u}(s, t) = \gamma_0(t)$, si $s \leq 0$. On en déduit que

$$\gamma_{-1/2,\varepsilon}(t) = \int_{-1/2-\varepsilon}^{-1/2+\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} \gamma_0(t) \varphi_\varepsilon(-\frac{1}{2}-x) \varphi_\varepsilon(t-y) dx dy = \int_{t-\varepsilon}^{t+\varepsilon} \gamma_0(t) \varphi_\varepsilon(t-y) dy = (\gamma_0 * \varphi_\varepsilon)(t),$$

et donc que $\gamma'_{-1/2,\varepsilon} = \gamma'_0 * \varphi_\varepsilon$. On a déjà montré que $\gamma_{-1/2,\varepsilon}(t) = u_\varepsilon(-\frac{1}{2}, t)$ tend vers $\tilde{u}(-\frac{1}{2}, t) = \gamma_0(t)$; les mêmes arguments montrent que $\gamma'_{-1/2,\varepsilon}(t) \rightarrow \gamma'_0(t)$, si t n'est pas un point anguleux de γ_0 . Finalement, si $t \in [0, 1]$, on a

$$|f(\gamma_{-1/2,\varepsilon}(t)) \gamma'_{-1/2,\varepsilon}(t)| \leq \sup_{z \in K_\delta} |f(z)| \cdot \sup_{t \in [0,1]} |\gamma'_0(t)|,$$

ce qui permet de déduire du théorème de convergence dominée que, quand $\varepsilon \rightarrow 0$,

$$\int_{\gamma_{-1/2,\varepsilon}} f(z) dz = \int_0^1 f(\gamma_{-1/2,\varepsilon}(t)) \gamma'_{-1/2,\varepsilon}(t) dt \rightarrow \int_0^1 f(\gamma_0(t)) \gamma'_0(t) dt = \int_{\gamma_0} f(z) dz.$$

On montre de même que $\int_{\gamma_{3/2,\varepsilon}} f(z) dz \rightarrow \int_{\gamma_1} f(z) dz$, et un passage à la limite permet d'en conclure que $\int_{\gamma_0} f(z) dz = \int_{\gamma_1} f(z) dz$. Ceci termine la démonstration du théorème.

3. Démonstration de la formule de Cauchy

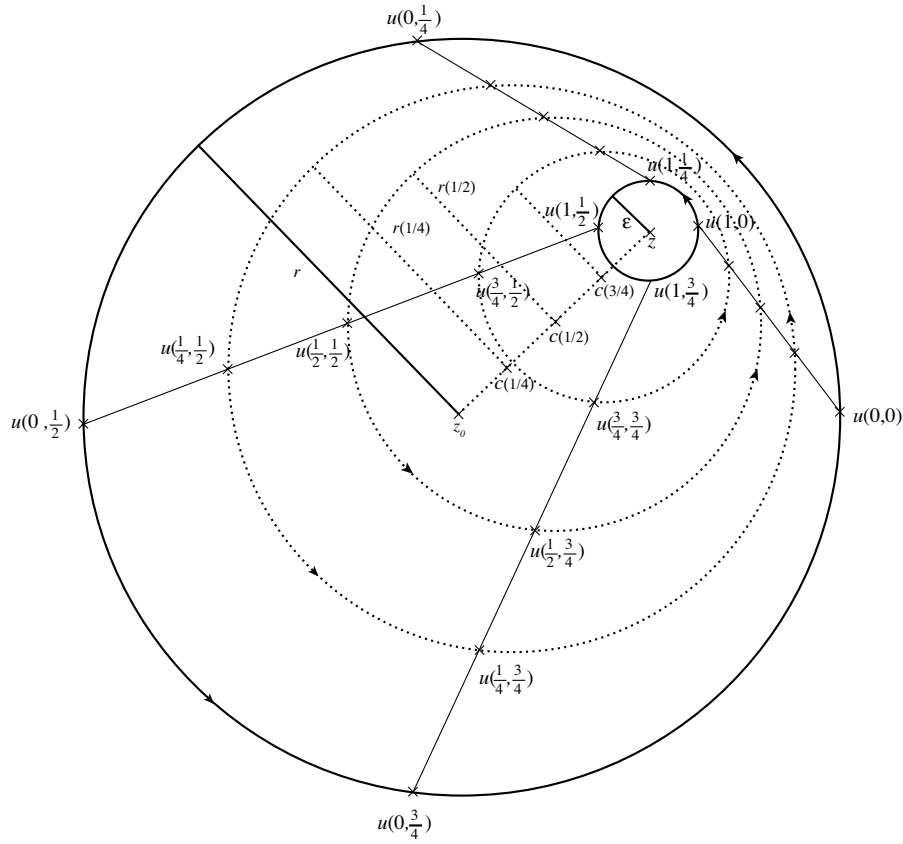
Reprenons les notations du th. V.2.5. Si $\varepsilon \in]0, r - |z - z_0|]$, soit

$$u(s, t) = (1-s)(z_0 + re^{2i\pi t}) + s(z + \varepsilon e^{2i\pi t}) = (1-s)z_0 + sz + ((1-s)r + s\varepsilon)e^{2i\pi t}.$$

Alors $\gamma_s(t) = u(s, t)$ est le cercle de centre $c(s) = (1 - s)z_0 + sz$ et de rayon $r(s) = (1 - s)r + s\varepsilon$, parcouru dans le sens direct. On a donc $\gamma_0 = C(z_0, r)$ et $\gamma_1 = C(z, \varepsilon)$. Par ailleurs, on a $u(s, 0) = u(s, 1)$ quel que soit $s \in [0, 1]$, et

$$\begin{aligned} |u(s, t) - z| &\geq |r(s) - |c(s) - z|| = s\varepsilon + (1 - s)(r - |z - z_0|) \geq \varepsilon, \\ |u(s, t) - z_0| &\leq |c(s) - z_0| + r(s) = r - s(r - |z - z_0| - \varepsilon) \leq r, \end{aligned}$$

ce qui prouve que u est une homotopie de lacets dans $D(z_0, r) - D(z, \varepsilon^-) \subset \Omega - \{z\}$, de $C(z_0, r)$ sur $C(z, \varepsilon)$.



L'homotopie de $C(z_0, r)$ sur $C(z, \varepsilon)$

Comme $\frac{f(w)}{w-z}$ est de classe \mathcal{C}^1 au sens complexe (en w) sur $\Omega - \{z\}$, on déduit du th. VI.1.3, que

$$\frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{w-z} dw = \frac{1}{2i\pi} \int_{C(z, \varepsilon)} \frac{f(w)}{w-z} dw, \quad \text{quel que soit } \varepsilon > 0.$$

Or $C(\varepsilon, r)$ est le chemin $t \mapsto z + \varepsilon e^{2i\pi t}$, et donc $\frac{1}{2i\pi} \int_{C(z, \varepsilon)} \frac{f(w)}{w-z} dw = \int_0^1 f(z + \varepsilon e^{2i\pi t}) dt$ tend vers $\int_0^1 f(z) dt = f(z)$ quand ε tend vers 0 (continuité d'une intégrale dépendant d'un paramètre, f étant continue en z puisque holomorphe dans un voisinage de z). En

passant à la limite, on en déduit la formule $\frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f(w)}{w-z} dw = f(z)$ que l'on cherchait à établir.

4. Primitives

Soit Ω un ouvert connexe de \mathbf{C} . Si f est holomorphe sur Ω , on dit que $F : \Omega \rightarrow \mathbf{C}$ est une primitive de f , si F est holomorphe sur Ω et si $F' = f$. Si $\gamma : [a, b] \rightarrow \Omega$ est un chemin de classe \mathcal{C}^1 , alors $(F \circ \gamma)'(t) = f(\gamma(t))\gamma'(t)$, quel que soit $t \in [a, b]$.

Une fonction holomorphe admet toujours *localement* une primitive. En effet, si f est holomorphe sur $D(z_0, r^-)$, on a $f(z) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n$, d'après le (i) de la rem. V.2.8, et f admet la fonction F , définie par $F(z) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(z_0)}{(n+1)!} (z - z_0)^{n+1}$, comme primitive sur $D(z_0, r^-)$. Par contre, une fonction holomorphe sur un ouvert Ω quelconque n'admet pas toujours une primitive sur tout Ω .

Proposition VI.1.5. — *Soit Ω un ouvert connexe de \mathbf{C} . Si f est holomorphe sur Ω , les conditions suivantes sont équivalentes :*

- (i) f admet une primitive F sur Ω ;
- (ii) $\int_{\gamma} f(z) dz = 0$, quel que soit le lacet γ de classe \mathcal{C}^1 par morceaux, contenu dans Ω .

Démonstration. — Si f admet une primitive F sur Ω , alors $f(z) dz = dF$. On en déduit, en utilisant le (iv) du th. V.2.2, que si $\gamma : [a, b] \rightarrow \Omega$ est \mathcal{C}^1 , alors $\int_{\gamma} f(z) dz = F(\gamma(b)) - F(\gamma(a))$. En particulier, si γ est un lacet, $\int_{\gamma} f(z) dz = 0$, puisque $\gamma(b) = \gamma(a)$. D'où l'implication (i) \Rightarrow (ii). Passons à la démonstration de la réciproque. Fixons $a \in \Omega$, et définissons $F : \Omega \rightarrow \mathbf{C}$, par $F(b) = \int_{\gamma} f(z) dz$, où γ est n'importe⁽⁴⁾ quel chemin \mathcal{C}^1 par morceaux dans Ω , d'extrémités a et b . Si γ_1 et γ_2 sont deux tels chemins, alors on obtient un lacet γ en composant γ_1 avec l'opposé de γ_2 ; on a donc $\int_{\gamma_1} f(z) dz - \int_{\gamma_2} f(z) dz = \int_{\gamma} f(z) dz = 0$, par hypothèse, ce qui prouve que $F(b)$ ne dépend pas du chemin reliant a à b que l'on a choisi, et donc que F est bien définie.

Il suffit de prouver que F est dérivable au sens complexe, et qu'en tout point de Ω sa dérivée est $f(z_0)$, car la prop. V.2.6 montre qu'alors F est holomorphe. Soit donc $z_0 \in \Omega$, et soit $r > 0$ tel que $D(z_0, r^-) \subset \Omega$. Si on fixe un chemin γ_0 joignant a à z_0 dans Ω , et si $b \in D(z_0, r^-)$, on fabrique un chemin γ_b joignant a à b dans Ω , en rajoutant le segment $[z_0, b]$ au chemin γ_0 . On a alors

$$\frac{F(b) - F(z_0)}{b - z_0} = \frac{1}{b - z_0} \left(\int_{\gamma_0} f(z) dz + \int_{[z_0, b]} f(z) dz - F(z_0) \right) = \int_0^1 f(z_0 + t(b - z_0)) dt,$$

⁽⁴⁾On déduit de la démonstration de ce qu'un ouvert connexe de \mathbf{R}^n est connexe par arcs qu'un tel ouvert est aussi connexe par lignes brisées (réunion finie de segments de droite), et donc aussi par chemins \mathcal{C}^1 par morceaux.

et la continuité de f en z_0 permet de montrer (par exemple en utilisant le théorème de continuité d'une intégrale dépendant d'un paramètre), que $\frac{F(b)-F(z_0)}{b-z_0} \rightarrow f(z_0)$ quand $b \rightarrow z_0$. Ceci permet de conclure.

Remarque VI.1.6. — (i) D'après le (ii) de la rem. VI.1.4, la condition (ii) est automatiquement vérifiée si Ω est un ouvert étoilé ou, plus généralement, contractile. Il en résulte qu'une fonction holomorphe sur un ouvert contractile possède une primitive.

(ii) Si $a \in \mathbf{C}$, et si $0 < R_1 < R_2$, la fonction $1/(z-a)$ est holomorphe sur la couronne $C(a, R_1, R_2)$, mais on a $\int_{C(a,r)} \frac{dz}{z-a} = 2i\pi \neq 0$ si $R_1 < r < R_2$. La fonction $1/(z-a)$ n'admet donc pas de primitive sur la couronne $C(a, R_1, R_2)$.

Proposition VI.1.7. — (logarithme d'une fonction holomorphe)

Soit Ω un ouvert contractile de \mathbf{C} , et soit f holomorphe sur Ω ne s'annulant pas sur Ω . Alors il existe g holomorphe sur Ω , telle que $f = e^g$, et on a $g' = \frac{f'}{f}$.

Démonstration. — Comme Ω est contractile, et comme $\frac{f'}{f}$ est holomorphe sur Ω , il existe, d'après le (i) de la rem. VI.1.6, h holomorphe sur Ω telle que $h' = \frac{f'}{f}$. Soit $z_0 \in \Omega$. Quitte à rajouter une constante à h , on peut supposer que $e^{h(z_0)} = f(z_0)$. Mais alors $(e^{-h}f)' = -h'e^{-h}f + e^{-h}f' = 0$, et donc $e^{-h}f$ est constante sur Ω , et comme elle vaut 1 en z_0 , on a $f = e^h$ sur Ω . Si g est une autre fonction holomorphe sur Ω vérifiant $e^g = f$, on a $e^{g-h} = 1$, et donc $g-h$ est holomorphe et à valeurs dans $2i\pi\mathbf{Z}$. Comme Ω est connexe, $g-h$ est constante et $g' = \frac{f'}{f}$.

Exercice VI.1.8. — (Théorème de Morera) Soient $z_0 \in \mathbf{C}$ et $r > 0$.

(i) Montrer que, si f est holomorphe sur $D(z_0, r^-)$, alors,

$$(1) \quad \int_{[a,b]} f(z) dz + \int_{[b,c]} f(z) dz + \int_{[c,a]} f(z) dz = 0 \quad \text{quels que soient } a, b, c \in D(z_0, r^-).$$

(ii) Soit $f : D(z_0, r^-) \rightarrow \mathbf{C}$ continue, et vérifiant la propriété (1). Soit $F(z) = \int_{[z_0, z]} f(z) dz$. Montrer que $\frac{1}{h}(F(z+h) - F(z))$ tend vers $f(z)$, quand h tend vers 0. En déduire que F et f sont holomorphes.

(iii) Soit Ω un ouvert de \mathbf{C} . Montrer que $f : \Omega \rightarrow \mathbf{C}$ est holomorphe si et seulement si f est continue et $\int_{[a,b]} f(z) dz + \int_{[b,c]} f(z) dz + \int_{[c,a]} f(z) dz = 0$ quels que soient $a, b, c \in \Omega$ tels que le triangle plein (i.e. avec son intérieur) de sommets a, b et c soit inclus dans Ω .

VI.2. La formule des résidus de Cauchy

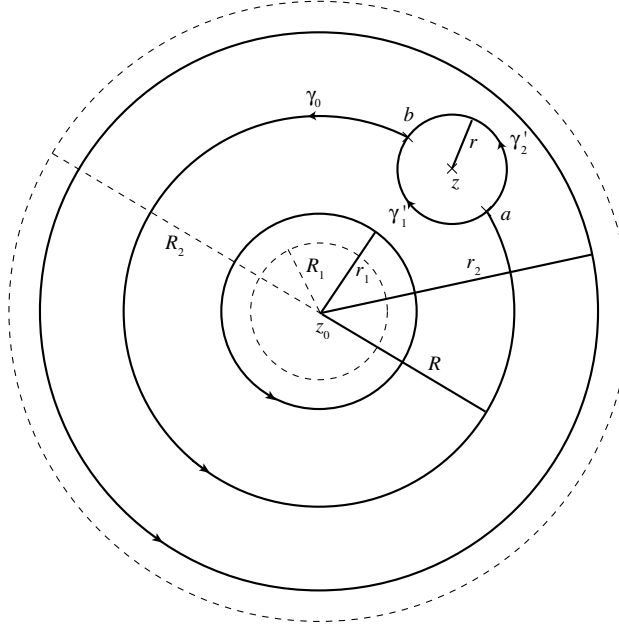
1. Fonctions holomorphes sur une couronne

Si $z_0 \in \mathbf{C}$ et si $0 \leq R_1 < R_2$, on note $C(z_0, R_1, R_2)$ la couronne ouverte des $z \in \mathbf{C}$, avec $R_1 < |z - z_0| < R_2$. Si $R_1 = 0$, on obtient le disque époinché ouvert de centre z_0 et de rayon R_2 (i.e. $D(z_0, R_2^-) - \{z_0\}$).

Lemme VI.2.1. — Si f est holomorphe sur la couronne $C(z_0, R_1, R_2)$, si $z \in C(z_0, R_1, R_2)$, et si $R_1 < r_1 < |z - z_0| < r_2 < R_2$, alors

$$f(z) = \frac{1}{2i\pi} \int_{C(z_0, r_2)} \frac{f(w)}{w - z} dw - \frac{1}{2i\pi} \int_{C(z_0, r_1)} \frac{f(w)}{w - z} dw.$$

Démonstration. — Le lecteur est invité à suivre les arguments qui suivent⁽⁵⁾ sur un dessin.



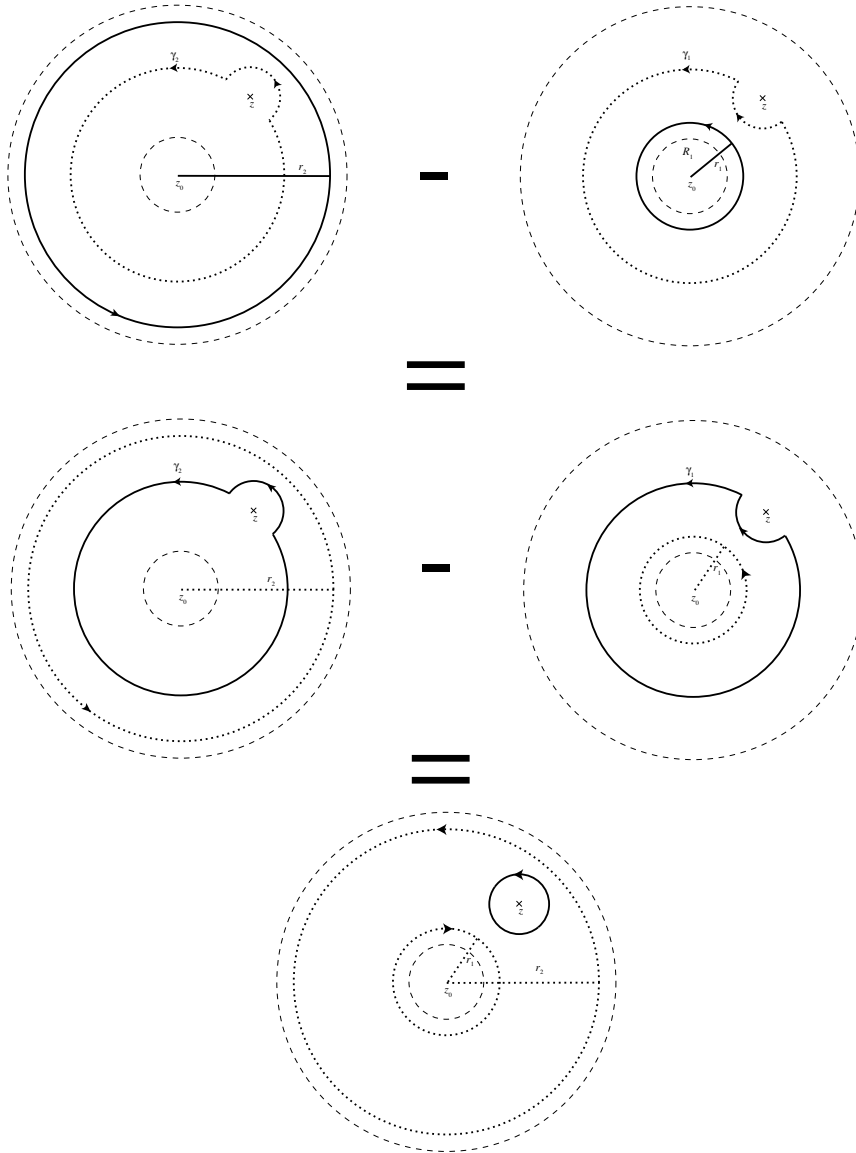
Soit $r < \inf(r_2 - |z - z_0|, |z - z_0| - r_1)$ de telle sorte que $D(z, r) \subset C(z_0, r_1, r_2)$. Soit $R \in]|z - z_0| - r, |z - z_0| + r[$ de telle sorte que les cercles $C(z_0, R)$ et $C(z, r)$ s'intersectent en deux points a et b . Soit γ_0 l'arc du cercle $C(z_0, R)$ à l'extérieur de $D(z, r)$, parcouru dans le sens direct; quitte à échanger les noms de a et b , c'est un chemin allant de b à a . Soit γ'_1 (resp. γ'_2) l'arc du cercle $C(z, r)$ à l'intérieur (resp. l'extérieur) de $D(z_0, R)$, parcouru dans le sens rétrograde (resp. direct); c'est un chemin allant de a à b . Si $i \in \{1, 2\}$, soit γ_i le lacet dans $C(z_0, r_1, r_2) - \{z\}$ obtenu en composant γ_0 , avec γ'_i ; c'est un lacet homotope⁽⁶⁾

⁽⁵⁾Il peut aussi préférer utiliser la forme générale de la formule de Stokes, dont on déduit directement que $\int_{C(z_0, r_i)} \frac{f(w)}{w - z} dw = \int_{\gamma_i} \frac{f(w)}{w - z} dw$, car f est holomorphe sur l'ouvert délimité par $C(z_0, r_i)$ et γ_i .

⁽⁶⁾C'est parfaitement évident sur un dessin, mais on peut aussi écrire une formule d'homotopie explicite. Par exemple, pour aller de γ_2 à $C(z_0, r_2)$, on paramètre γ_0 par $\gamma_0(t) = z_0 + R e^{2i\pi(t-\beta)}$, où β est choisit pour que $\gamma_0(t) = b$, et $t \in [0, \alpha]$, avec $\alpha < 1$ vérifiant $\gamma_0(\alpha) = a$, puis on paramètre γ'_2 sous la forme $\gamma'_2(t) = z + r e^{2i\pi(\lambda t + \mu)}$, avec $t \in [\alpha, 1]$, ce qui nous fournit un paramétrage $t \mapsto \gamma_2(t)$, avec $t \in [0, 1]$, de γ_2 . Finalement, on paramètre $C(z_0, r_2)$ sous la forme $t \mapsto \delta(t) = z_0 + r_2 e^{2i\pi(t-\beta)}$, avec $t \in [0, 1]$, et on fabrique une homotopie $u(s, t)$ de γ_2 sur $C(z_0, r_2)$, en posant $u(s, t) = (1 - s)\gamma_2(t) + s\delta(t)$. Il reste à vérifier que cette homotopie a bien lieu dans $C(z_0, R_1, R_2) - \{z\}$, ce qui est évident sur un dessin (on s'est débrouillé pour que l'angle entre $\gamma_2(t) - z_0$ et $\delta(t) - z_0$ soit petit, et $u(s, t)$ parcourt le segment $[\gamma_2(t), \delta(t)]$ quand s varie de 0 à 1).

dans $C(z_0, R_1, R_2) - \{z\}$ à $C(z_0, r_i)$. On a donc

$$\begin{aligned} \int_{C(z_0, R_2)} \frac{f(w)}{w-z} dw - \int_{C(z_0, R_1)} \frac{f(w)}{w-z} dw &= \int_{\gamma_2} \frac{f(w)}{w-z} dw - \int_{\gamma_1} \frac{f(w)}{w-z} dw \\ &= \int_{\gamma'_2} \frac{f(w)}{w-z} dw - \int_{\gamma'_1} \frac{f(w)}{w-z} dw = \int_{C(z, r)} \frac{f(w)}{w-z} dw. \end{aligned}$$



On conclut en utilisant la formule de Cauchy pour la fonction f , qui est holomorphe sur $D(z, r^-)$.

Corollaire VI.2.2. — Si f est holomorphe sur $C(z_0, R_1, R_2)$, il existe une unique suite $(a_n)_{n \in \mathbf{Z}}$ d'éléments de \mathbf{C} vérifiant

- (i) $\sum_{n=0}^{+\infty} a_n z^n$ converge si $|z| < R_2$ et $\sum_{n=-\infty}^{-1} a_n z^n$ converge si $|z|^{-1} < R_1^{-1}$;
 - (ii) $f(z)$ est somme de la série (de Laurent) $\sum_{n \in \mathbf{Z}} a_n (z - z_0)^n$, si $z \in C(z_0, R_1, R_2)$.
- De plus, quel que soient $r \in]R_1, R_2[$ et $n \in \mathbf{Z}$, on a

$$a_n = \frac{1}{2i\pi} \int_{C(z_0, r)} (w - z_0)^{-n-1} f(w) dw.$$

Démonstration. — Si $r, r' \in]R_1, R_2[$, les cercles $C(z_0, r)$ et $C(z_0, r')$ sont homotopes dans $C(z_0, R_1, R_2)$. Comme $(z - z_0)^{-n-1} f(z)$ est holomorphe dans $C(z_0, R_1, R_2)$, cela permet de montrer que $\frac{1}{2i\pi} \int_{C(z_0, r)} (z - z_0)^{-n-1} f(z) dz$ ne dépend pas du choix de $r \in]R_1, R_2[$; notons le a_n . Maintenant, d'après le lemme VI.2.1, si $R_1 < r_1 < |z - z_0| < r_2 < R_2$, on a

$$f(z) = \frac{1}{2i\pi} \int_{C(z_0, r_2)} \frac{f(w)}{w - z} dw - \frac{1}{2i\pi} \int_{C(z_0, r_1)} \frac{f(w)}{w - z} dw.$$

Par ailleurs,

$$\frac{1}{w - z} = \begin{cases} \sum_{n=0}^{+\infty} \frac{(z - z_0)^n}{(w - z_0)^{n+1}}, & \text{si } |w - z_0| = r_2, \\ -\sum_{n=0}^{+\infty} \frac{(w - z_0)^n}{(z - z_0)^{n+1}} = -\sum_{n=-\infty}^{-1} \frac{(z - z_0)^n}{(w - z_0)^{n+1}}, & \text{si } |w - z_0| = r_1. \end{cases}$$

On en déduit, en reportant les développements ci-dessus dans l'intégrale comme dans la démonstration de la prop. V.2.6, la convergence normale (pour la norme uniforme) de la série $\sum_{n \in \mathbf{Z}} a_n (z - z_0)^n$ vers $f(z)$ sur la couronne $r_1 < |z - z_0| < r_2$. On conclut en faisant tendre r_1 vers R_1 et r_2 vers R_2 .

2. Fonctions holomorphes sur un disque épointé; résidus

Soient Ω un ouvert de \mathbf{C} , $z_0 \in \Omega$, et f holomorphe sur $\Omega - \{z_0\}$. Soit $r > 0$ tel que $D(z_0, r^-) \subset \Omega$. La fonction f est donc holomorphe sur le disque épointé $D(z_0, r^-) - \{z_0\}$ qui peut aussi être vu comme la couronne $C(z_0, 0, r)$. On peut donc lui appliquer les résultats du n° 1, et on en déduit l'existence d'une suite $(a_n)_{n \in \mathbf{Z}}$ de nombres complexes vérifiant les propriétés suivantes :

- la série $g(z) = \sum_{n=0}^{+\infty} a_n (z - z_0)^n$ converge sur $D(z_0, r^-)$;
- la série $h(z) = \sum_{n \leq -1} a_n (z - z_0)^n$ converge, et définit une fonction holomorphe, sur $\mathbf{C} - \{z_0\}$;
- $f(z) = \sum_{n \in \mathbf{Z}} a_n (z - z_0)^n$, quel que soit $z \in D(z_0, r^-) - \{z_0\}$.

La série $h(z)$ est la *partie singulière* de f en z_0 , et le coefficient a_{-1} de $(z - z_0)^{-1}$ est le *résidu* $\text{Res}(f, z_0)$ de f en z_0 .

On dit que f est *holomorphe*⁽⁷⁾ en z_0 , si $h = 0$. Plus généralement, on dit que f est *méromorphe* en z_0 , s'il existe $k \in \mathbf{Z}$, avec $a_n = 0$, si $n \leq k$. Si k est un entier ≥ 1 , on dit

⁽⁷⁾C'est un abus de langage, mais cela veut dire que la fonction f peut se prolonger par continuité en z_0 , et que la fonction obtenue est holomorphe sur $D(z_0, r^-)$.

que f a un pôle d'ordre k en z_0 , si $a_{-k} \neq 0$ et $a_n = 0$, quel que soit $n < -k$ ⁽⁸⁾. On dit que f a une singularité essentielle en z_0 si elle n'est pas méromorphe en z_0 .

On peut reformuler les définitions ci-dessus en terme de la valuation $v_{z_0}(f)$ de f en z_0 définie par

$$v_{z_0}(f) = \inf_{n \in \mathbf{Z}, a_n \neq 0} n \in \mathbf{Z} \cup \{\pm\infty\}.$$

On a alors :

- $v_{z_0}(f) = -\infty \iff f$ a une singularité essentielle en z_0 ;
- $v_{z_0}(f) > -\infty \iff f$ est méromorphe en z_0 ;
- $v_{z_0}(f) = -k, k \in \mathbf{N} - \{0\} \iff f$ a un pôle d'ordre k en z_0 ;
- $v_{z_0}(f) \geq 0 \iff f$ est holomorphe en z_0 ;
- $v_{z_0}(f) = 0 \iff f$ est holomorphe et ne s'annule pas en z_0 ;
- $v_{z_0}(f) = k, k \in \mathbf{N} \iff f$ a un zéro d'ordre k en z_0 ;
- $v_{z_0}(f) = +\infty \iff f$ est nulle dans la composante connexe de z_0 dans Ω .

Dans les équivalences précédentes, la seule qui ne soit pas une reformulation de la définition est la dernière, qui est une reformulation du théorème des zéros isolés (th. V.1.15).

Si Ω est un ouvert de \mathbf{C} , on dit que $f : \Omega \rightarrow \mathbf{C}$ est *méromorphe* sur Ω , si f est méromorphe en tout point de Ω .

Pour un certain nombre d'applications, il est important de savoir calculer explicitement le résidu $\text{Res}(f, z_0)$ de f en z_0 . L'exercice ci-dessous fournit quelques recettes ⁽⁹⁾.

Exercice VI.2.3. — Soient Ω un ouvert de \mathbf{C} et $z_0 \in \Omega$.

- (i) Si f est holomorphe en z_0 , alors $\text{Res}(f, z_0) = 0$.
- (ii) Si $f = \frac{g}{h}$, où g et h sont holomorphes sur Ω , si $g(z_0) \neq 0$, et si h a un zéro simple en z_0 , alors $\text{Res}(f, z_0) = \frac{g(z_0)}{h'(z_0)}$.
- (iii) Si f a un pôle simple en z_0 , et si g est holomorphe en z_0 , alors $\text{Res}(gf, z_0) = g(z_0)\text{Res}(f, z_0)$.
- (iv) Si $k \geq 1$, et $f = (z - z_0)^{-k}g$, où g est holomorphe sur Ω , alors $\text{Res}(f, z_0) = \frac{g^{(k-1)}(z_0)}{(k-1)!}$.
- (v) Si f est méromorphe sur Ω , alors $\frac{f'}{f}$ est méromorphe sur Ω , avec des pôles simples aux pôles et zéros de f , et on a

$$\text{Res}\left(\frac{f'}{f}, z_0\right) = v_{z_0}(f), \quad \text{quel que soit } z_0 \in \Omega.$$

Exercice VI.2.4. — Soit $\lambda \in \mathbf{R}_+^*$. Calculer $\text{Res}\left(\frac{1}{z^2 + \lambda^2} e^{1/z}, 0\right)$.

⁽⁸⁾Une fonction est donc méromorphe si et seulement si elle est holomorphe ou a un pôle d'ordre k , avec $k \in \mathbf{N}$

⁽⁹⁾Si f a une singularité essentielle, il est en général impossible de trouver une expression « finie » du résidu. C'est une des raisons qui fait que certaines intégrales résistent à la méthode des résidus (cf. exercices du n° 5).

Exercice VI.2.5. — Montrer que f est méromorphe sur Ω , si et seulement si tout point z_0 de Ω a un voisinage⁽¹⁰⁾ ouvert U sur lequel f peut s'écrire sous la forme $f = \frac{g}{h}$, avec g et h holomorphes sur U .

Exercice VI.2.6. — (i) Montrer qu'une fonction méromorphe bornée à l'infini (il existe M et R tels que $|f(z)| \leq M$, si $|z| \geq R$) est une fraction rationnelle.

(ii) Montrer qu'une fonction méromorphe tendant vers l'infini à l'infini est une fraction rationnelle.

Exercice VI.2.7. — Soit $\mathcal{H} = \{z \in \mathbf{C}, \operatorname{Im}(z) > 0\}$ le demi-plan de Poincaré, et soit $f : \mathcal{H} \rightarrow \mathbf{C}$ une fonction holomorphe, périodique de période 1.

(i) Montrer qu'il existe $\tilde{f} : D(0, 1^-) - \{0\} \rightarrow \mathbf{C}$, holomorphe, telle que $f(z) = \tilde{f}(e^{2i\pi} z)$.

(ii) En déduire qu'il existe une suite $(a_n(f))_{n \in \mathbf{Z}}$ de nombres complexes, telle que l'on ait $f(z) = \sum_{n \in \mathbf{Z}} a_n(f) e^{2i\pi n z}$, quel que soit $z \in \mathcal{H}$, la série étant normalement convergente sur toute bande horizontale $a \leq \operatorname{Im}(z) \leq b$, avec $0 < a < b < +\infty$.

Exercice VI.2.8. — Soit $(x_n)_{n \in \mathbf{N}}$ une suite d'éléments de \mathbf{C}^* tendant vers l'infini, et soit $(k_n)_{n \in \mathbf{N}}$ une suite d'éléments de \mathbf{N} .

(i) Montrer qu'il existe $a_n \in \mathbf{N}$ tel que, si $|z| < \frac{|x_n|}{2}$, alors

$$\left| \left(\left(1 - \frac{z}{x_n} \right) e^{\frac{z}{x_n} + \frac{z^2}{2x_n^2} + \dots + \frac{z^{a_n}}{a_n x_n^{a_n}}} \right)^{k_n} - 1 \right| \leq 2^{-n}.$$

(ii) Construire une fonction holomorphe sur \mathbf{C} dont les zéros sont les x_n avec multiplicité k_n .

(iii) Montrer que toute fonction méromorphe sur \mathbf{C} est quotient de deux fonctions holomorphes sur \mathbf{C} .

(iv) Montrer, en adaptant la méthode, que toute fonction méromorphe sur $D(0, 1^-)$ est quotient de deux fonctions holomorphes sur $D(0, 1^-)$.

Exercice VI.2.9. — Soient $R > 0$ et f holomorphe sur $D(z_0, R^-) - \{z_0\}$. Montrer que :

(i) f est holomorphe en z_0 , si et seulement si f est bornée dans $D(z_0, r) - \{z_0\}$, quel que soit $r \in]0, R[$ (on pourra s'intéresser à $\frac{1}{2i\pi} \int_{C(z_0, r)} (z - z_0)^{-1-n} f(z) dz$);

(ii) f est méromorphe non holomorphe en z_0 si et seulement si $|f(z)| \rightarrow +\infty$ quand $z \rightarrow z_0$;

(iii) f a une singularité essentielle en z_0 si et seulement si l'image de $D(z_0, r^-) - \{z_0\}$ par f est dense⁽¹¹⁾ dans \mathbf{C} , quel que soit $r \in]0, R[$.

Exercice VI.2.10. — (i) Soit $f : \mathbf{C} \rightarrow \mathbf{C}$ holomorphe. Montrer, en utilisant l'exercice précédent, que si f n'est pas un polynôme, alors $f(\mathbf{C} - D(0, n))$ est un ouvert dense de \mathbf{C} , quel que soit $n \in \mathbf{N}$. En déduire que f n'est pas injective.

(ii) Montrer que si $f : \mathbf{C} \rightarrow \mathbf{C}$ est holomorphe bijective, alors il existe $a \in \mathbf{C}^*$ et $b \in \mathbf{C}$ tels que $f(z) = az + b$, quel que soit $z \in \mathbf{C}$.

3. Indice d'un lacet par rapport à un point

Soit $z_0 \in \mathbf{C}$, et soit γ un lacet ne passant pas par z_0 . Notre but est de définir mathématiquement le nombre de tours que fait γ autour de z_0 .

⁽¹⁰⁾On peut montrer que l'on peut écrire $f = \frac{g}{h}$, avec g et h holomorphes sur Ω tout entier, mais c'est loin d'être évident si Ω est un ouvert quelconque.

⁽¹¹⁾Le « grand théorème de Picard » affirme qu'en fait l'image de $D(z_0, r^-) - \{z_0\}$ par f contient \mathbf{C} à au plus un point près (l'exemple de $e^{1/z}$ montre que ce résultat est optimum). La démonstration repose sur des techniques un peu plus sophistiquées que celles introduites dans ce cours.

3.1. Définition

Lemme VI.2.11. — Soit $\gamma : [a, b] \rightarrow \mathbf{C} - \{z_0\}$ un chemin \mathcal{C}^1 par morceaux. Si $t \in [a, b]$, soit γ_t la restriction de γ à $[a, t]$, et soit $f(t) = \int_{\gamma_t} \frac{dz}{z - z_0}$. Alors $e^{-f(t)}(\gamma(t) - z_0)$ est constante sur $[a, b]$.

Démonstration. — Par définition, on a $f(t) = \int_0^t \frac{\gamma'(u)}{\gamma(u) - z_0} du$, et donc $f'(t) = \frac{\gamma'(t)}{\gamma(t) - z_0}$, si t n'est pas un point anguleux de γ . La dérivée de $e^{-f(t)}(\gamma(t) - z_0)$ est donc

$$-f'(t)e^{-f(t)}(\gamma(t) - z_0) + e^{-f(t)}\gamma'(t) = e^{-f(t)}\left(-\frac{\gamma'(t)}{\gamma(t) - z_0}(\gamma(t) - z_0) + \gamma'(t)\right) = 0,$$

et f est constante par morceaux; comme elle est continue cela permet de conclure.

Corollaire VI.2.12. — Si $\gamma : [a, b] \rightarrow \mathbf{C}$ est un lacet \mathcal{C}^1 par morceaux, ne rencontrant pas z_0 , alors $I(\gamma, z_0) = \frac{1}{2i\pi} \int_{\gamma} \frac{dz}{z - z_0}$ est un entier.

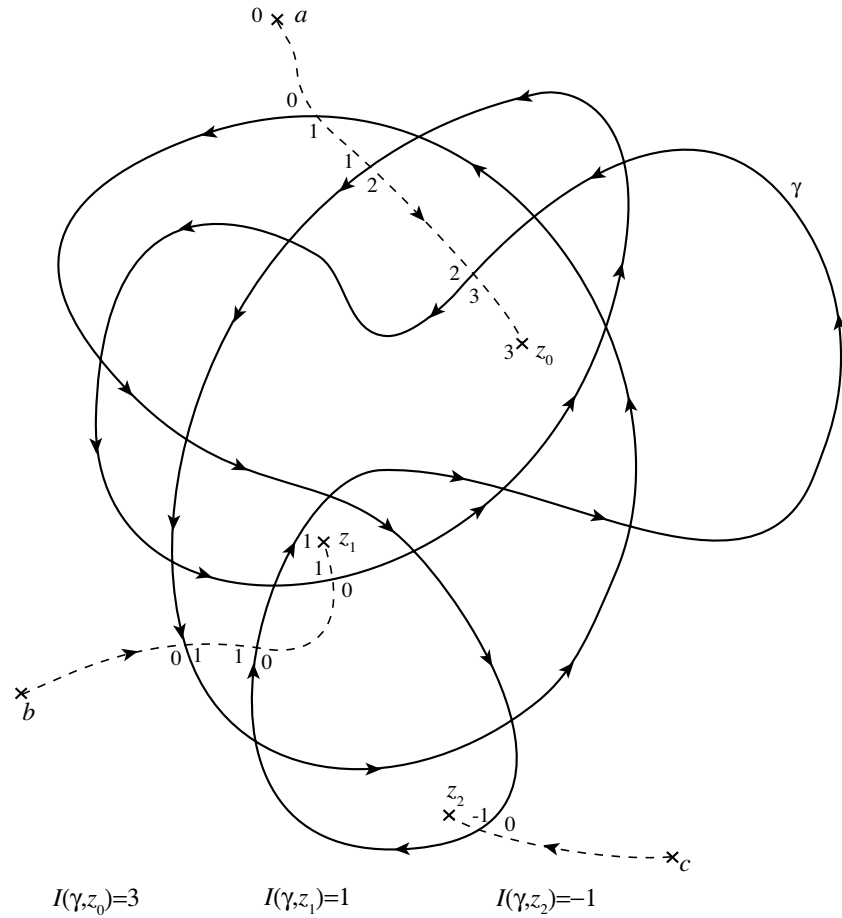
Démonstration. — En reprenant les notations du lemme, on voit que $e^{f(b)} = e^{f(a)} \frac{\gamma(b) - z_0}{\gamma(a) - z_0}$, et, comme γ est un lacet, que $\exp\left(\int_{\gamma} \frac{dz}{z - z_0}\right) = e^{f(b)} = e^{f(a)} = 1$, ce qui permet de conclure.

L'entier $I(\gamma, z_0)$ défini ci-dessus est l'indice de γ par rapport à z_0 . Par exemple, si γ est le chemin $C(a, r)$, et si $|z_0 - a| < r$, la formule de Cauchy utilisée pour la fonction constante 1, montre que $I(C(a, r), z_0) = 1$. Par contre, si $|z_0 - a| > r$, le cercle $C(a, r)$ est homotope à $\{a\}$ dans $\mathbf{C} - \{z_0\}$, et comme $\frac{1}{z - z_0}$ est holomorphe sur $\mathbf{C} - \{z_0\}$, on a $I(C(a, r), z_0) = 0$. Autrement dit, si z_0 est à l'intérieur du cercle parcouru dans le sens direct, alors l'indice du cercle par rapport à z_0 est 1, alors que si z_0 est à l'extérieur, cet indice est 0, ce qui est en accord avec l'idée selon laquelle $I(\gamma, z_0)$ représente le nombre de tours que fait γ autour de z_0 . On fera attention au fait que, si le cercle est parcouru dans le sens rétrograde, son indice par rapport à un point à l'intérieur est -1 .

3.2. Détermination visuelle de l'indice d'un lacet par rapport à un point

Dans les applications que l'on verra dans ce cours, les lacets considérés seront sans point double, et un théorème de Jordan affirme qu'un tel lacet découpe le plan en deux régions, l'une à l'intérieur du lacet et l'autre à l'extérieur du lacet, auquel cas on se retrouve dans le même cas de figure que pour le cercle. Dans le cas général, déterminer sur un dessin l'indice d'un lacet par rapport à un chemin peut donner le tournis, mais la prop. VI.2.13 ci-dessous fournit un procédé permettant un calcul en regardant toujours droit devant soi. La recette est la suivante. On choisit un point $a \in \mathbf{C}$ assez grand pour que $\gamma \subset D(0, |a|^-)$, et on trace un chemin u allant de a à z_0 . Alors $I(\gamma, z_0) = n_g - n_d$, où n_g (resp. n_d) est le nombre de points d'intersections de u et γ , où γ arrive de la gauche⁽¹²⁾ (resp. de la droite), quand on va de a à z_0 .

⁽¹²⁾Pour que ceci ait un sens, il vaut mieux prendre quelques précautions; par exemple, imposer que u et γ ne s'intersectent qu'en des points simples non anguleux de γ , et que les tangentes à γ et à u en un point d'intersection ne soient pas colinéaires.



Soit γ un lacet dans \mathbf{C} . Comme γ est compact (en tant qu'image d'un intervalle compact par une application continue), son complémentaire est ouvert, et chacune des composantes connexes du complémentaire est un ouvert. D'autre part, si R est assez grand, γ est inclus dans $D(0, R)$, et comme $\mathbf{C} - D(0, R)$ est connexe, une (et une seule) des composantes connexes de $\mathbf{C} - \gamma$ contient $\mathbf{C} - D(0, R)$, pour R assez grand. Cette composante connexe est la composante connexe de l'infini.

Proposition VI.2.13. — Soit γ un lacet \mathcal{C}^1 par morceaux.

- (i) $z \mapsto I(\gamma, z)$ est constant sur chacune des composantes connexes de $\mathbf{C} - \gamma$.
- (ii) $I(\gamma, z) = 0$ si z est dans la composante connexe de l'infini de $\mathbf{C} - \gamma$.
- (iii) Soit $t \mapsto \gamma(t)$, $t \in [a, b]$ un paramétrage de γ , \mathcal{C}^1 par morceaux, et soit $t_0 \in]a, b[$ tel que $\gamma(t_0)$ soit un point simple de γ , et $\gamma'(t_0) \neq 0$. Si $h \in \mathbf{C}$ vérifie $\text{Im}(h) > 0$, et si $\varepsilon > 0$ est assez petit, alors
 - $\gamma(t_0) + \varepsilon h \gamma'(t_0)$ et $\gamma(t_0) - \varepsilon h \gamma'(t_0)$ n'appartiennent pas à γ ,
 - $I(\gamma, \gamma(t_0) + \varepsilon h \gamma'(t_0)) - I(\gamma, \gamma(t_0) - \varepsilon h \gamma'(t_0)) = 1$.

Démonstration. — Le théorème de continuité d'une intégrale dépendant d'un paramètre montre que $z_0 \mapsto I(\gamma, z_0)$ est continue sur $\mathbf{C} - \gamma$, et comme $I(\gamma, z_0)$ est à valeurs dans \mathbf{Z} , qui est discret, cela démontre le (i).

Si $\gamma \subset D(0, R)$, et si $|z_0| > R$, alors γ est homotope à un point dans $\mathbf{C} - \{z_0\}$ (et même dans $D(0, R)$, puisque $D(0, R)$ est contractile), et comme $\frac{1}{z - z_0}$ est holomorphe sur $\mathbf{C} - \{z_0\}$, on a $\int_{\gamma} \frac{dz}{z - z_0} = 0$. On en déduit le (ii) pour $|z_0| > R$, et le (i) permet de terminer la démonstration du (ii).

Passons à la démonstration du (iii). Choisissons $h \in \mathbf{C}$, avec $\operatorname{Im}(h) > 0$. On a donc $0 < \arg(h) < \pi$. Quitte à faire une translation sur la variable, on peut supposer $t_0 = 0$, et quitte à transformer la situation par la similitude $z \mapsto \gamma'(t_0)^{-1}(z - \gamma(t_0))$, on peut supposer que $\gamma(t_0) = 0$, et $\gamma'(t_0) = 1$, ce qui permet de simplifier un peu les formules. On a alors $\lim_{t \rightarrow 0} t^{-1}\gamma(t) = 1$, ce qui implique qu'il existe $\delta > 0$ tel que, quel que soit $t \in [-\delta, \delta] - \{0\}$,

- $|\arg(t^{-1}\gamma(t))| < \frac{1}{2} \inf(\arg(h), \pi - (\arg(h))),$
- $|t^{-1}\gamma(t) - 1| \leq 1/2.$

Par ailleurs, comme $0 = \gamma(0)$ est un point simple de γ , il existe $\eta > 0$ tel que $d(0, \gamma(t)) \geq \eta$, si $t \in [a, b] -]-\delta, \delta[$. On en déduit que, si $0 < \varepsilon < |h|^{-1}\eta$, alors $\pm\varepsilon h$ n'est pas de la forme $\gamma(t)$, avec $t \in [a, b] -]-\delta, \delta[$. Par ailleurs, $\pm\varepsilon h$ n'est pas non plus de la forme $\gamma(t)$, avec $t \in [-\delta, \delta]$, puisque $\pm\varepsilon h \neq 0$ et $\arg(\frac{\pm\varepsilon h}{\gamma(t)}) \neq 0$, si $t \in [-\delta, \delta] - \{0\}$, d'après le choix de δ . En résumé, si $0 < \varepsilon < |h|^{-1}\eta$, alors $\pm\varepsilon h$ n'appartient pas à γ .

Maintenant, on a $I(\gamma, \varepsilon h) - I(\gamma, -\varepsilon h) = I_1(\varepsilon) + I_2(\varepsilon)$, où l'on a posé

$$I_1(\varepsilon) = \frac{1}{2i\pi} \int_{\gamma([a, -\delta] \cup \gamma([-\delta, b])} \frac{dz}{z - \varepsilon h} - \frac{dz}{z + \varepsilon h} \quad \text{et} \quad I_2(\varepsilon) = \frac{1}{2i\pi} \int_{\gamma([-\delta, \delta])} \frac{dz}{z - \varepsilon h} - \frac{dz}{z + \varepsilon h}.$$

On a $\lim_{\varepsilon \rightarrow 0^+} I_1(\varepsilon) = 0$, puisque $\frac{1}{z - \varepsilon h} - \frac{1}{z + \varepsilon h}$ tend vers 0 et est majoré en module par $\frac{2}{\eta - \varepsilon|h|}$. Comme $I(\gamma, \varepsilon h) - I(\gamma, -\varepsilon h) \in \mathbf{Z}$, il suffit donc, pour montrer que $I(\gamma, \varepsilon h) - I(\gamma, -\varepsilon h) = 1$, si $\varepsilon > 0$ est assez petit, de prouver que $\lim_{\varepsilon \rightarrow 0^+} I_2(\varepsilon) = 1$. Pour cela, écrivons $2i\pi I_2(\varepsilon)$ sous la forme

$$2i\pi I_2(\varepsilon) = \int_{\gamma([-\delta, \delta])} \frac{2\varepsilon h}{z^2 - \varepsilon^2 h^2} dz = \int_{-\delta}^{\delta} \frac{2\varepsilon h \gamma'(t)}{\gamma(t)^2 - \varepsilon^2 h^2} dt = \int_{-\delta/\varepsilon}^{\delta/\varepsilon} \frac{2h \gamma'(\varepsilon t)}{\varepsilon^{-2} \gamma(\varepsilon t)^2 - h^2} dt.$$

L'expression sous l'intégrale tend vers $\frac{2h}{t^2 - h^2}$ quand ε tend vers 0 et $t \in \mathbf{R}$. Par ailleurs, d'après le choix de δ , il existe $c > 0$ tel que $2\pi - c > \arg(h^2 \gamma(u)^{-2}) > c$, si $u \in [-\delta, \delta]$; on en déduit l'existence de $C > 0$ (avec $C = 1$, si $\cos c \leq 0$, et $C = |\sin c|$, si $\cos c \geq 0$) tel que $|\lambda \gamma(t)^2 - \mu h^2| \geq C \sup(|\lambda \gamma(t)^2|, |\mu h^2|)$, quels que soient $\mu, \lambda \in \mathbf{R}_+$. Comme $|\gamma(u)| \geq \frac{1}{2}|u|$, et comme il existe $M \geq 0$ tel que $|\gamma'(u)| \leq M$, si $u \in [-\delta, \delta]$, on peut majorer, en module, $\frac{2h \gamma'(\varepsilon t)}{\varepsilon^{-2} \gamma(\varepsilon t)^2 - h^2}$ sur $[-\delta/\varepsilon, \delta/\varepsilon]$, par $\frac{2|h|M}{C \sup(t^2/4, |h|^2)}$, qui est une fonction sommable sur \mathbf{R} , indépendante de ε . On obtient donc, via le théorème de continuité d'une intégrale dépendant d'un paramètre,

$$\lim_{\varepsilon \rightarrow 0^+} 2i\pi I_2(\varepsilon) = \int_{-\infty}^{+\infty} \frac{2h}{t^2 - h^2} dt.$$

Finalement, en posant $h = \sigma + i\tau$, avec $\tau > 0$ par hypothèse, cette dernière intégrale devient.

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{dt}{t - h} - \frac{dt}{t + h} &= \int_{-\infty}^{+\infty} \frac{(t - \sigma) + i\tau}{(t - \sigma)^2 + \tau^2} - \frac{(t + \sigma) - i\tau}{(t + \sigma)^2 + \tau^2} dt \\ &= \left[\frac{1}{2} \log \frac{(t - \sigma)^2 + \tau^2}{(t + \sigma)^2 + \tau^2} \right]_{-\infty}^{+\infty} + i \left[\operatorname{artg} \frac{t - \sigma}{\tau} + \operatorname{artg} \frac{t + \sigma}{\tau} \right]_{-\infty}^{+\infty} = 2i\pi, \end{aligned}$$

ce qui permet de conclure. (On peut fabriquer (exercice) une démonstration de ce que $I_2(\varepsilon) \rightarrow 1$ en remarquant qu'au voisinage de 0, γ est le graphe d'une fonction, ce qui permet d'écrire $2i\pi I_2(\varepsilon)$ comme l'intégrale de $\frac{dz}{z}$ sur un contour C_ε ressemblant à un parallélogramme de sommets $\pm\delta \pm \varepsilon h$, privé des cotés verticaux (dont la longueur tend vers 0). La formule des résidus montre que l'intégrale sur C_ε est égale à $2i\pi$. On évite de cette manière les majorations ci-dessus.)

4. La formule des résidus

Théorème VI.2.14. — Soient Ω un ouvert de \mathbf{C} , F un ensemble fini de points de Ω , f une fonction holomorphe sur $\Omega - F$, et γ un lacet de Ω , contractile dans Ω , et ne

rencontrant pas F . Alors

$$\int_{\gamma} f(z) dz = 2i\pi \sum_{a \in F} I(\gamma, a) \operatorname{Res}(f, a).$$

Démonstration. — Si $a \in F$, et si $r_a > 0$ est tel que $D(a, r_a^-)$ est inclus dans Ω et ne contient aucun autre point de F , alors il existe une suite $(c_{a,n})_{n \in \mathbf{Z}}$ d'éléments de \mathbf{C} , telle que $f(z) = \sum_{n \in \mathbf{Z}} c_{a,n} (z - a)^n$, si $z \in D(a, r_a^-)$. De plus, la série $\sum_{n \leq -2} c_{a,n} (z - a)^n$ définit une fonction holomorphe g_a dans $\mathbf{C} - \{a\}$, et g_a admet $G_a(z) = \sum_{n \leq -2} c_{a,n} \frac{(z-a)^{n+1}}{n+1}$ comme primitive sur $\mathbf{C} - \{a\}$. Soit alors $h = f - \sum_{a \in F} (g_a + \frac{c_{a,-1}}{z-a})$. Par construction, h est holomorphe sur $\Omega - F$, et a une singularité fictive en tous les points de F ; elle se prolonge donc, par continuité, en une fonction holomorphe sur Ω tout entier. On a alors

- $\int_{\gamma} f(z) dz = \int_{\gamma} h(z) dz + \sum_{a \in F} \int_{\gamma} g_a(z) dz + \sum_{a \in F} c_{a,-1} \int_{\gamma} \frac{dz}{z-a}$;
- $\int_{\gamma} h(z) dz = 0$ puisque h est holomorphe dans Ω et γ est contractile dans Ω ;
- $\int_{\gamma} g_a(z) dz = 0$, si $a \in F$, puisque g_a a une primitive sur $\Omega - \{a\}$ qui contient γ ;
- $c_{a,-1} = \operatorname{Res}(f, a)$ et $\int_{\gamma} \frac{dz}{z-a} = 2i\pi I(\gamma, a)$, par définition.

Ceci permet de conclure.

La formule des résidus permet de localiser les zéros d'une fonction holomorphe.

Corollaire VI.2.15. — Soit $\Omega \subset \mathbf{C}$ un ouvert connexe, et soient $z_0 \in \Omega$ et $r > 0$ tels que $D(z_0, r) \subset \Omega$. Soit f holomorphe sur Ω , non identiquement nulle. Si $C(z_0, r)$ ne contient aucun zéro de f , alors le nombre de zéros de f dans $D(z_0, r^-)$, comptés avec multiplicité, est égal à $\frac{1}{2i\pi} \int_{C(z_0, r)} \frac{f'(z)}{f(z)} dz$.

Démonstration. — On sait (cf. (v) de l'ex. VI.2.3) que $\frac{f'(z)}{f(z)}$ est méromorphe sur Ω , avec des pôles simples aux zéros de f , le résidu en chacun de ces pôles étant l'ordre du zéro de f . La formule des résidus permet de conclure.

5. Exercices

Exercice VI.2.16. — Soit f définie sur \mathbf{R} par $f(t) = e^{-\pi t^2}$. Rappel : $\int_{\mathbf{R}} e^{-\pi t^2} dt = 1$.

(i) Montrer que f se prolonge analytiquement en une fonction (encore notée f) holomorphe sur \mathbf{C} .

(ii) Si $a \in \mathbf{R}$ et $R \in \mathbf{R}_+$, soit $\gamma_{a,R}$ le lacet composé des segments $[-R, R]$, $[R, R + ai]$, $[R + ai, -R + ai]$ et $[-R + ai, -R]$. Que vaut $\int_{\gamma_{a,R}} f(z) dz$?

(iii) En déduire, en faisant tendre R vers $+\infty$, que $\hat{f}(a) = e^{-\pi a^2}$.

Exercice VI.2.17. — Calculer les intégrales suivantes par la méthode des résidus :

(a) $\int_{\mathbf{R}} \frac{dx}{(x^2+1)(x^2+2)(x^2+3)}$ et $\int_{\mathbf{R}} \frac{dx}{(x+i)(x-i)(x-2i)\dots(x-ni)}$. (On prendra un lacet γ formé du segment $[-R, R]$ et d'un demi-cercle convenable.)

(b) $\lim_{R \rightarrow +\infty} \int_{-R}^R \frac{x^5}{(x^2+1)(x^2+2)(x^2+x+1)} dx$.

(c) $\int_0^{+\infty} \frac{x^a}{x^b+1} dx$, avec $a, b \in \mathbf{N}$ et $b \geq a + 2$. (On prendra un lacet γ formé du segment $[0, R]$, d'un arc de cercle convenable, et d'un segment $[Re^{i\alpha}, 0]$, avec α bien choisi.)

Exercice VI.2.18. — (transformée de Fourier de fonctions rationnelles)

(i) Calculer $\int_{\mathbf{R}} \frac{e^{-2i\pi t\xi}}{t^2+1} dt$ et $\int_{\mathbf{R}} \frac{e^{-2i\pi t\xi}}{(t+i)^2} dt$ par la méthode des résidus. (On prendra un lacet γ formé du segment $[-R, R]$ et d'un demi-cercle convenable⁽¹³⁾)

(ii) Calculer de même $\lim_{R \rightarrow +\infty} \int_{-R}^R \frac{te^{-2i\pi t\xi}}{t^2+t+1} dt$.

Exercice VI.2.19. — Soit $\gamma_{\varepsilon, R}$ le lacet composé du segment $[\varepsilon, R]$, du demi-cercle $C^+(0, R) : [0, \pi] \rightarrow \mathbf{C}$, donné par $t \mapsto Re^{it}$, du segment $[-R, -\varepsilon]$, et demi-cercle $C^+(0, \varepsilon)^{\text{opp}}$, donné par $t \mapsto \varepsilon e^{i(\pi-t)}$. (Faire un dessin.)

(i) Calculer $\int_{\gamma_{\varepsilon, R}} \frac{e^{iz}}{z} dz$ via la formule des résidus.

(ii) Montrer que $\int_{C^+(0, R)} \frac{e^{iz}}{z} dz \rightarrow 0$ quand $R \rightarrow +\infty$.

(iii) Calculer la limite de $\int_{C^+(0, \varepsilon)^{\text{opp}}} \frac{e^{iz}}{z} dz$ quand $\varepsilon \rightarrow 0$.

(iv) En déduire que $\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$.

Exercice VI.2.20. — Nous nous proposons de calculer $\int_0^{+\infty} \frac{t^s}{t^2+t+1} dt$, pour $-1 < \text{Re}(s) < 1$ par la méthode des résidus (on peut adapter ce qui suit pour calculer les intégrales du type $\int_0^{+\infty} f(t)t^s dt$, où f est une fraction rationnelle). Si $r > 0$, on note $C^+(0, r)$ le demi-cercle $\theta \mapsto re^{i\theta}$, pour $\theta \in [0, \pi]$ et $C^-(0, r)$ le demi-cercle $\theta \mapsto re^{i\theta}$, pour $\theta \in [-\pi, 0]$. Si $0 < \varepsilon < R$, soit $\gamma_{\varepsilon, R}^+$ le lacet obtenu en composant $C^+(0, R)$, $[-R, -\varepsilon]$, $C^+(0, \varepsilon)^{\text{opp}}$ et $[\varepsilon, R]$ (faire un dessin!), et soit $\gamma_{\varepsilon, R}^-$ le lacet obtenu en composant $C^-(0, R)$, $[R, \varepsilon]$, $C^-(0, \varepsilon)^{\text{opp}}$ et $[-\varepsilon, -R]$.

(i) Calculer $\int_{\gamma_{\varepsilon, R}^+} \frac{(-z)^s}{z^2+z+1} dz$ et $\int_{\gamma_{\varepsilon, R}^-} \frac{(-z)^s}{z^2+z+1} dz$ en utilisant la formule des résidus, avec $\log(-z) \in \mathbf{R}$, si $z \in \mathbf{R}_+^*$. (Attention à la détermination du logarithme!).

(ii) Montrer que les intégrales sur les demi-cercles tendent vers 0 quand ε tend vers 0 et R vers $+\infty$.

(iii) Conclure, si $s \neq 0$, en considérant $\int_{\gamma_{\varepsilon, R}^+} \frac{(-z)^s}{z^2+z+1} dz + \int_{\gamma_{\varepsilon, R}^-} \frac{(-z)^s}{z^2+z+1} dz$.

(iv) En déduire le cas $s = 0$. (Ce n'est sûrement pas la méthode la plus rapide si on doit tout justifier, mais une fois qu'on a compris comment elle marche, on peut se contenter du calcul des résidus, et c'est très efficace).

Exercice VI.2.21. — Soit $P \in \mathbf{C}[X]$ unitaire de degré n . Calculer

$$\lim_{R \rightarrow +\infty} \frac{1}{2i\pi} \int_{C(0, R)} \frac{P'(z)}{P(z)} dz.$$

En déduire que \mathbf{C} est algébriquement clos.

Exercice VI.2.22. — Soit Ω un ouvert connexe de \mathbf{C} , Soit f une fonction holomorphe non identiquement nulle sur Ω , et soit f_n une suite de fonctions holomorphes tendant vers f uniformément sur tout compact de Ω .

(i) Soient $z_0 \in \Omega$ et $r > 0$ tels que $D(z_0, r) \subset \Omega$. On suppose que $C(z_0, r)$ ne contient aucun zéro de f . Montrer qu'il existe $N(z_0, r)$ tel que, si $n \geq N(z_0, r)$, alors f et f_n ont le même nombre de zéros, comptés avec multiplicité, dans $D(z_0, r^-)$.

(ii) Montrer que, si f_n est injective sur Ω , pour tout n assez grand, il en est de même de f .

⁽¹³⁾Une manière de voir que l'on a pris le bon demi-cercle est de vérifier que ce qu'on obtient tend vers 0 quand $|\xi| \rightarrow +\infty$; il faut aussi faire attention à l'indice du lacet par rapport aux pôles...

Exercice VI.2.23. — (Théorème de Rouché et applications).

Soient $R > 0$, $D = D(0, R)$ et $C = \partial D$ le cercle de centre 0 et de rayon R . Soient Ω un ouvert contenant D , f holomorphe sur Ω , ne s'annulant pas sur C , et g holomorphe sur Ω , telle que $|f(z) - g(z)| < |f(z)|$, si $z \in C$.

(i) Montrer qu'il existe $\Omega' \subset \Omega$ ouvert contenant C et h holomorphe sur Ω' tels que $\frac{g}{f} = e^h$ sur Ω' . Que vaut h' ?

(ii) Montrer que f et g ont le même nombre de zéros (comptés avec multiplicité) dans D (théorème de Rouché).

(iii) Soit G holomorphe sur Ω telle que $|G(z)| < R$, si $z \in C$. Montrer que $G(D) \subset D$ et que G a un unique point fixe dans D .

(iv) Montrer que toutes les solutions de $z \sin z = 1$ sont réelles.

Exercice VI.2.24. — Soit Γ la fonction méromorphe sur \mathbf{C} , holomorphe en dehors de pôles simples en les entiers négatifs, définie par $\Gamma(z) = \int_0^{+\infty} e^{-tz} \frac{dt}{t}$, si $\operatorname{Re}(z) > 0$. On rappelle (cf. ex. V.2.20) que l'on a $\Gamma(z) = \frac{\Gamma(z+n+1)}{z(z+1)\cdots(z+n)}$, quel que soit $n \in \mathbf{N}$.

(i) Montrer que $\lim_{z \rightarrow -n} (z+n)\Gamma(z) = \frac{(-1)^n}{n!}$.

(ii) Montrer que, si $x > 0$, et si $c \notin -\mathbf{N}$, alors $I_c(x) = \int_{c-i\infty}^{c+i\infty} x^{-z} \Gamma(z) dz$ converge.

(iii) Si $n \in \mathbf{N}$, calculer $I_1(x) - I_{\frac{1}{2}-n}(x)$ par la méthode des résidus.

(iv) Montrer que $I_{\frac{1}{2}-n}(x) \rightarrow 0$. En déduire que $\frac{1}{2i\pi} \int_{1-i\infty}^{1+i\infty} x^{-z} \Gamma(z) dz = e^{-x}$.

(v) Retrouver le résultat en utilisant la transformée de Fourier.

Exercice VI.2.25. — Si $N \in \mathbf{N}$, soit C_N le carré de sommets $(2N+1)\pi(\pm 1 \pm i)$ parcouru dans le sens direct, et soit $I_N = \int_{C_N} \frac{dz}{z^2(e^z-1)}$.

(i) Calculer I_N en utilisant la formule des résidus.

(ii) Montrer que $I_N \rightarrow 0$ quand $N \rightarrow +\infty$.

(iii) En déduire la formule $\zeta(2) = \frac{\pi^2}{6}$.

(iv) Soient B_n , pour $n \in \mathbf{N}$, définis par $\frac{z}{e^z-1} = \sum_{n=0}^{+\infty} B_n \frac{z^n}{n!}$. Adapter ce qui précède pour calculer $\zeta(2k)$, pour $k \in \mathbf{N} - \{0\}$, en fonction des B_n . En déduire que $\pi^{-2k} \zeta(2k) \in \mathbf{Q}$.

Exercice VI.2.26. — Calculer $\int_{\mathbf{R}} e^{2i\pi t\xi} \frac{e^{\pi t}}{e^{2\pi t}+1} dt$ de deux manières : d'une part en intégrant $e^{2i\pi z\xi} \frac{e^{\pi z}}{e^{2\pi z}+1} dz$ sur un rectangle convenable, en faisant tendre les sommets vers l'infini, d'autre part en écrivant $\frac{e^{\pi t}}{e^{2\pi t}+1}$ comme une série en $e^{\pi t}$ ou $e^{-\pi t}$ suivant que t est négatif ou positif. Comparer avec l'ex. V.2.14.

VI.3. Loi d'addition sur une courbe elliptique, problème corrigé

Soit (ω_1, ω_2) une base directe ($\operatorname{Im}(\frac{\omega_2}{\omega_1}) > 0$) de \mathbf{C} sur \mathbf{R} , et soit $\Lambda = \{m\omega_1 + n\omega_2, m, n \in \mathbf{Z}\}$. On dit que $f : \mathbf{C} \rightarrow \mathbf{C}$ est Λ -périodique si on a $f(z + \omega) = f(z)$, quels que soient $z \in \mathbf{C}$ et $\omega \in \Lambda$.

Partie I

(o) Montrer qu'il existe $C > 0$ tel que $|a\omega_1 + b\omega_2| \geq C \sup(|a|, |b|)$, quels que soient $a, b \in \mathbf{R}$, et montrer que $r(\Lambda) = \inf_{\omega \in \Lambda - \{0\}} |\omega|$ est non nul.

L'application $(a, b) \mapsto a\omega_1 + b\omega_2$ est un isomorphisme de \mathbf{R}^2 sur \mathbf{C} et $(a, b) \mapsto |a\omega_1 + b\omega_2|$ est une norme sur \mathbf{R}^2 ; elle est donc équivalente à la norme $\|(a, b)\| = \sup(|a|, |b|)$ (on est en dimension finie); on en déduit l'existence de $C > 0$ tel que $|a\omega_1 + b\omega_2| \geq C \sup(|a|, |b|)$, quels que soient $a, b \in \mathbf{R}$. Finalement, on a $r(\Lambda) \geq C > 0$.

(i) Soit $A = \{\alpha\omega_1 + \beta\omega_2, \alpha, \beta \in [0, 1]\}$. Montrer que A est compact et que, si $z \in \mathbf{C}$, il existe $\omega \in \Lambda$ et $u \in A$ tels que $z = \omega + u$. En déduire que si $f : \mathbf{C} \rightarrow \mathbf{R}_+$ est continue et Λ -périodique, alors f est bornée et atteint son maximum.

A est compact car c'est l'image du compact $[0, 1] \times [0, 1]$ par l'application $(\alpha, \beta) \mapsto \alpha\omega_1 + \beta\omega_2$, qui est continue. On peut écrire z sous la forme $a\omega_1 + b\omega_2$, avec $a, b \in \mathbf{R}$, et il suffit de poser $\omega = [a]\omega_1 + [b]\omega_2$ et $u = \{a\}\omega_1 + \{b\}\omega_2$ pour obtenir une décomposition de z sous la forme $z = \omega + u$ voulue. On en déduit que, si f est Λ -périodique, on a $f(\mathbf{C}) = f(A)$, et comme A est compact, si f est de plus continue, cela implique que f est bornée et atteint son maximum sur A .

(ii) Montrer qu'une fonction Λ -périodique, holomorphe sur \mathbf{C} , est constante.

Si F est Λ -périodique, holomorphe sur \mathbf{C} , alors $f = |F|$ est continue et Λ -périodique. D'après la question (i), cela implique que $|F|$ atteint son maximum, et d'après le th. V.1.20, cela implique que F est constante.

(iii) Montrer que $\sum_{\omega \in \Lambda - \{0\}} \frac{1}{|\omega|^k}$ converge si k est un entier ≥ 3 .

D'après la question (o), on a $\sum_{\omega \in \Lambda - \{0\}} \frac{1}{|\omega|^k} \leq C^{-k} \sum_{(n,m) \in \mathbf{Z}^2 - \{(0,0)\}} \frac{1}{\sup(|m|, |n|)^k}$. Or il y a $8N$ couples (m, n) vérifiant $\sup(|m|, |n|) = N$, ce qui nous fournit la majoration

$$\sum_{\omega \in \Lambda - \{0\}} \frac{1}{|\omega|^k} \leq C^{-k} \sum_{N=1}^{+\infty} \frac{8N}{N^k} \leq 8C^{-k} \zeta(k-1) < +\infty, \quad \text{si } k > 2.$$

(iv) Montrer que, si $R > 0$ et si $|\omega| \geq 2R$, alors $z + \omega$ ne s'annule pas sur $D(0, R)$ et les séries

$$\sum_{\omega \in \Lambda, |\omega| \geq 2R} \left(\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right) \quad \text{et} \quad \sum_{\omega \in \Lambda, |\omega| \geq 2R} \frac{1}{(z + \omega)^3}$$

convergent normalement sur $D(0, R)$.

On a $\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} = \frac{2\omega z + z^2}{\omega^2(\omega + z)^2}$ et $|\omega + z| \geq |\omega| - |z| \geq \frac{|\omega|}{2}$, si $|\omega| \geq 2R$ et $z \in D(0, R)$. En particulier, $z + \omega$ ne s'annule pas sur $D(0, R)$, et on a

$$\left| \frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right| \leq \frac{2|\omega|R + R^2}{|\omega|^4/4} = \frac{8R}{|\omega|^3} + \frac{4R^2}{|\omega|^4},$$

si $z \in D(0, R)$ et $|\omega| \geq 2R$. On déduit la convergence normale de la première série sur $D(0, R)$ de la question (iii). Pour démontrer celle de la seconde, on remarque que $\left| \frac{1}{(z + \omega)^3} \right| \leq \frac{8}{|\omega|^3}$, si $z \in D(0, R)$ et $|\omega| \geq 2R$, et on conclut de la même manière.

(v) En déduire que, si $z \in \mathbf{C} - \Lambda$, alors la série $F(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda - \{0\}} \left(\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right)$ converge, et que $z \mapsto F(z)$ est holomorphe⁽¹⁴⁾ sur $\mathbf{C} - \Lambda$.

$F(z)$ est la somme, sur $D(0, R^-)$ de la somme finie $\frac{1}{z^2} + \sum_{\omega \in \Lambda - \{0\}, |\omega| < 2R} \left(\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right)$, dont chacun des termes est une fonction holomorphe en dehors de Λ , et de la série normalement convergente $\sum_{\omega \in \Lambda, |\omega| \geq 2R} \left(\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right)$ qui est holomorphe sur $D(0, R^-)$ d'après le th. V.2.12. On en déduit l'holomorphie de F sur $D(0, R^-) - \Lambda$ et, ceci étant vrai pour tout R , l'holomorphie de F sur $\mathbf{C} - \Lambda$.

(vi) Montrer que F' est Λ -périodique et que F est paire. En déduire que F est Λ -périodique.

La convergence de la série $\frac{1}{z^2} + \sum_{\omega \in \Lambda - \{0\}} \left(\frac{1}{(z + \omega)^2} - \frac{1}{\omega^2} \right)$ étant uniforme sur tout compact de $\mathbf{C} - \Lambda$, on peut calculer la dérivée de F en dérivant la série terme à terme (th. V.2.12); on a donc $F'(z) = \sum_{\omega \in \Lambda} \frac{-2}{(z + \omega)^3}$. Maintenant, si $\alpha \in \Lambda$, on a $F'(z + \alpha) = \sum_{\omega \in \Lambda} \frac{-2}{(z + \alpha + \omega)^3}$, et la série étant absolument convergente, on peut utiliser le fait que $\omega \mapsto \alpha + \omega$ est une bijection de Λ pour en déduire que $F'(z + \alpha) = F'(z)$.

La parité de F suit de ce que $\left(\frac{1}{(-z + \omega)^2} - \frac{1}{\omega^2} \right) = \left(\frac{1}{(z - \omega)^2} - \frac{1}{(-\omega)^2} \right)$, et $\omega \mapsto -\omega$ est une bijection de Λ .

Maintenant, si $\omega \in \Lambda$, la fonction $F(z + \omega) - F(z)$ a une dérivée nulle et donc est constante sur $\mathbf{C} - \Lambda$. Notons $c(\omega)$ cette constante. On a $c(-\omega) = c(\omega)$ car F est paire, et

$$c(\omega_1 + \omega_2) = F(z + \omega_1 + \omega_2) - F(z + \omega_1) + F(z + \omega_1) - F(z) = c(\omega_2) + c(\omega_1).$$

⁽¹⁴⁾La notation standard est $\wp(z)$ ou $\wp(z, L)$; c'est la fonction \wp de Weierstrass.

En prenant $\omega_1 = \omega$ et $\omega_2 = -\omega$, on en déduit que $2c(\omega) = 0$, et donc que F est Λ -périodique.

(vii) Si $k \geq 3$, soit $G_k = \sum_{\omega \in \Lambda - \{0\}} \frac{1}{\omega^k}$. Montrer que la série $\sum_{n=1}^{+\infty} (-1)^n (n+1) G_{n+2} z^n$ est de rayon de convergence exactement $r(\Lambda)$, et que sa somme est $G(z) = F(z) - \frac{1}{z^2}$, si $|z| < r(\Lambda)$.

La fonction $G(z)$ est holomorphe sur $\mathbf{C} - (\Lambda - \{0\})$, mais a un pôle en tous les éléments de $\Lambda - \{0\}$. Le plus grand disque ouvert de centre 0 contenu dans $\mathbf{C} - (\Lambda - \{0\})$ étant $D(0, r(\Lambda)^-)$ par définition de $r(\Lambda)$, la série de Taylor de G en 0 a pour rayon de convergence $r(\Lambda)$ d'après le (i) de la rem. V.2.8. Maintenant, on a $G(0) = 0$ d'après la formule définissant F , et, d'après le th. V.2.12, on peut calculer $G^{(n)}(0)$ comme la somme de la série des dérivées; on obtient, si $n \geq 1$,

$$G^{(n)}(0) = \sum_{\omega \in \Lambda - \{0\}} \frac{(-1)^n (n+1)!}{\omega^{n+2}} = (-1)^n (n+1)! G_{n+2}.$$

Comme, d'après le (i) de la rem. V.2.8 on a $G(z) = \sum_{n=0}^{+\infty} \frac{G^{(n)}(0)}{n!} z^n$, si $|z| < r(\Lambda)$, cela permet de conclure.

(viii) Montrer que $G_k = 0$, si k est impair. En déduire que, si l'on pose $g_2 = 60G_4$ et $g_3 = 140G_6$, alors $H(z) = F'(z)^2 - 4F(z)^3 + g_2F(z) + g_3$ est holomorphe et nulle en 0.

La fonction G étant paire, on a $G_k = 0$, si k est impair, et donc

$$\begin{aligned} F(z) &= \frac{1}{z^2} + 3G_4 z^2 + 5G_6 z^4 + O(z^5) \quad \text{et} \quad F'(z) = \frac{-2}{z^3} + 6G_4 z + 20G_6 z^3 + O(z^4) \\ F(z)^3 &= \frac{1}{z^6} + \frac{9G_4}{z^2} + 15G_6 + O(z) \quad \text{et} \quad (F'(z))^2 = \frac{4}{z^6} - \frac{24G_4}{z^2} - 80G_6 + O(z) \end{aligned}$$

On en déduit que $H(z)$ se prolonge en une fonction holomorphe nulle en 0.

(ix) Montrer que $F'(z)^2 = 4F(z)^3 - g_2F(z) - g_3$, quel que soit $z \in \mathbf{C} - \Lambda$.

La fonction $H(z)$ est Λ -périodique, holomorphe sur $\mathbf{C} - \Lambda$ et holomorphe en 0. Par Λ -périodicité, elle est aussi holomorphe en tous les points de Λ , et donc est holomorphe sur \mathbf{C} . Elle est donc constante, d'après la question (ii), et comme elle vaut 0 en 0, elle est identiquement nulle.

Partie II

Soit \mathbf{C}/Λ le quotient de \mathbf{C} par son sous-groupe Λ . Rappelons que cela signifie que l'on dispose d'un morphisme de groupes $\pi : \mathbf{C} \rightarrow \mathbf{C}/\Lambda$ surjectif et tel que $\pi(z) = 0$ si et seulement si $z \in \Lambda$ (où l'on a noté 0 l'élément neutre du groupe commutatif \mathbf{C}/Λ). L'équation⁽¹⁵⁾ $u = -u$ a 4 solutions 0, $e_1 = \pi(\frac{\omega_1}{2})$, $e_2 = \pi(\frac{\omega_2}{2})$ et $e_3 = \pi(\frac{\omega_1 + \omega_2}{2})$ dans \mathbf{C}/Λ .

Si $a \in \mathbf{C}$, soit $S_a = \{a + \alpha\omega_1 + \beta\omega_2, \alpha, \beta \in [0, 1[\}$. On peut écrire tout élément z de \mathbf{C} de manière unique sous la forme $z = \omega + u$, avec⁽¹⁶⁾ $u \in S_a$ et $\omega \in \Lambda$. On en déduit que π induit une bijection de S_a sur \mathbf{C}/Λ , et donc que S_a est un système de représentants de \mathbf{C}/Λ dans \mathbf{C} . On note $s_a : (\mathbf{C}/\Lambda) \rightarrow S_a$ la bijection réciproque, et donc $s_a(u)$ est le représentant de u dans S_a , si $u \in \mathbf{C}/\Lambda$.

On note $\Omega_a = \{a + \alpha\omega_1 + \beta\omega_2, \alpha, \beta \in]0, 1[\}$ l'intérieur de S_a , et γ_a le bord de Ω_a parcouru dans le sens direct. C'est le composé des segments $[a, a + \omega_1]$, $[a + \omega_1, a + \omega_1 + \omega_2]$, $[a + \omega_1 + \omega_2, a + \omega_2]$ et $[a + \omega_2, a]$. On remarquera que $S_a \subset \Omega_a \cup \gamma_a$. On notera simplement⁽¹⁷⁾ $\int_{\alpha}^{\beta} f(z) dz$ l'intégrale $\int_{[\alpha, \beta]} f(z) dz$.

Une fonction $f : \mathbf{C} \rightarrow \mathbf{C}$ qui est Λ -périodique peut être considérée comme une fonction de \mathbf{C}/Λ dans \mathbf{C} . Si f est une fonction méromorphe sur \mathbf{C} , Λ -périodique, on dit que les zéros (resp. les pôles) de f dans \mathbf{C}/Λ sont les u_i , pour $i \in I$, avec multiplicités $m_i \geq 1$, si les zéros (resp. les pôles) de f dans S_a sont les

⁽¹⁵⁾ Elle est équivalente à $2u = 0$ dans \mathbf{C}/Λ , ce qui se traduit, en choisissant $\tilde{u} \in \mathbf{C}$ avec $\pi(\tilde{u}) = u$, par $2\tilde{u} \in \Lambda$, ou encore $\tilde{u} \in \frac{1}{2}\Lambda$, et tout élément de $\frac{1}{2}\Lambda$ peut s'écrire de manière unique sous la forme ω ou $\frac{\omega_1}{2} + \omega$ ou $\frac{\omega_2}{2} + \omega$ ou $\frac{\omega_1 + \omega_2}{2} + \omega$, avec $\omega \in \Lambda$.

⁽¹⁶⁾ On écrit $z - a$ sous la forme $x\omega_1 + y\omega_2$, avec $x, y \in \mathbf{R}$, et alors $\omega = [x]\omega_1 + [y]\omega_2$, et $u = a + \{x\}\omega_1 + \{y\}\omega_2$.

⁽¹⁷⁾ Ça prend moins de place si α et β sont compliqués.

$s_a(u_i)$, pour $i \in I$, et si la valuation de f en $s_a(u_i)$ est m_i (resp. $-m_i$), ce qui ne dépend pas du choix de $a \in \mathbf{C}$. Si tel est le cas, on note $N_0(f) = \sum_{i \in I} m_i$ (resp. $N_\infty(f) = \sum_{i \in I} m_i$) le nombre de zéros (resp. de pôles) de f dans \mathbf{C}/Λ , comptés avec multiplicité.

(i) Soit f une fonction méromorphe sur \mathbf{C} non identiquement nul. Montrer que, si $a \in \mathbf{C}$, il existe $r_a > 0$ tel que f n'ait aucun zéro ni pôle dans $D(a, r_a^-) - \{a\}$. En déduire que si K est un compact de \mathbf{C} , alors f n'a qu'un nombre fini de zéros et de pôles dans K .

Soit $k = v_a(f)$. Alors $g(z) = (z - a)^{-k} f(z)$ est holomorphe dans un voisinage de a et non nulle en a ; il existe donc $r_a > 0$ tel que g ne s'annule pas sur $D(a, r_a^-)$, et alors f n'a ni zéro ni pôle sur $D(a, r_a^-) - \{a\}$. Maintenant, si K est compact, on peut extraire un recouvrement fini⁽¹⁸⁾ du recouvrement de K par les $D(a, r_a^-)$, pour $a \in K$; autrement dit, il existe un sous-ensemble fini A de K tel que $K \subset \cup_{a \in A} D(a, r_a^-)$. Par construction de r_a , l'ensemble des zéros et pôles de f sur K est alors inclus dans A , et donc est fini.

(ii) Soit f une fonction Λ -périodique, non identiquement nulle, méromorphe sur \mathbf{C} . Montrer qu'il existe $a \in \mathbf{C}$ tel que f n'ait ni zéro ni pôle sur le chemin γ_a .

Soit $B = \{\alpha\omega_1 + \beta\omega_2, \alpha, \beta \in [-1, 1]\}$. Comme B est compact, f n'a qu'un nombre fini de zéros et de pôles dans B d'après la question (i). Si les zéros et pôles de f dans B sont les $\alpha_i\omega_1 + \beta_i\omega_2$, pour $i \in I$ fini, il suffit de prendre a de la forme $\alpha\omega_1 + \beta\omega_2$, où $\alpha \in [-1, 0]$ n'est pas de la forme α_i ou $\alpha_i - 1$, et $\beta \in [-1, 0]$ n'est pas de la forme β_i ou $\beta_i - 1$, pour $i \in I$.

(iii) Montrer que $\int_{\gamma_a} \frac{f'(z)}{f(z)} dz = 2i\pi(N_0(f) - N_\infty(f))$. En déduire que $N_0(f) = N_\infty(f)$.

Si $w \in \Omega_a$, le résidu de $\frac{f'(z)}{f(z)}$ en w est $v_w(f)$; on déduit donc de la formule des résidus que $\frac{1}{2i\pi} \int_{\gamma_a} \frac{f'(z)}{f(z)} dz = \sum_{w \in \Omega_a} v_w(f)$. Par ailleurs, comme f n'a ni zéro ni pôle sur $S_a - \Omega_a$, cette dernière somme est aussi égale à $\sum_{w \in S_a} v_w(f) = N_0(f) - N_\infty(f)$.

Comme $\frac{f'(z)}{f(z)}$ est Λ -périodique, on a $\int_{a+\omega_1+\omega_2}^a \frac{f'(z)}{f(z)} dz = \int_{a+\omega_1}^a \frac{f'(z+\omega_2)}{f(z+\omega_2)} dz = - \int_a^{a+\omega_1} \frac{f'(z)}{f(z)} dz$. On a donc $\int_a^{a+\omega_1} \frac{f'(z)}{f(z)} dz + \int_{a+\omega_1+\omega_2}^a \frac{f'(z)}{f(z)} dz = 0$ et $\int_{a+\omega_1}^{a+\omega_1+\omega_2} \frac{f'(z)}{f(z)} dz + \int_a^{a+\omega_2} \frac{f'(z)}{f(z)} dz = 0$. On en déduit que $\int_{\gamma_a} \frac{f'(z)}{f(z)} dz = 0$, et donc que $N_0(f) = N_\infty(f)$.

(iv) Dans tout ce qui suit, F est la fonction de Weierstrass définie à la question (v) de la partie I. Calculer $N_0(F')$, et montrer que F' est impaire. En déduire que les zéros de F' dans \mathbf{C}/Λ sont e_1, e_2 et e_3 , et que ce sont des zéros simples.

F' a un pôle d'ordre 3 en 0 et est holomorphe en dehors de Λ ; on a donc $N_\infty(F') = 3$. Par ailleurs, F étant paire, F' est impaire et on a $F'(z) = 0$, si $z = -z$ dans \mathbf{C}/Λ . On en déduit que e_1, e_2 et e_3 sont des zéros de F' dans \mathbf{C}/Λ . Comme $N_0(F') = N_\infty(F') = 3$, ce sont les seuls zéros de F' et ils sont simples.

(v) Calculer $N_\infty(F - b)$, si $b \in \mathbf{C}$. En déduire que $F : (\mathbf{C}/\Lambda) - \{0\} \rightarrow \mathbf{C}$ est surjective, et que $F(a) = F(a')$ si et seulement si $a' = \pm a$. (On fera attention au cas $F(a) \in \{F(e_1), F(e_2), F(e_3)\}$.)

Si $b \in \mathbf{C}$, la fonction $F(z) - b$ a un pôle double en 0 et est holomorphe en dehors de Λ . On a donc $N_\infty(F - b) = 2$, et aussi $N_0(F - b) = 2$, ce qui implique en particulier que l'ensemble des solutions de l'équation $F(z) = b$ dans $(\mathbf{C}/\Lambda) - \{0\}$ n'est pas vide. On en déduit la surjectivité de $F : (\mathbf{C}/\Lambda) - \{0\} \rightarrow \mathbf{C}$.

Maintenant, la fonction F étant paire, si a est une solution de $F(z) = b$, alors $-a$ aussi. Comme $N_0(F - b) = 2$, ce sont les seules solutions si $a \neq -a$, c'est-à-dire si $a \notin \{e_1, e_2, e_3\}$. On en déduit que, si $F(a) \notin \{F(e_1), F(e_2), F(e_3)\}$, et si $F(a') = F(a)$, alors $a' = \pm a$.

Si $a \in \{e_1, e_2, e_3\}$, comme $F'(a) = 0$ d'après le (i), la fonction $F(z) - F(a)$ a un zéro double en $z = a$, et comme $N_0(F(z) - F(a)) = 2$, ce zéro est l'unique zéro de $F(z) - F(a)$ dans \mathbf{C}/Λ ; on en déduit encore dans ce cas que $F(a') = F(a)$ si et seulement si $a' = \pm a$ (et $a = -a$).

⁽¹⁸⁾ On peut aussi raisonner en termes de suites, en disant que, si l'ensemble Z des zéros et pôles de f dans K est infini, et si $(a_n)_{n \in \mathbf{N}}$ est une suite d'éléments distincts de Z , alors on peut extraire de $(a_n)_{n \in \mathbf{N}}$ une sous-suite ayant une limite a dans K , mais alors $D(a, r_a^-)$ contient une infinité de a_n , ce qui est contraire à la définition de r_a .

(vi) En déduire que $z \mapsto P(z) = (F(z), F'(z))$ est une bijection de $(\mathbf{C}/\Lambda) - \{0\}$ sur l'ensemble $E(\mathbf{C})$ des solutions (X, Y) dans \mathbf{C}^2 de l'équation $Y^2 = 4X^3 - g_2X - g_3$.

L'appartenance de $P(z)$ à $E(\mathbf{C})$ résulte de la question (ix) de la partie I. Si $P(z_1) = P(z_2)$, on a en particulier $F(z_1) = F(z_2)$, et donc $z_1 = \pm z_2$ d'après la question (v). Si $z_1 = -z_2$, alors $F'(z_2) = -F'(z_1)$, et donc $P(z_1) = P(z_2)$ et $z_1 \neq z_2$ impliquent $F'(z_1) = 0$. Or ceci implique $z_1 \in \{e_1, e_2, e_3\}$, et donc $z_1 = -z_1$ et $z_1 = z_2$. On en déduit l'injectivité de $z \mapsto P(z)$. Pour prouver la surjectivité, il suffit de constater que si $(a, b) \in E(\mathbf{C})$, il existe z tel que $F(z) = a$, et on a $F'(z) = \pm b$, ce qui fait que l'on a $(a, b) = P(z)$ ou $(a, b) = P(-z)$.

(vii) Soient $\Omega \subset \mathbf{C}$ un ouvert contractile et $f : \Omega \rightarrow \mathbf{C}$, holomorphe et ne s'annulant pas sur Ω . Montrer qu'il existe g holomorphe sur Ω telle que $f = e^g$, et que $g' = \frac{f'}{f}$.

Comme Ω est contractile, et comme $\frac{f'}{f}$ est holomorphe sur Ω , il existe, d'après le (i) de la rem. VI.1.6, h holomorphe sur Ω telle que $h' = \frac{f'}{f}$. Soit $z_0 \in \Omega$. Quitte à rajouter une constante à h , on peut supposer que $e^{h(z_0)} = f(z_0)$. Mais alors $(e^{-h}f)' = -h'e^{-h}f + e^{-h}f' = 0$, et donc $e^{-h}f$ est constante sur Ω , et comme elle vaut 1 en z_0 , on a $f = e^h$ sur Ω . Si g est une autre fonction holomorphe sur Ω vérifiant $e^g = f$, on a $e^{g-h} = 1$, et donc $g - h$ est holomorphe et à valeurs dans $2i\pi\mathbf{Z}$. Comme Ω est connexe, $g - h$ est constante et $g' = \frac{f'}{f}$.

(viii) Si $a, b \in \mathbf{C}$, et si $r > 0$, soit $\Omega_r(a, b) = \{z \in \mathbf{C}, \exists c \in [a, b], |z - c| < r\}$. Montrer que $\Omega_r(a, b)$ est un ouvert convexe.

$\Omega_r(a, b)$ est la réunion des $D(c, r^-)$, pour $c \in [a, a + \omega]$, et donc est un ouvert en tant que réunion d'ouverts. Si $z_1, z_2 \in \Omega_r$, il existe $c_1, c_2 \in [a, a + \omega]$ tels que $|z_i - c_i| < r$, si $i = 1, 2$. Maintenant, si $t \in [0, 1]$, alors $tc_1 + (1-t)c_2 \in [a, a + \omega]$ et $|t(z_1 + (1-t)z_2) - (tc_1 + (1-t)c_2)| = |t(z_1 - c_1) + (1-t)(z_2 - c_2)| < r(t + (1-t)) = r$, ce qui prouve que $tz_1 + (1-t)z_2 \in \Omega_r(a, b)$, et que $\Omega_r(a, b)$ est convexe.

(ix) Soit f une fonction Λ -périodique, non identiquement nulle, méromorphe sur \mathbf{C} , et soient $a \in \mathbf{C}$ et $\omega \in \Lambda$ tels que f n'ait ni zéro ni pôle sur $[a, a + \omega]$. Montrer que, si $r > 0$ est assez petit, f n'a ni zéro ni pôle dans $\Omega_r(a, a + \omega)$. En déduire que $\frac{1}{2i\pi} \int_a^{a+\omega} \frac{f'(z)}{f(z)} dz \in \mathbf{Z}$.

$\Omega_1(a, a + \omega)$ est un ouvert borné contenant $[a, a + \omega]$. Son adhérence K est un compact et donc ne contient qu'un nombre fini de zéros et de pôles de f , d'après la question (i). Il suffit de prendre pour r le minimum des distances de ces zéros et pôles au segment $[a, a + \omega]$ pour être sûr que $\Omega_r(a, a + \omega)$ ne contient ni zéro ni pôle de f .

Un ouvert convexe étant contractile, il existe g holomorphe sur $\Omega_r(a, a + \omega)$ telle que $e^g = f$ et $g' = \frac{f'}{f}$. On a alors $\int_a^{a+\omega} \frac{f'(z)}{f(z)} dz = g(a + \omega) - g(a)$, et donc $\exp(\int_a^{a+\omega} \frac{f'(z)}{f(z)} dz) = f(a + \omega)/f(a) = 1$, puisque f est Λ -périodique. On en déduit que $\int_a^{a+\omega} \frac{f'(z)}{f(z)} dz \in 2i\pi\mathbf{Z}$.

(x) Soit f une fonction Λ -périodique, non identiquement nulle, méromorphe sur \mathbf{C} .

(a) Montrer que l'ensemble X_a des éléments de S_a tels que $v_z(f) \neq 0$ est fini et que l'image de $\sum_{z \in X_a} v_z(f) z$ dans \mathbf{C}/Λ ne dépend pas de $a \in \mathbf{C}$. On la note $\pi(f)$.

Comme S_a est borné, l'ensemble des zéros et des pôles de f dans l'adhérence $\overline{S_a}$ de S_a est fini ; il en est donc a fortiori de même dans S_a . Par ailleurs, si on change a en b , alors $z \mapsto s_b(\pi(z))$ est une bijection de S_a sur S_b qui induit une bijection de X_a sur X_b , et on a $z - s_b(\pi(z)) \in \Lambda$ et $v_z(f) = v_{s_b(\pi(z))}(f)$. On en déduit que

$$\sum_{z \in X_a} v_z(f) z - \sum_{z \in X_b} v_z(f) z = \sum_{z \in X_a} v_z(f)(z - s_b(\pi(z))) \in \Lambda,$$

et donc que $\pi(f)$ est bien indépendant de a .

(b) Soit $a \in \mathbf{C}$ tel que f n'ait ni zéro ni pôle sur le chemin γ_a . Montrer que

$$I_1 = \frac{1}{2i\pi} \left(\int_a^{a+\omega_1} z \frac{f'(z)}{f(z)} dz + \int_{a+\omega_1+\omega_2}^{a+\omega_2} z \frac{f'(z)}{f(z)} dz \right) \quad \text{et} \quad I_2 = \frac{1}{2i\pi} \left(\int_{a+\omega_1}^{a+\omega_1+\omega_2} z \frac{f'(z)}{f(z)} dz + \int_{a+\omega_2}^a z \frac{f'(z)}{f(z)} dz \right)$$

appartiennent à Λ . (On ne traitera que I_1 ou I_2 .)

$\int_{a+\omega_1+\omega_2}^{a+\omega_2} z \frac{f'(z)}{f(z)} dz = \int_{a+\omega_1}^a (z + \omega_2) \frac{f'(z+\omega_2)}{f(z+\omega_2)} dz = \int_{a+\omega_1}^a (z + \omega_2) \frac{f'(z)}{f(z)} dz$ puisque f est Λ -périodique. On en déduit que $I_1 = \frac{-1}{2i\pi} \int_{a+\omega_1}^{a+\omega_2} \omega_2 \frac{f'(z)}{f(z)} dz$, et comme $\frac{1}{2i\pi} \int_a^{a-\omega_1} \frac{f'(z)}{f(z)} dz \in \mathbf{Z}$ d'après la question (ix), on a $I_1 \in \mathbf{Z}\omega_2 \subset \Lambda$. L'argument est le même pour I_2 .

(c) En déduire que $\pi(f) = 0$.

Il résulte du (b) que $\frac{1}{2i\pi} \int_{\gamma_a} z \frac{f'(z)}{f(z)} dz = I_1 + I_2 \in \Lambda$. Par ailleurs, il résulte de la formule des résidus que $\frac{1}{2i\pi} \int_{\gamma_a} z \frac{f'(z)}{f(z)} dz = \sum_{u \in \Omega_a} v_u(f) u$, et comme f n'a ni zéro ni pôle sur $S_a - \Omega_a$, on a $\sum_{u \in S_a} v_u(f) u = \sum_{u \in \Omega_a} v_u(f) u \in \Lambda$. L'image $\pi(f)$ de $\sum_{u \in S_a} v_u(f) u$ dans \mathbf{C}/Λ est donc nulle.

(xi) Soit $\bar{E}(\mathbf{C}) = E(\mathbf{C}) \cup \{\infty\}$. On étend l'application $z \mapsto P(z)$ de la question (vi) en une bijection de \mathbf{C}/Λ sur $\bar{E}(\mathbf{C})$, en posant $P(0) = \infty$, et on note $\iota : \bar{E}(\mathbf{C}) \rightarrow \mathbf{C}/\Lambda$ la bijection réciproque. On définit $Q_1 \oplus Q_2$, si $Q_1, Q_2 \in \bar{E}(\mathbf{C})$ par $Q_1 \oplus Q_2 = P(\iota(Q_1) + \iota(Q_2))$. Montrer que \oplus est une loi de groupe commutatif d'élément neutre ∞ sur $\bar{E}(\mathbf{C})$, et que, si $Q_1, Q_2, Q_3 \in E(\mathbf{C})$ sont distincts, alors $Q_1 \oplus Q_2 \oplus Q_3 = \infty$ si et seulement si Q_1, Q_2, Q_3 sont sur une même droite complexe de \mathbf{C}^2 .

(On pourra s'intéresser aux zéros de la fonction $G(z) = \begin{vmatrix} F(z) & F(z_1) & F(z_2) \\ F'(z) & F'(z_1) & F'(z_2) \\ 1 & 1 & 1 \end{vmatrix}$, où $z_i = \iota(Q_i)$.)

Par construction, on a $Q_1 \oplus Q_2 = Q_3$ si et seulement si $\iota(Q_1) + \iota(Q_2) = \iota(Q_3)$. Comme $(\mathbf{C}/\Lambda, +)$ est un groupe commutatif d'élément neutre 0, et comme ι est une bijection de $\bar{E}(\mathbf{C})$ sur \mathbf{C}/Λ , cela implique que $(\bar{E}(\mathbf{C}), \oplus)$ est un groupe commutatif d'élément neutre $\iota^{-1}(0) = \infty$.

Maintenant, $Q_1, Q_2, Q_3 \in E(\mathbf{C})$ sont distincts et vérifient $Q_1 \oplus Q_2 \oplus Q_3 = \infty$ si et seulement si $z_1 = \iota(Q_1)$, $z_2 = \iota(Q_2)$ et $z_3 = \iota(Q_3)$ sont distincts, non nuls, et vérifient $z_1 + z_2 + z_3 = 0$. En particulier, on a $z_1 \neq \pm z_2$, et donc $F(z_1) \neq F(z_2)$. Ceci implique que si on écrit $G(z)$ sous la forme $\alpha F(z) + \beta F'(z) + \gamma$, alors $\beta \neq 0$.

Comme F' a un pôle d'ordre 3 en $z = 0$, et F a un pôle d'ordre 2, le pôle en $z = 0$ de $\beta F' + \alpha F + \gamma$ est d'ordre 3 exactement. On a donc $N_\infty(G) = 3$, et on déduit de la question (iii) que G a 3 zéros dans \mathbf{C}/Λ comptés avec multiplicité.

Il est clair que z_1 et z_2 sont deux zéros de G , et si z' est le troisième, alors $\pi(G) = z_1 + z_2 + z' - 3 \cdot 0$. Il résulte alors du (c) de la question (x) que $z_1 + z_2 + z' = 0$, et donc $z' = z_3$. On a donc prouvé que $G(z) = 0$ si et seulement si $z \in \{z_1, z_2, z_3\}$.

Par ailleurs $\alpha X + \beta Y + \gamma = 0$ est l'équation de la droite passant par $P(z_1) = Q_1$ et $P(z_2) = Q_2$; on en déduit que $G(z) = 0$ si et seulement si $P(z)$ appartient à la droite (Q_1, Q_2) . Ceci permet de conclure.

CHAPITRE VII

SÉRIES DE DIRICHLET

Une série de Dirichlet générale est une expression de la forme $\sum_{n=1}^{+\infty} a_n e^{-\lambda_n s}$, où les a_n sont des nombres complexes, et les λ_n sont des nombres complexes dont la partie réelle tend vers $+\infty$. (Si $\lambda_n = n - 1$ pour tout n , on retombe, modulo le changement de variable $e^{-s} = z$, sur le cas des séries entières (avec une indexation bizarre), ce qui permet de voir les séries de Dirichlet générales comme une généralisation des séries entières.) Dans ce chapitre, nous ne nous intéresserons qu'au cas où $\lambda_n = \log n$, pour tout n , qui est le cas originellement considéré par Dirichlet, mais le cas général intervient naturellement quand, par exemple, on essaie de définir le déterminant d'un opérateur en dimension infinie, ce qui a de multiples applications en physique⁽¹⁾ et en mathématiques. Comme illustration de ce procédé « de zêta-régularisation », mentionnons la formule

$$1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdots = \sqrt{2\pi},$$

équivalente à $-\sum_{n=1}^{+\infty} \log n = -\frac{1}{2} \log 2\pi$, dans laquelle le membre de gauche s'interprète comme la dérivée en 0 de la fonction ζ définie par $\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$, si $\operatorname{Re}(s) > 1$, et prolongée analytiquement⁽²⁾ à $\mathbf{C} - \{1\}$.

VII.1. Séries de Dirichlet

1. Abscisse de convergence absolue

On appelle *série de Dirichlet* une série de la forme $L(\mathbf{a}, s) = \sum_{n=1}^{+\infty} a_n n^{-s}$, où $s \in \mathbf{C}$ et $\mathbf{a} = (a_n)_{n \geq 1}$ est une suite de nombres complexes (et $n^{-s} = \exp(-s \log n)$, où $\log n \in \mathbf{R}_+$). Une série de Dirichlet peut ne converger pour aucune valeur de s , mais si elle converge

⁽¹⁾Les fréquences d'une membrane vibrante sont les valeurs propres du laplacien sur l'ouvert de \mathbf{C} représentant cette membrane. La *fonction zêta du laplacien* $\zeta_\Delta(s) = \sum \lambda^{-s}$, où la somme porte sur les valeurs propres non nulles, encode des relations entre ces fréquences et la géométrie de la surface. Le déterminant du laplacien est $\exp(-\zeta'_\Delta(0))$; il intervient par exemple dans des questions de renormalisation.

⁽²⁾L'existence de ce prolongement analytique fait l'objet du th. VII.3.4, et une démonstration de la formule $\zeta'(0) = -\frac{1}{2} \log 2\pi$ est proposée dans l'ex. VII.3.8.

pour s_0 , on a en particulier, $a_n n^{-s_0} \rightarrow 0$, et donc $|a_n| = o(n^{\operatorname{Re}(s_0)})$. Réciproquement, si $|a_n| = O(n^\alpha)$, pour un certain $\alpha \in \mathbf{R}$, alors $\sum_{n=1}^{+\infty} a_n n^{-s}$ converge normalement sur tout demi-plan de la forme $\operatorname{Re}(s) > \alpha + 1 + \delta$, avec $\delta > 0$; elle définit donc une fonction holomorphe sur le demi-plan $\operatorname{Re}(s) > \alpha + 1$. De même, si $\sum_{n=1}^{+\infty} a_n n^{-s}$ converge absolument pour $s = s_0$, alors elle converge normalement sur le demi-plan $\operatorname{Re}(s - s_0) \geq 0$, puisque $|a_n n^{-s}| \leq |a_n n^{-s_0}|$ sur ce demi-plan; elle définit donc une fonction holomorphe sur le demi-plan ouvert $\operatorname{Re}(s - s_0) > 0$, qui se prolonge par continuité au demi-plan fermé $\operatorname{Re}(s - s_0) \geq 0$.

La discussion précédente amène naturellement à définir les éléments suivants de $\overline{\mathbf{R}}$:

- $\sigma_{\text{conv}} = \inf\{\operatorname{Re}(s), L(\mathbf{a}, s) \text{ converge}\}$ *abscisse de convergence*;
- $\sigma_{\text{abs}} = \inf\{\operatorname{Re}(s), L(\mathbf{a}, s) \text{ converge absolument}\}$ *abscisse de convergence absolue*;
- $\sigma_{\text{hol}} = \inf\{\sigma \in \mathbf{R}, L(\mathbf{a}, s) \text{ admet un prolongement holomorphe sur } \operatorname{Re}(s) > \sigma\}$.
- $\tau = \inf\{\alpha \in \mathbf{R}, a_n = O(n^\alpha)\}$.

Le nombre τ n'a pas de signification particulière, mais des quantités précédentes, c'est la plus facile à calculer; de plus la discussion ci-dessus nous fournit les encadrements :

$$\tau \leq \sigma_{\text{conv}} \leq \sigma_{\text{abs}} \leq \tau + 1.$$

Par ailleurs, contrairement au cas des séries entières (cf. (i) de la rem. V.2.8), on n'a pas forcément $\sigma_{\text{hol}} = \sigma_{\text{abs}}$, comme le montre l'exemple des fonctions L de Dirichlet (th. VII.4.4).

On suppose dans tout ce qui suit que les séries de Dirichlet que l'on considère convergent quelque part (i.e. $\sigma_{\text{abs}} \neq +\infty$), le cas contraire n'ayant qu'un intérêt limité...

Théorème VII.1.1. — (Landau) Soit $L(\mathbf{a}, s) = \sum_{n \geq 1} \frac{a_n}{n^s}$ une série de Dirichlet à coefficients positifs (i.e. $a_n \in \mathbf{R}_+$, quel que soit $n \geq 1$). Alors σ_{abs} n'a aucun voisinage dans \mathbf{C} sur lequel $L(\mathbf{a}, s)$ admet un prolongement analytique⁽³⁾.

Corollaire VII.1.2. — Si $L(\mathbf{a}, s) = \sum_{n \geq 1} \frac{a_n}{n^s}$ est une série de Dirichlet à coefficients positifs, alors $\sigma_{\text{hol}} = \sigma_{\text{abs}}$.

Démonstration. — Le corollaire est immédiat. Passons à la démonstration du théorème. Posons $\sigma = \sigma_{\text{abs}}$, et supposons, par l'absurde, qu'il existe $\varepsilon > 0$ tel que $L(\mathbf{a}, s)$ admette un prolongement analytique au disque $D(\sigma, 3\varepsilon^-)$. Alors $L(\mathbf{a}, s)$ est holomorphe sur l'ouvert Ω réunion de $D(\sigma, 3\varepsilon^-)$ et du demi-plan $\operatorname{Re}(s) > \sigma$. Comme Ω contient $D(\sigma + \varepsilon, 3\varepsilon^-)$, $L(\mathbf{a}, s)$ est somme de sa série de Taylor en $\sigma + \varepsilon$ sur ce disque, et en particulier, on a $L(\mathbf{a}, \sigma - \varepsilon) = \sum_{k=0}^{+\infty} \frac{L^{(k)}(\mathbf{a}, \sigma + \varepsilon)}{k!} (-2\varepsilon)^k$. Par ailleurs, comme $\sigma + \varepsilon$ est dans le demi-plan de convergence de $L(\mathbf{a}, s)$, on peut, d'après le th. V.2.12, calculer $L^{(k)}(\mathbf{a}, \sigma + \varepsilon)$ en dérivant

⁽³⁾Il arrive que $L(\mathbf{a}, s)$ admette un prolongement méromorphe comme le montre le cas de la fonction zêta de Riemann (cf. th. VII.3.4), mais alors σ_{abs} est un pôle de ce prolongement.

terme à terme la série de Dirichlet. On obtient donc

$$L(\mathbf{a}, \sigma - \varepsilon) = \sum_{k=0}^{+\infty} L^{(k)}(\mathbf{a}, \sigma + \varepsilon) \frac{(-2\varepsilon)^k}{k!} = \sum_{k=0}^{+\infty} \left(\sum_{n=1}^{+\infty} \frac{a_n (-\log n)^k}{n^{\sigma+\varepsilon}} \right) \frac{(-2\varepsilon)^k}{k!}.$$

En faisant rentrer le $\frac{(-2\varepsilon)^k}{k!}$ à l'intérieur de la parenthèse, on obtient une série double à termes positifs, ce qui permet d'échanger l'ordre des sommations, et d'obtenir :

$$L(\mathbf{a}, \sigma - \varepsilon) = \sum_{n=1}^{+\infty} \frac{a_n}{n^{\sigma+\varepsilon}} \left(\sum_{k=0}^{+\infty} \frac{(2\varepsilon \log n)^k}{k!} \right) = \sum_{n=1}^{+\infty} \frac{a_n}{n^{\sigma+\varepsilon}} n^{2\varepsilon} = \sum_{n=1}^{+\infty} \frac{a_n}{n^{\sigma-\varepsilon}}.$$

On en déduit que $L(\mathbf{a}, s)$ converge en $\sigma - \varepsilon$, ce qui est contraire à la définition de σ . Ceci permet de conclure.

2. Demi-plan de convergence d'une série de Dirichlet

Contrairement au cas des séries entières, on a, en général⁽⁴⁾, $\sigma_{\text{conv}} \neq \sigma_{\text{abs}}$, et la détermination de σ_{conv} peut cacher des difficultés redoutables... D'autre part, d'après le cor. VII.1.6 ci-dessous, on a $\sigma_{\text{hol}} \leq \sigma_{\text{conv}}$, et le cas des fonctions L de Dirichlet (th. VII.4.4) montre que, contrairement au cas des séries entières, on n'a pas toujours égalité. La quantité σ_{hol} est très souvent extrêmement difficile à calculer. Une grande partie du programme de Langlands (cf. annexe D) est destinée à prouver que pour beaucoup de séries de Dirichlet issues de la théorie des nombres, on a $\sigma_{\text{hol}} = -\infty$ (autrement dit que ces séries ont un prolongement analytique à tout le plan complexe).

L'étude de la convergence des séries de Dirichlet repose sur lemme suivant.

Lemme VII.1.3. — Soient $\alpha, \beta \in \mathbf{R}$, avec $0 < \alpha < \beta$. Soit $z = x + iy$, avec $x, y \in \mathbf{R}$, et $x > 0$. Alors

$$|e^{-\alpha z} - e^{-\beta z}| \leq \left| \frac{z}{x} \right| (e^{-\alpha x} - e^{-\beta x}).$$

Démonstration. — On a $e^{-\alpha z} - e^{-\beta z} = z \int_{\alpha}^{\beta} e^{-tz} dt$, et donc

$$|e^{-\alpha z} - e^{-\beta z}| \leq |z| \int_{\alpha}^{\beta} e^{-tx} dt = \left| \frac{z}{x} \right| (e^{-\alpha x} - e^{-\beta x}).$$

En appliquant ce lemme à $\alpha = \log n$, $\beta = \log(n+1)$ et $z = s - s_0$, on obtient le résultat suivant.

Corollaire VII.1.4. — Si $n \geq 1$, et si $\text{Re}(s - s_0) > 0$, alors

$$|n^{-(s-s_0)} - (n+1)^{-(s-s_0)}| \leq \frac{|s - s_0|}{\text{Re}(s - s_0)} (n^{-\text{Re}(s-s_0)} - (n+1)^{-\text{Re}(s-s_0)}).$$

⁽⁴⁾ Supposons par exemple que $a_n \in \{\pm 1\}$, pour tout $n \geq 1$. On a alors $\sigma_{\text{abs}} = 1$. Pour étudier σ_{conv} , on peut utiliser la formule sommatoire d'Abel : on pose $A_n = \sum_{k=1}^n a_k$ de telle sorte que

$$\sum_{n=1}^N \frac{a_n}{n^s} = \frac{A_N}{(N+1)^s} + \sum_{n=1}^N A_n (n^{-s} - (n+1)^{-s}) = \frac{A_N}{(N+1)^s} + \sum_{n=1}^N \frac{A_n}{n^{s+1}} \left(n \left(1 - \left(1 + \frac{1}{n} \right)^{-s} \right) \right).$$

On en déduit le fait que, si $A_n = O(n^\alpha)$, avec $\alpha < 1$, alors $\sigma_{\text{conv}} \leq \alpha$. Or Hausdorff (1913) a démontré que, presque sûrement (en considérant les a_n comme des variables aléatoires indépendantes à valeurs dans $\{\pm 1\}$ muni de l'équiprobabilité), $A_n = O(n^{1/2+\varepsilon})$, quel que soit $\varepsilon > 0$, et donc $\sigma_{\text{conv}} \leq 1/2$ presque sûrement (et $1/2$ est sûrement différent de 1).

Théorème VII.1.5. — Soit $L(\mathbf{a}, s) = \sum_{n=1}^{+\infty} a_n n^{-s}$ une série de Dirichlet, et soit $s_0 \in \mathbf{C}$.

(i) Si la suite des sommes partielles $A_n(s_0) = \sum_{k=1}^n a_k k^{-s_0}$ est bornée, alors $L(\mathbf{a}, s)$ converge uniformément dans tout compact du demi-plan $\operatorname{Re}(s - s_0) > 0$.

(ii) Si la série $\sum_{n=1}^{+\infty} a_n n^{-s_0}$ converge, alors $L(\mathbf{a}, s)$ converge uniformément dans tout secteur angulaire $|\arg(s - s_0)| \leq \alpha < \frac{\pi}{2}$ du demi-plan $\operatorname{Re}(s - s_0) \geq 0$.

Démonstration. — La démonstration est la même dans les deux cas et repose sur la formule sommatoire d'Abel. Si $p, q \in \mathbf{N}$ vérifient $p \leq q$, soit

$$M_{p,q} = \sup_{p \leq n \leq q} |B_{n,p}|, \quad \text{avec } B_{n,p} = A_n(s_0) - A_{p-1}(s_0)$$

(et $A_0(s_0) = 0$, puisqu'une somme vide est nulle par convention). On a

$$\begin{aligned} \sum_{n=p}^q a_n n^{-s} &= \sum_{n=p}^q (B_{n,p} - B_{n-1,p}) n^{-(s-s_0)} \\ &= B_{q,p} q^{-(s-s_0)} + \sum_{n=p}^{q-1} B_{n,p} (n^{-(s-s_0)} - (n+1)^{-(s-s_0)}). \end{aligned}$$

Si $\operatorname{Re}(s - s_0) > 0$, en utilisant la majoration du cor. VII.1.4, l'inégalité $1 \leq \frac{|s-s_0|}{\operatorname{Re}(s-s_0)}$, et la définition de $M_{p,q}$, on en déduit la majoration

$$\begin{aligned} \left| \sum_{n=p}^q a_n n^{-s} \right| &\leq M_{p,q} \left(q^{-\operatorname{Re}(s-s_0)} + \frac{|s-s_0|}{\operatorname{Re}(s-s_0)} \sum_{n=p}^{q-1} (n^{-\operatorname{Re}(s-s_0)} - (n+1)^{-\operatorname{Re}(s-s_0)}) \right) \\ &\leq M_{p,q} \frac{|s-s_0|}{\operatorname{Re}(s-s_0)} p^{-\operatorname{Re}(s-s_0)}. \end{aligned}$$

Maintenant, si $|A_n(s_0)| \leq M$, quel que soit $n \in \mathbf{N}$, on a $M_{p,q} \leq 2M$ quels que soient $p \leq q$. Par ailleurs, si K est un compact du demi-plan $\operatorname{Re}(s - s_0) > 0$, alors il existe $a > 0$ et $b > 0$ tels que $|s - s_0| \leq a$ et $\operatorname{Re}(s - s_0) \geq b$, quel que soit $s \in K$. On a donc $\left| \sum_{n=p}^q a_n n^{-s} \right| \leq 2Mab^{-1}p^{-b}$, si $s \in K$, ce qui prouve que la série $\sum_{n \geq 1} a_n n^{-s}$ satisfait au critère de Cauchy uniforme sur K . On en déduit le (i).

Pour démontrer le (ii), il suffit de constater que l'hypothèse selon laquelle $\sum_{n \geq 1} a_n n^{-s_0}$ converge est équivalente à $M_{p,q} \rightarrow 0$ quand $p \rightarrow +\infty$ d'après le critère de Cauchy. Comme $\frac{|s-s_0|}{\operatorname{Re}(s-s_0)}$ est majoré par $\operatorname{tg} \alpha$ sur le secteur angulaire $|\arg(s - s_0)| \leq \alpha < \frac{\pi}{2}$ du demi-plan $\operatorname{Re}(s - s_0) \geq 0$, et comme $p^{-\operatorname{Re}(s-s_0)} \leq 1$ sur le demi-plan, on en déduit que la série $\sum_{n \geq 1} a_n n^{-s}$ satisfait au critère de Cauchy uniforme sur le secteur angulaire, ce qui démontre le (ii).

Corollaire VII.1.6. — Soit $L(\mathbf{a}, s) = \sum_{n \geq 1} \frac{a_n}{n^s}$ une série de Dirichlet, et soit $\sigma_{\text{conv}} \in \overline{\mathbf{R}}$ l'abscisse de convergence de $L(\mathbf{a}, s)$. Alors $L(\mathbf{a}, s)$ converge en tout s du demi-plan $\operatorname{Re}(s) > \sigma_{\text{conv}}$ et définit une fonction holomorphe sur ce demi-plan, et donc $\sigma_{\text{hol}} \leq \sigma_{\text{conv}}$.

Démonstration. — Par définition de σ_{conv} , pour tout $\varepsilon > 0$, il existe $s_\varepsilon \in \mathbf{C}$, avec $\operatorname{Re}(s_\varepsilon) \leq \sigma_{\text{conv}} + \varepsilon$, tel que $L(\mathbf{a}, s_\varepsilon)$ soit convergente. D'après le th. VII.1.5, la série $L(\mathbf{a}, s)$ converge pour tout s du demi-plan $\operatorname{Re}(s) > \sigma_{\text{conv}} + \varepsilon$, car ce demi-plan est une réunion de secteurs angulaires de sommet s_ε , et comme ceci est vrai pour tout $\varepsilon > 0$, cela prouve que $L(\mathbf{a}, s)$ converge en tout s du demi-plan $\operatorname{Re}(s) > \sigma_{\text{conv}}$. L'holomorphie de $L(\mathbf{a}, s)$ sur ce demi-plan suit de ce qu'une fonction qui est limite uniforme sur tout compact de fonctions holomorphes est elle-même holomorphe (th. V.2.12).

VII.2. Séries de Dirichlet et transformée de Mellin

1. La fonction Γ dans le plan complexe

On rappelle que la *constante d'Euler* γ est la limite de $-\log n + \sum_{k=1}^n \frac{1}{k}$, quand $n \rightarrow +\infty$.

Théorème VII.2.1. — (i) *Le produit*

$$f(z) = z e^{\gamma z} \prod_{k=1}^{+\infty} \left(\left(1 + \frac{z}{k}\right) e^{-z/k} \right)$$

est uniformément convergent sur tout compact de \mathbf{C} , et coïncide avec $\frac{1}{\Gamma}$ sur \mathbf{R}_+^* .

(ii) *La fonction Γ complexe définie par $\Gamma(z) = \frac{1}{f(z)}$ est méromorphe sur \mathbf{C} , holomorphe en dehors de pôles simples aux entiers négatifs, de résidu $\frac{(-1)^n}{n!}$ en $-n$, si $n \in \mathbf{N}$. De plus, on a $\Gamma(z+1) = z\Gamma(z)$, quel que soit $z \in \mathbf{C} - (-\mathbf{N})$.*

Démonstration. — La fonction $h(z) = z^{-2}((1+z)e^{-z} - 1)$ est holomorphe sur \mathbf{C} , et donc est bornée sur tout compact. En particulier, il existe M tel que $|(1+z)e^{-z} - 1| \leq M|z|^2$, si $|z| \leq 1$. Maintenant, si K est un compact, il existe $R(K)$ tel que $|z| \leq R(K)$, quel que soit $z \in K$, et on a $|(1 + \frac{z}{k})e^{-z/k} - 1| \leq \frac{MR(K)^2}{k^2}$, si $k \geq R(K)$ et $z \in K$. On en déduit la convergence uniforme du produit sur K , et le th. V.2.15, montre que f est une fonction holomorphe sur \mathbf{C} avec des zéros simples en 0 et les $-k$, pour $k \in \mathbf{N} - \{0\}$. En passant à l'inverse, cela montre que Γ est méromorphe sur \mathbf{C} , holomorphe en dehors de pôles simples aux entiers négatifs.

Maintenant, si $x \in \mathbf{R}_+^*$, on a, d'après la formule de Gauss (cf. ex.IV.1.5),

$$\begin{aligned} \frac{1}{\Gamma(x)} &= \lim_{n \rightarrow +\infty} \frac{x(x+1) \cdots (x+n)}{n! n^x} = x \lim_{n \rightarrow +\infty} e^{-x \log n} \prod_{k=1}^n \left(1 + \frac{x}{k}\right) \\ &= x e^{\gamma x} \left(\lim_{n \rightarrow +\infty} e^{-(1+\cdots+1/n)x} \prod_{k=1}^n \left(1 + \frac{x}{k}\right) \right) \\ &= x e^{\gamma x} \lim_{n \rightarrow +\infty} \prod_{k=1}^n \left(\left(1 + \frac{x}{k}\right) e^{-x/k} \right), \end{aligned}$$

ce qui permet de montrer que f coïncide avec $\frac{1}{\Gamma}$ sur \mathbf{R}_+^* .

Finalement, les fonctions $\Gamma(z+1)$ et $z\Gamma(z)$ sont holomorphes sur $\mathbf{C} - (-\mathbf{N})$, et coïncident sur \mathbf{R}_+^* ; elle coïncident donc sur $\mathbf{C} - (-\mathbf{N})$. Comme $\Gamma(1) = 1$, cela permet de montrer, par récurrence sur n , que l'on a $\lim_{z \rightarrow -n} (z+n)\Gamma(z) = \frac{(-1)^n}{n!}$. Ceci termine la démonstration du théorème.

Exercice VII.2.2. — (i) Établir la *formule des compléments* $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$ (on prendra la *dérivée logarithmique*⁽⁵⁾ des deux membres, et on utilisera l'ex. V.2.14). En déduire une démonstration de l'identité $\int_{-\infty}^{+\infty} e^{-\pi t^2} dt = 1$.

(ii) Établir de même la *formule de multiplication* : si $p \in \mathbf{N} - \{0\}$, et si $z \in \mathbf{C} - (-\mathbf{N})$, alors

$$\prod_{j=0}^{p-1} \Gamma\left(\frac{z+j}{p}\right) = (2\pi)^{(p-1)/2} p^{-z+1/2} \Gamma(z).$$

Proposition VII.2.3. — *Sur tout secteur angulaire de la forme $|\arg(z)| < \alpha$, avec $\alpha < \pi$, on a les développements suivants au voisinage de $|z| = \infty$:*

(i) $\frac{\Gamma'(z)}{\Gamma(z)} = \log z - \frac{1}{2z} + O\left(\frac{1}{z^2}\right)$;

(ii) $\log \Gamma(z) = z \log z - z + \frac{1}{2}(\log(2\pi) - \log z) + O\left(\frac{1}{z}\right)$, où $\log \Gamma(z) = \int_{[1,z]} \frac{\Gamma'(z)}{\Gamma(z)} dz$ est le logarithme de $\Gamma(z)$, holomorphe dans le secteur angulaire, prenant la valeur 0 en $z = 1$. (formule de Stirling complexe).

Démonstration. — On part de la formule $\frac{\Gamma'(z)}{\Gamma(z)} = -\gamma - \frac{1}{z} + \sum_{k=1}^{+\infty} \left(\frac{1}{k} - \frac{1}{z+k}\right)$, qui découle (th. V.2.15) de la définition de $\frac{1}{\Gamma}$ comme un produit convergent. Si $z \notin \mathbf{R}_-$, on peut réécrire cette formule sous la forme (où $[t]$ désigne la partie entière de t et $\{t\} = t - [t]$, sa partie fractionnaire)

$$\begin{aligned} \frac{\Gamma'(z)}{\Gamma(z)} &= -\gamma - \frac{1}{z} + \int_1^{+\infty} \left(\frac{1}{[t]} - \frac{1}{z+[t]}\right) dt \\ &= -\gamma - \frac{1}{z} + \int_1^{+\infty} \left(\frac{1}{[t]} - \frac{1}{t} - \frac{1}{z+[t]} + \frac{1}{z+t}\right) dt + \int_1^{+\infty} \left(\frac{1}{t} - \frac{1}{z+t}\right) dt \\ &= -\gamma - \frac{1}{z} + \int_1^{+\infty} \frac{\{t\}}{t[t]} dt - \int_1^{+\infty} \frac{\{t\}}{(z+t)(z+[t])} dt + \left[\log \frac{t}{z+t}\right]_1^{+\infty}. \end{aligned}$$

Maintenant, $\left[\log \frac{t}{z+t}\right]_1^{+\infty} = \log(z+1) = \log z + \frac{1}{z} + O\left(\frac{1}{z^2}\right)$ et $C = \int_1^{+\infty} \frac{\{t\}}{t[t]} dt$ est une constante.

Finalement, on peut écrire $\int_1^{+\infty} \frac{\{t\}}{(z+t)(z+[t])} dt$ sous la forme

$$\begin{aligned} \int_1^{+\infty} \frac{\{t\}}{(z+t)(z+[t])} dt &= \int_1^{+\infty} \frac{\{t\}}{(z+t)^2} dt + \int_1^{+\infty} \frac{\{t\}^2}{(z+t)^2(z+[t])} dt \\ &= \frac{1}{2(z+1)} + \int_1^{+\infty} \frac{\{t\} - 1/2}{(z+t)^2} dt + \int_1^{+\infty} \frac{\{t\}^2}{(z+t)^2(z+[t])} dt, \end{aligned}$$

et une intégration par parties nous donne

$$\int_1^{+\infty} \frac{\{t\} - 1/2}{(z+t)^2} dt = \left[\frac{\{t\}^2 - \{t\}}{2(z+t)^2}\right]_1^{+\infty} + \int_1^{+\infty} \frac{\{t\}^2 - \{t\}}{(z+t)^3} dt = \int_1^{+\infty} \frac{\{t\}^2 - \{t\}}{(z+t)^3} dt.$$

En utilisant la minoration

$$|z+u|^2 \geq \frac{1+\cos \alpha}{2} (|z|+u)^2, \quad \text{si } u \in \mathbf{R}^+ \text{ et } |\arg(z)| \leq \alpha,$$

cela permet de majorer en module les quantités $\int_1^{+\infty} \frac{\{t\}^2}{(z+t)^2(z+[t])} dt$ et $\int_1^{+\infty} \frac{\{t\}^2 - \{t\}}{(z+t)^3} dt$ par $\int_1^2 \frac{1}{(|z|-2)^3} dt + \int_2^{+\infty} \left(\frac{2}{1+\cos \alpha}\right)^{3/2} \frac{1}{(|z|+t-1)^3} dt$. On en déduit que ces deux intégrales sont des $O\left(\frac{1}{z^2}\right)$ et comme il en est de

⁽⁵⁾La dérivée logarithmique d'une fonction f est la fonction $\frac{f'}{f}$; c'est la dérivée de $\log f$.

même de $\frac{1}{2(z+1)} - \frac{1}{2z}$, on obtient finalement :

$$\frac{\Gamma'(z)}{\Gamma(z)} = \log z + C - \gamma - \frac{1}{2z} + O\left(\frac{1}{z^2}\right).$$

On en déduit, en intégrant, que

$$\log \Gamma(z) = \left[w \log w - w + (C - \gamma)w - \frac{1}{2} \log w \right]_1^z + \int_1^{+\infty} O\left(\frac{1}{w^2}\right) dw + \int_{+\infty}^z O\left(\frac{1}{w^2}\right) dw.$$

D'où l'existence de $C' \in \mathbf{C}$, tel que l'on ait $\log \Gamma(z) = z \log z - z + (C - \gamma)z - \frac{1}{2} \log z + C' + O\left(\frac{1}{z}\right)$ au voisinage de l'infini. La formule de Stirling réelle :

$$\log(\Gamma(x+1)) = x \log x - x + \frac{1}{2} \log(2\pi x) + o(1), \quad \text{au voisinage de } +\infty$$

permet alors d'en déduire que $C - \gamma = 0$ et $C' = \frac{1}{2} \log 2\pi$, ce qui permet de conclure.

Corollaire VII.2.4. — *Quand $|\tau|$ tend vers $+\infty$, on a*

$$|\Gamma(\sigma + i\tau)| = \sqrt{2\pi} |\tau|^{\sigma - \frac{1}{2}} e^{-\pi|\tau|/2} \left(1 + O\left(\frac{1}{\tau}\right)\right),$$

uniformément dans toute bande verticale de largeur finie.

Démonstration. — Si $s = \sigma + i\tau$, et $a \leq \sigma \leq b$, où $a, b \in \mathbf{R}$ sont fixés, alors

$$\log s = \log i\tau + \frac{\sigma}{i\tau} + O\left(\frac{1}{\tau^2}\right) = \log |\tau| + \frac{i\pi}{2} \cdot \frac{\tau}{|\tau|} + \frac{\sigma}{i\tau} + O\left(\frac{1}{\tau^2}\right).$$

En utilisant le (ii) de la prop. VII.2.3 et la formule ci-dessus, on en déduit que

$$\operatorname{Re}(\log \Gamma(\sigma + i\tau)) = \sigma \log |\tau| - \frac{\pi|\tau|}{2} + \sigma - \sigma - \frac{1}{2} \log |\tau| + \frac{1}{2} \log 2\pi + O\left(\frac{1}{\tau}\right).$$

Ceci permet de conclure.

2. Une formule intégrale pour les séries de Dirichlet

Le changement de variable $u = \lambda t$ montre que, si $\lambda \in \mathbf{R}_+^*$, et si $\operatorname{Re}(s) > 0$, alors

$$\frac{1}{\lambda^s} = \frac{1}{\Gamma(s)} \int_0^{+\infty} e^{-\lambda t} t^s \frac{dt}{t}.$$

Cette simple remarque va se révéler extrêmement utile pour étudier le prolongement analytique de certaines séries de Dirichlet.

Soit $L(\mathbf{a}, s) = \sum_{n \geq 1} \frac{a_n}{n^s}$ une série de Dirichlet convergeant quelque part.

Lemme VII.2.5. — (i) *La série entière $F_{\mathbf{a}}(z) = \sum_{n=1}^{+\infty} a_n z^n$ est de rayon de convergence au moins 1.*

(ii) *$f_{\mathbf{a}}(t) = F_{\mathbf{a}}(e^{-t})$ est \mathcal{C}^∞ sur \mathbf{R}_+^* et à décroissance rapide à l'infini, ainsi que toutes ses dérivées.*

Démonstration. — Il existe $\tau \in \mathbf{R}$ tel que $a_n = O(n^\tau)$; on en déduit le (i) et le fait que $F_{\mathbf{a}}(z) = O(|z|)$ au voisinage de 0. Ceci implique que $f_{\mathbf{a}}$ est \mathcal{C}^∞ sur \mathbf{R}_+^* et $O(e^{-t})$ au voisinage de $+\infty$ (et donc à décroissance rapide, ce qui, pour une fonction $f : \mathbf{R}_+ \rightarrow \mathbf{C}$, signifie, rappelons le, que $t^N f(t)$ est bornée quand $t \rightarrow +\infty$, pour tout $N \in \mathbf{N}$). Pour passer au cas d'une dérivée d'ordre quelconque de $f_{\mathbf{a}}$, il suffit de constater que $f_{\mathbf{a}}^{(k)} = f_{\mathbf{a}^{(k)}}$, où $\mathbf{a}^{(k)} = ((-n)^k a_n)_{n \in \mathbf{N}}$.

Lemme VII.2.6. — Si $\operatorname{Re}(s) > \sup(\sigma_{\text{abs}}, 0)$, la fonction $f_{\mathbf{a}}(t)t^{s-1}$ est sommable sur \mathbf{R}_+^* , et on a

$$L(\mathbf{a}, s) = \frac{1}{\Gamma(s)} \int_0^{+\infty} f_{\mathbf{a}}(t)t^s \frac{dt}{t}.$$

Démonstration. — Si $\operatorname{Re}(s) = \sigma > \sup(\sigma_{\text{abs}}, 0)$, on a

$$\begin{aligned} \int_0^{+\infty} |f_{\mathbf{a}}(t)t^{s-1}| dt &\leq \int_0^{+\infty} \left(\sum_{n=1}^{+\infty} |a_n| e^{-nt} \right) t^\sigma \frac{dt}{t} \\ &= \sum_{n=1}^{+\infty} |a_n| \int_0^{+\infty} e^{-nt} t^\sigma \frac{dt}{t} = \Gamma(\sigma) \sum_{n=1}^{+\infty} \frac{|a_n|}{n^\sigma} < +\infty, \end{aligned}$$

ce qui prouve que $f_{\mathbf{a}}(t)t^{s-1}$ est sommable sur \mathbf{R}_+^* et que la série $\sum_{n=1}^{+\infty} a_n e^{-nt} t^{s-1}$ converge dans $L^1(\mathbf{R}_+^*)$ vers $f_{\mathbf{a}}(t)t^{s-1}$. On a donc

$$\int_0^{+\infty} f_{\mathbf{a}}(t)t^{s-1} dt = \sum_{n=1}^{+\infty} \int_0^{+\infty} a_n e^{-nt} t^{s-1} dt = \sum_{n=1}^{+\infty} a_n \frac{\Gamma(s)}{n^s}.$$

Ceci permet de conclure.

Remarque VII.2.7. — (i) Si $f : \mathbf{R}_+^* \rightarrow \mathbf{C}$, la fonction $s \mapsto \operatorname{Mel}(f, s) = \int_0^{+\infty} f(t)t^s \frac{dt}{t}$ est la *transformée de Mellin* de f . Le changement de variable $t = e^u$ montre que, sur une droite verticale $\operatorname{Re}(s) = \sigma$, la transformée de Mellin coïncide, à homothétie près, avec la transformée de Fourier de $f(e^u)e^{\sigma u}$, ce qui permet de déduire un grand nombre de ses propriétés de celles de la transformée de Fourier.

(ii) La fonction $f_{\mathbf{a}}(t) = \sum_{n=1}^{+\infty} a_n e^{-nt}$ est \mathcal{C}^∞ sur \mathbf{R}_+^* , mais n'a aucune raison, a priori, d'être très sympathique en 0. De fait, les propriétés de prolongement analytique de la série de Dirichlet $\sum_{n=1}^{+\infty} \frac{a_n}{n^s}$ sont étroitement liées à la régularité de $f_{\mathbf{a}}$ en 0. La prop. VII.2.8 ci-dessous donne une illustration de ce phénomène.

Proposition VII.2.8. — Soit f une fonction \mathcal{C}^∞ sur \mathbf{R}_+ , à décroissance rapide à l'infini ainsi que toutes ses dérivées.

(i) La fonction $M(f, s) = \frac{1}{\Gamma(s)} \int_0^{+\infty} f(t) t^s \frac{dt}{t}$, définie pour $\operatorname{Re}(s) > 0$, admet un prolongement holomorphe à \mathbf{C} tout entier⁽⁶⁾

(ii) Si $k \in \mathbf{N}$, alors $M(f, -k) = (-1)^k f^{(k)}(0)$.

(iii) Si $a \leq b$ sont deux réels, alors pour tout $k \in \mathbf{N}$, il existe $C_{a,b,k}(f)$ tel que

$$|L(\mathbf{a}, s)| \leq C_{a,b,k}(f) (1 + |\tau|)^{-k} e^{\frac{\pi}{2}|\tau|}, \text{ si } s = \sigma + i\tau \text{ et } a \leq \sigma \leq b.$$

Démonstration. — Sur la bande $b > \operatorname{Re}(s) > a > 0$, on a $|f(t)t^{s-1}| \leq |f(t) \sup(t^{a-1}, t^{b-1})|$. Or $\int_0^{+\infty} |f(t) \sup(t^{a-1}, t^{b-1})| dt < +\infty$, grâce à la décroissance rapide de f à l'infini, et à l'hypothèse $a > 0$ pour ce qui se passe au voisinage de 0. On est sous les conditions d'application du th. V.2.18, ce qui permet de montrer que $M(f, s)$ est holomorphe sur le demi-plan $\operatorname{Re}(s) > 0$.

Par ailleurs, une intégration par partie nous donne $M(f, s) = -M(f', s+1)$, si $\operatorname{Re}(s) > 0$. On a donc, plus généralement, $M(f, s) = (-1)^n M(f^{(n)}, s+n)$, si $\operatorname{Re}(s) > 0$ et $n \in \mathbf{N}$. En appliquant ce qui précède à $f^{(n)}$, au lieu de f , on en déduit que $(-1)^n M(f^{(n)}, s+n)$ est holomorphe sur le demi-plan $\operatorname{Re}(s) > -n$. Comme cette fonction coïncide avec $M(f, s)$ sur le demi-plan $\operatorname{Re}(s) > 0$, cela prouve que $M(f, s)$ admet un prolongement holomorphe au demi-plan $\operatorname{Re}(s) > -n$, et comme ceci est vrai quel que soit $n \in \mathbf{N}$, cela permet de démontrer le (i).

Le (ii) suit de ce que $M(f, -k) = (-1)^{k+1} M(f^{(k+1)}, 1)$ est aussi égal à

$$(-1)^{k+1} \int_0^{+\infty} f^{(k+1)}(t) dt = (-1)^{k+1} [f^{(k)}]_0^{+\infty} = (-1)^k f^{(k)}(0).$$

Passons à la démonstration du (iii), et choisissons $n \in \mathbf{N}$ tel que $n + a > k + \frac{1}{2}$. Si $s = \sigma + i\tau$, avec $a \leq \sigma \leq b$, alors

$$|M(f, s)| \leq \frac{1}{|\Gamma(s+n)|} \int_0^{+\infty} |f^{(n)}(t)| t^{\sigma+n} \frac{dt}{t} \leq \frac{C_n}{|\Gamma(s+n)|},$$

où $C_n = \int_0^{+\infty} |f^{(n)}(t)| \sup(t^{a+n}, t^{b+n}) \frac{dt}{t}$. Par ailleurs, d'après le cor. VII.2.4, il existe $T > 0$ tel que, si $s = \sigma + i\tau$, alors

$$|\Gamma(s+n)|^{-1} \leq \frac{2}{\sqrt{2\pi}} (1 + |\tau|)^{\frac{1}{2} - \sigma - n} e^{\frac{\pi}{2}|\tau|} \leq \frac{2}{\sqrt{2\pi}} (1 + |\tau|)^{-k} e^{\frac{\pi}{2}|\tau|}, \text{ si } |\tau| \geq T \text{ et } a \leq \sigma \leq b.$$

Donc $(1 + |\tau|)^k e^{-\frac{\pi}{2}|\tau|} |M(f, s)|$ est bornée sur $\{s = \sigma + i\tau, |\tau| \geq T, a \leq \sigma \leq b\}$, et comme elle est continue sur la bande $a \leq \operatorname{Re}(s) \leq b$, elle est aussi bornée sur la bande toute entière. Ceci permet de conclure.

Corollaire VII.2.9. — Si $f_{\mathbf{a}}$ est \mathcal{C}^∞ sur \mathbf{R}_+ , alors $L(\mathbf{a}, s)$ admet un prolongement holomorphe à \mathbf{C} tout entier. De plus, si $k \in \mathbf{N}$, on a $L(\mathbf{a}, -k) = (-1)^k f_{\mathbf{a}}^{(k)}(0)$.

⁽⁶⁾La distribution $f \mapsto M(f, s)$ que ceci permet de définir, est une *partie finie de Hadamard*. Le (ii) montre que si $k \in \mathbf{N}$, alors $M(\cdot, -k)$ est la dérivée k -ième de la masse de Dirac en 0.

Démonstration. — C'est une conséquence directe de la prop. VI.2.7 puisque $f_{\mathbf{a}}$ est à décroissance rapide à l'infini (lemme VII.2.5).

Remarque VII.2.10. — La formule $L(\mathbf{a}, -k) = (-1)^k f_{\mathbf{a}}^{(k)}(0)$ montre que l'on obtient la même valeur pour la somme de la série divergente $\sum_{n \geq 1} a_n n^k$, en prenant la valeur de $L(\mathbf{a}, s)$ en $-k$ ou la limite quand $x \rightarrow 1^-$ de $\sum_{n=1}^{+\infty} a_n n^k x^n$.

VII.3. La fonction zêta de Riemann

1. Séries de Dirichlet attachées à des fonctions multiplicatives

On note \mathcal{P} l'ensemble des nombres premiers.

Une fonction $n \mapsto a(n)$ de $\mathbf{N} - \{0\}$ dans \mathbf{C} est *multiplicative*, si $a(1) = 1$, et si $a(nm) = a(n)a(m)$ pour tous n et m premiers entre eux; elle est *strictement multiplicative*, si $a(1) = 1$, et $a(nm) = a(n)a(m)$, quels que soient $m, n \geq 1$. On remarquera qu'une fonction strictement multiplicative est déterminée par les $a(p)$, pour $p \in \mathcal{P}$ puisque, si $n = \prod_{i \in I} p_i^{k_i}$ est la décomposition de n en facteurs premiers, on a $a(n) = \prod_{i \in I} a(p_i)^{k_i}$. Par contre, pour une fonction multiplicative, on a besoin de connaître les $a(p^k)$, pour $p \in \mathcal{P}$ et $k \geq 1$. On a alors $a(n) = \prod_{i \in I} a(p_i^{k_i})$.

Proposition VII.3.1. — Si $n \mapsto a(n)$ est multiplicative, et si $L(a, s) = \sum_{n \geq 1} \frac{a(n)}{n^s}$ converge quelque part, alors pour tout s dans le demi-plan $\operatorname{Re}(s) > \sigma_{\text{abs}}$, on a :

- (i) $1 + \frac{a(p)}{p^s} + \frac{a(p^2)}{p^{2s}} + \dots$ est absolument convergent quel que soit $p \in \mathcal{P}$;
- (ii) le produit $\prod_{p \in \mathcal{P}} (1 + \frac{a(p)}{p^s} + \frac{a(p^2)}{p^{2s}} + \dots)$ converge uniformément sur tout demi-plan $\operatorname{Re}(s) > c > \sigma_{\text{abs}}$, et sa valeur est $L(a, s)$.

Démonstration. — Soit $c > \sigma_{\text{abs}}$. Si $\operatorname{Re}(s) > c$, on a

$$\sum_{p \in \mathcal{P}} \left| \frac{a(p)}{p^s} + \frac{a(p^2)}{p^{2s}} + \dots \right| \leq \sum_{p \in \mathcal{P}} \left(\left| \frac{a(p)}{p^s} \right| + \left| \frac{a(p^2)}{p^{2s}} \right| + \dots \right) \leq \sum_{n=2}^{+\infty} \frac{|a(n)|}{n^c} < +\infty.$$

On en déduit le (i) et, en utilisant le th. V.2.15 (ou plutôt la démonstration du (i) de ce théorème), la convergence uniforme du produit sur $\operatorname{Re}(s) > c$.

Si $X \in \mathbf{R}_+$, soit $\mathcal{P}(X)$ l'ensemble des nombres premiers $\leq X$, et soit $I(X)$ l'ensemble des entiers dont tous les diviseurs premiers sont dans $\mathcal{P}(X)$. Si $k \in \mathbf{N}$, soit $I(X, k)$ l'ensemble des entiers de la forme $\prod_{p \in \mathcal{P}(X)} p^{k_p}$, avec $0 \leq k_p \leq k$. L'ensemble $I(X, k)$ est donc un ensemble fini. Par ailleurs, la multiplicativité de $n \mapsto a(n)$ fait que l'on a

$$\prod_{p \in \mathcal{P}(X)} \left(1 + \frac{a(p)}{p^s} + \dots + \frac{a(p^k)}{p^{ks}} \right) = \sum_{n \in I(X, k)} \frac{a(n)}{n^s}.$$

Comme toutes les séries qui interviennent sont majorées, en valeur absolue, par la série sommable $\sum_{n \geq 1} \left| \frac{a(n)}{n^s} \right|$, le théorème de convergence dominée pour les séries montre, en

faisant tendre k vers $+\infty$, que

$$\prod_{p \in \mathcal{P}(X)} \left(1 + \frac{a(p)}{p^s} + \frac{a(p^2)}{p^{2s}} + \cdots\right) = \sum_{n \in I(X)} \frac{a(n)}{n^s},$$

puis, en faisant tendre X vers $+\infty$, que

$$\prod_{p \in \mathcal{P}} \left(1 + \frac{a(p)}{p^s} + \frac{a(p^2)}{p^{2s}} + \cdots\right) = \sum_{n \geq 1} \frac{a(n)}{n^s},$$

ce que l'on cherchait à établir.

Remarque VII.3.2. — (i) Le facteur $1 + \frac{a(p)}{p^s} + \frac{a(p^2)}{p^{2s}} + \cdots$ du produit est le *facteur d'Euler* en p de la fonction $L(a, s)$, et la formule $L(a, s) = \prod_{p \in \mathcal{P}} \left(1 + \frac{a(p)}{p^s} + \frac{a(p^2)}{p^{2s}} + \cdots\right)$ est la *décomposition de $L(a, s)$ en produit de facteurs d'Euler* (ou en *produit eulérien*).

(ii) Si a est complètement multiplicative, les facteurs d'Euler sont donnés par des séries géométriques et on a $L(a, s) = \prod_{p \in \mathcal{P}} \frac{1}{1 - a(p)p^{-s}}$.

2. Prolongement analytique de la fonction ζ

La série de Dirichlet $L(\mathbf{1}, s) = \sum_{n \geq 1} \frac{1}{n^s}$ a 1 comme abscisse de convergence absolue (et comme abscisse de convergence puisqu'elle est à coefficients positifs (cor. VII.1.2)). De plus $n \mapsto 1$ est on ne peut plus strictement multiplicative; on déduit donc de la théorie générale le résultat suivant (dû à Euler).

Proposition VII.3.3. — Si $\operatorname{Re}(s) > 1$, alors $\sum_{n \geq 1} \frac{1}{n^s} = \prod_{p \in \mathcal{P}} \frac{1}{1 - p^{-s}}$, et le produit est uniformément convergent sur tout demi-plan $\operatorname{Re}(s) > c > 1$.

Théorème VII.3.4. — Il existe une unique fonction, notée ζ (fonction zêta de Riemann), vérifiant :

- ζ est méromorphe sur \mathbf{C} tout entier, holomorphe en dehors d'un pôle simple en $s = 1$, de résidu 1 ;
- $\zeta(s) = L(\mathbf{1}, s)$, si $\operatorname{Re}(s) > 1$.

Démonstration. — L'unicité suit est une conséquence de l'unicité du prolongement analytique (cor. V.1.17). Si $f_1(t) = \sum_{n=1}^{+\infty} e^{-nt} = \frac{1}{e^t - 1}$ et si $g(t) = t f_1(t)$, on déduit du lemme VII.2.6 que, quel que soit s , avec $\operatorname{Re}(s) > 1$,

$$L(\mathbf{1}, s) = \frac{1}{\Gamma(s)} \int_0^{+\infty} f_1(t) t^s \frac{dt}{t} = \frac{1}{(s-1)\Gamma(s-1)} \int_0^{+\infty} t f_1(t) t^{s-1} \frac{dt}{t} = \frac{1}{s-1} M(g, s-1).$$

Comme $g(t)$ est \mathcal{C}^∞ sur \mathbf{R}_+^* , holomorphe au voisinage de 0 (et donc \mathcal{C}^∞ sur \mathbf{R}_+), et à décroissance rapide à l'infini, la prop. VII.2.8 s'applique. On en déduit que $M(g, s)$ a un prolongement analytique à \mathbf{C} tout entier, avec $M(g, 0) = g(0) = 1$. Le résultat s'en déduit.

Exercice VII.3.5. — Soit $\sum_{n=0}^{+\infty} B_n \frac{t^n}{n!}$ le développement de Taylor⁽⁷⁾ de $g(t) = \frac{t}{e^t - 1}$ en 0.

(i) Calculer $g(t) - g(-t)$. En déduire $B_{2k+1} = 0$, si $k \geq 1$.

(ii) Montrer que $\zeta(-n) = (-1)^n \frac{B_{n+1}}{n+1}$, si $n \in \mathbf{N}$. En déduire que ζ prend des valeurs rationnelles aux entiers⁽⁸⁾ négatifs, et a des zéros en les entiers pairs < 0 .

(iii) Montrer que $g(z) - \frac{2i\pi}{z-2i\pi} + \frac{2i\pi}{z+2i\pi}$ est holomorphe sur $D(0, (4\pi)^-)$. En déduire des équivalents de $\frac{B_{2n}}{(2n)!}$ et de $\zeta(1-2n)$ quand $n \rightarrow +\infty$.

3. Équation fonctionnelle de la fonction zêta

Théorème VII.3.6. — La fonction ζ vérifie l'équation fonctionnelle

$$\zeta(s) = 2 \cdot (2\pi)^{s-1} \cdot \Gamma(1-s) \cdot \sin \frac{\pi s}{2} \cdot \zeta(1-s).$$

Remarque VII.3.7. — (i) Soit $\xi(s) = \pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) \zeta(s)$. Modulo les formules classiques (ex. VII.2.2)

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}, \quad \Gamma\left(\frac{s}{2}\right)\Gamma\left(\frac{s+1}{2}\right) = 2^{1-s}\Gamma\left(\frac{1}{2}\right)\Gamma(s), \quad \text{et} \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi},$$

on peut déduire de l'équation fonctionnelle de ζ que ξ vérifie l'équation fonctionnelle

$$\xi(s) = \xi(1-s).$$

(ii) L'équation fonctionnelle de la fonction ξ peut s'établir directement (cf. ex. VII.6.6), mais la démonstration ci-dessous a l'avantage de se généraliser plus facilement aux fonctions L de Dirichlet. De plus, c'est sous la forme du th. VII.3.6 que nous utiliserons l'équation fonctionnelle de ζ dans la démonstration du théorème des nombres premiers (annexe A).

Démonstration. — Si $c > 0$, soit γ_c le contour obtenu en composant la demi-droite $(+\infty, c]$ suivi du carré C_c de sommets $c(\pm 1 \pm i)$ parcouru dans le sens direct et de la demi-droite $[c, +\infty)$. Soit

$$F_c(s) = \frac{1}{2i\pi} \int_{\gamma_c} f_1(z) (-z)^s \frac{dz}{z},$$

⁽⁷⁾Les B_n sont des nombres rationnels, appelés nombres de Bernoulli, et qu'on retrouve dans toutes les branches des mathématiques. On a en particulier

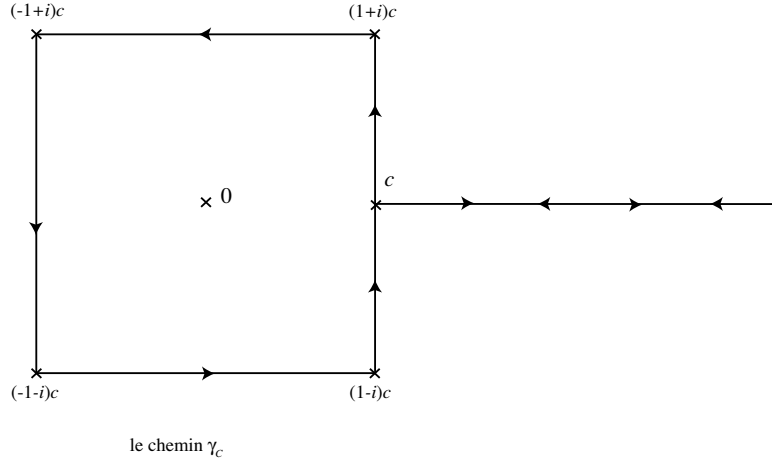
$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad \dots, \quad B_{12} = \frac{-691}{2730}, \dots$$

Un test presque infaillible pour savoir si une suite de nombres a un rapport avec les nombres de Bernoulli est de regarder si 691 apparaît dans les premiers termes de cette suite.

⁽⁸⁾Les valeurs aux entiers de la fonction zêta, ou plus généralement des fonctions L de la géométrie arithmétique, recellent une quantité impressionnante d'informations arithmétiques. Kummer fut l'un des premiers à exploiter cette information, ce qui lui a permis de montrer (1852) que, si p est un nombre premier ≥ 3 ne divisant pas le numérateur de $\zeta(-1)$, $\zeta(-3)$, \dots , $\zeta(2-p)$, alors l'équation $a^p + b^p = c^p$ n'a pas de solution en nombres entiers avec $abc \neq 0$ (i.e. le théorème de Fermat est vrai pour un tel p (dit *régulier*)). Jusqu'à 100, les seuls nombres premiers irréguliers sont 37, 59 et 67.

où $f_1(z) = \frac{1}{e^z - 1}$, et $(-z)^s = \exp(s \log(-z))$, la détermination du logarithme choisie étant celle dont la partie imaginaire est comprise entre $-\pi$ et π ; en particulier, on a $(-z)^s = e^{-i\pi s} z^s$ de $+\infty$ à c et $(-z)^s = e^{i\pi s} z^s$ de c à $+\infty$ (après avoir parcouru le carré).

La démonstration consiste à utiliser la formule des résidus pour évaluer $F_d(s) - F_c(s)$, ce qui fait apparaître la fonction $\zeta(1-s)$. La fonction $\zeta(s)$ s'obtient en faisant tendre c vers 0, et un passage à la limite quand $c \rightarrow +\infty$ donne le lien cherché entre $\zeta(s)$ et $\zeta(1-s)$.



• *Calcul de $F_d(s) - F_c(s)$.* (Un peu plus de familiarité avec les fonctions holomorphes permettrait de se passer de la cuisine peu ragoûtante qui suit et de démontrer directement que $F_d(s) - F_c(s)$ est la somme des résidus de la fonction $f_1(z) \frac{(-z)^s}{z}$ (qui est méromorphe sur $\mathbf{C} - [0, +\infty)$) en les points à l'intérieur du chemin $\gamma_{c,d}$ composé de C_d , $[d, c]$, C_c parcouru en sens opposé, et $[c, d]$.)

Les chemins sur lesquels on intègre ne sont pas vraiment contenus dans un ouvert sur lequel la fonction qu'on intègre est méromorphe; pour se ramener à ce cas, on va être forcé de tout découper en morceaux.

On note g_s^+ (resp. g_s^-) la fonction $g_s^+(z) = f_1(z) \frac{(-z)^s}{z}$ sur l'ouvert Ω^+ (resp. Ω^-) obtenu en enlevant la demi-droite $[0, -i\infty)$ (resp. $[0, +i\infty)$) à \mathbf{C} , la détermination de $\log(-z)$ étant celle prenant des valeurs réelles sur la demi-droite $[0, -\infty)$. Les fonctions g_s^+ et g_s^- sont méromorphes sur Ω^+ et Ω^- respectivement, et coïncident sur le demi-plan $\operatorname{Re}(s) < 0$; par contre, sur le demi-plan $\operatorname{Re}(s) > 0$, on a $g_s^-(z) = e^{2i\pi s} g_s^+(z)$.

On note C_c^+ (resp. C_c^-) le morceau de C_c contenu dans le demi-plan $\operatorname{Im}(s) \geq 0$ (resp. $\operatorname{Im}(s) \leq 0$). On a donc $C_c^+ = [c, (1+i)c] \cdot [(1+i)c, (-1+i)c] \cdot [(-1+i)c, -c]$, et $C_c^- = [-c, (-1-i)c] \cdot [(-1-i)c, (1-i)c] \cdot [(1-i)c, c]$. Soit γ_c^+ (resp. γ_c^-) le chemin composé de $(+\infty, c]$ et C_c^+ (resp. de C_c^- et $[c, +\infty)$), et soient $F_c^+(s) = \frac{1}{2i\pi} \int_{\gamma_c^+} g_s^+(z) dz$ et $F_c^-(s) = \frac{1}{2i\pi} \int_{\gamma_c^-} g_s^-(z) dz$, de telle sorte que $F_c(s) = F_c^+(s) + F_c^-(s)$.

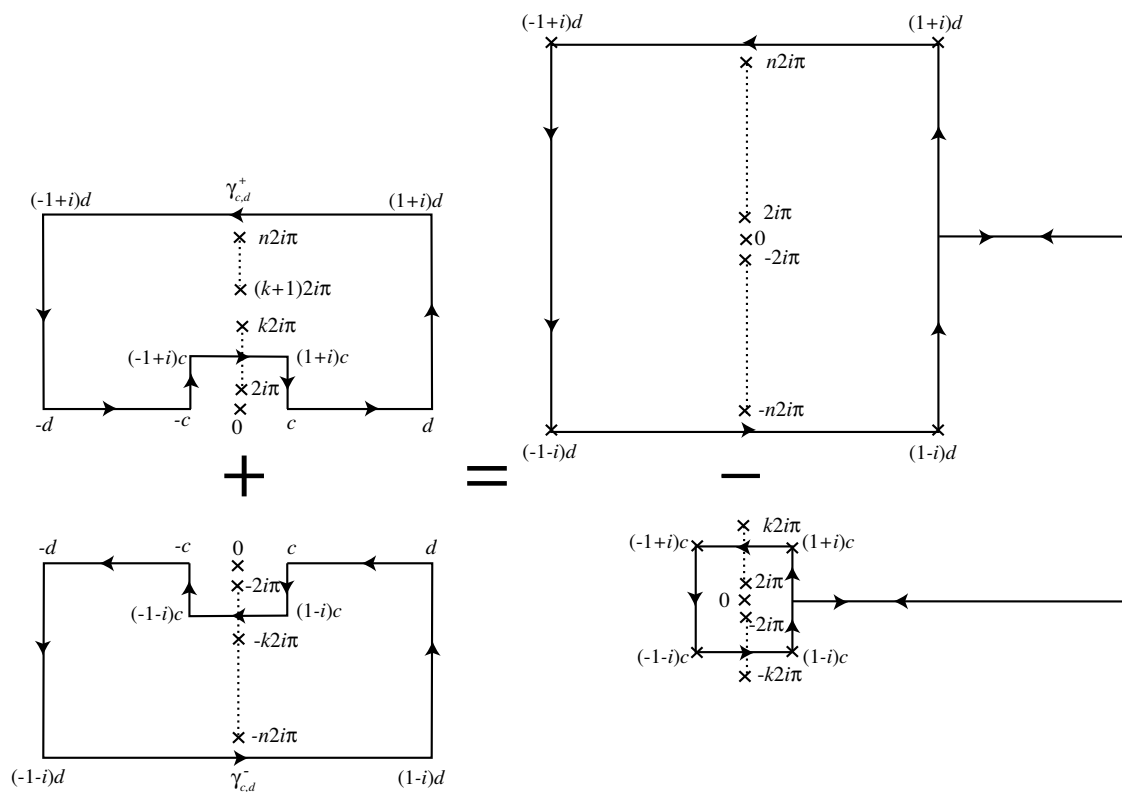
Si $c < d$, soient $\gamma_{c,d}^+$ le lacet $C_d^+ \cdot [-d, -c] \cdot (C_c^+)^{\text{opp}} \cdot [c, d]$ et $\gamma_{c,d}^-$ le lacet $C_d^- \cdot [d, c] \cdot (C_c^-)^{\text{opp}} \cdot [-c, -d]$, où $(C_c^+)^{\text{opp}}$ et $(C_c^-)^{\text{opp}}$ désignent les chemins C_c^+ et C_c^- parcourus dans le sens opposé. On a

$$F_d^+(s) - F_c^+(s) = \frac{1}{2i\pi} \int_{\gamma_{c,d}^+} g_s^+(z) dz - \frac{1}{2i\pi} \int_{[-d, -c]} g_s^+(z) dz$$

$$F_d^-(s) - F_c^-(s) = \frac{1}{2i\pi} \int_{\gamma_{c,d}^-} g_s^-(z) dz - \frac{1}{2i\pi} \int_{[-c, -d]} g_s^-(z) dz,$$

comme on peut le voir sur le dessin ci-dessus. Comme $g_s^+(z) = g_s^-(z)$ sur $[-c, -d]$, l'intégrale sur $[-c, -d]$ disparaît quand on fait la somme des deux égalités ci-dessus, et on obtient

$$F_d(s) - F_c(s) = \frac{1}{2i\pi} \int_{\gamma_{c,d}^+} g_s^-(z) dz + \frac{1}{2i\pi} \int_{\gamma_{c,d}^-} g_s^+(z) dz.$$



$$\int_{\gamma_d} - \int_{\gamma_c} = \int_{\gamma_{c,d}^+} + \int_{\gamma_{c,d}^-}$$

Le membre de droite peut se calculer grâce à la formule des résidus. La fonction g_s^+ est méromorphe sur Ω^+ , avec des pôles simples aux $2i\pi k$, pour $k \in \mathbf{N} - \{0\}$, et comme $e^z - 1$ a une dérivée égale à 1 en $2i\pi k$, on a $\text{Res}(g_s^+, 2i\pi k) = \frac{(-2i\pi k)^s}{2i\pi k} = -(2k\pi)^{s-1} e^{-i\pi \frac{s-1}{2}}$. De même, g_s^- est méromorphe sur Ω^- , avec des pôles simples aux $-2i\pi k$, pour $k \in \mathbf{N} - \{0\}$, et on a $\text{Res}(g_s^-, -2i\pi k) = \frac{(2i\pi k)^s}{-2i\pi k} = -(2k\pi)^{s-1} e^{i\pi \frac{s-1}{2}}$. Par ailleurs, l'indice de $\gamma_{c,d}^+$ par rapport à $2i\pi k$ est 1, si $c < 2k < d$, et 0 sinon ; l'indice de $\gamma_{c,d}^-$ par rapport à $-2i\pi k$ est 1, si $c < 2k < d$, et 0 sinon. On obtient donc, en utilisant la formule des résidus,

$$F_d(s) - F_c(s) = - \sum_{c < 2k\pi < d} 2 \cos \pi \frac{s-1}{2} \cdot (2k\pi)^{s-1}.$$

• *Une majoration pour $|(-z)^s|$.* On a $|(-z)^s| = e^{\operatorname{Re}(s \log(-z))} = |z|^{\operatorname{Re}(s)} e^{-\operatorname{Im}(s) \arg(-z)}$. Comme $\arg(-z)$ varie entre $-\pi$ et π sur les domaines considérés, cela montre que, si s est fixé, il existe une constante $c(s)$ telle que $|(-z)^s| \leq c(s) |z|^{\operatorname{Re}(s)}$, pour tout z .

• *Lien entre F_π et $\zeta(s)$.* La formule obtenue pour $F_d(s) - F_c(s)$ montre, en particulier, que F_c ne dépend pas de c dans l'intervalle $]0, 2\pi[$, et donc que $F_\pi(s) = \lim_{c \rightarrow 0^+} F_c(s)$. Or la majoration ci-dessus pour $|(-z)^s|$ implique que, quand c tend vers 0, l'intégrale sur C_c tend vers 0, si $\operatorname{Re}(s) > 1$. On obtient donc, en passant à la limite si $\operatorname{Re}(s) > 1$,

$$F_\pi(s) = \frac{1}{2i\pi} \left(e^{-i\pi s} \int_{+\infty}^0 \frac{1}{e^t - 1} t^s \frac{dt}{t} + e^{i\pi s} \int_0^{+\infty} \frac{1}{e^t - 1} t^s \frac{dt}{t} \right) = \frac{\sin \pi s}{\pi} \cdot \Gamma(s) \cdot \zeta(s).$$

• *Lien entre F_π et $\zeta(1-s)$.* On a $|\frac{1}{e^z - 1}| \leq \frac{1}{1 - e^{-(2N+1)\pi}} \leq \frac{1}{2}$ sur le carré $C_{(2N+1)\pi}$ (sur un bord vertical, on minore $|e^z - 1|$ par $||e^z| - 1|$, et on remarque que e^z est réel négatif sur les deux bords horizontaux). On en déduit, en utilisant la majoration $|(-z)^s| \leq c(s) |z|^{\operatorname{Re}(s)}$, que, si $\operatorname{Re}(s) < 0$, alors $F_{(2N+1)\pi}(s) \rightarrow 0$ quand $N \rightarrow +\infty$. En utilisant la formule ci-dessus pour $F_d(s) - F_c(s)$, avec $d = (2N+1)\pi$ et $c = \pi$, et en passant à la limite quand $N \rightarrow +\infty$, on obtient

$$F_\pi(s) = 2 \cdot \cos\left(\pi \frac{s-1}{2}\right) \cdot (2\pi)^{s-1} \cdot \zeta(1-s) \quad \text{si } \operatorname{Re}(s) < 0.$$

• *Holomorphie de F_π .* On a $F_\pi(s) = \frac{1}{2i\pi} \left(\int_{C_\pi} \frac{(-z)^s}{e^z - 1} \frac{dz}{z} + (e^{i\pi s} - e^{-i\pi s}) \int_\pi^{+\infty} \frac{t^s}{e^t - 1} \frac{dt}{t} \right)$. Sur un demi-plan $\operatorname{Re}(s) > a$, on a $|t^s| \leq t^a$, si $t \in [\pi, +\infty[$, et comme $\frac{t^{a-1}}{e^t - 1}$ est sommable sur $[\pi, +\infty[$, on déduit du th. V.2.18 que $\int_\pi^{+\infty} \frac{t^s}{e^t - 1} \frac{dt}{t}$ est holomorphe en s , sur tout demi-plan de la forme $\operatorname{Re}(s) > a$, et donc aussi sur \mathbf{C} tout entier. Pour prouver que F_π est holomorphe sur \mathbf{C} , il suffit donc de vérifier que $\int_{C_\pi} \frac{(-z)^s}{e^z - 1} \frac{dz}{z}$ l'est, ce qui suit du cor. V.2.19.

• *Conclusion.* Comme F_π , Γ , \sin , \cos et ζ sont méromorphes sur \mathbf{C} tout entier, on en déduit que

$$\sin \pi s \cdot \Gamma(s) \cdot \zeta(s) = \pi F_\pi(s) = \cos \pi \frac{s-1}{2} \cdot (2\pi)^s \cdot \zeta(1-s),$$

pour tout s qui n'est pas un pôle d'une des fonctions ci-dessus. En utilisant la formule $\sin \pi s = -\sin \pi(s-1) = -2 \sin \pi \frac{s-1}{2} \cdot \cos \pi \frac{s-1}{2}$, on peut réécrire cette équation fonctionnelle sous la forme

$$\zeta(1-s) = -2(2\pi)^{-s} \cdot \sin \pi \frac{s-1}{2} \cdot \Gamma(s) \cdot \zeta(s),$$

et on obtient l'équation fonctionnelle du théorème en appliquant l'équation fonctionnelle ci-dessus à $1-s$ au lieu de s .

Exercice VII.3.8. — On note γ la constante d'Euler comme dans le th. VII.2.1.

(i) Montrer que $\Gamma'(1) = -\gamma$.

(ii) Montrer que $\zeta(s) - \frac{1}{s-1} = \sum_{n=1}^{+\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{t^s} \right) dt$, si $\operatorname{Re}(s) > 1$. En déduire $\lim_{s \rightarrow 1} \zeta(s) - \frac{1}{s-1} = \gamma$, puis $\frac{\zeta'(0)}{\zeta(0)} = \log 2\pi$ et $\zeta'(0) = -\frac{1}{2} \log 2\pi$.

4. Les zéros de la fonction ζ

Le théorème VII.3.6 permet une bonne localisation des zéros de la fonction ζ dans le plan complexe.

Corollaire VII.3.9. — *Les seuls zéros de la fonction ζ qui ne sont pas dans la bande verticale $0 \leq \operatorname{Re}(s) \leq 1$ sont des zéros simples aux entiers pairs < 0 .*

Démonstration. — $\zeta(s)$ est donnée (prop. VII.3.3) par le produit absolument convergent $\prod_{p \in \mathcal{P}} \frac{1}{1-p^{-s}}$ sur le demi-plan $\operatorname{Re}(s) > 1$. Comme aucun des termes du produit ne s'annule sur ce demi-plan, la fonction ζ ne s'annule pas sur le demi-plan $\operatorname{Re}(s) > 1$. D'autre part, si $\operatorname{Re}(s) < 0$, on a $\Gamma(1-s) \neq 0$, et, d'après ce qui précède, $\zeta(1-s) \neq 0$. L'équation fonctionnelle du th. VII.3.6 montre donc que les zéros de ζ sur le demi-plan $\operatorname{Re}(s) < 0$ sont les mêmes que ceux de $\sin \frac{\pi s}{2}$. Ceci permet de conclure.

On appelle *zéros triviaux* les zéros aux entiers pairs < 0 . La *bande critique* est la bande verticale $0 \leq \operatorname{Re}(s) \leq 1$; c'est un monde mystérieux, où il est difficile de voir clair. Riemann, à qui on doit la démonstration ci-dessus de l'équation fonctionnelle de la fonction ζ , le premier (en 1858), dans un mémoire qui est un des grands classiques des mathématiques, montré comment la répartition des zéros dans cette bande critique est reliée à la répartition des nombres premiers (cf. annexe A). Ce mémoire contient aussi *l'hypothèse de Riemann*, toujours non résolue à ce jour, et dont la tête est mise à prix pour un million de dollar, selon laquelle *tous les zéros de ζ dans la bande critique sont sur la droite $\operatorname{Re}(s) = \frac{1}{2}$.*

VII.4. Fonctions L de Dirichlet

1. Caractères de Dirichlet et Fonctions L de Dirichlet

Si D est un entier, un *caractère de Dirichlet modulo D* est un morphisme de groupes $\chi : (\mathbf{Z}/D\mathbf{Z})^* \rightarrow \mathbf{C}^*$. L'image d'un caractère de Dirichlet est un sous-groupe fini de \mathbf{C}^* , et donc est incluse dans le groupe des racines de l'unité. On note $\mathbf{1}_D$ le *caractère trivial*, défini par $\mathbf{1}_D(a) = 1$, quel que soit $a \in (\mathbf{Z}/D\mathbf{Z})^*$.

Si χ est un caractère de Dirichlet modulo D , on considère aussi souvent χ comme une fonction périodique sur \mathbf{Z} de période D , en composant χ avec la projection naturelle de \mathbf{Z} sur $\mathbf{Z}/D\mathbf{Z}$, et en étendant χ par 0 sur les entiers non premiers à D . La fonction $\chi \mapsto \chi(n)$ est alors strictement multiplicative : en effet, si m et n sont premiers à D , alors $\chi(mn) = \chi(m)\chi(n)$ par multiplicativité de la réduction modulo D et celle de χ , tandis que si m ou n n'est pas premier à D , alors mn non plus, et on a $\chi(mn) = 0 = \chi(m)\chi(n)$. La *fonction L de Dirichlet* attachée à χ est alors la série de Dirichlet $L(\chi, s) = \sum_{n \geq 1} \frac{\chi(n)}{n^s}$. Comme $|\chi(n)| = 1$, si $(n, D) = 1$, l'abscisse de convergence absolue de $L(\chi, s)$ est 1, et on a la proposition suivante.

Proposition VII.4.1. — Soit χ un caractère de Dirichlet modulo D .

(i) Si $\operatorname{Re}(s) > 1$, alors $L(\chi, s) = \prod_{p \in \mathcal{P}} \frac{1}{1 - \chi(p)p^{-s}}$, et le produit est uniformément convergent sur tout demi-plan de la forme $\operatorname{Re}(s) > c > 1$.

(ii) Si $\chi \neq \mathbf{1}_D$, l'abscisse de convergence de $L(\chi, s)$ est 0.

Démonstration. — Le (i) suit juste de la théorie générale (cf. prop. VII.3.1). Maintenant, si $\chi \neq \mathbf{1}_D$, il existe $a \in (\mathbf{Z}/D\mathbf{Z})^*$, tel que $\chi(a) \neq 1$. On a alors

$$\chi(a) \sum_{x \in (\mathbf{Z}/D\mathbf{Z})^*} \chi(x) = \sum_{x \in (\mathbf{Z}/D\mathbf{Z})^*} \chi(ax) = \sum_{x \in (\mathbf{Z}/D\mathbf{Z})^*} \chi(x),$$

puisque $x \mapsto ax$ est une bijection de $(\mathbf{Z}/D\mathbf{Z})^*$. On en déduit $\sum_{x \in (\mathbf{Z}/D\mathbf{Z})^*} \chi(x) = 0$, et donc $\sum_{n=kD+1}^{(k+1)D} \chi(n) = 0$. Ceci implique que la suite des sommes partielles $\sum_{k=1}^n \chi(k)$ est bornée (par D en valeur absolue), et permet d'utiliser le (i) du th. VII.1.5 pour démontrer le (ii).

2. Conducteur et sommes de Gauss

Si D' est un diviseur de D et χ est un caractère de Dirichlet modulo D' , on peut aussi voir χ comme un caractère de Dirichlet modulo D en composant χ avec la projection $(\mathbf{Z}/D\mathbf{Z})^* \rightarrow (\mathbf{Z}/D'\mathbf{Z})^*$. On dit que χ est *primitif*, si on ne peut pas trouver de diviseur D' de D distinct de D , tel que χ provienne d'un caractère modulo D' . On dit que χ est de *conducteur* D , si c'est un caractère de Dirichlet modulo D qui est primitif. Si χ est de conducteur D et si N est un multiple de D , on note χ_N le caractère modulo N obtenu en composant χ avec la projection $(\mathbf{Z}/N\mathbf{Z})^* \rightarrow (\mathbf{Z}/D\mathbf{Z})^*$.

Lemme VII.4.2. — On a $L(\chi_N, s) = L(\chi, s) \prod_{p|N} (1 - \chi(p)p^{-s})$.

Démonstration. — Il suffit d'utiliser la décomposition en produit de facteurs d'Euler.

Comme $1 - \chi(p)p^{-s}$ est une fonction holomorphe sur \mathbf{C} ayant tous ses zéros sur la droite $\operatorname{Re}(s) = 0$, on voit que l'étude des propriétés analytiques des fonctions L de Dirichlet se ramène à celle des fonctions L associées aux caractères primitifs.

Si χ est un caractère de Dirichlet modulo D , on note $\bar{\chi}$ le caractère de Dirichlet modulo D défini par $\bar{\chi}(n) = \overline{\chi(n)}$ si $n \in (\mathbf{Z}/D\mathbf{Z})^*$. Comme $\chi(n)$ est une racine de l'unité, on a aussi $\bar{\chi}(n) = \chi(n)^{-1}$.

Si D est un entier, si χ est un caractère de Dirichlet de conducteur D et si $n \in \mathbf{Z}$, on définit la somme de Gauss tordue $G(\chi, n)$ par la formule

$$G(\chi, n) = \sum_{a \bmod D} \chi(a) e^{2i\pi \frac{na}{D}},$$

et on pose $G(\chi) = G(\chi, 1)$.

Lemme VII.4.3. — Si $n \in \mathbf{N}$, alors $G(\chi, n) = \bar{\chi}(n)G(\chi)$

Démonstration. — Si $(n, D) = 1$, alors n est inversible dans $(\mathbf{Z}/D\mathbf{Z})^*$, ce qui permet d'écrire

$$G(\chi, n) = \sum_{a \bmod D} \chi(a) e^{2i\pi \frac{na}{D}} = \bar{\chi}(n) \sum_{an \bmod D} \chi(an) e^{2i\pi \frac{na}{D}} = \bar{\chi}(n) G(\chi).$$

Si $(n, D) = d > 1$, on peut écrire $D = dD'$ et $n = dn'$. Comme χ est de conducteur D , il existe $b \equiv 1 \bmod D/d$ tel que $\chi(b) \neq 1$ (sinon χ serait de conducteur divisant D'). On a alors

$$e^{2i\pi \frac{nab}{D}} = e^{2i\pi \frac{na}{D}} e^{2i\pi \frac{n(b-1)a}{D}} = e^{2i\pi \frac{na}{D}},$$

puisque n est divisible par d et $b - 1$ par D/d . On en déduit que

$$\chi(b) G(\chi, n) = \sum_{a \bmod D} \chi(ab) e^{2i\pi \frac{nab}{D}} = \sum_{a \bmod D} \chi(a) e^{2i\pi \frac{na}{D}} = G(\chi, n),$$

et donc, comme $\chi(b) \neq 1$, que $G(\chi, n) = 0 = \bar{\chi}(n) G(\chi)$. Ceci permet de conclure.

Théorème VII.4.4. — Si χ est un caractère de Dirichlet de conducteur $D \neq 1$, alors $L(\chi, s)$ admet un prolongement analytique à \mathbf{C} tout entier. De plus, $L(\chi, s) = M(f_\chi, s)$, où $f_\chi : \mathbf{R}_+ \rightarrow \mathbf{C}$ est donnée par la formule

$$f_\chi(t) = \frac{1}{G(\bar{\chi})} \sum_{b=1}^{D-1} \frac{\bar{\chi}(b)}{e^{-\frac{2i\pi b}{D}} e^t - 1}.$$

Démonstration. — Il résulte du lemme VII.2.6, que $L(\chi, s) = M(f, s)$, où f est définie par $f(t) = \sum_{n=1}^{+\infty} \chi(n) e^{-nt}$. Si on utilise l'identité⁽⁹⁾ $\chi(n) = \frac{G(\bar{\chi}, n)}{G(\bar{\chi})}$ du lemme VII.4.3, on obtient

$$f(t) = \sum_{n=1}^{+\infty} \left(\frac{1}{G(\bar{\chi})} \sum_{b \bmod D} \bar{\chi}(b) e^{2i\pi \frac{nb}{D}} \right) e^{-nt} = \frac{1}{G(\bar{\chi})} \sum_{b=1}^{D-1} \frac{\bar{\chi}(b)}{e^{-\frac{2i\pi b}{D}} e^t - 1} = f_\chi(t).$$

(La dernière égalité venant de ce que $\chi(n) = 0$ si n est un multiple de D , ce qui fait que $0 \bmod D$ ne contribue pas à la somme). Maintenant, f_χ est à décroissance rapide à l'infini, et $e^{2i\pi \frac{nb}{D}} \neq 1$, si $1 \leq b \leq D-1$, ce qui fait que f_χ est \mathcal{C}^∞ sur \mathbf{R}_+ . On conclut en utilisant la prop. VII.2.8.

3. Équation fonctionnelle des fonctions L de Dirichlet

Théorème VII.4.5. — Si χ est un caractère de Dirichlet de conducteur $D \neq 1$, alors $L(\chi, s)$ vérifie l'équation fonctionnelle

$$L(\chi, s) = \begin{cases} 2 \cdot G(\chi) \cdot D^{-s} \cdot (2\pi)^{s-1} \cdot \Gamma(1-s) \cdot \sin \frac{\pi s}{2} \cdot L(\bar{\chi}, 1-s) & \text{si } \chi(-1) = 1, \\ -2i \cdot G(\chi) \cdot D^{-s} \cdot (2\pi)^{s-1} \cdot \Gamma(1-s) \cdot \cos \frac{\pi s}{2} \cdot L(\bar{\chi}, 1-s) & \text{si } \chi(-1) = -1. \end{cases}$$

⁽⁹⁾On a quand même besoin de vérifier que $G(\bar{\chi}) \neq 0$ (cf. ex. VII.4.7).

Démonstration. — La démonstration est très semblable à celle de l'équation fonctionnelle de la fonction ζ , et nous reprenons les notations de cette dernière en indiquant les points où l'argument diffère. Soit $F_c(\chi, s) = \frac{1}{2i\pi} \int_{\gamma_c} f_\chi(z)(-z)^s \frac{dz}{z}$. Comme $f_\chi(z)$ est à décroissance rapide à l'infini, la fonction $F_c(s)$ est holomorphe sur \mathbf{C} pour tout c qui n'est pas de la forme $\frac{2\pi b}{D} + 2\pi k$, avec $k \in \mathbf{N}$ et $b \leq D$ premier à D (pour éviter les pôles de f_χ). Comme f_χ n'a pas de pôle à l'intérieur du carré de sommets $\frac{2\pi}{D}(\pm 1 \pm i)$, on a $F_c(\chi, s) = F_{\pi/D}(\chi, s)$, quel que soit $c \in]0, \frac{2\pi}{D}[$. En faisant tendre c vers 0, on en déduit, si $\operatorname{Re}(s) > 1$, la formule

$$F_{\pi/D}(\chi, s) = \frac{1}{2i\pi} \left(e^{-i\pi s} \int_{+\infty}^0 f_\chi(t)t^s \frac{dt}{t} + e^{i\pi s} \int_0^{+\infty} f_\chi(t)t^s \frac{dt}{t} \right) = \frac{\sin \pi s}{\pi} \cdot \Gamma(s) \cdot L(\chi, s).$$

Maintenant, quand N tend vers $+\infty$, la fonction $F_{N\pi}(\chi, s)$ tend vers 0 quand $\operatorname{Re}(s) < 0$. La différence entre $F_{\pi/D}(s)$ et $F_{N\pi}(s)$ peut se calculer grâce au théorème des résidus. La fonction $f_\chi(z) \frac{(-z)^s}{z}$ a des pôles en les $z = \pm \frac{2i\pi k}{D}$, avec $1 \leq k \leq ND - 1$ dans le contour délimité par la différence entre $\gamma_{N\pi}$ et $\gamma_{\pi/D}$. Si $1 \leq k \leq ND - 1$, on a

$$\begin{aligned} \operatorname{Res}\left(f_\chi(z) \frac{(-z)^s}{z}, \frac{2i\pi k}{D}\right) &= \frac{\bar{\chi}(k)}{G(\bar{\chi})} \frac{(-2i\pi k/D)^s}{(2i\pi k/D)} = -\frac{\bar{\chi}(k)}{G(\bar{\chi})} \cdot \left(\frac{2\pi k}{D}\right)^{s-1} e^{-i\pi \frac{s-1}{2}} \\ \operatorname{Res}\left(f_\chi(z) \frac{(-z)^s}{z}, -\frac{2i\pi k}{D}\right) &= \frac{\bar{\chi}(-k)}{G(\bar{\chi})} \frac{(2i\pi k/D)^s}{(-2i\pi k/D)} = -\frac{\bar{\chi}(-k)}{G(\bar{\chi})} \cdot \left(\frac{2\pi k}{D}\right)^{s-1} e^{i\pi \frac{s-1}{2}} \end{aligned}$$

On obtient donc

$$\frac{1}{2i\pi} (F_{N\pi}(\chi, s) - F_{\pi/D}(\chi, s)) = \frac{-1}{G(\bar{\chi})} \sum_{k=1}^{ND-1} \left(\frac{2\pi k}{D}\right)^{s-1} (\bar{\chi}(k) e^{-i\pi \frac{s-1}{2}} + \bar{\chi}(-k) e^{i\pi \frac{s-1}{2}}).$$

En faisant tendre N vers $+\infty$, on en déduit, si $\operatorname{Re}(s) < 0$, les formules

$$F_{\pi/D}(\chi, s) = \frac{1}{G(\bar{\chi})} \begin{cases} (2i\pi) \cdot \left(\frac{2\pi}{D}\right)^{s-1} \cdot (2 \cos \pi \frac{s-1}{2}) \cdot L(\bar{\chi}, 1-s) & \text{si } \chi(-1) = 1, \\ (2i\pi) \cdot \left(\frac{2\pi}{D}\right)^{s-1} \cdot (-2i \sin \pi \frac{s-1}{2}) \cdot L(\bar{\chi}, 1-s) & \text{si } \chi(-1) = -1. \end{cases}$$

On en tire l'équation fonctionnelle

$$\sin \pi s \cdot \Gamma(s) \cdot L(\chi, s) = \frac{1}{G(\bar{\chi})} \begin{cases} (2\pi)^s \cdot D^{1-s} \cdot \cos \pi \frac{s-1}{2} \cdot L(\bar{\chi}, 1-s) & \text{si } \chi(-1) = 1, \\ -i(2\pi)^s \cdot D^{1-s} \cdot \sin \pi \frac{s-1}{2} \cdot L(\bar{\chi}, 1-s) & \text{si } \chi(-1) = -1, \end{cases}$$

qui peut aussi se mettre sous la forme

$$L(\bar{\chi}, 1-s) = \begin{cases} 2G(\bar{\chi}) \cdot (2\pi)^{-s} \cdot D^{s-1} \cdot \sin \pi \frac{s-1}{2} \cdot \Gamma(s) \cdot L(\chi, s) & \text{si } \chi(-1) = 1, \\ -2iG(\bar{\chi}) \cdot (2\pi)^{-s} \cdot D^{s-1} \cdot \cos \pi \frac{s-1}{2} \cdot \Gamma(s) \cdot L(\chi, s) & \text{si } \chi(-1) = -1. \end{cases}$$

L'équation fonctionnelle du théorème s'obtient en appliquant l'équation fonctionnelle ci-dessus à $\bar{\chi}$ au lieu de χ et $1-s$ au lieu de s .

Remarque VII.4.6. — (i) Si χ est un caractère de Dirichlet de conducteur $D \neq 1$, posons

$$w(\chi) = \begin{cases} \frac{G(\chi)}{\sqrt{D}} & \text{si } \chi(-1) = 1, \\ \frac{G(\chi)}{i\sqrt{D}} & \text{si } \chi(-1) = -1, \end{cases}$$

$$\Lambda(\chi, s) = \begin{cases} \Gamma\left(\frac{s}{2}\right) \cdot \left(\frac{D}{\pi}\right)^{s/2} \cdot L(\chi, s) & \text{si } \chi(-1) = 1, \\ \Gamma\left(\frac{s+1}{2}\right) \cdot \left(\frac{D}{\pi}\right)^{(s+1)/2} \cdot L(\chi, s) & \text{si } \chi(-1) = -1. \end{cases}$$

Un petit calcul permet de déduire de l'équation fonctionnelle de $L(\chi, s)$ que $\Lambda(\chi, s)$ vérifie l'équation fonctionnelle

$$\Lambda(\chi, s) = w(\chi) \Lambda(\bar{\chi}, 1-s).$$

(ii) La fonction $L(\chi, s)$ est holomorphe sur \mathbf{C} tout entier et ne s'annule pas sur le demi-plan $\operatorname{Re}(s) > 1$ sur lequel elle est donnée (prop. VII.4.1) par un produit absolument convergent dont aucun des termes ne s'annule. Elle ne s'annule pas non plus sur la droite $\operatorname{Re}(s) = 1$ (cf. ex. A.4.4), ce qui est nettement plus profond. L'équation fonctionnelle du th. VII.4.5 montre donc que, en dehors de la bande $0 < \operatorname{Re}(s) < 1$, les seuls zéros de $L(\chi, s)$ sont des zéros (dit triviaux) aux entiers négatifs pairs (resp. impairs), si $\chi(-1) = 1$ (resp. $\chi(-1) = -1$). On conjecture (*hypothèse de Riemann généralisée*, GRH en abréviation anglaise) que tous les zéros de $L(\chi, s)$ dans la bande $0 < \operatorname{Re}(s) < 1$ se trouvent sur la droite $\operatorname{Re}(s) = \frac{1}{2}$. Ceci aurait des implications profondes sur la répartition des nombres premiers entre les différentes progressions arithmétiques⁽¹⁰⁾.

Exercice VII.4.7. — Soit χ un caractère de conducteur $D \neq 1$. Montrer que :

- (i) $\overline{G(\chi)} = \chi(-1)G(\bar{\chi})$;
- (ii) $G(\chi)G(\bar{\chi}) = \chi(-1)D$;
- (iii) $|w(\chi)| = 1$.

VII.5. Autres exemples

1. La fonction de Moebius

Soit μ la *fonction de Moebius*. Elle est définie par $\mu(n) = 0$ si n est divisible par le carré d'au moins un nombre premier, et $\mu(n) = (-1)^r$, si $n = p_1 \cdots p_r$, où les p_i sont des nombres premiers distincts. C'est une fonction multiplicative, et on a

$$\zeta(s)^{-1} = \prod_{p \in \mathcal{P}} (1 - p^{-s}) = \sum_{n \in \mathbf{N}} \mu(n) n^{-s} = L(\mu, s).$$

On en déduit que $L(\mu, s)$ a un prolongement méromorphe à tout le plan complexe. Il est facile de voir que son abscisse de convergence absolue σ_{abs} est 1, mais son abscisse de convergence σ_{conv} est inconnue. On conjecture que $\sigma_{\text{conv}} = \frac{1}{2}$, mais c'est équivalent à l'hypothèse de Riemann (cf. n° 2 du § A.5).

⁽¹⁰⁾En particulier, sur la taille du plus petit nombre premier dans une progression arithmétique, ce qui intervient naturellement dans beaucoup de questions ; par exemple en cryptographie.

Exercice VII.5.1. — (Formule d'inversion de Moebius)

- (i) Montrer que $\sum_{d|n} \mu(d) = 0$, si $n \geq 2$.
- (ii) Soit $F : [0, +\infty[\rightarrow \mathbf{C}$, vérifiant $F(x) = 0$, si $x < 1$, et soit $G : [0, +\infty[\rightarrow \mathbf{C}$ définie par $G(x) = \sum_{n=1}^{+\infty} F(\frac{x}{n})$. Montrer que $G(x) = 0$, si $x < 1$, et que $F(x) = \sum_{n=1}^{+\infty} \mu(n)G(\frac{x}{n})$.
- (iii) Montrer que $\sum_{n=1}^{+\infty} \mu(n)[\frac{x}{n}] = 1$, si $x \geq 1$.

Exercice VII.5.2. — (Critère de Báez-Duarte)

Soit E l'espace de Hilbert, séparé de l'espace $\mathcal{L}^2([1, +\infty[, \frac{dt}{t^2})$ des $\phi : [1, +\infty[\rightarrow \mathbf{C}$ telles que $t \mapsto t^{-1}\phi(t)$ soit de carré sommable, muni du produit scalaire $\langle f, g \rangle = \int_0^{+\infty} \overline{f(t)}g(t) \frac{dt}{t^2}$.

- (i) Vérifier que $t \mapsto t^{1-s}$ appartient à E , si $\operatorname{Re}(s) > \frac{1}{2}$, et qu'il en est de même de ϕ_n , définie par $\phi_n(t) = [\frac{t}{n}] - \frac{[t]}{n}$, si n est un entier ≥ 2 .
- (ii) Montrer que $\int_1^{+\infty} \{t\} t^{-s-1} dt = \frac{1}{s-1} - \frac{\zeta(s)}{s}$, si $\operatorname{Re}(s) > 1$.
- (iii) Montrer que $s \mapsto \int_1^{+\infty} \{t\} t^{-s-1} dt$ est holomorphe sur $\operatorname{Re}(s) > 0$. En déduire que $\int_0^{+\infty} \{t\} t^{-s-1} dt = -\frac{\zeta(s)}{s}$, si $0 < \operatorname{Re}(s) < 1$.
- (iv) Montrer que $\langle t^{1-s}, \phi_n \rangle = (\frac{1}{n^s} - \frac{1}{n}) \frac{\zeta(s)}{s}$, si $n \geq 2$, et si $\operatorname{Re}(s) > \frac{1}{2}$.
- (v) En déduire que si l'adhérence dans E du sous-espace engendré par les ϕ_n , pour $n \geq 2$, contient les fonctions constantes sur $[1, +\infty[$, alors l'hypothèse de Riemann est vraie⁽¹¹⁾.

2. La fonction τ de Ramanujan

La fonction τ de Ramanujan. Elle est définie par l'identité

$$q \prod_{n=1}^{+\infty} (1 - q^n)^{24} = \sum_{n=1}^{+\infty} \tau(n) q^n.$$

S. Ramanujan a fait deux conjectures à son sujet. La première, démontrée peu après par L. Mordell (1917) peut s'énoncer sous la forme

$$L(\tau, s) = \prod_{p \in \mathcal{P}} \frac{1}{1 - \tau(p)p^{-s} + p^{11-2s}}.$$

En particulier, τ est une fonction multiplicative! Sa démonstration repose sur le fait que $\Delta(z) = q \prod_{n=1}^{+\infty} (1 - q^n)^{24}$, avec $q = e^{2i\pi z}$, est une forme modulaire de poids 12 pour $\mathbf{SL}_2(\mathbf{Z})$ (cf. ex. VII.6.3 et VII.6.11).

La seconde conjecture de Ramanujan affirme que les pôles du facteur d'Euler en p de la fonction $L(\tau, s)$ sont tous sur la droite $\operatorname{Re}(s) = \frac{11}{2}$. Elle peut aussi s'énoncer sous la forme $|\tau(p)| \leq 2p^{11/2}$, si $p \in \mathcal{P}$. Il a fallu attendre 1973 pour que P. Deligne démontre la conjecture de Ramanujan (généralisée sous le nom de conjecture de Ramanujan-Petersson aux formes modulaires quelconques) comme conséquence de sa démonstration (qui lui a valu la médaille Fields en 1978) de l'hypothèse de Riemann pour les variétés sur les corps finis, conjecturée par A. Weil vers la fin des années 1940. La démonstration de Deligne est

⁽¹¹⁾On peut montrer, mais c'est plus difficile, que si l'hypothèse de Riemann est vraie, alors l'adhérence dans E du sous-espace engendré par les ϕ_n , pour $n \geq 2$, contient les fonctions constantes sur $[1, +\infty[$. Il s'agit d'une variante du critère de Nyman-Beurling (1955) due à Báez-Duarte (2003). Il est à noter que la question (iii) de l'ex. VII.5.1 et la formule $\sum_{n=1}^{+\infty} \frac{\mu(n)}{n} = \zeta(1)^{-1} = 0$, permettent de montrer que $\sum_{n=1}^{+\infty} \frac{\mu(n)}{n} \phi_n(t) = -1$, si $t \geq 1$, mais la série ne converge probablement pas dans E . La convergence de la série $\sum_{n=1}^{+\infty} \frac{\mu(n)}{n}$ est plus délicate qu'il n'y paraît; elle est plus ou moins équivalente à la non annulation de la fonction ζ sur la droite $\operatorname{Re}(s) = 1$.

l'aboutissement d'un énorme programme mis sur pied par A. Grothendieck entre 1958 et 1964, et qui a totalement révolutionné la géométrie algébrique.

VII.6. Exercices

Les exercices qui suivent explorent certaines propriétés des formes modulaires, qui possèdent tellement de symétries qu'elles ne devraient pas exister, mais se retrouvent, un peu inexplicablement, jouer un rôle dans les questions les plus variées (cf. note 3 du chap. I, ou note 1 de l'annexe B, par exemple). Ils sont l'occasion d'utiliser les résultats des chapitres précédents en démontrant une série de jolis résultats comme le théorème des 4 carrés.

Soit $\mathcal{H} = \{z, \operatorname{Im}(z) > 0\}$ le demi-plan de Poincaré. On rappelle (ex. VI.2.7) que si $f : \mathcal{H} \rightarrow \mathbf{C}$ est holomorphe et périodique de période 1, alors il existe une suite $(a_n(f))_{n \in \mathbf{Z}}$ de nombres complexes telle que l'on ait $f(z) = \sum_{n \in \mathbf{Z}} a_n(f) e^{2i\pi n z}$, quel que soit $z \in \mathcal{H}$. Il est d'usage de poser $e^{2i\pi z} = q$ et d'appeler q -développement de f la série $\sum_{n \in \mathbf{Z}} a_n(f) q^n$. On définit alors $v_\infty(f) \in \mathbf{Z} \cup \{\pm\infty\}$ comme l'inf. de l'ensemble des n tels que $a_n(f) \neq 0$ (en particulier, $v_\infty(f) = +\infty$ si et seulement si $f = 0$).

Si $k \in \mathbf{Z}$, une *fonction modulaire de poids k* est une fonction holomorphe f sur \mathcal{H} , et qui vérifie :

- $f(z+1) = f(z)$, quel que soit $z \in \mathcal{H}$,
- $f(z) = z^{-k} f(-1/z)$, quel que soit $z \in \mathcal{H}$,
- $v_\infty(f) > -\infty$.

Si $v_\infty(f) \geq 0$, on dit que f est une *forme modulaire de poids k* , et si $v_\infty(f) > 0$, on dit que f est parabolique (ou cuspidale).

Soit Ω l'ouvert $\{z \in \mathcal{H}, |z| > 1 \text{ et } -\frac{1}{2} < \operatorname{Re}(z) < \frac{1}{2}\}$. Si $a, b \in \mathcal{H}$ vérifient $|a| = |b| = 1$, et si C^+ est le demi-cercle de centre 0 et rayon 1 contenu dans \mathcal{H} , on note $A(a, b)$ l'arc de C^+ allant de a à b . Soit $\alpha = e^{i\pi/3}$. Le bord $\partial\Omega$ de Ω est alors la réunion des demi-droites verticales $[\alpha, \alpha + i\infty)$ et $[\alpha^2, \alpha^2 + i\infty)$, et de l'arc de cercle $A(\alpha^2, \alpha)$. On note D la réunion de Ω , de la demi-droite $[\alpha, \alpha + i\infty)$ et de l'arc de cercle $A(i, \alpha)$.

Exercice VII.6.1. — (formule $\frac{k}{12}$) Soit f une forme modulaire de poids k non nulle. Le but de cet exercice est de prouver la formule suivante.

$$v_\infty(f) + \frac{1}{2}v_i(f) + \frac{1}{3}v_\alpha(f) + \sum_{z \in D - \{i, \alpha\}} v_z(f) = \frac{k}{12}.$$

(i) On suppose que f ne s'annule pas sur $\partial\Omega$. Si $T \geq 2$, soit γ_T le lacet composé des segments $[\alpha, \alpha + iT]$, $[\alpha + iT, \alpha^2 + iT]$, $[\alpha^2 + iT, \alpha^2]$, et de l'arc de cercle $A(\alpha^2, \alpha)$. On note $\frac{df}{f}$ la 1-forme $\frac{f'(z)}{f(z)} dz$.

- Montrer que $\frac{1}{2i\pi} \int_{[\alpha + iT, \alpha^2 + iT]} \frac{df}{f}$ tend vers $-v_\infty(f)$ quand $T \rightarrow +\infty$.
- Montrer que $\int_{[\alpha, \alpha + iT]} \frac{df}{f} + \int_{[\alpha^2 + iT, \alpha^2]} \frac{df}{f} = 0$.
- Montrer que $\int_{A(\alpha^2, i)} \frac{df}{f} = -\int_{A(i, \alpha)} \left(\frac{df}{f} + \frac{k}{z} dz\right)$. En déduire que $\frac{1}{2i\pi} \int_{A(\alpha^2, \alpha)} \frac{df}{f} = \frac{k}{12}$.
- Montrer que $v_\infty(f) + \sum_{z \in \Omega} v_z(f) = \frac{k}{12}$.

(ii) Montrer que, si on ne suppose pas que f ne s'annule pas sur $\partial\Omega$, alors

$$v_\infty(f) + \sum_{z \in \Omega} v_z(f) + \frac{1}{2} \sum_{z \in \partial\Omega - \{\alpha, \alpha^2\}} v_z(f) + \frac{1}{6}(v_\alpha(f) + v_{\alpha^2}(f)) = \frac{k}{12}.$$

(Si f s'annule en $z \in \partial\Omega$, modifier le chemin γ_T au voisinage de z en le remplaçant par un arc de cercle, à l'intérieur de Ω , de centre z et de rayon tendant vers 0.)

(iii) Conclure.

La « formule $\frac{k}{12}$ » a beaucoup de conséquences intéressantes. En voici quelques-unes ; d'autres se trouvent dans les ex. VII.6.8 et VII.6.10.

Exercice VII.6.2. — (Dimension des espaces de formes modulaires)

- (i) Montrer qu'une forme modulaire de poids 0 est constante.
- (ii) Montrer qu'il n'y a pas⁽¹²⁾ de forme modulaire de poids 2 ou de poids impair.
- (iii) Montrer que l'ensemble M_k des formes modulaires de poids k est un espace vectoriel et que $f \mapsto (a_n(f))_{0 \leq n \leq \frac{k}{12}}$ est une application linéaire injective de M_k dans $\mathbf{C}^{d(k)}$, avec $d(k) = 1 + [\frac{k}{12}]$. En déduire que M_k est de dimension finie et que $\dim M_k \leq 1 + [\frac{k}{12}]$.

Exercice VII.6.3. — (fonction L d'une forme modulaire)

Soit $Y = \{z \in \mathbf{C}, \operatorname{Im}(z) \geq \frac{1}{2}, |\operatorname{Re}(z)| \leq \frac{1}{2}\}$. Soit $f = \sum_{n=1}^{+\infty} a_n q^n$ une forme modulaire parabolique de poids $2k$, et soit $F(z) = \operatorname{Im}(z)^k |f(z)|$.

- (i) Montrer qu'il existe $M > 0$ tel que $|F(z)| \leq M$, quel que soit $z \in Y$.
- (ii) Montrer que $F(z+1) = F(z)$ et $F(-1/z) = F(z)$.
- (iii) Montrer que, si $|\operatorname{Re}(z_0)| \leq \frac{1}{2}$ et si $0 < \operatorname{Im}(z_0) \leq \frac{1}{2}$, alors il existe z_1 vérifiant $|\operatorname{Re}(z_1)| \leq \frac{1}{2}$, $\operatorname{Im}(z_1) \geq 2\operatorname{Im}(z_0)$, et $F(z_1) = F(z_0)$. En déduire que $F(z) \leq M$, quel que soit z vérifiant $|\operatorname{Re}(z)| \leq \frac{1}{2}$.
- (iv) Montrer que $a_n = \int_{-1/2}^{1/2} f(x + \frac{i}{n}) e^{-2i\pi(nx+i)} dx$. En déduire que $|a_n| \leq e^{2\pi} M n^k$.
- (v) Soit $L(f, s) = \sum_{n=1}^{+\infty} \frac{a_n}{n^s}$. Montrer que $L(f, s)$ a une abscisse de convergence finie, possède un prolongement analytique à \mathbf{C} tout entier, que $L(f, -n) = 0$, si $n \in \mathbf{N}$, et que $\Lambda(f, s) = \frac{\Gamma(s)}{(2\pi)^s} L(f, s)$ vérifie l'équation fonctionnelle $\Lambda(f, s) = (-1)^k \Lambda(f, 2k - s)$. (On s'intéressera à $\int_0^{+\infty} f(iy) y^s \frac{dy}{y}$.)

L'exercice suivant montre que l'on n'est pas en train de faire la théorie de l'ensemble vide (dont les éléments ont, comme chacun sait, beaucoup de propriétés mirifiques...).

Exercice VII.6.4. — (Séries d'Eisenstein)

- (i) Montrer que $|mz + n| \geq \inf(y, \frac{y}{|z|}) \sup(|m|, |n|)$, si $z = x + iy \in \mathcal{H}$ et $m, n \in \mathbf{Z}$.
- (ii) Montrer que, si $k \geq 3$, la série $\sum_{(m,n) \in \mathbf{Z}^2 - (0,0)} \frac{1}{(mz+n)^k}$ converge uniformément sur tout compact de \mathcal{H} .
- (iii) Si $z \in \mathcal{H}$, soit $G_k(z) = \frac{\Gamma(k)}{2(-2i\pi)^k} \sum_{(m,n) \in \mathbf{Z}^2 - (0,0)} \frac{1}{(mz+n)^k}$. Montrer que G_k est une fonction holomorphe sur \mathcal{H} et que

$$G_k\left(\frac{az+b}{cz+d}\right) = (cz+d)^k G_k(z), \quad \text{quel que soit } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbf{Z}).$$

- (iv) Montrer que, si k est impair, alors $G_k = 0$, et que si k est pair, G_k est une forme modulaire de poids k non nulle.

Exercice VII.6.5. — (q -développement des séries d'Eisenstein)

Le but de cet exercice est de prouver que, si k est un entier pair ≥ 3 , le q -développement de G_k est donné par la formule suivante⁽¹³⁾ :

$$G_k = \frac{\Gamma(k)}{(-2i\pi)^k} \zeta(k) + \sum_{n=1}^{+\infty} \sigma_{k-1}(n) q^n, \quad \text{où } \sigma_{k-1}(n) = \sum_{d|n, d \geq 1} d^{k-1}.$$

⁽¹²⁾Ce résultat intervient de manière cruciale dans la démonstration de Wiles du théorème de Fermat.

⁽¹³⁾On remarquera que tous les termes de ce q -développement sont visiblement, à l'exception du terme constant, des nombres rationnels. On peut utiliser ceci pour (modulo un certain travail) en déduire qu'il en est de même du terme constant, ce qui permet de donner une démonstration du résultat d'Euler sur les valeurs aux entiers pairs de la fonction zêta.

Soit k un entier ≥ 2 .

(i) Montrer que la série $\sum_{n \in \mathbf{Z}} \frac{1}{(z+n)^k}$ converge uniformément sur tout compact de \mathcal{H} . En déduire que sa somme $A_k(z)$ est une fonction holomorphe sur \mathcal{H} .

(ii) Montrer que A_k est périodique de période 1. Déduire du (i) que $x \mapsto A_k(x + iy)$ est somme de sa série de Fourier $\sum_{n \in \mathbf{Z}} a_n(y) e^{2i\pi nx}$ pour tout $x \in \mathbf{R}$.

(iii) Calculer $a_n(y)$ par la formule des résidus.

(iv) En déduire que $\frac{\Gamma(k)}{(-2i\pi)^k} A_k(z) = \sum_{n \geq 1} n^{k-1} e^{2i\pi nz}$.

(v) Montrer que $G_k(z) = \frac{\Gamma(k)}{(-2i\pi)^k} (\zeta(k) + \sum_{m \geq 1} A_k(mz))$, et en déduire le résultat.

Exercice VII.6.6. — (La fonction thêta de Jacobi)

(i) Montrer que $\theta(z) = \sum_{n \in \mathbf{Z}} e^{i\pi n^2 z}$ converge normalement sur tout compact de \mathcal{H} . En déduire que θ est holomorphe sur \mathcal{H} .

(ii) Montrer, en utilisant le fait que la transformée de Fourier de $e^{-\pi t^2}$ est $e^{-\pi x^2}$ (cf. ex. IV.3.28), que $\theta(iu) = \frac{1}{\sqrt{u}} \theta(\frac{i}{u})$, si $u \in \mathbf{R}_+^*$.

(iii) En déduire que l'on a $\theta(z) = \sqrt{\frac{i}{z}} \theta(\frac{-1}{z})$, si $z \in \mathcal{H}$ (où \sqrt{z} est la racine carrée de z holomorphe sur $\mathbf{C} - \mathbf{R}_-$, et valant 1 en 1).

(iv) On pose $\xi(s) = \frac{\Gamma(s/2)}{\pi^{s/2}} \zeta(s)$, où ζ est la fonction zêta de Riemann. Montrer que, si $\operatorname{Re}(s) > 1$, alors $\xi(s) = \frac{1}{2} \int_0^{+\infty} (\theta(iy) - 1) y^{s/2} \frac{dy}{y}$.

(v) En déduire que, si $\operatorname{Re}(s) > 1$, alors

$$\xi(s) = -\frac{1}{s} + \frac{1}{s-1} + \frac{1}{2} \int_1^{+\infty} (\theta(iy) - 1) (y^{s/2} + y^{(1-s)/2}) \frac{dy}{y},$$

puis que ξ admet un prolongement méromorphe à \mathbf{C} , holomorphe en dehors de pôles simples en $s = 0$ et $s = 1$, et vérifie l'équation fonctionnelle $\xi(s) = \xi(1-s)$.

(vi) Montrer que $\xi(s) = O(\frac{1}{\operatorname{Im}(s)})$ dans toute bande verticale de largeur finie.

Exercice VII.6.7. — (La série d'Eisenstein de poids 2)

Soit G_2 la fonction holomorphe sur \mathcal{H} , périodique de période 1, dont le q -développement est donné par⁽¹⁴⁾

$$G_2 = \frac{-1}{24} + \sum_{n=1}^{+\infty} \sigma(n) q^n, \quad \text{où } \sigma(n) = \sum_{d|n, d \geq 1} d.$$

Nous allons montrer⁽¹⁵⁾ que G_2 est presque modulaire; plus précisément, G_2 vérifie l'équation fonctionnelle $z^{-2} G_2(-1/z) = G_2(z) - \frac{1}{4i\pi z}$.

On rappelle (ex. VI.2.24) que $\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \Gamma(s) y^{-s} = e^{-y}$, si $y \in \mathbf{R}_+^*$ et $c > 0$, l'intégrale étant absolument convergente.

⁽¹⁴⁾ Si on reprend l'exercice VII.6.5, et qu'on utilise la formule $\zeta(2) = \frac{\pi^2}{6}$, on voit que $G_2(z)$ est la somme des $\frac{\Gamma(2)}{(-2i\pi)^2 (mz+n)^2}$ en sommant d'abord sur n puis sur m .

⁽¹⁵⁾ La méthode de l'exercice consiste à déduire une équation fonctionnelle reliant $\varphi(-1/z)$ et $\varphi(z)$, à partir d'une équation fonctionnelle reliant $\Lambda(2-s)$ et $\Lambda(s)$, où $\Lambda(s) = \int_0^{+\infty} \varphi(iy) y^s \frac{dy}{y}$. Cette méthode a beaucoup d'autres applications. On peut par exemple montrer que la fonction dont le q -développement est celui de G_k est une forme modulaire, sans passer par la construction de G_k et le calcul de son q -développement.

(i) Montrer que, si $L(\mathbf{a}, s) = \sum_{n=1}^{+\infty} \frac{a_n}{n^s}$ a une abscisse de convergence finie, alors $F_{\mathbf{a}}(z) = \sum_{n=1}^{+\infty} a_n e^{2i\pi n z}$ est holomorphe sur \mathbf{C} , et si $c > \sup(0, \sigma_{\text{abs}})$, alors

$$\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(s)}{(2\pi)^s} L(\mathbf{a}, s) y^{-s} ds = F_{\mathbf{a}}(iy), \quad \text{si } y \in \mathbf{R}_+^*.$$

(ii) Soit $\sigma(n) = \sum_{d|n, d \geq 1} d$. Montrer que $L(\sigma, s) = \zeta(s)\zeta(s-1)$.

(iii) Soit $H(s) = \frac{\Gamma(s)}{(2\pi)^s} \zeta(s)\zeta(s-1)$. Montrer que $H(s) = \frac{s-1}{4\pi} \xi(s)\xi(s-1)$. En déduire que $H(2-s) = -H(s)$, que $H(s)$ tend vers 0 à l'infini dans toute bande verticale de largeur finie, et que H est holomorphe sur \mathbf{C} en dehors de pôles simples en 0, 1 et 2 de résidus respectifs $\frac{1}{24}$, $\frac{-1}{4\pi}$ et $\frac{1}{24}$. (On pourra utiliser les formules $\zeta(0) = -\frac{1}{2}$, $\zeta(-1) = -\frac{1}{12}$ et $\zeta(2) = \frac{\pi^2}{6}$.)

(iv) Montrer que, si $y > 0$, alors $G_2(iy) + y^{-2}G_2(i/y) = \frac{1}{24} - \frac{1}{4\pi y} + \frac{1}{24y^2}$. (On intégrera $H(s)y^{-s}$ sur le rectangle de sommets $3-iT$, $3+iT$, $-1+iT$ et $-1-iT$.)

(v) Conclure.

Exercice VII.6.8. — Cet exercice est un préliminaire pour le théorème des 4 carrés. Son but est de démontrer que si $(a_n)_{n \geq 2}$ est une suite de nombres complexes vérifiant :

- il existe $c \in \mathbf{N}$ tel que $a_n = O(n^c)$,
- $F(z) = \sum_{n=2}^{+\infty} a_n e^{i\pi n z}$ vérifie l'équation fonctionnelle $z^{-4}F(-1/z) = F(z)$, si $z \in \mathcal{H}$,

alors $a_n = 0$ pour tout $n \geq 2$.

(i) Montrer que $F(z) = O(e^{-2\pi \text{Im}(z)})$ dans $Y = \{z \in \mathbf{C}, |\text{Re}(z)| \leq 1, \text{Im}(z) \geq 1\}$.

(ii) Montrer que $F(1-1/z) = O(\text{Im}(z)^{c+1})$ dans Y .

(iii) Soit k un entier pair. Si $f : \mathcal{H} \rightarrow \mathbf{C}$ est une fonction, et si $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbf{R})$, on définit $f|_k \gamma : \mathcal{H} \rightarrow \mathbf{C}$ par la formule $f|_k \gamma(z) = (cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right)$. Vérifier que ceci est bien défini et que $(f|_k \gamma_1)|_k \gamma_2 = f|_k \gamma_1 \gamma_2$, pour tous $\gamma_1, \gamma_2 \in \mathbf{SL}_2(\mathbf{R})$.

(iv) Soient $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Vérifier que $S^2 = (TS)^3 = -I$; en déduire que pour toute fonction $f : \mathcal{H} \rightarrow \mathbf{C}$, on a $f|_k S^2 = f|_k (TS)^3 = f$.

(v) Montrer que $F|_4 TS = F|_4 TST$. (On pourra s'intéresser à $F|_4 T^2 STST S^2$.)

(vi) En déduire⁽¹⁶⁾ que $F|_4 (TS)^2 = F|_4 T$ et que, si $G = F \cdot F|_4 TS \cdot F|_4 (TS)^2$, alors $G|_{12} S = G|_{12} T = G$ et $G = O(\text{Im}(y)^{c+1-4} e^{-4\pi \text{Im}(y)})$ dans Y .

(vii) En déduire que G est une forme modulaire de poids 12, que $v_{\infty}(G) \geq 2$, et conclure.

Exercice VII.6.9. — (Sommes de 4 carrés)

Le but de cet exercice est de démontrer la formule suivante (C. Jacobi, 1829), dont on déduit une forme effective du théorème de Lagrange (1770) : *tout nombre entier positif est somme d'au plus 4 carrés de nombres entiers*,

$$|\{(a, b, c, d) \in \mathbf{Z}^4, a^2 + b^2 + c^2 + d^2 = n\}| = 8 \sum_{d|n, 4 \nmid d} d.$$

On note $r(n)$ la quantité $|\{(a, b, c, d) \in \mathbf{Z}^4, a^2 + b^2 + c^2 + d^2 = n\}|$; autrement dit $r(n)$ est le nombre de décompositions de n en somme de quatre carrés. On note $\theta = \sum_{n \in \mathbf{Z}} e^{i\pi n^2 z}$ la fonction thêta de Jacobi de l'ex. VII.6.6.

(i) Montrer que $\sum_{n=0}^{+\infty} r(n) e^{i\pi n z} = \theta(z)^4$.

⁽¹⁶⁾ Ces calculs cachent les résultats suivants. Le sous-groupe de $\mathbf{SL}_2(\mathbf{R})$ engendré par S et T est $\mathbf{SL}_2(\mathbf{Z})$, et celui engendré par S et T^2 est le sous-groupe Γ de $\mathbf{SL}_2(\mathbf{Z})$ des matrices dont l'image dans $\mathbf{SL}_2(\mathbf{Z}/2\mathbf{Z})$ est I ou S . Comme $|\mathbf{SL}_2(\mathbf{Z}/2\mathbf{Z})| = 6$, l'indice de Γ dans $\mathbf{SL}_2(\mathbf{Z})$ est 3, et $I, TS, (TS)^2$ forment un système de représentants de $\Gamma \backslash \mathbf{SL}_2(\mathbf{Z})$.

(ii) Soit $F(z) = \theta(z)^4 - 8(G_2(\frac{z}{2}) - 4G_2(2z))$. Montrer que $F|_2S = -F$ et $F|_2T^2 = F$. (On utilisera l'ex. VII.6.7.)

(iii) Montrer que $r(n) \leq (1 + 2\sqrt{n})^4$ et $\sigma(n) \leq \frac{n(n+1)}{2}$. En déduire que, si $F(z) = \sum_{n=1}^{+\infty} a_n e^{i\pi n z}$, alors $a_n = O(n^2)$.

(iv) En déduire, en utilisant l'ex. VII.6.8, que $\theta^4(z) = 8(G_2(\frac{z}{2}) - 4G_2(2z))$, et conclure.

On a en particulier (car $\zeta(4) = \frac{\pi^4}{90}$ et $\zeta(6) = \frac{\pi^6}{3^3 \cdot 5 \cdot 7}$, cf. ex. V.2.14)

$$G_4 = \frac{1}{240} + \sum_{n \geq 1} \sigma_3(n) q^n \quad \text{et} \quad G_6 = \frac{-1}{504} + \sum_{n \geq 1} \sigma_5(n) q^n.$$

On définit alors la *forme discriminant* Δ et l'*invariant modulaire*⁽¹⁷⁾ j par

$$\Delta = \frac{1}{1728} ((240G_4)^3 - (504G_6)^2) \quad \text{et} \quad j = \frac{(240G_4)^3}{\Delta}.$$

Exercice VII.6.10. — (La forme modulaire Δ et l'invariant modulaire j)

(i) Montrer que Δ est une forme parabolique non nulle de poids 12 et, en utilisant la « formule $\frac{k}{12}$ », que Δ ne s'annule pas sur D ou sur $\bar{\Omega}$.

(ii) Montrer que j induit une bijection de D sur \mathbf{C} .

(iii) Montrer que le seul zéro de G_4 dans D est α et le seul zéro de G_6 dans D est i .

(iv) Montrer que, si $\Delta(z) = 0$, $\Delta(z+n) = 0$ quel que soit $n \in \mathbf{Z}$ et $\Delta(-1/z) = 0$. En déduire que si $\Delta(z_0) = 0$, et si $|\operatorname{Re}(z_0)| \leq \frac{1}{2}$ et $|z_0| < 1$, alors il existe $z_1 \in \mathcal{H}$, vérifiant $\Delta(z_1) = 0$, $\operatorname{Im}(z_1) > \operatorname{Im}(z_0)$ et $|\operatorname{Re}(z_1)| \leq \frac{1}{2}$.

(v) Montrer que Δ ne s'annule pas sur \mathcal{H} , et que j est une fonction modulaire de poids 0.

Exercice VII.6.11. — (La formule de Jacobi pour Δ)

Soit $F(q) = q \prod_{n=1}^{+\infty} (1 - q^n)^{24}$. Notre but est de prouver que $\Delta(z) = F(e^{2i\pi z})$. On utilisera pleinement le résultat de l'ex. VII.6.7.

(i) Soit $f(z) = \log(F(e^{2i\pi z}))$. Montrer que $\frac{1}{2i\pi} f'(z) = -24G_2(z)$.

(ii) Soit $g(z) = f(-1/z) - 12 \log z$. Montrer que $g'(z) = f'(z)$.

(iii) En déduire que, si on pose $H(z) = F(e^{2i\pi z})$, alors $\frac{z^{-12} H(-1/z)}{H(z)}$ est constante sur \mathcal{H} , puis que H est une forme modulaire de poids 12.

(iv) Conclure.

⁽¹⁷⁾ Son q -développement $j = q^{-1} + 744 + 196884q + 21493760q^2 + \dots$ recèle des trésors (cf. note 3 du chap. I).

ANNEXE A

LE THÉORÈME DES NOMBRES PREMIERS

A.1. Introduction

Ce chapitre est consacré à la démonstration du *théorème des nombres premiers* et du *théorème de la progression arithmétique*. On sait depuis les grecs qu'il existe une infinité de nombres premiers. Leur répartition n'a cessé depuis de fasciner les mathématiciens. Voici, par exemple, ce qu'écrivait Euler en 1747 : « *Les mathématiciens ont tâché jusqu'ici en vain à découvrir un ordre quelconque dans la progression des nombres premiers, et on a lieu de croire, que c'est un mystère auquel l'esprit humain ne saurait jamais pénétrer. Pour s'en convaincre, on n'a qu'à jeter les yeux sur les tables des nombres premiers, que quelques personnes se sont donné la peine de continuer au-delà de cent-mille : et on s'apercevra d'abord qu'il n'y règne aucun ordre ni règle. Cette circonstance est d'autant plus surprenante, que l'arithmétique nous fournit des règles sûres, par le moyen desquelles on est en état de continuer la progression de ces nombres aussi loin que l'on souhaite, sans pourtant nous y laisser apercevoir la moindre marque d'un ordre quelconque.* ». Le même Euler a, entre autres :

- démontré que $\sum_{k=1}^n \frac{1}{k}$ diverge comme $\log n$, (et même que $-\log n + \sum_{k=1}^n \frac{1}{k}$ tend vers une limite γ appelée depuis « constante d'Euler »),
- factorisé $\sum_{n \in \mathbf{N}} \frac{1}{n}$ sous la forme $\prod_{p \in \mathcal{P}} (1 + \frac{1}{p} + \frac{1}{p^2} \cdots)$,
- remarqué que le logarithme du produit était $\sum_p \frac{1}{p}$ à une somme convergente près.

Ceci lui a fourni une nouvelle démonstration de l'existence d'une infinité de nombres premiers (puisque la somme de leurs inverses diverge). De plus, en partant de la formule $\sum_p \frac{1}{p} \sim \log(\sum_n \frac{1}{n})$, il en avait déduit que $\sum_{p \leq x} \frac{1}{p} \sim \log \log x$. Maintenant, si Δx est petit devant x , les nombres premiers entre x et $x + \Delta x$ sont tous de taille x . Si on note $\pi(x)$ le nombre de nombres premiers $\leq x$, on a donc $\log \log(x + \Delta x) - \log \log x \sim \frac{\pi(x + \Delta x) - \pi(x)}{x}$. Comme la dérivée de $\log \log x$ est $\frac{1}{x \log x}$, on en « déduit » que la densité des nombres premiers autour de x est de l'ordre de $\frac{1}{\log x}$, et donc que⁽¹⁾ $\pi(x) \sim \text{Li}(x) = \int_2^x \frac{dt}{\log t}$. Il a

⁽¹⁾La fonction Li est le *logarithme intégral*; au voisinage de l'infini, on a $\text{Li}(x) \sim \frac{x}{\log x}$ (faire une intégration par partie, en intégrant 1 et en dérivant $\frac{1}{\log t}$). On peut donc reformuler le théorème des nombres premiers

fallu attendre plus d'un siècle pour que ce résultat soit rigoureusement⁽²⁾ démontré, ce qui fut fait en 1896 par J. Hadamard et de la Vallée Poussin indépendamment.

Théorème A.1.1. — (des nombres premiers) $\pi(x) \sim \frac{x}{\log x}$.

Les démonstrations de J. Hadamard et C. de la Vallée Poussin reposent sur la stratégie suggérée par B. Riemann, utilisant le lien entre les zéros de la fonction ζ et la répartition des nombres premiers. L'ingrédient fondamental est la non existence de zéros de la fonction ζ sur la droite $\operatorname{Re}(s) = 1$, où la convergence du produit eulérien cesse, ce qui ne permet pas de conclure quoi que ce soit de la non annulation de chacun de ses facteurs. De fait, le théorème des nombres premiers est équivalent (cf. ex. A.4.6) à la non annulation de ζ sur la droite $\operatorname{Re}(s) = 1$.

Si P est un polynôme, et s'il n'y a aucune obstruction arithmétique à ce que les valeurs de P aux entiers puissent être des nombres premiers⁽³⁾, on peut partir du principe que $P(n)$ a autant de chance d'être premier qu'un nombre de même taille pris au hasard, soit $\frac{1}{\log P(n)} \sim \frac{1}{\deg P} \cdot \frac{1}{\log n}$. C'est ce genre d'heuristique qui mène à la conjecture de Bateman et Horn ci-dessous.

Soient P_1, \dots, P_k des polynômes distincts, à coefficients entiers, irréductibles dans $\mathbf{Q}[X]$, dont le coefficient dominant est > 0 , et soit $P = P_1 \cdots P_k$. Si $p \in \mathscr{P}$, on note $N_p(P)$ le nombre de solutions de l'équation $P(x) = 0$ dans le corps $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, et on définit une constante⁽⁴⁾ $C(P)$ par

$$C(P) = \prod_{p \in \mathscr{P}} \left(\left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{N_p(P)}{p}\right) \right).$$

Conjecture A.1.2. — (Bateman-Horn, 1962) Si $C(P) \neq 0$, alors l'ensemble des $n \in \mathbf{N}$ tels que $P_1(n), \dots, P_k(n)$ soient simultanément premiers est infini, et on a

$$|\{n \leq x, P_1(n) \in \mathscr{P}, \dots, P_k(n) \in \mathscr{P}\}| \sim \frac{C(P)}{\deg P_1 \cdots \deg P_k} \cdot \frac{x}{(\log x)^k}.$$

sous la forme plus parlante $\pi(x) \sim \frac{x}{\log x}$; c'est sous cette forme que nous le démontrerons, mais $\operatorname{Li}(x)$ est une bien meilleure approximation (cf. prop. A.5.1) de $\pi(x)$ que $\frac{x}{\log x}$.

⁽²⁾Le lecteur pourra vérifier que, si $M = \cup_{k \geq 1} [2^{2^k}, 2^{2^k+1}[$, alors $\sum_{n \in \mathbf{N} \cap M, n \leq x} \frac{1}{n} \sim \log \log x$, mais que $\frac{|\{n \in \mathbf{N} \cap M, n \leq x\}|}{x/\log x}$ n'a pas de limite.

⁽³⁾Par exemple $12n + 9$, $n(n^2 + 1)$ ou $n(n - 1) + 2$ ne peuvent prendre qu'un nombre fini de valeurs premières pour des raisons évidentes.

⁽⁴⁾La convergence du produit définissant $C(P)$ n'est pas du tout évidente (ce n'est pas une convergence absolue), mais peut se démontrer; en particulier, $C(P) \neq 0$ si $N_p(P) \neq p$ pour tout $p \in \mathscr{P}$. Cette convergence traduit le fait que le nombre de solutions modulo p de l'équation $P(x) = 0$, pour $P \in \mathbf{Z}[X]$ sans racine double, est, en moyenne, le nombre de facteurs de la décomposition de P en produit de facteurs irréductibles.

Le théorème des nombres premiers correspond au cas $k = 1$, $P = X$ de la conjecture. Le seul autre résultat que l'on ait confirmant cette conjecture est celui où $k = 1$, et $P = P_1$ est de degré 1. Dans ce cas, P est de la forme $Dx + a$, où $D \geq 2$, et a est premier à D , sinon il existe un nombre premier p pour lequel $Dx + a$ est identiquement nul modulo p , et $C(P) = 0$. Si $p \mid D$, alors $N_p(P) = 0$, et si $p \nmid D$, alors $N_p(P) = 1$. La constante $C(P)$ de la conjecture de Bateman et Horn est donc

$$\prod_{p \mid D} \left(1 - \frac{1}{p}\right)^{-1} = \prod_{p \mid D} \frac{p}{p-1} = \prod_{p \mid D} \frac{p^{v_p(D)}}{(p-1)p^{v_p(D)-1}} = \frac{D}{\varphi(D)},$$

où $\varphi(D) = |(\mathbf{Z}/D\mathbf{Z})^*|$ est la fonction indicatrice d'Euler. La conjecture de Bateman et Horn est donc, dans ce cas, conséquence du résultat suivant (appliqué à Dx au lieu de x).

Théorème A.1.3. — (de la progression arithmétique) *Si $D \geq 2$, si a est premier à D , et si $\pi(D, a, x) = |\{p \in \mathcal{P}, p \equiv a \pmod{D}, p \leq x\}|$, alors $\pi(D, a, x) \sim \frac{1}{\varphi(D)} \cdot \frac{x}{\log x}$*

On peut paraphraser ce théorème en disant que les nombres premiers s'équirépartissent dans les progressions arithmétiques dans lesquelles ils peuvent exister⁽⁵⁾. Sous la forme du théorème, le résultat est dû à de la Vallée Poussin, mais n'est qu'une petite extension du théorème des nombres premiers, si on utilise les idées de Dirichlet (1837) qui avait déjà démontré un résultat d'équirépartition des nombres premiers dans les progressions arithmétiques (nettement moins fin, mais quand même spectaculaire : démontrer, à la main, l'existence d'une infinité de nombres premiers de la forme $7n + 3$ (par exemple), n'est pas si facile...). La démonstration du th. A.1.3 suit celle du théorème des nombres premiers, en utilisant toutes les fonctions L de Dirichlet (que celui-ci avait précisément introduites pour démontrer son théorème), au lieu de la seule fonction ζ . Nous l'avons transformée en une série d'exercices.

Le théorème de la progression arithmétique est aussi le seul cas où on sache prouver l'existence d'une infinité de n tel que $P_1(n), \dots, P_k(n)$ soient simultanément premiers (ici $k = 1$). Par exemple, on ne sait pas démontrer l'existence d'une infinité de $p \in \mathcal{P}$ tel que $p + 2$ soit premier (problème des *nombres premiers jumeaux* ⁽⁶⁾ ; il correspond à $k = 2$,

⁽⁵⁾Dans la rubrique « nombres premiers et progressions arithmétiques », mentionnons que B. Green et T. Tao ont démontré en 2004 qu'il existait des progressions arithmétiques de longueur arbitraire (mais pas infinie...) dans l'ensemble des nombres premiers, ce qui joua un rôle certain dans l'attribution de la médaille Fields à T. Tao en 2006. P. Erdős avait espéré que l'on pourrait démontrer ce résultat en utilisant juste que la somme des inverses des nombres premiers diverge. Malgré les efforts de pas mal de gens réputés pour leur astuce (dont Roth, Gowers, Bourgain et Tao), on ne sait toujours pas démontrer (ou infirmer...) que si $X \subset \mathbf{N} - \{0\}$ ne contient pas de progression arithmétique de longueur 3 (i.e. il n'existe pas $a, b, c \in X$ distincts, tels que $a + c = 2b$), alors $\sum_{n \in X} \frac{1}{n} < +\infty$. Le meilleur résultat dans cette direction est un résultat de J. Bourgain (1999) selon lequel $|\{n \in X, n \leq N\}| = O\left(\frac{N\sqrt{\log \log N}}{\sqrt{\log N}}\right)$.

⁽⁶⁾Il a fallu attendre 2005 pour que D. Goldston et C. Yıldırım démontrent que $\liminf \frac{p_{n+1} - p_n}{\log p_n} = 0$, si p_n désigne le n -ième nombre premier, ce qui constitue un pas encourageant en direction du problème des

$P_1 = X$, $P_2 = X + 2$). De même, on ne sait pas démontrer qu'il existe une infinité de nombres premiers de la forme $n^2 + 1$ ⁽⁷⁾.

A.2. Les fonctions ψ et ψ_1

1. Théorème des nombres premiers et comportement de ψ_1 en $+\infty$

La fonction $\pi(x)$ n'est pas reliée directement à la fonction ζ , ce qui nous conduit à introduire les fonctions auxiliaires suivantes.

• La fonction Λ de von Mangoldt définie par $\sum_{n=1}^{+\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$. Si $\operatorname{Re}(s) > 1$, on a $\zeta(s) = \prod_{p \in \mathcal{P}} \frac{1}{1-p^{-s}}$, d'après la prop. VII.3.3, et donc (th. V.2.15),

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{p \in \mathcal{P}} \frac{p^{-s} \log p}{1-p^{-s}} = \sum_{p \in \mathcal{P}} \sum_{\nu=1}^{+\infty} \frac{\log p}{p^{\nu s}}, \quad \text{si } \operatorname{Re}(s) > 1.$$

On en déduit que $\Lambda(n) = 0$, si n n'est pas une puissance d'un nombre premier, et $\Lambda(n) = \log p$, si $n = p^\nu$, et $\nu \geq 1$.

- La fonction ψ définie par $\psi(x) = \sum_{n \leq x} \Lambda(n)$.
- La fonction ψ_1 définie par $\psi_1(x) = \int_0^x \psi(t) dt$.

Lemme A.2.1. — *Les énoncés suivants sont équivalents au voisinage de $+\infty$.*

- (i) $\pi(x) \sim \frac{x}{\log x}$.
- (ii) $\psi(x) \sim x$.
- (iii) $\psi_1(x) \sim \frac{1}{2}x^2$.

Démonstration. — Par définition, $\psi(x) = \sum_{p^\nu \leq x} \log p$. Or, $p^\nu \leq x$ implique en particulier $\nu \leq \frac{\log x}{\log 2}$, et $\nu \geq 2$ implique $p \leq \sqrt{x}$. On en déduit l'encadrement

$$\sum_{p \leq x} \log p \leq \psi(x) \leq \sum_{p \leq x} \log p + (\sqrt{x} \log x) / \log 2 \leq \pi(x) \log x + (\sqrt{x} \log x) / \log 2.$$

Par ailleurs, si $\beta < 1$, et $x^\beta \leq p$, on a $\log p \geq \beta \log x$. On en déduit que

$$\sum_{p \leq x} \log p \geq \beta \log x (\pi(x) - \pi(x^\beta)),$$

et comme $\pi(x^\beta) \leq x^\beta$, on obtient, quel que soit $\beta < 1$, l'encadrement

$$\beta(\pi(x) \log x - x^\beta \log x) \leq \psi(x) \leq \pi(x) \log x + (\sqrt{x} \log x) / \log 2.$$

nombres premiers jumeaux. (La division par $\log n$ se justifie par le fait que p_n est de l'ordre de $n \log n$ et l'écart moyen entre $p_{n+1} - p_n$ est $\log n$.)

⁽⁷⁾On sait, depuis Fermat, que tout nombre premier de la forme $4k + 1$ est somme de deux carrés; on en déduit qu'il existe une infinité de nombres premiers de la forme $n^2 + m^2$, avec $n, m \in \mathbf{N}$. Il a fallu attendre 1998 pour que J. Friedlander et H. Iwaniec démontrent l'existence d'une infinité de nombres premiers de la forme $n^2 + m^4$; on est encore loin de $n^2 + 1$...

On en déduit l'équivalence entre (i) et (ii).

L'implication (ii) \Rightarrow (iii) est immédiate par intégration. Pour démontrer la réciproque, constatons que ψ est une fonction croissante, et donc que l'on a

$$\frac{\psi_1(x) - \psi_1((1-\varepsilon)x)}{\varepsilon x} \leq \psi(x) \leq \frac{\psi_1((1+\varepsilon)x) - \psi_1(x)}{\varepsilon x},$$

quel que soit $\varepsilon > 0$. Maintenant, si $\psi_1(x) = \frac{1}{2}x^2 + x^2\eta(x)$, avec $\lim_{x \rightarrow +\infty} \eta(x) = 0$, on obtient, en divisant l'encadrement ci-dessus par x , l'encadrement suivant :

$$1 - \frac{\varepsilon}{2} + \frac{\eta(x) - \eta((1-\varepsilon)x)}{\varepsilon} \leq \frac{\psi(x)}{x} \leq 1 + \frac{\varepsilon}{2} + \frac{\eta((1+\varepsilon)x) - \eta(x)}{\varepsilon}.$$

En faisant tendre x vers $+\infty$, on en déduit que $\liminf \frac{\psi(x)}{x} \geq 1 - \frac{\varepsilon}{2}$ et $\limsup \frac{\psi(x)}{x} \leq 1 + \frac{\varepsilon}{2}$. Comme ceci est vrai pour tout $\varepsilon > 0$, cela prouve que $\psi(x) \sim x$, ce qui permet de conclure.

Exercice A.2.2. — Si $D \geq 2$, si a est premier à D , et si $x \geq 1$, soient

$$\psi(D, a, x) = \sum_{n \leq x, n \equiv a \pmod{D}} \Lambda(n) \quad \text{et} \quad \psi_1(D, a, x) = \int_0^x \psi(D, a, t) dt.$$

Montrer que les deux énoncés suivants sont équivalents au théorème de la progression arithmétique : « $\varphi(D)\psi(D, a, x) \sim x$ » et « $\varphi(D)\psi_1(D, a, x) \sim \frac{x^2}{2}$ ».

Exercice A.2.3. — Montrer que le théorème des nombres premier équivaut à $\text{ppcm}(1, \dots, n) = e^{n+o(n)}$.

2. Une formule intégrale pour ψ_1

Lemme A.2.4. — Si $c > 0$ et $x > 0$, alors

$$\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} ds = \begin{cases} 0 & \text{si } x < 1, \\ x-1 & \text{si } x \geq 1. \end{cases}$$

Démonstration. — C'est un calcul de résidus parfaitement standard. La fonction $\frac{x^{s+1}}{s(s+1)}$ est méromorphe sur \mathbf{C} , holomorphe en dehors de deux pôles simples en $s = 0$ et $s = -1$, de résidus respectifs $\text{Res}(\frac{x^{s+1}}{s(s+1)}, 0) = x$ et $\text{Res}(\frac{x^{s+1}}{s(s+1)}, -1) = -1$.

• Si $x \geq 1$, on intègre sur le lacet γ_T constitué du segment $[c-iT, c+iT]$, et de $C^+(T)$, arc de cercle $c + Te^{i\theta}$, avec θ variant de $\frac{\pi}{2}$ à $\frac{3\pi}{2}$. Si $T > c+1$, le lacet γ_T a pour indice 1 par rapport aux deux points 0 et -1 . On déduit de la formule des résidus que $\frac{1}{2i\pi} \int_{\gamma_T} \frac{x^{s+1}}{s(s+1)} ds = x-1$, si $T > c+1$. Par ailleurs, sur $C^+(T)$, on a $|x^{s+1}| \leq x^{c+1}$ et $|\frac{1}{s(s+1)}| \leq \frac{1}{(T-c)(T-c-1)}$. Donc $\int_{C^+(T)} \frac{x^{s+1}}{s(s+1)} ds \rightarrow 0$ quand $T \rightarrow +\infty$, et

$$\begin{aligned} \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{x^{s+1}}{s(s+1)} ds &= \lim_{T \rightarrow +\infty} \frac{1}{2i\pi} \int_{[c-iT, c+iT]} \frac{x^{s+1}}{s(s+1)} ds \\ &= x-1 - \lim_{T \rightarrow +\infty} \int_{C^+(T)} \frac{x^{s+1}}{s(s+1)} ds = x-1. \end{aligned}$$

• Si $x < 1$, on intègre sur le lacet γ_T constitué du segment $[c - iT, c + iT]$, et de $C^-(T)$, arc de cercle $c + Te^{i\theta}$, avec θ variant de $\frac{\pi}{2}$ à $-\frac{\pi}{2}$. Dans ce cas γ_T a pour indice 0 par rapport aux deux points 0 et -1 , et la formule des résidus nous donne $\frac{1}{2i\pi} \int_{\gamma_T} \frac{x^{s+1}}{s(s+1)} ds = 0$, quel que soit $T > 0$. Le reste de l'argument est le même que ci-dessus.

Corollaire A.2.5. — Soit $L(\mathbf{a}, s) = \sum_{n=1}^{+\infty} \frac{a_n}{n^s}$ une série de Dirichlet d'abscisse de convergence absolue $\sigma_{\text{abs}} \neq +\infty$. Alors, si $c > \sup(0, \sigma_{\text{abs}})$ et $x > 0$, on a

$$\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{L(\mathbf{a}, s)x^{s+1}}{s(s+1)} ds = \int_0^x \left(\sum_{n \leq t} a_n \right) dt.$$

Démonstration. — Par hypothèse, la série $\sum_{n=1}^{+\infty} \frac{a_n}{n^s}$ converge normalement sur la droite $c + i\mathbf{R}$, et comme la fonction $\frac{1}{s(s+1)}$ est sommable sur cette droite, on peut échanger somme et intégrale. Par ailleurs, on a

$$\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{1}{n^s} \frac{x^{s+1}}{s(s+1)} ds = \begin{cases} 0 & \text{si } x < n, \\ n\left(\frac{x}{n} - 1\right) = x - n & \text{si } x \geq n. \end{cases}$$

On en déduit que

$$\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{L(\mathbf{a}, s)x^{s+1}}{s(s+1)} ds = \sum_{n \leq x} a_n(x - n) = \int_0^x \left(\sum_{n \leq t} a_n \right) dt,$$

ce qui permet de conclure.

Corollaire A.2.6. — Si $c > 1$, et si $x > 1$, alors $\psi_1(x) = -\frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{\zeta'(s)}{\zeta(s)} \frac{x^{s+1}}{s(s+1)} ds$.

Exercice A.2.7. — Si $D \geq 2$, notons $\text{Dir}(D)$ l'ensemble des caractères de Dirichlet modulo D . On utilisera le résultat suivant démontré dans le n° 5 du § I.2 :

$$\text{Si } a \text{ est premier à } D, \text{ alors } \sum_{\chi \in \text{Dir}(D)} \overline{\chi(a)} \chi(n) = \begin{cases} \varphi(D) & \text{si } n \equiv a \pmod{D}, \\ 0 & \text{sinon.} \end{cases}$$

Établir, si $c > 1$ et $x > 1$, la formule

$$\varphi(D) \psi_1(D, a, x) = \frac{1}{2i\pi} \sum_{\chi \in \text{Dir}(D)} \overline{\chi(a)} \int_{c-i\infty}^{c+i\infty} \frac{-L'(\chi, s)}{L(\chi, s)} \frac{x^{s+1}}{s(s+1)} ds.$$

A.3. Formules explicites

La formule du cor. A.2.6 est, en elle-même, assez peu utile en ce qui concerne la démonstration du théorème des nombres premiers. La seule information qu'on puisse en tirer est, semble-t-il, en majorant brutalement ce qu'on intègre, l'existence, pour tout $c > 1$, d'une constante $C(c)$ telle que $\psi_1(x) \leq C(c)x^{1+c}$, ce qui était évident dès le départ. Pour obtenir des résultats plus fins, on va déplacer la ligne d'intégration vers la gauche (de manière

à diminuer le c). La traversée de la bande critique est un peu périlleuse à cause de la présence des zéros de ζ qui entraîne l'existence de nombreux pôles pour la fonction $\frac{\zeta'}{\zeta}$, ce qui fait que la fonction $\frac{\zeta'}{\zeta}$ est loin d'être assez petite pour qu'on puisse se dispenser de prendre des précautions en se déplaçant. On va donc être forcé de majorer cette fonction $\frac{\zeta'}{\zeta}$ dans la bande critique, et le miracle des fonctions holomorphes fait que la connaissance de la fonction ζ des deux côtés de la bande critique permet de contrôler un peu ce qui se passe à l'intérieur (on maîtrise bien la fonction ζ dans le demi-plan $\operatorname{Re}(s) < 0$ grâce à l'équation fonctionnelle et à la formule de Stirling).

1. Énoncé du résultat

Afin de pouvoir appliquer les résultats de ce § aux fonctions L de Dirichlet, nous allons partir d'une fonction L méromorphe sur \mathbf{C} , holomorphe en dehors d'un pôle simple éventuel en $s = 1$, et qui vérifie les conditions (L1)–(L3) ci-dessous :

(L1) Si $a > 1$, il existe $c(a)$, tel que, si $\operatorname{Re}(s) \geq a$, on ait

$$|L(s)| \leq c(a), \quad |L(s)^{-1}| \leq c(a), \quad \left| \frac{L'(s)}{L(s)} \right| \leq c(a).$$

En particulier, L ne s'annule pas sur le demi-plan $\operatorname{Re}(s) > 1$.

(L2) Il existe $A \in \mathbf{C}^*$, $B \in \mathbf{R}_+^*$ et $c \in [0, 2[$ tels que L vérifie l'équation fonctionnelle

$$L(s) = A \cdot B^s \cdot \Gamma(1-s) \cdot \sin \frac{\pi(s-c)}{2} \cdot \bar{L}(1-s), \quad \text{où } \bar{L} \text{ est définie par } \bar{L}(s) = \overline{L(\bar{s})}.$$

(L3) Quels que soient $a \leq b$ réels, il existe $C(a, b) > 0$ et $c(a, b) > 0$ tels que, si $a \leq \sigma \leq b$ et $|\tau| \geq 1$, alors

$$|L(\sigma + i\tau)| \leq C(a, b)e^{c(a, b)|\tau|}.$$

Lemme A.3.1. — *La fonction ζ vérifie les propriétés (L1)–(L3).*

Démonstration. — Si $\operatorname{Re}(s) > a$, on a

$$\begin{aligned} |\zeta(s)| &= \prod_{p \in \mathcal{P}} \frac{1}{|1 - p^{-s}|} \leq \prod_{p \in \mathcal{P}} \frac{1}{1 - p^{-\operatorname{Re}(s)}} \leq \prod_{p \in \mathcal{P}} \frac{1}{1 - p^{-a}} \leq \zeta(a), \\ |\zeta(s)^{-1}| &= \prod_{p \in \mathcal{P}} |1 - p^{-s}| \leq \prod_{p \in \mathcal{P}} (1 + p^{-a}) \leq \frac{\zeta(a)}{\zeta(2a)}, \\ \left| \frac{\zeta'(s)}{\zeta(s)} \right| &\leq \sum_{n \in \mathbf{N}} \Lambda(n) |n^{-s}| \leq \sum_{n \in \mathbf{N}} \Lambda(n) n^{-a} = \frac{\zeta'(a)}{\zeta(a)}, \end{aligned}$$

ce qui montre que ζ vérifie (L1), avec $c(a) = \sup(\zeta(a), \frac{\zeta'(a)}{\zeta(a)})$. Les propriétés (L2) et (L3) sont nettement plus délicates mais ont déjà été démontrées (la propriété (L2) fait l'objet du th. VII.3.6, et le (iii) de la prop. VII.2.8 (allié à la formule $\zeta(s) = \frac{1}{s-1} M(\frac{t}{e^t-1}, s-1)$

de la démonstration du th. VII.3.4) montre que l'on peut prendre n'importe quoi de $> \frac{\pi}{2}$ pour $c(a, b)$.

Exercice A.3.2. — (i) Montrer que, si χ est primitif, alors $L(\chi, s)$ vérifie les propriétés (L1)–(L3).

(ii) Montrer que, si χ est primitif de conducteur divisant D , si $c > 1$, et si $x > 1$, alors

$$\frac{1}{2i\pi} \int_{2-i\infty}^{2+i\infty} \left(\frac{L'(\chi_D, s)}{L(\chi_D, s)} - \frac{L'(\chi, s)}{L(\chi, s)} \right) \frac{x^{s+1}}{s(s+1)} ds = \sum_{p|D} \sum_{\nu \leq \frac{\log x}{\log p}} \chi(\nu)(x - p^\nu).$$

En déduire que $\frac{1}{2i\pi} \int_{2-i\infty}^{2+i\infty} \left(\frac{L'(\chi_D, s)}{L(\chi_D, s)} - \frac{L'(\chi, s)}{L(\chi, s)} \right) \frac{x^{s+1}}{s(s+1)} ds$ est un $O(x \log x)$.

Soit donc L une fonction méromorphe sur \mathbf{C} , holomorphe en dehors d'un pôle simple éventuel (ce pôle simple éventuel pouvant en fait être un zéro) en $s = 1$, et vérifiant les propriétés (L1)–(L3). Si $x > 1$, on note F_x la fonction $F_x(s) = -\frac{L'(s)}{L(s)} \frac{x^{s+1}}{s(s+1)}$. C'est une fonction méromorphe sur \mathbf{C} , holomorphe en dehors de pôles en 0 , -1 , éventuellement en 1 , et en les zéros de L . Comme L ne s'annule pas sur $\operatorname{Re}(s) > 0$ et vérifie l'équation fonctionnelle de la propriété (L2), les seuls zéros de L en dehors de la bande $0 \leq \operatorname{Re}(s) \leq 1$ (bande critique) sont les $c - 2k$, pour $k \in \mathbf{N} - \{0\}$, ce qui peut inclure -1 . On note $Y(L)$ l'ensemble des zéros dans la bande critique, distincts de 0 et 1 .

Soient $a_0 = x^{-1} \operatorname{Res}(F_x, 0)$ et $a_{-1} = \operatorname{Res}(F_x, -1)$ ⁽⁸⁾. Le résultat que nous avons en vue est la « formule explicite » suivante (dans laquelle $v_z(L) \in \mathbf{Z}$ désigne la valuation (i.e. l'ordre du zéro) de L au point z).

Théorème A.3.3. — (i) La série $\sum_{\rho \in Y(L)} \frac{v_\rho(L)}{\rho(\rho+1)}$ est absolument convergente.

(ii) Si $x > 1$, alors $\frac{1}{2i\pi} \int_{2-i\infty}^{2+i\infty} \frac{-L'(s)}{L(s)} \frac{x^{s+1}}{s(s+1)} ds$ est aussi égal à

$$-v_1(L) \frac{x^2}{2} + a_0 x + a_{-1} - \sum_{\rho \in Y(L)} \frac{v_\rho(L) x^{\rho+1}}{\rho(\rho+1)} - \sum_{k \geq 1, c-2k \neq -1} \frac{x^{1+c-2k}}{(2k-c)(2k-c-1)},$$

les deux séries ci-dessus étant absolument convergentes.

La démonstration va demander un peu de préparation, mais on peut tout de même remarquer que l'expression finale peut se réécrire sous la forme plus compacte ⁽⁹⁾

$$\frac{1}{2i\pi} \int_{2-i\infty}^{2+i\infty} \frac{-L'(s)}{L(s)} \frac{x^{s+1}}{s(s+1)} ds = \sum_{\operatorname{Re}(\rho) < 2} \operatorname{Res}(F_x, \rho),$$

⁽⁸⁾ Si 0 n'est pas un zéro de L , alors $a_0 = \frac{-L'(0)}{L(0)}$. Si 0 est un zéro de L , alors 0 est un pôle double de F , et a_0 est de la forme $a_{0,0} + a_{0,1} \log x$.

Si $c \neq 1$, alors -1 est un pôle simple de F , et $a_{-1} = \frac{L'(1)}{L(1)}$. Si $c = 1$, alors -1 est un pôle double de F , et a_{-1} est de la forme $a_{-1,0} + a_{-1,1} \log x$.

⁽⁹⁾ Formellement, cette formule n'est autre que la formule des résidus si on considère la droite $\operatorname{Re}(s) = 2$ comme un lacet entourant le demi-plan $\operatorname{Re}(s) < 2$.

l'expression du théorème étant obtenue en explicitant le résidu de F_x en tous ses pôles (qui sont simples sauf peut-être 0 et -1). De plus, comme $|x^{\rho+1}| \leq x^3$, si $\operatorname{Re}(\rho) < 2$, la convergence des séries suit juste de la convergence absolue des séries $\sum_{\rho \in Y(L)} \frac{v_\rho(L)}{\rho(\rho+1)}$ et $\sum_{k \geq 1} \frac{1}{(2k-c)(2k-c-1)}$. Il suffit donc de démontrer le (i) et la forme compacte ci-dessus.

2. Les fonctions L et $\frac{L'}{L}$ en dehors de la bande critique

Lemme A.3.4. — Si $a < 0$, la fonction $(s-a+1)^{a-\frac{3}{2}}(s-1)L(s)$ est bornée sur la droite $\operatorname{Re}(s) = a$.

Démonstration. — Sur la droite $\operatorname{Re}(s) = a$, la fonction $F(s) = (s-a+1)^{a-\frac{3}{2}}(s-1)L(s)$ est continue. Par ailleurs, si $t \in \mathbf{R}$, on a

$$L(a+it) = A B^{a+it} \sin\left(\frac{\pi(a-c+it)}{2}\right) \Gamma(1-a-it) \bar{L}(1-a-it).$$

Or $A B^{a+it} \bar{L}(1-a-it)$ est bornée pour $t \in \mathbf{R}$, et $e^{-\pi|t|/2} \left| \sin \frac{\pi(a-c+it)}{2} \right|$ tend vers $\frac{1}{2}$ quand $|t|$ tend vers $+\infty$, ce qui permet de déduire du cor. VII.2.4 que

$$\left| \sin\left(\frac{\pi(a-c+it)}{2}\right) \Gamma(1-a-it) \right| |t|^{-\frac{1}{2}+a}$$

est borné quand $|t|$ tend vers $+\infty$. Donc $|F(a+it)| = |(it+1)^{a-\frac{3}{2}}(a+it-1)L(a+it)|$ est borné quand $|t|$ tend vers $+\infty$, et comme $F(a+it)$ est continue sur \mathbf{R} , elle est bornée sur \mathbf{R} . Ceci permet de conclure.

Lemme A.3.5. — Soit $(t_n^+)_{n \geq 2}$ (resp. $(t_n^-)_{n \geq 2}$) une suite de réels vérifiant $n \leq t_n^+ \leq n+1$ (resp. $-n-1 \leq t_n^- \leq -n$). Soient $b_n^+ = -1+it_n^+$ (resp. $b_n^- = -1+it_n^-$) et $c_n^+ = c+1-2n+it_n^+$ (resp. $c_n^- = c+1-2n+it_n^-$). Alors il existe $C > 0$ tels que, $\left| \frac{L'(s)}{L(s)} \right| \leq C + \log n$, quels que soient $n \geq 2$ et $s \in [b_n^+, c_n^+] \cup [c_n^+, c_n^-] \cup [c_n^-, b_n^-]$.

Démonstration. — On part de l'équation fonctionnelle (L2) dont on déduit l'identité

$$\frac{L'(s)}{L(s)} = \log B + \frac{i\pi}{2} \cdot \frac{e^{i\pi(s-c)/2} + e^{-i\pi(s-c)/2}}{e^{i\pi(s-c)/2} - e^{-i\pi(s-c)/2}} - \frac{\Gamma'(1-s)}{\Gamma(1-s)} - \frac{\bar{L}'(1-s)}{\bar{L}(1-s)}.$$

Maintenant, quand s décrit $[b_n^+, c_n^+] \cup [c_n^+, c_n^-] \cup [c_n^-, b_n^-]$, on a $\operatorname{Re}(1-s) \geq 2$.

• Or $\left| \frac{\bar{L}'(1-s)}{\bar{L}(1-s)} \right| \leq c(2)$, si $\operatorname{Re}(1-s) \geq 2$, d'après la propriété (L1).

• D'après la prop. VII.2.3, il existe $C_1 > 0$ tel que $\left| \frac{\Gamma'(1-s)}{\Gamma(1-s)} - \log(1-s) \right| \leq C_1$, si $\operatorname{Re}(1-s) \geq 2$. On en déduit, si $s \in [b_n^+, c_n^+] \cup [c_n^+, c_n^-] \cup [c_n^-, b_n^-]$, que

$$\left| \frac{\Gamma'(1-s)}{\Gamma(1-s)} \right| \leq \pi + C_1 + \log |1-s| \leq \pi + C_1 + \log \sqrt{(2n-c)^2 + (n+1)^2} \leq \pi + C_1 + \log 3 + \log n.$$

- Si s appartient à $[b_n^+, c_n^+]$ ou $[c_n^-, b_n^-]$, on a $|\operatorname{Im}(s)| \geq n$, et donc

$$\left| \frac{e^{i\pi(s-c)/2} + e^{-i\pi(s-c)/2}}{e^{i\pi(s-c)/2} - e^{-i\pi(s-c)/2}} \right| \leq \frac{e^{\pi n/2} + e^{-\pi n/2}}{e^{\pi n/2} - e^{-\pi n/2}} \leq \frac{1 + e^{-\pi}}{1 - e^{-\pi}}.$$

Si $s \in [c_n^+, c_n^-]$, et si $s = c + 1 - 2n + it$, alors

$$\left| \frac{e^{i\pi(s-c)/2} + e^{-i\pi(s-c)/2}}{e^{i\pi(s-c)/2} - e^{-i\pi(s-c)/2}} \right| = \left| \frac{e^{-\pi t/2} - e^{\pi t/2}}{e^{-\pi t/2} + e^{\pi t/2}} \right| \leq 1 \leq \frac{1 + e^{-\pi}}{1 - e^{-\pi}}.$$

On en déduit que $\left| \frac{L'(s)}{L(s)} \right| \leq \log B + c(2) + \pi + C_1 + \log 3 + \frac{\pi}{2} \frac{1+e^{-\pi}}{1-e^{-\pi}} + \log n$, si $s \in [b_n^+, c_n^+] \cup [c_n^+, c_n^-] \cup [c_n^-, b_n^-]$. Ceci permet de conclure.

3. La fonction L dans la bande critique

Lemme A.3.6. — Si $a < 0$, la fonction $(s - a + 1)^{a-\frac{3}{2}}(s - 1)L(s)$ est bornée sur le demi-plan $\operatorname{Re}(s) \geq a$.

Démonstration. — Comme $\frac{3}{2} - a \geq 1$ et comme L est bornée sur le demi-plan $\operatorname{Re}(s) \geq 1 - a$, la fonction $F(s) = (s - a + 1)^{a-\frac{3}{2}}(s - 1)L(s)$ est bornée sur le demi-plan $\operatorname{Re}(s) \geq 1 - a$. Par ailleurs, la fonction F est bornée sur la droite $\operatorname{Re}(s) = a$, d'après le lemme A.3.4. Finalement, la fonction F est holomorphe dans la bande $a \leq \operatorname{Re}(s) \leq 1 - a$ et la propriété (L3) montre que, si $\varepsilon > 0$, alors $F(s)e^{\varepsilon s^2}$ tend vers 0 quand s tend vers l'infini dans cette bande, puisque $\operatorname{Re}(\varepsilon s^2) \sim -\varepsilon \operatorname{Im}(s)^2$ tend vers $-\infty$ beaucoup plus vite que $c(a, b)|\operatorname{Im}(s)|$. Le principe du maximum permet d'en déduire que, si T est assez grand, alors le maximum de $|F(s)e^{\varepsilon s^2}|$ sur le rectangle de sommets $a \pm iT$ et $1 - a \pm iT$ est atteint sur un des segments verticaux, et donc que le maximum de $|F(s)e^{\varepsilon s^2}|$ dans la bande est atteint sur la droite $\operatorname{Re}(s) = a$ ou sur la droite $\operatorname{Re}(s) = 1 - a$. En notant M le maximum de $|F|$ sur les deux droites $\operatorname{Re}(s) = a$ et $\operatorname{Re}(s) = 1 - a$, on en déduit que $|F(s)e^{\varepsilon s^2}| \leq M \sup(e^{\varepsilon a^2}, e^{\varepsilon(1-a)^2})$, quels que soient $\varepsilon > 0$, et s dans la bande $a \leq \operatorname{Re}(s) \leq 1 - a$. En faisant tendre ε vers 0, cela montre que $|F|$ est majorée par M sur la bande, ce qui permet de conclure.

Lemme A.3.7. — Il existe $C_1 > 0$ tel que $\sup_{s \in D(2+it, 12)} \left| \frac{L(s)}{L(2+it)} \right| \leq C_1 |t|^{21/2}$, si $t \in \mathbf{R}$ et $|t| \geq 15$.

Démonstration. — On déduit du lemme A.3.6 (avec $a = -10$) l'existence d'une constante C'_1 telle que

$$|L(s)| \leq C'_1 \frac{|s + 11|^{23/2}}{|s - 1|}, \quad \text{si } \operatorname{Re}(s) \geq -10.$$

Maintenant, si $s \in D(2 + it, 12)$, on a $|s + 11| \leq |t| + 25$, et $|s - 1| \geq |s| - 1 \geq |t| - 13$. Comme de plus, $|L(2 + it)^{-1}| \leq c(2)$, d'après la propriété (L1), on obtient, si $|t| > 13$,

$$\sup_{s \in D(2+it, 12)} \left| \frac{L(s)}{L(2 + it)} \right| \leq c(2) C'_1 \frac{(|t| + 25)^{23/2}}{|t| - 13}.$$

Le résultat s'en déduit sans problème.

4. La fonction $\frac{L'}{L}$ dans la bande critique

Lemme A.3.8. — Soient Ω un ouvert de \mathbf{C} , $R > 0$ tel que $D(0, R) \subset \Omega$, et f holomorphe sur Ω , avec $f(0) = 0$. Si $A = \sup_{|s|=R} \operatorname{Re}(f(s))$, alors

$$\left| \frac{f^{(k)}(s)}{k!} \right| \leq \frac{4AR}{(R - |s|)^{k+1}}, \quad \text{quels que soient } k \in \mathbf{N} \text{ et } s \in D(0, R^-).$$

Démonstration. — On a $F(s) = \sum_{n=1}^{+\infty} a_n s^n$ sur $D(0, R)$. En écrivant a_n sous la forme $|a_n|e^{i\theta_n}$, on obtient $\operatorname{Re}(f(Re^{i\theta})) = \sum_{m=1}^{+\infty} |a_m|R^m \cos(m\theta + \theta_m)$ et

$$\begin{aligned} \int_0^{2\pi} (1 + \cos(n\theta + \theta_n)) \operatorname{Re}(f(Re^{i\theta})) d\theta \\ = \sum_{m=1}^{+\infty} |a_m|R^m \int_0^{2\pi} (1 + \cos(n\theta + \theta_n)) \cos(m\theta + \theta_m) d\theta = \pi |a_n|R^n. \end{aligned}$$

Comme $0 \leq 1 + \cos(n\theta + \theta_n) \leq 2$, on en tire la majoration $\pi |a_n|R^n \leq 4\pi A$ et donc $|a_n| \leq 4AR^{-n}$. (En particulier, $A \geq 0$.) Maintenant, si $|s| < R$, on a

$$\left| \frac{f^{(k)}(s)}{k!} \right| = \left| \sum_{n=0}^{+\infty} \binom{n+k}{k} a_{n+k} s^n \right| \leq 4A \sum_{n=0}^{+\infty} \binom{n+k}{k} \frac{|s|^n}{R^{n+k}} = \frac{4AR}{(R - |s|)^{k+1}}.$$

Ceci permet de conclure.

Lemme A.3.9. — Soient Ω un ouvert de \mathbf{C} , $R > 0$ tel que $D(0, 3R) \subset \Omega$, et f holomorphe sur Ω , avec $f(0) = 1$. Soit $M = \sup_{s \in D(0, 3R)} |f(s)|$, et soit Y l'ensemble des zéros de f dans $D(0, R)$. Alors :

$$\sum_{\rho \in Y} v_\rho(f) \leq \frac{\log M}{\log 2} \quad \text{et} \quad \left| \frac{f'(s)}{f(s)} - \sum_{\rho \in Y} \frac{v_\rho(f)}{s - \rho} \right| \leq \frac{4R \log M}{(R - |s|)^2}, \quad \text{quel que soit } s \in D(0, R^-).$$

Démonstration. — Soit $g(s) = f(s) \prod_{\rho \in Y} (1 - s/\rho)^{-v_\rho(f)}$. Par construction, g ne s'annule pas sur $D(0, R)$, et $g(0) = 1$; il existe donc (cf. ex VI.1.7) un ouvert $\Omega' \subset \Omega$ contenant $D(0, R)$ et h holomorphe sur Ω' , avec $h(0) = 0$, tels que $g = e^h$ sur Ω' . Soit $N = \sum_{\rho \in Y} v_\rho(f)$. Comme $|1 - s/\rho| \geq 2$, si $|s| = 3R$, on a $|g(s)| \leq 2^{-N}M$, si $|s| = 3R$. Par le principe du maximum (rem. V.1.21) cela implique que $1 = |g(1)| \leq 2^{-N}M$, et donc que $N \leq \frac{\log M}{\log 2}$. Finalement, on a $|g(s)| \leq M$, si $|s| = 3R$ et donc, d'après le principe du maximum, $\sup_{|s|=R} |g(s)| \leq M$ et $\sup_{|s|=R} \operatorname{Re}(h(s)) \leq \log M$. Le lemme A.3.8 permet d'en déduire que $|h'(s)| \leq \frac{4R \log M}{(R - |s|)^2}$, si $|s| < R$, et comme $h'(s) = \frac{f'(s)}{f(s)} - \sum_{\rho \in Y} \frac{v_\rho(f)}{s - \rho}$, cela permet de conclure.

Si $n \in \mathbf{N}$, notons Z_n l'ensemble des zéros de L dans le disque $D(2 + in, 4)$.

Corollaire A.3.10. — Il existe des constantes C_2, C_3 , telles que, si $|n|$ est assez grand :

- (i) $\sum_{\rho \in Z_n} v_\rho(L) \leq C_2 \log |n|$;
- (ii) $\left| \frac{L'(s)}{L(s)} \right| \leq C_3 \log |n| + C_2 \log |n| (\inf_{\rho \in Z_n} |s - \rho|)^{-1}$, si $s \in D(2 + in, \sqrt{10})$.

Démonstration. — On applique le lemme A.3.9 à $f(s) = \frac{L(s+(2+in))}{L(2+in)}$, et $R = 4$; dans les notations de ce lemme, on peut prendre $M = C_1|n|^{21/2}$, d'après le lemme A.3.7. On en déduit que

$$\sum_{\rho \in Z_n} v_\rho(L) \leq \frac{\frac{21}{2} \log n + \log C_1}{\log 2} \quad \text{et} \quad \left| \frac{L'(s)}{L(s)} \right| \leq \frac{16(\frac{21}{2} \log n + \log C_1)}{4 - \sqrt{10}} + \sum_{\rho \in Z_n} \frac{v_\rho(L)}{|s - \rho|},$$

si $|s - (2 + in)| \leq \sqrt{10}$. Le résultat s'en déduit.

Corollaire A.3.11. — *Il existe $C_4 > 0$, tel que, si $n \in \mathbf{N}$ est assez grand, il existe $t_n^+ \in [n, n+1]$ et $t_n^- \in [-n-1, -n]$, tels que*

$$\sup_{s \in [2+it_n^+, -1+it_n^+]} \left| \frac{L'(s)}{L(s)} \right| \leq C_4(\log n)^2 \quad \text{et} \quad \sup_{s \in [-1+it_n^-, 2+it_n^-]} \left| \frac{L'(s)}{L(s)} \right| \leq C_4(\log n)^2.$$

Démonstration. — Comme $|Z_n| \leq C_2 \log n$, les segments $]\text{Im}(\rho) - \frac{1}{2C_2 \log n}, \text{Im}(\rho) + \frac{1}{2C_2 \log n}[$, pour $\rho \in Z_n$ ne recouvrent pas complètement $[n, n+1]$, puisque leur réunion est un ouvert de longueur ≤ 1 . Il existe donc t_n^+ tel que $|\text{Im}(\rho) - t_n^+| \geq \frac{1}{2C_2 \log n}$, quel que soit $\rho \in Z_n$. On a alors $\inf_{\rho \in Z_n} |s - \rho| \geq \inf_{\rho \in Z_n} |\text{Im}(s - \rho)| \geq \frac{1}{2C_2 \log n}$, quel que soit $s \in [2 + it_n^+, -1 + it_n^+]$. De même il existe $t_n^- \in [-n-1, -n]$ tel que $\inf_{\rho \in Z_{-n}} |s - \rho| \geq \frac{1}{2C_2 \log n}$, quel que soit $s \in [-1 + it_n^-, 2 + it_n^-]$. Le (ii) du lemme A.3.9 permet de conclure (avec $C_4 = 2C_2^2 + 1$ par exemple).

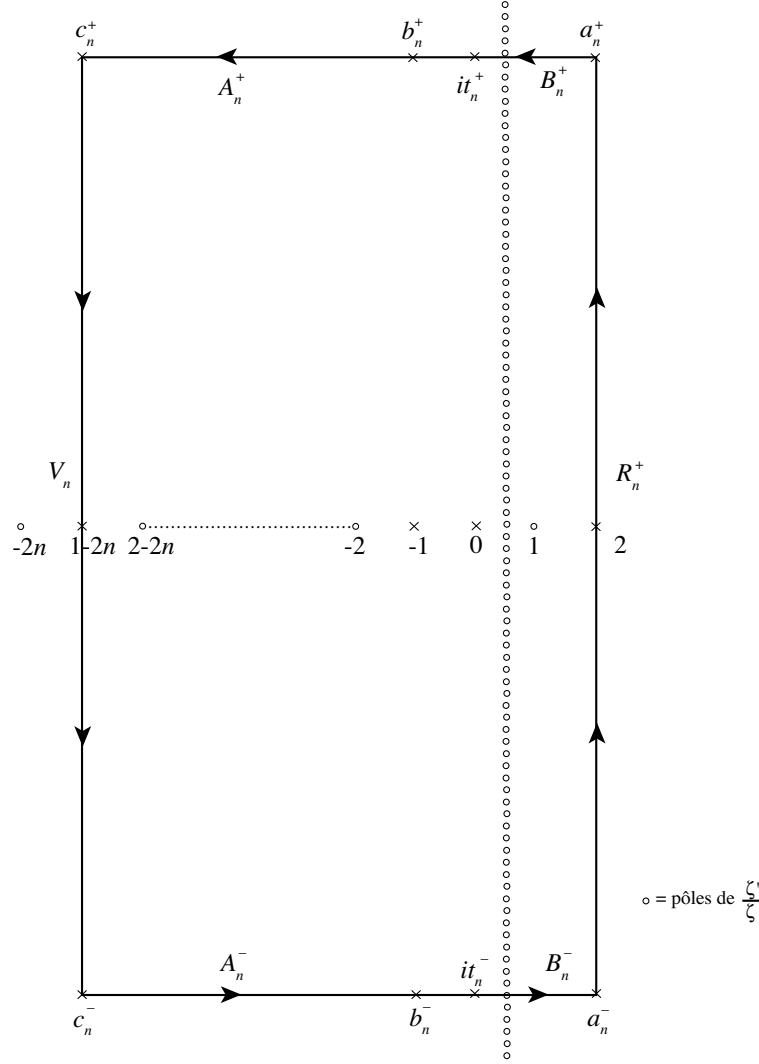
5. Conclusion

Passons à la démonstration du th. A.3.3. Pour prouver la convergence absolue de la série $\sum_{\rho \in Y(L)} \frac{v_\rho(L)}{\rho(\rho+1)}$, constatons que l'ensemble Z'_n des éléments de $Y(L)$ dont la partie imaginaire est comprise entre n et $n+1$ est inclus dans le disque $D(2+in, 4)$, et donc est un sous-ensemble de Z_n . Comme $\sum_{\rho \in Z_n} v_\rho(L) \leq C_2 \log |n|$, si n est assez grand, on déduit la convergence absolue de $\sum_{\rho \in Y(L)} \frac{v_\rho(L)}{\rho(\rho+1)}$ de celle de $\sum_{n=1}^{+\infty} \frac{\log n}{n^2}$.

Maintenant, choisissons, pour tout n assez grand, des réels t_n^+ et t_n^- vérifiant les conclusions du cor. A.3.11. Soient a_n^+, b_n^+, c_n^+ les points $2 + it_n^+, -1 + it_n^+, c+1 - 2n + it_n^+$, et a_n^-, b_n^-, c_n^- les points $2 + it_n^-, -1 + it_n^-, c+1 - 2n + it_n^-$. On note

- V_n le segment vertical $[c_n^+, c_n^-]$,
- A_n^+ et A_n^- les segments horizontaux $[b_n^+, c_n^+]$ et $[c_n^-, b_n^-]$,
- B_n^+ et B_n^- les segments horizontaux $[a_n^+, b_n^+]$ et $[b_n^-, a_n^-]$,
- R_n^- le chemin composé $B_n^+ \cdot A_n^+ \cdot V_n \cdot A_n^- \cdot B_n^-$,
- R_n^+ le segment vertical $[a_n^-, a_n^+]$.

Alors $R_n = R_n^+ \cdot R_n^-$ est le rectangle de sommets a_n^-, a_n^+, c_n^+ et c_n^- parcouru dans le sens direct. On note Y_n l'ensemble des pôles de F_x se trouvant à l'intérieur du rectangle R_n .

Le chemin R_n dans le cas de ζ

La formule des résidus nous donne alors

$$\int_{R_n} F_x(s) ds = \sum_{\rho \in Y_n} \text{Res}(F_x, \rho).$$

Maintenant, $\int_{R_n} F_x(s) ds = \int_{R_n^+} F_x(s) ds + \int_{B_n^+} F_x(s) ds + \int_{A_n^+ \cdot V_n \cdot A_n^-} F_x(s) ds + \int_{B_n^-} F_x(s) ds$.

- $\int_{R_n^+} F_x(s) ds$ tend vers $\int_{2-i\infty}^{2+i\infty} \frac{-L'(s)}{L(s)} \frac{x^{s+1}}{s(s+1)} ds$.
- Sur B_n^+ et B_n^- , $F_x(s)$ est, d'après le cor. A.3.11, majorée en valeur absolue, si $x \geq 1$, par $C_4(\log n)^2 \frac{x^3}{|\text{Im}(s)| \cdot |\text{Im}(s+1)|} \leq \frac{x^3 C_4 (\log n)^2}{n^2}$. Donc $\int_{B_n^+} F_x(s) ds$ et $\int_{B_n^-} F_x(s) ds$ sont majorés en valeur absolue par $3 \frac{x^3 C_4 (\log n)^2}{n^2}$ et tendent vers 0 quand n tend vers $+\infty$.

• Sur $A_n^+ \cdot V_n \cdot A_n^-$, $F_x(s)$ est, d'après le lemme A.3.5, majorée en valeur absolue, si $x \geq 1$, par $(C_0 + \log n) \frac{x^3}{|s(s+1)|} \leq (C_0 + \log n) \frac{x^3}{n^2}$, et comme la longueur de $A_n^+ \cdot V_n \cdot A_n^-$ est $2(2n-2) + t_n^+ - t_n^- \leq 6n-2$, on en déduit que $\int_{A_n^+ \cdot V_n \cdot A_n^-} F_x(s) ds$ tend vers 0 quand n tend vers $+\infty$.

On en déduit le théorème en faisant tendre n vers $+\infty$, et en remarquant que l'ensemble des pôles de F_x est la réunion croissante de Y_n , et donc que $\sum_{\rho \in Y_n} \text{Res}(F_x, \rho)$ tend vers $\sum_{\text{Re}(\rho) < 2} \text{Res}(F_x, \rho)$.

Exercice A.3.12. — On note $\mathcal{N}(T)$ le nombre de zéros s de la fonction ζ dans la bande critique, vérifiant $0 \leq \text{Im}(s) \leq T$. Montrer que $\mathcal{N}(T) = \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T)$. (Intégrer $\frac{\xi'(s)}{\xi(s)} ds$ sur le rectangle de sommets $2+i$, a_n^+ , b_n^+ et $-1+i$, où $\xi(s) = \frac{\Gamma(s/2)}{\pi^{s/2}} \zeta(s)$, et utiliser l'équation fonctionnelle satisfaite par ξ (rem. VII.3.7 ou ex. VII.6.6), la formule de Stirling dans le plan complexe (prop. VII.2.3) et les majorations du cor. A.3.11.)

A.4. Démonstration du théorème des nombres premiers

1. Non annulation sur la droite $\text{Re}(s) = 1$

Lemme A.4.1. — Si $\alpha, \beta \in \mathbf{C}$, et si $|z| < \inf(1, |\alpha|^{-1}, |\beta|^{-1}, |\alpha\beta|^{-1})$, alors

$$\frac{1 - \alpha\beta z^2}{(1-z)(1-\alpha z)(1-\beta z)(1-\alpha\beta z)} = \sum_{n=0}^{+\infty} (1 + \alpha + \cdots + \alpha^n)(1 + \beta + \cdots + \beta^n) z^n.$$

Démonstration. — Si on multiplie par $(1-z)(1-\alpha z)(1-\beta z)(1-\alpha\beta z)$ les deux membres de l'identité à démontrer, on obtient des séries du type $\sum_{n=0}^{+\infty} P_n(\alpha, \beta) z^n$ et $\sum_{n=0}^{+\infty} Q_n(\alpha, \beta) z^n$, où les P_n et Q_n sont des polynômes (en deux variables). On est ramené à prouver que $P_n = Q_n$ pour tout n . Or pour prouver que $P_n = Q_n$, il suffit de prouver que $P_n = Q_n$ sur un ouvert, le reste s'en déduisant par prolongement analytique. On peut donc se restreindre au cas où $1, \alpha, \beta$ et $\alpha\beta$ sont tous distincts. Dans ce cas, la fraction rationnelle $F(z) = \frac{1-\alpha\beta z^2}{(1-z)(1-\alpha z)(1-\beta z)(1-\alpha\beta z)}$ n'a que des pôles simples, ce qui rend sa décomposition en éléments simples particulièrement facile à calculer. Soient

$$\begin{aligned} a_1 &= \lim_{z \rightarrow 1} (1-z)F(z) = \frac{1-\alpha\beta}{(1-\alpha)(1-\beta)(1-\alpha\beta)} = \frac{1}{(1-\alpha)(1-\beta)}, \\ a_\alpha &= \lim_{z \rightarrow \alpha^{-1}} (1-\alpha z)F(z) = \frac{1-\alpha^{-1}\beta}{(1-\alpha^{-1})(1-\alpha^{-1}\beta)(1-\beta)} = -\frac{\alpha}{(1-\alpha)(1-\beta)}, \\ a_\beta &= \lim_{z \rightarrow \beta^{-1}} (1-\beta z)F(z) = \frac{1-\alpha\beta^{-1}}{(1-\beta^{-1})(1-\alpha\beta^{-1})(1-\alpha)} = -\frac{\beta}{(1-\alpha)(1-\beta)}, \\ a_{\alpha\beta} &= \lim_{z \rightarrow \alpha^{-1}\beta^{-1}} (1-\alpha\beta z)F(z) = \frac{1-\alpha^{-1}\beta^{-1}}{(1-\alpha^{-1}\beta^{-1})(1-\beta^{-1})(1-\alpha^{-1})} = \frac{\alpha\beta}{(1-\alpha)(1-\beta)}. \end{aligned}$$

Alors

$$\begin{aligned}
 F(z) &= \frac{a_1}{1-z} + \frac{a_\alpha}{1-\alpha z} + \frac{a_\beta}{1-\beta z} + \frac{a_{\alpha\beta}}{1-\alpha\beta z} \\
 &= \frac{1}{(1-\alpha)(1-\beta)} \left(\frac{1}{1-z} - \frac{\alpha}{1-\alpha z} - \frac{\beta}{1-\beta z} + \frac{\alpha\beta}{1-\alpha\beta z} \right) \\
 &= \frac{1}{(1-\alpha)(1-\beta)} \sum_{n=0}^{+\infty} (1 - \alpha^{n+1} - \beta^{n+1} + (\alpha\beta)^{n+1}) z^n \\
 &= \sum_{n=0}^{+\infty} \frac{(1 - \alpha^{n+1})(1 - \beta^{n+1})}{(1-\alpha)(1-\beta)} z^n = \sum_{n=0}^{+\infty} (1 + \alpha + \dots + \alpha^n)(1 + \beta + \dots + \beta^n) z^n.
 \end{aligned}$$

Proposition A.4.2. — Si $t \in \mathbf{R}$, alors $F(s) = \zeta(2s)^{-1} \zeta(s)^2 \zeta(s+it) \zeta(s-it)$ est une série de Dirichlet à coefficients positifs.

Démonstration. — En utilisant la factorisation de ζ en facteurs d'Euler, on obtient

$$F(s) = \prod_{p \in \mathcal{P}} \frac{(1 - p^{-2s})}{(1 - p^{-s})^2 (1 - p^{-s-it}) (1 - p^{-s+it})}.$$

On peut alors utiliser le lemme A.4.1 avec $z = p^{-s}$, $\alpha = p^{it}$ et $\beta = p^{-it}$ pour obtenir la formule

$$F(s) = \prod_{p \in \mathcal{P}} \left(\sum_{k=1}^{+\infty} |1 + p^{it} + \dots + p^{ikt}|^2 p^{-ks} \right),$$

ce qui permet de conclure.

Théorème A.4.3. — La fonction ζ ne s'annule pas sur la droite $\operatorname{Re}(s) = 1$.

Démonstration. — Supposons, par l'absurde qu'il existe $t \in \mathbf{R}$, avec $\zeta(1+it) = 0$. On a alors $\zeta(1-it) = \overline{\zeta(1+it)} = 0$. La fonction $\zeta(s+it)\zeta(s-it)$ a donc un zéro double en $s = 1$, ce qui fait que $F(s) = \zeta(2s)^{-1} \zeta(s)^2 \zeta(s+it)\zeta(s-it)$ est holomorphe en $s = 1$, le zéro double contrebalançant le pôle double de $\zeta(s)^2$. De même, les zéros de $\zeta(s)$ en $1+it$ et $1-it$ contrebalancent les pôles de $\zeta(s+it)$ en $1-it$ et $\zeta(s-it)$ en $1+it$. On en déduit que $F(s)$ est holomorphe dans le demi-plan $\operatorname{Re}(s) > \frac{1}{2}$. De plus, comme $\zeta(2s)^{-1}$ est holomorphe au voisinage de $s = \frac{1}{2}$, et comme F est à coefficients positifs, il suit du théorème de Landau (th. VII.1.1) que l'abscisse de convergence absolue de F est $< \frac{1}{2}$. On aboutit à une contradiction car, d'une part $F(\frac{1}{2}) = 0$ puisque $\zeta(2s)^{-1} = 0$, et d'autre part $F(\frac{1}{2}) > 0$ comme somme d'une série à termes positifs non tous nuls. Ceci permet de conclure.

Exercice A.4.4. — Soit χ un caractère de Dirichlet de conducteur D .

(i) Montrer que $\zeta_D(2s)^{-1}\zeta_D(s)^2L(\chi, s+it)L(\bar{\chi}, s-it)$ est une série de Dirichlet à coefficients positifs (ζ_D désigne la fonction ζ privée de ses facteurs d'Euler en les p divisant D , i.e. $\zeta_D(s) = \prod_{p \in \mathcal{P}, p \nmid D} \frac{1}{1-p^{-s}}$).

(ii) En déduire que $L(\chi, s)$ ne s'annule pas sur la droite $\operatorname{Re}(s) = 1$.

2. Conclusion

D'après le lemme A.2.1, pour démontrer le théorème des nombres premiers, il suffit de prouver que $\psi_1(x) \sim \frac{x^2}{2}$. Maintenant, comme ζ vérifie les propriétés (L1)–(L3) (cf. lemme A.3.1), le cor. A.2.6 et le th. A.3.3 nous fournissent la formule explicite

$$\psi_1(x) = \frac{x^2}{2} - \frac{\zeta'(0)}{\zeta(0)}x + \frac{\zeta'(-1)}{\zeta(-1)} - \sum_{\rho \in Y(\zeta)} \frac{v_\rho(\zeta)x^{\rho+1}}{\rho(\rho+1)} - \sum_{r=1}^{+\infty} \frac{x^{1-2r}}{2r(2r-1)}.$$

Il en résulte que « $\psi_1(x) \sim \frac{x^2}{2}$ » est équivalent à « $\lim_{x \rightarrow +\infty} \sum_{\rho \in Y(\zeta)} \frac{x^{\rho-1}}{\rho(\rho+1)} = 0$ ». Comme ζ ne s'annule pas sur la droite $\operatorname{Re}(s) = 1$, on a $\operatorname{Re}(\rho-1) < 0$, et donc $\lim_{x \rightarrow +\infty} \frac{x^{\rho-1}}{\rho(\rho+1)} = 0$, quel que soit $\rho \in Y(\zeta)$. Comme de plus $|\frac{v_\rho(\zeta)x^{\rho-1}}{\rho(\rho+1)}| \leq |\frac{v_\rho(\zeta)}{\rho(\rho+1)}|$, et comme la somme des $|\frac{v_\rho(\zeta)}{\rho(\rho+1)}|$ est convergente, on peut utiliser le théorème de convergence dominée pour les séries pour intervertir sommes et limites, ce qui permet de conclure.

Exercice A.4.5. — Adapter les arguments ci-dessus pour démontrer le théorème de la progression arithmétique.

Exercice A.4.6. — On se propose de montrer, en partant de la formule explicite ci-dessus, et en utilisant la convergence absolue de la série $\sum_{\rho \in Y(\zeta)} \frac{v_\rho(\zeta)}{\rho(\rho+1)}$, que le théorème des nombres premiers implique la non annulation de ζ sur la droite $\operatorname{Re}(s) = 1$. On numérote les zéros de ζ sur la droite $\operatorname{Re}(s) = 1$ sous la forme $\rho = 1 + i\tau_n$, avec $n \in \mathbb{I}$, \mathbb{I} sous-ensemble de \mathbb{N} , et on pose $a_n = \frac{v_{1+i\tau_n}(\zeta)}{(1+i\tau_n)(2+i\tau_n)}$.

(i) Montrer que, si $\psi_1(x) \sim \frac{x^2}{2}$, alors $\lim_{x \rightarrow +\infty} \sum_{n \in \mathbb{I}} a_n x^{i\tau_n} = 0$.

(ii) En déduire que $\lim_{U \rightarrow +\infty} \frac{1}{U} \int_0^U |f(u)|^2 du = 0$, où $f(u) = \sum_{n \in \mathbb{I}} a_n e^{i\tau(n)u}$.

(iii) Exprimer $\lim_{U \rightarrow +\infty} \frac{1}{U} \int_0^U |f(u)|^2 du = 0$ en fonction des a_n .

(iv) Conclure.

A.5. Compléments

1. L'hypothèse de Riemann et ses conséquences

Si on suppose que l'hypothèse de Riemann est vraie, tous les ρ apparaissant dans la démonstration ci-dessus ont une partie réelle égale à $\frac{1}{2}$, ce qui permet de majorer $|\frac{v_\rho(\zeta)x^{\rho+1}}{\rho(\rho+1)}|$ par $x^{3/2}|\frac{v_\rho(\zeta)}{\rho(\rho+1)}|$, et comme la série $\sum_{\rho \in Y(\zeta)} \frac{v_\rho(\zeta)}{\rho(\rho+1)}$ est absolument convergente, on en déduit que $\psi_1(x) = \frac{1}{2}x^2 + O(x^{3/2})$.

Un peu plus de travail montre, en partant de l'encadrement $\psi_1(x) - \psi_1(x-1) \leq \psi(x) \leq \psi(x+1) - \psi(x)$, et de la formule explicite pour $\psi_1(x)$, que l'on a $\psi(x) = x + O(x^{1/2}(\log x)^2)$.

Finalement, en utilisant l'exercice A.5.2 ci-dessous, on en déduit le résultat suivant, qui montre que l'hypothèse de Riemann a des implications profondes sur la répartition des nombres premiers.

Proposition A.5.1. — *Si l'hypothèse de Riemann est vraie, alors*

$$\pi(x) = \int_2^x \frac{1}{\log t} dt + O(x^{1/2} \log x).$$

Exercice A.5.2. — On suppose l'hypothèse de Riemann vraie, ce que l'on utilisera sous la forme $\psi(x) = x + O(x^{1/2}(\log x)^2)$. Soient $A(x) = \sum_{p \leq x} \log p$ et $B(x) = A(x) - x$.

(i) Montrer que $\psi(x) - A(x) = O(x^{1/2} \log x)$; en déduire qu'il existe $C > 0$ tel que $B(x) \leq C x^{1/2}(\log x)^2$, si $x \geq 2$.

(ii) Montrer que $\pi(x) = \sum_{2 \leq n \leq x} \frac{A(n) - A(n-1)}{\log n}$.

(iii) En déduire que

$$\pi(x) = \left(\sum_{2 \leq n \leq x} \frac{1}{\log n} \right) + \frac{B([x])}{\log([x] + 1)} + \left(\sum_{n=2}^{[x]} B(n) \left(\frac{1}{\log n} - \frac{1}{\log(n+1)} \right) \right).$$

(iv) Conclure.

2. L'hypothèse de Riemann et la fonction M de Mertens

On rappelle que l'on note μ la fonction de Moebius définie par

$$\sum_{n=1}^{+\infty} \frac{\mu(n)}{n^s} = \zeta(s)^{-1} = \prod_{p \in \mathcal{P}} (1 - p^{-s}),$$

si $\operatorname{Re}(s) > 0$. On a donc $\mu(n) = 0$ si n est divisible par le carré d'un nombre premier et $\mu(n) = (-1)^r$, si $n = p_1 \cdots p_r$, où les p_i sont des nombres premiers distincts. La fonction M de Mertens est définie par $M(x) = \sum_{n \leq x} \mu(n)$, et F. Mertens a conjecturé que $|M(x)| \leq \sqrt{x}$, si $x > 1$. Cette conjecture, si elle avait le bon goût d'être vraie, impliquerait l'hypothèse de Riemann. En effet, la formule de sommation d'Abel nous donne

$$\zeta(s)^{-1} = \sum_{n=1}^{+\infty} M(n)(n^{-s} - (n+1)^{-s}),$$

et l'hypothèse $|M(n)| \leq \sqrt{n}$ implique que la série est absolument convergente pour $\operatorname{Re}(s) > \frac{1}{2}$, et donc que ζ ne s'annule pas pour $\operatorname{Re}(s) > \frac{1}{2}$.

On ne connaît aucun contreexemple explicite à la conjecture de Mertens, mais celle-ci n'est pas très raisonnable à cause de la loi du logarithme itéré (Khinchine 1924), selon

laquelle, si $a_n \in \{\pm 1\}$ et $A_n = \sum_{k=1}^n a_k$, alors presque sûrement

$$\limsup \frac{A_n}{\sqrt{2n \log \log n}} = 1 \quad \text{et} \quad \liminf \frac{A_n}{\sqrt{2n \log \log n}} = -1.$$

De fait, A. Odlyzko et H. te Riele ont démontré en 1985 que la conjecture de Mertens est fautive et on sait qu'il existe des contre-exemples plus petits que $3,21 \cdot 10^{64}$. Par contre, la même loi du logarithme itéré, ou le résultat de Hausdorff mentionné dans la note 4 du chap. VII rendent l'hypothèse de Riemann tout à fait plausible. D'un autre côté, le résultat de Odlyzko et te Riele relativise quelque peu les confirmations numériques de l'hypothèse de Riemann (on a vérifié que les 10^{13} premiers zéros de la fonction ζ dans la bande critique sont effectivement sur la droite $\operatorname{Re}(s) = \frac{1}{2}$).

3. L'hypothèse de Lindelöf

C'est une conjecture impliquée par l'hypothèse de Riemann, mais qui est a priori plus faible. Elle est équivalente à un des énoncés suivants (il faut travailler pas mal pour démontrer l'équivalence de ces énoncés).

- On a $|\zeta(\frac{1}{2} + it)| = O(t^\varepsilon)$ au voisinage de $+\infty$, quel que soit $\varepsilon > 0$.
- Si $\sigma > \frac{1}{2}$, et si $\mathcal{N}(\sigma, T)$ est le nombre de zéros de la fonction ζ vérifiant $\operatorname{Re}(s) \geq \sigma$ et $0 \leq \operatorname{Im}(s) \leq T$, alors $\mathcal{N}(\sigma, T+1) - \mathcal{N}(\sigma, T) = o(\log T)$ quand $T \rightarrow +\infty$.
- $\sum_{n \leq t} n^{-(\frac{1}{2} + 2i\pi t)} = O(t^\varepsilon)$, au voisinage de $+\infty$, quel que soit $\varepsilon > 0$.

Le deuxième de ces énoncés, que l'on peut comparer avec l'ex. A.3.12, est une conséquence de l'hypothèse de Riemann qui peut s'énoncer sous la forme $\mathcal{N}(\sigma, T) = 0$, quels que soient $\sigma > \frac{1}{2}$ et $T \geq 0$. Le troisième ne fait intervenir que des sommes finies et donc a l'air parfaitement innocent.

Contrairement à l'hypothèse de Riemann, il y a des progrès à intervalles réguliers concernant l'hypothèse de Lindelöf. Par exemple, G. Kolesnik a prouvé en 1982 que

$$\zeta(1/2 + it) = O(t^{1/6 - 1/216 + \varepsilon}), \quad \text{quel que soit } \varepsilon > 0.$$

F. Bombieri et H. Iwaniec ont démontré en 1986 que

$$\zeta(1/2 + it) = O(t^{1/6 - 1/28 + \varepsilon}), \quad \text{quel que soit } \varepsilon > 0.$$

ANNEXE B

GROUPES FINIS ET REPRÉSENTATIONS : EXEMPLES

B.1. p -Groupes

1. Généralités sur les p -groupes

Soit p un nombre premier. Un p -groupe est un groupe d'ordre une puissance de p .

Proposition B.1.1. — (i) Si G est un p -groupe, alors le centre de G est non trivial.

(ii) Si G est un p -groupe non commutatif, alors G contient un sous-groupe commutatif distingué contenant strictement le centre de G .

Démonstration. — (i) On fait agir G sur lui même par conjugaison. Les orbites sont alors les classes de conjugaison et le centre Z est l'ensemble des $c \in G$ dont la classe de conjugaison est réduite à un point. On sait que Z contient l'élément neutre, et notre problème est de montrer qu'il existe d'autres classes de conjugaison réduites à un point. Maintenant, si O est une classe de conjugaison, et si $x \in O$, alors $O = G/Z_x$, où Z_x est l'ensemble des éléments de G commutant à x ; on a alors $|O| = \frac{|G|}{|Z_x|}$, ce qui prouve que $|O|$ est une puissance de p ; en particulier, $|O|$ est divisible par p sauf si $|O| = 1$. Comme $|G|$ est divisible par p , et est aussi la somme des cardinaux des orbites, on en déduit que l'ensemble des orbites réduites à un point a un cardinal divisible par p . Autrement dit, $|Z|$ est divisible par p , et comme $|Z| \geq 1$, on a $|Z| \geq p$, ce qui démontre le (i).

(ii) Supposons maintenant G non commutatif (et donc $G \neq Z$), et soit $H = G/Z$. Alors H est un p -groupe, et son centre Y est non trivial. Soit y un élément du centre de H distinct de l'élément neutre, soit $\langle y \rangle$ le sous-groupe de Y engendré par y , et soit A l'image inverse de $\langle y \rangle$ dans G . Comme $\langle y \rangle$ est distingué (car central) dans H , cela implique que A est un sous-groupe distingué de G , et comme A contient strictement Z , il suffit de vérifier que A est commutatif. Soit a l'ordre de y , et soit $\tilde{y} \in A$ ayant pour image y dans H . Alors les éléments de $\langle y \rangle$ sont les y^i , avec $0 \leq i \leq a-1$, et tout élément de A peut s'écrire sous la forme $\tilde{y}^i z$, avec $0 \leq i \leq a-1$ et $z \in Z$. Si $x_1 = \tilde{y}^{i_1} z_1$ et $x_2 = \tilde{y}^{i_2} z_2$, on a, en tenant compte du fait que les éléments de Z commutent à tout,

$$x_1 x_2 = \tilde{y}^{i_1} z_1 \tilde{y}^{i_2} z_2 = \tilde{y}^{i_1} \tilde{y}^{i_2} z_1 z_2 = \tilde{y}^{i_1+i_2} z_2 z_1 = \tilde{y}^{i_2} z_2 \tilde{y}^{i_1} z_1 = x_2 x_1,$$

ce qui permet de conclure.

2. Représentations des p -groupes

Remarque B.1.2. — Soit G un groupe fini. Si A est un sous-groupe distingué de G , si W est une représentation de A , et si $g \in G$, on peut fabriquer une représentation W^g de A en posant $\rho_{W^g}(a) = \rho_W(g^{-1}ag)$, ce qui a un sens puisque $g^{-1}ag \in A$ si $a \in A$ et $g \in G$, par définition d'un sous-groupe distingué. Il est immédiat que W^g est irréductible si et seulement si W l'est, et on obtient de la sorte une action $(g, \chi) \mapsto \chi^g$ de G sur $\text{Irr}(A)$. [Si $\chi \in \text{Irr}(A)$ est un caractère irréductible, on a $\chi^g(a) = \chi(g^{-1}ag)$.]

Proposition B.1.3. — *Soit G un groupe fini, et soit A un sous-groupe distingué de G . Si V est une représentation irréductible de G dont la restriction $\text{Res}_G^A V$ à A n'est pas isotypique, alors il existe un sous-groupe H de G contenant A et strictement inclus dans G , et une représentation irréductible W de H , tels que $V = \text{Ind}_H^G W$.*

Démonstration. — Décomposons $\text{Res}_G^A V$ en la somme directe $\bigoplus_{\chi \in X} V_\chi$, avec $X \subset \text{Irr}(A)$, de ses composantes isotypiques. Par hypothèse, on a $|X| \geq 2$. Si $\chi \in X$, si $g \in G$, et si $a \in A$, on a $a \cdot (g \cdot V_\chi) = g \cdot (g^{-1}ag \cdot V_\chi)$, ce qui prouve que $g \cdot V_\chi$ est stable par a , et que $\rho_{g \cdot V_\chi}(a) = \rho_{V_\chi}(g^{-1}ag)$. On en déduit que G envoie la composante isotypique V_χ sur la composante isotypique V_{χ^g} , et donc permute les éléments de X .

Soit $\chi_0 \in X$, et soient W la composante isotypique de $\text{Res}_G^A V$ correspondant à χ_0 , et $H \subset G$ le stabilisateur de χ_0 . Alors W est stable par H et peut donc être considérée comme une représentation de H . Maintenant $\sum_{g \in G} g \cdot W = \bigoplus_{g \in G/H} V_{\chi_0^g}$ est stable par G , et donc est égal à V puisqu'on a supposé V irréductible. On en déduit que $X = G/H$, et donc que $H \neq G$, et que V et $\text{Ind}_H^G W$ ont même dimension $|X| \cdot \dim W$. Par ailleurs, $\text{Hom}_H(W, V)$ est non nul, et donc $\text{Hom}_G(\text{Ind}_H^G W, V)$ est aussi non nul (ex. I.3.13). Comme V est irréductible, et comme $\text{Ind}_H^G W$ et V ont même dimension, tout élément non nul de $\text{Hom}_G(\text{Ind}_H^G W, V)$ induit un isomorphisme de $\text{Ind}_H^G W$ sur V . Ceci permet de conclure.

Proposition B.1.4. — *Si G est un p -groupe, et si V est une représentation irréductible de G , alors il existe un sous-groupe H de G et un caractère linéaire $\chi \in \widehat{H}$, tels que $V = \text{Ind}_H^G \chi$.*

Démonstration. — La démonstration se fait par récurrence sur $|G|$. Si G est commutatif, alors toute représentation irréductible est de dimension 1, et il n'y a rien à démontrer. Supposons donc G non commutatif, et soit V une représentation irréductible de G .

- Si le noyau N de ρ_V est non trivial, alors ρ_V se factorise à travers $G' = G/N$, et on peut appliquer l'hypothèse de récurrence à G' : il existe $H' \subset G'$ et $\chi \in \widehat{H'}$ tel que $V = \text{Ind}_{H'}^{G'} \chi$, en tant que représentation de G' . Si H est l'image inverse de H' dans G , on peut voir χ comme un élément de \widehat{H} en composant avec la projection $H \rightarrow H'$, et on a $V = \text{Ind}_H^G \chi$, ce qui permet de conclure dans ce cas.

• Si le noyau de ρ_V est trivial, considérons un sous-groupe commutatif A de G , distingué, et non contenu dans le centre de G (cf. (ii) de la prop. B.1.1). En particulier, il existe $a \in A$ et $g \in G$, avec $ag \neq ga$, et comme le noyau de ρ_V est trivial, on a $\rho_V(a)\rho_V(g) \neq \rho_V(g)\rho_V(a)$. Ceci implique que $\rho_V(a)$ n'est pas une homothétie et, A étant commutatif, que $\text{Res}_G^A V$ n'est pas isotypique. On en déduit, en utilisant la prop. B.1.3, l'existence d'un sous-groupe H de G , contenant A et distinct de G , et d'une représentation irréductible W de H , tels que $V = \text{Ind}_H^G W$. L'hypothèse de récurrence appliquée à H nous fournit un sous-groupe H' de H et $\chi \in \widehat{H'}$, tels que $W = \text{Ind}_{H'}^H \chi$. On a donc

$$V = \text{Ind}_H^G (\text{Ind}_{H'}^H \chi) = \text{Ind}_{H'}^G \chi,$$

ce qui permet de conclure.

Exercice B.1.5. — Soient p un nombre premier et G un groupe non commutatif d'ordre p^3 . Montrer que G a p^2 représentations irréductibles de dimension 1 et $p - 1$ de dimension p .

B.2. Représentations du groupe symétrique S_n

La théorie des représentations de S_n fait intervenir une combinatoire assez amusante, qui est loin d'être triviale. Nous nous contenterons de décrire les résultats, et de renvoyer le lecteur à des ouvrages plus spécialisés (comme le livre de Fulton et Harris cité dans l'introduction) pour les démonstrations (qu'il peut aussi voir comme une série de défis...).

1. Partitions de n et représentations de S_n

Soit n un entier ≥ 1 . Une *partition* $\ell = (\ell_1, \dots, \ell_r)$ de n est une décomposition de n sous la forme $n = \ell_1 + \dots + \ell_r$, avec $\ell_1 \geq \ell_2 \geq \dots \geq \ell_r \geq 1$. Le nombre de partitions⁽¹⁾ de n est traditionnellement noté $p(n)$. Par exemple, les partitions de 1, 2, 3, 4, 5 et 6 sont

⁽¹⁾ L'étude de la fonction $p(n)$ a donné lieu à de nombreux travaux; la plupart ont pour point de départ la formule

$$\sum_{n=1}^{+\infty} p(n) T^n = \prod_{\ell=1}^{+\infty} \frac{1}{1 - T^\ell},$$

qui se démontre en remarquant que $\frac{1}{1-T} = \sum_{i=0}^{+\infty} T^i$, et en développant brutalement le second membre. Cette formule fait intervenir l'inverse de la fonction η de Dedekind définie par $\eta(q) = q^{1/24} \prod_{\ell=1}^{+\infty} (1 - q^\ell)$, avec $q = e^{2i\pi z}$ et z variant dans le demi-plan de Poincaré. En utilisant les propriétés de la fonction η , qui se trouve être une forme modulaire de poids $\frac{1}{2}$, on peut par exemple obtenir l'équivalent asymptotique suivant de $p(n)$: on a $p(n) \sim \frac{1}{4\sqrt{3n}} e^{\pi\sqrt{2n/3}}$, ce qui montre que $p(n)$ croît assez vite avec n .

données par

1	4	5	6
	3 + 1	4 + 1	5 + 1
2	2 + 2	3 + 2	4 + 2
1 + 1	2 + 1 + 1	3 + 1 + 1	4 + 1 + 1
	1 + 1 + 1 + 1	2 + 2 + 1	3 + 3
3		2 + 1 + 1 + 1	3 + 2 + 1
2 + 1		1 + 1 + 1 + 1 + 1	3 + 1 + 1 + 1
1 + 1 + 1			2 + 2 + 2
			2 + 2 + 1 + 1
			2 + 1 + 1 + 1 + 1
			1 + 1 + 1 + 1 + 1 + 1

et on a $p(1) = 1$, $p(2) = 2$, $p(3) = 3$, $p(4) = 5$, $p(5) = 7$ et $p(6) = 11$.

On munit l'ensemble des partitions de n de l'ordre *lexicographique* défini par $\ell = (\ell_1, \dots, \ell_r) > \mathbf{k} = (k_1, \dots, k_s)$ si et seulement si $\ell_i > k_i$, si i est le plus petit indice tel que $\ell_i \neq k_i$. Les partitions de 1, 2, 3, 4, 5 et 6 données ci-dessus sont rangées en ordre décroissant pour l'ordre lexicographique. (C'est le meilleur moyen pour ne pas en oublier.)

Il y a une bijection naturelle $\ell \mapsto C_\ell$ entre les partitions de n et les classes de conjugaison de S_n : si $\ell = (\ell_1, \dots, \ell_r)$, alors C_ℓ est la classe de conjugaison de S_n constituée des permutations σ dont la décomposition en cycles est (ℓ_1, \dots, ℓ_r) (ce qui, rappelons-le, signifie que σ est le produit de r cycles τ_1, \dots, τ_r de longueurs respectives ℓ_1, \dots, ℓ_r , et dont les supports sont disjoints).

De manière remarquable, il y a aussi une bijection naturelle entre les partitions de n et les représentations irréductibles de S_n . Si $\lambda = (\lambda_1, \dots, \lambda_s)$ est une partition de n , soit $S_\lambda \cong S_{\lambda_1} \times \dots \times S_{\lambda_s}$ le sous-groupe de S_n constitué des permutations laissant globalement stable chacun des sous-ensembles

$$\{1, \dots, \lambda_1\}, \{\lambda_1 + 1, \dots, \lambda_1 + \lambda_2\}, \dots, \{\lambda_1 + \dots + \lambda_{s-1} + 1, \dots, \lambda_1 + \dots + \lambda_s\}$$

de $\{1, \dots, n\}$, et soit $U_\lambda = \text{Ind}_{S_\lambda}^{S_n} \mathbf{1}$. Par exemple :

- Si $\lambda = (n)$, alors U_λ est la représentation triviale.
- Si $\lambda = (n-1, 1)$, alors U_λ est la représentation standard de S_n sur \mathbf{C}^n obtenue en permutant les éléments de la base canonique ; elle se décompose sous la forme $\mathbf{1} \oplus V_{(n-1,1)}$, où $\mathbf{1}$ est la droite de \mathbf{C}^n engendrée par $(1, \dots, 1)$ et $V_{(n-1,1)}$ est l'hyperplan d'équation $\sum_{i=1}^n x_i = 0$.
- Si $\lambda = (1, \dots, 1)$, on a $S_\lambda = \{1\}$, et donc $U_{(1, \dots, 1)}$ est la représentation régulière de S_n .

Comme le montre déjà l'exemple de $U_{(n-1,1)}$, la représentation U_λ n'est pas irréductible, mais sa décomposition en représentations irréductibles fait apparaître une unique

représentation irréductible V_λ qui n'apparaît pas déjà dans les U_μ , pour $\mu > \lambda$. On peut d'ailleurs donner une construction directe de V_λ (voir plus bas).

2. Diagrammes de Young et représentations de S_n

Si $\lambda = (\lambda_1, \dots, \lambda_r)$ est une partition de n , on note Y_λ le *diagramme de Young attaché à λ* : c'est le sous-ensemble du carré $n \times n$ obtenu en prenant les λ_1 premières cases de la 1^{ère} ligne, les λ_2 premières cases de la 2^{ème} ligne, ..., les λ_r premières cases de la r -ième ligne. Si $\lambda^* = (\lambda_1^*, \dots, \lambda_s^*)$ est la partition de n *conjuguée de λ* , obtenue en définissant λ_i^* comme le nombre de λ_j qui sont $\geq i$ (et donc $\lambda_1^* = r$, et $s = \lambda_1$), alors Y_λ est aussi obtenu en prenant les λ_1^* premières cases de la 1^{ère} colonne, les λ_2^* premières cases de la 2^{ème} colonne, ... En particulier, les diagrammes de Young de λ et λ^* sont symétriques par rapport à la diagonale.

La représentation V_λ admet la construction directe suivante : on remplit le diagramme de Young associé à λ avec les nombres de 1 à n (on obtient de la sorte un *tableau de Young*). Ceci permet de définir deux sous-groupes A et B de S_n , en prenant pour A (resp. pour B) l'ensemble des permutations de $\{1, \dots, n\}$ qui préserve globalement chaque ligne (resp. chaque colonne). Alors V_λ est la sous-représentation de la représentation régulière $\bigoplus_{\sigma \in S_n} \mathbf{C} e_\sigma$ de S_n engendrée par le vecteur $\sum_{a \in A, b \in B} \text{sign}(b) e_{ab}$. Le sous-espace engendré par $\sum_{a \in A, b \in B} \text{sign}(b) e_{ab}$ dépend du tableau de Young, mais il est facile de voir que, à isomorphisme près, la représentation obtenue n'en dépend pas ; elle ne dépend que de la partition λ .

Par exemple, si $\lambda = (1, \dots, 1)$, on a $A = \{1\}$ et $B = S_n$. La représentation V_λ est engendrée par $v = \sum_{\sigma \in S_n} \text{sign}(\sigma) e_\sigma$, et comme $g \cdot v = \text{sign}(g) v$, si $g \in S_n$ (ne pas oublier que $\text{sign}(g) \in \{\pm 1\}$), cette représentation est la représentation de dimension 1 correspondant au caractère linéaire sign de S_n . Plus généralement, il est facile, sur la construction ci-dessus, de voir que $V_{\lambda^*} = V_\lambda \otimes \text{sign}$.

La dimension de V_λ est donnée par la *formule des équerres*. Si a est une case du diagramme de Young de λ , l'équerre E_a de sommet a , est la réunion des cases se trouvant à droite, sur la même ligne que a , ou en dessous, sur la même colonne que a (en incluant a) ; la longueur $\text{lg}(E_a)$ est le nombre de ses éléments. On a alors

$$\dim V_\lambda = \frac{n!}{\prod_{a \in Y_\lambda} \text{lg}(E_a)}.$$

Exercice B.2.1. — Calculer les dimensions des représentations irréductibles de S_3 , S_4 et S_5 , et comparer les résultats avec les ex. I.3.19, I.3.22 et I.3.23.

3. Caractères de S_n

Les caractères de U_{λ} et V_{λ} ont des expressions compactes assez miraculeuses : si $\ell = (\ell_1, \dots, \ell_r)$, et si $\lambda = (\lambda_1, \dots, \lambda_s)$, alors

$$\begin{aligned}\chi_{U_{\lambda}}(C_{\ell}) &= \text{coefficient de } X_1^{\lambda_1} \cdots X_n^{\lambda_n} \text{ dans } P_{\ell}, \\ \chi_{V_{\lambda}}(C_{\ell}) &= \text{coefficient de } X_1^{\lambda_1+n-1} \cdots X_n^{\lambda_n} \text{ dans } P_{\ell} \cdot \prod_{i < j \leq n} (X_i - X_j),\end{aligned}$$

où $P_{\ell}(X_1, \dots, X_n) = (X_1^{\ell_1} + \cdots + X_n^{\ell_1}) \cdots (X_1^{\ell_r} + \cdots + X_n^{\ell_r})$, et où on a posé $\lambda_i = 0$ si $s+1 \leq i \leq n$.

La formule pour $\chi_{U_{\lambda}}$ est assez facile à établir car U_{λ} est une représentation induite, ce qui permet d'utiliser le th. I.3.10. Le passage de U_{λ} à V_{λ} utilise la combinatoire des polynômes de Schur.

Si $\mu = (\mu_1, \dots, \mu_s)$ est une partition de n , on note encore μ la famille (μ_1, \dots, μ_n) , avec $\mu_i = 0$ si $s+1 \leq i \leq n$. On définit le *polynôme de Schur* Sch_{μ} par

$$Sch_{\mu}(X_1, \dots, X_n) = \frac{\det \begin{pmatrix} X_1^{\mu_1+n-1} & \cdots & X_n^{\mu_n+n-1} \\ \vdots & X_j^{\mu_j+n-i} & \vdots \\ X_1^{\mu_1} & \cdots & X_n^{\mu_n} \end{pmatrix}}{\det \begin{pmatrix} X_1^{n-1} & \cdots & X_n^{n-1} \\ \vdots & X_j^{n-i} & \vdots \\ 1 & \cdots & 1 \end{pmatrix}}$$

Si $\lambda = (\lambda_1, \dots, \lambda_r)$ est une autre partition de n , on définit le *nombre de Kostka* $K_{\mu, \lambda}$ comme le coefficient de $X_1^{\lambda_1} \cdots X_n^{\lambda_n}$ dans $Sch_{\mu}(X_1, \dots, X_n)$.

On a $U_{\lambda} = \sum_{\mu} K_{\mu, \lambda} V_{\mu}$, et $K_{\mu, \lambda} = 0$, si $\mu < \lambda$, $K_{\lambda, \lambda} = 1$. C'est comme ça que l'on vérifie que V_{λ} n'apparaît pas dans les U_{μ} , pour $\mu > \lambda$. Comme de plus $U_{(1, \dots, 1)}$ est la représentation régulière, on déduit du cor. I.2.19, et de la formule ci-dessus, que $\dim V_{\mu} = K_{\mu, (1, \dots, 1)}$.

Le nombre $K_{\mu, \lambda}$ a aussi une interprétation combinatoire : c'est le nombre de manières de remplir le diagramme de Young de λ avec λ_1 fois le nombre 1, λ_2 fois le nombre 2, ..., λ_r fois le nombre r , de telle sorte que chaque ligne soit croissante (au sens large) et chaque colonne soit strictement croissante.

En utilisant la formule ci-dessus pour la dimension de V_{μ} , on en déduit que $\dim V_{\mu}$ est le nombre de tableaux de Young pour la partition μ dans lesquels les lignes et les colonnes sont croissantes ; un tel tableau de Young est dit *standard*⁽²⁾. Il n'est pas totalement évident de prouver, combinatoirement, que ceci coïncide avec la formule des équerres.

⁽²⁾ C. Schensted (1961) a établi une bijection entre les paires de tableaux de Young standard de même forme et les permutations, ce qui fournit une démonstration combinatoire de la formule de Burnside ((ii) du cor. I.2.19) dans le cas de S_n (modulo la formule des équerres).

B.3. Représentations de $\mathbf{GL}_2(\mathbf{F})$

1. Le groupe $\mathbf{GL}_2(\mathbf{F})$

Soit \mathbf{F} un corps fini de cardinal q . Il existe alors un nombre premier p tel que l'on ait $p = 0$ dans \mathbf{F} , ce qui fait que \mathbf{F} est un \mathbf{F}_p -espace vectoriel de dimension finie (avec $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$), et donc que q est une puissance de p .

Soit $G = \mathbf{GL}_2(\mathbf{F})$ le groupe des matrices 2×2 , à coefficients dans \mathbf{F} , de déterminant non nul. Si $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{M}_2(\mathbf{F})$, alors $g \in G$, si et seulement si les vecteurs colonnes de la matrice sont linéairement indépendants sur \mathbf{F} , ce qui signifie que le premier vecteur est non nul et que le second n'est pas dans la droite engendrée par le premier. On en déduit que $|\mathbf{GL}_2(\mathbf{F})| = (q^2 - 1)(q^2 - q)$.

Le corps \mathbf{F} a une unique extension \mathbf{K} de degré 2 ⁽³⁾. Si $p \neq 2$ ⁽⁴⁾, on choisit $\Delta \in \mathbf{F}^*$ qui n'est pas un carré dans \mathbf{F}^* , et une racine carrée δ de Δ dans \mathbf{K} . On peut alors associer à $z = x + \delta y$ la matrice $C_z = \begin{pmatrix} x & \Delta y \\ y & x \end{pmatrix}$ de la multiplication par z dans la base $1, \delta$ de \mathbf{K} sur \mathbf{F} ⁽⁵⁾. La conjugaison par $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ envoie C_z , avec $z = x + \delta y$, sur $C_{\bar{z}}$, avec $\bar{z} = x - \delta y$, comme le montre un calcul immédiat.

2. Construction de représentations de $\mathbf{GL}_2(\mathbf{F})$

On cherche à faire la liste des représentations irréductibles de G . Le théorème de Brauer (th. I.3.16) montre que l'on peut les obtenir à partir d'induites de caractères linéaires de sous-groupes. Par ailleurs, comme les dimensions des représentations irréductibles ont tendance à être assez petites, on a intérêt à induire à partir des sous-groupes les plus gros possibles. Toutes les représentations ci-dessous sont obtenues avec ces considérations en tête.

- *Représentations de dimension 1.* Si $\eta \in \widehat{\mathbf{F}^*}$ est un caractère linéaire de \mathbf{F}^* , on fabrique un caractère linéaire $\eta \circ \det$ de G en composant avec le déterminant.

- *La steinberg et ses tordues.* On note $\mathbf{P}^1(\mathbf{F})$ l'ensemble des droites vectorielles de \mathbf{F}^2 ; c'est un ensemble à $q + 1$ éléments sur lequel G agit. Il lui correspond donc une

⁽³⁾ Il y a q^2 polynômes unitaires de degré 2 à coefficients dans \mathbf{F} parmi lesquels q ont une racine double et $\frac{q(q-1)}{2}$ ont deux racines distinctes. Il y a donc $q^2 - q - \frac{q(q-1)}{2} = \frac{q(q-1)}{2} > 0$ polynômes irréductibles de degré 2, ce qui prouve que \mathbf{F} a au moins une extension de degré 2. Maintenant, une telle extension \mathbf{K} a q^2 éléments, et \mathbf{K}^* est d'ordre $q^2 - 1$, ce qui prouve que tout élément inversible x de \mathbf{K} vérifie $x^{q^2-1} = 1$, et donc que tout élément de \mathbf{K} est racine de $X^{q^2} - X = 0$. Comme un polynôme de degré q^2 a au plus q^2 racines dans un corps commutatif, on en déduit l'unicité (si on avait deux extensions distinctes \mathbf{K}_1 et \mathbf{K}_2 , $X^{q^2} - X$ aurait trop de racines dans l'extension composée $\mathbf{K}_1\mathbf{K}_2$).

⁽⁴⁾ Si $p = 2$, il faut modifier un peu ce qui suit, car un polynôme de la forme $X^2 - \Delta$ n'est jamais irréductible. On choisit $\Delta \in \mathbf{F}$ qui n'est pas dans l'image de $x \mapsto x^2 + x$, et $\delta \in \mathbf{K}$ vérifiant $\delta^2 + \delta = \Delta$. On peut alors associer à $z = x + \delta y$ la matrice $C_z = \begin{pmatrix} x & \Delta y \\ y & x + y \end{pmatrix}$ de la multiplication par z dans la base $1, \delta$ de \mathbf{K} sur \mathbf{F} . La conjugaison par $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ envoie C_z , avec $z = x + \delta y$, sur $C_{\bar{z}}$, avec $\bar{z} = x + (\delta + 1)y$.

⁽⁵⁾ On remarquera la similarité avec la représentation matricielle des nombres complexes.

représentation de permutation $V_{\mathbf{P}^1(\mathbf{F})} = \bigoplus_{x \in \mathbf{P}^1(\mathbf{F})} \mathbf{C}e_x$, qui n'est pas irréductible puisque la droite engendrée par $\sum_{x \in \mathbf{P}^1(\mathbf{F})} e_x$ est fixe par G . On note St la représentation de G sur l'hyperplan $\{\sum_{x \in \mathbf{P}^1(\mathbf{F})} \lambda_x e_x, \sum_{x \in \mathbf{P}^1(\mathbf{F})} \lambda_x = 0\}$; c'est une représentation de dimension q .

Plus généralement, si $\eta \in \widehat{\mathbf{F}^*}$, on peut considérer la représentation $\text{St} \otimes (\eta \circ \det)$ obtenue en tordant l'action de G sur St par le caractère linéaire $\eta \circ \det$.

• *La série principale.* Soit B le sous-groupe de Borel de G ; c'est l'ensemble des matrices triangulaires supérieures inversibles. Si $\eta_1, \eta_2 \in \widehat{\mathbf{F}^*}$ sont deux caractères linéaires de \mathbf{F}^* , on note $\eta_1 \otimes \eta_2$ le caractère linéaire de B défini par

$$(\eta_1 \otimes \eta_2) \left(\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \right) = \eta_1(a) \eta_2(d).$$

Si $\eta_1 \neq \eta_2$, la représentation $\text{Ind}_B^G \eta_1 \otimes \eta_2$ est dite *de la série principale*.⁽⁶⁾

• *Autres représentations.* Comme nous le verrons les représentations précédentes sont irréductibles, mais ne suffisent pas à remplir la liste des représentations irréductibles de G . Les représentations qui suivent ne sont pas irréductibles (on induit à partir d'un sous-groupe trop petit), mais contiennent les représentations qui nous manquent.

– Soit $C \subset G$ l'image de \mathbf{K}^* par l'application $z \mapsto C_z$. C'est un sous-groupe⁽⁷⁾ de G de cardinal $q^2 - 1$. Si $\eta \in \widehat{\mathbf{K}^*}$, on peut aussi considérer η comme un caractère de C , en utilisant l'isomorphisme $z \mapsto C_z$ de \mathbf{K}^* sur C , et nous aurons à considérer la représentation $\text{Ind}_C^G \eta$.

– Soit $ZU \subset B$ le sous-groupe des matrices ayant une seule valeur propre (Z est là pour désigner le centre, qui est l'ensemble des matrices d'homothéties de rapport non nul, et U est le groupe des matrices triangulaires supérieures avec des 1 sur la diagonale; une telle matrice est *unipotente*). On fixe un caractère linéaire non trivial $\psi \in \widehat{\mathbf{F}}$ de \mathbf{F} , et si $\eta \in \widehat{\mathbf{F}^*}$ est un caractère linéaire de \mathbf{F}^* , on note $\eta \otimes \psi$ le caractère linéaire de ZU défini par

$$(\eta \otimes \psi) \left(\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} \right) = \eta(a) \psi(a^{-1}b).$$

Nous aurons à considérer la représentation $\text{Ind}_{ZU}^G \eta \otimes \psi$. En fait, le cas qui nous intéressera est celui où on part d'un élément η de $\widehat{\mathbf{K}^*}$, que l'on restreint à \mathbf{F}^* (nous nous permettrons de garder le même nom pour la restriction).

3. Les classes de conjugaison de $\text{GL}_2(\mathbf{F})$

Comme \mathbf{F} a une unique extension de degré 2, la classification, à conjugaison près par une matrice inversible, des matrices 2×2 à coefficients dans \mathbf{F} est identique à ce qui se passe sur \mathbf{R} . Il y a quatre types distincts :

• type I : A est la matrice d'une homothétie, et donc de la forme $a_x = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$, avec $x \in \mathbf{F}$;

⁽⁶⁾Le lecteur pourra vérifier que $\text{Ind}_B^G \eta \otimes \eta = V_{\mathbf{P}^1(\mathbf{F})} \otimes (\eta \circ \det)$, ce qui explique que l'on ne s'intéresse pas au cas $\eta_1 = \eta_2$.

⁽⁷⁾Un tel sous-groupe est un *sous-groupe de Cartan non déployé*, un sous-groupe de Cartan déployé étant un sous-groupe conjugué au sous-groupe des matrices diagonales.

- type II : A a une unique valeur propre $x \in \mathbf{F}$, mais n'est pas diagonalisable ; elle est alors conjuguée à $b_x = \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$;
- type III : A a deux valeurs propres $x \neq y$ dans \mathbf{F} , et donc est conjuguée à $b_{x,y} = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$, et donc aussi à $b_{y,x}$;
- type IV : A n'a pas de valeurs propres dans \mathbf{F} , et donc a deux valeurs propres $z \neq \bar{z}$ dans \mathbf{K} , et A est conjuguée à c_z et $c_{\bar{z}}$.

Comme on s'intéresse aux classes de conjugaison de G, il faut ne garder que les matrices inversibles dans l'énumération ci-dessus. On aboutit à la liste de paramètres de la fig. 1.

	description	cardinal
E _I	\mathbf{F}^*	$q - 1$
E _{II}	\mathbf{F}^*	$q - 1$
E _{III}	$\{(x, y) \in \mathbf{F}^* \times \mathbf{F}^*, x \neq y\} / (x, y) \sim (y, x)$	$\frac{1}{2}(q-1)(q-2)$
E _{IV}	$\{z \in \mathbf{K}^*, \bar{z} \neq z\} / \bar{z} \sim z$	$\frac{1}{2}q(q-1)$

FIG. 1. Paramètres des classes de conjugaison de $\mathbf{GL}_2(\mathbf{F})$

Proposition B.3.1. — La liste des classes de conjugaison de $\mathbf{GL}_2(\mathbf{F})$ est celle de la figure 2

type	nombre	classe	représentant	cardinal
I	$q - 1$	$a_x, x \in E_I$	$A_x = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$	1
II	$q - 1$	$b_x, x \in E_{II}$	$B_x = \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$	$q^2 - 1$
III	$\frac{(q-1)(q-2)}{2}$	$b_{x,y}, (x, y) \in E_{III}$	$B_{x,y} = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$	$q(q+1)$
IV	$\frac{q^2-q}{2}$	$c_z, z = x + \delta y \in E_{IV}$	$C_z = \begin{pmatrix} x & \Delta y \\ y & x \end{pmatrix}$	$q^2 - q$

FIG. 2. Classes de conjugaison de $\mathbf{GL}_2(\mathbf{F})$

Démonstration. — La seule chose qui ne suive pas directement de la discussion précédant la proposition, est la détermination du cardinal de chacune des classes. De manière générale, si G est un groupe, et $x \in G$, la classe de conjugaison de x est l'orbite O_x de x sous l'action de G par conjugaison intérieure ($g \cdot x = gxg^{-1}$) ; si Z_x est le centralisateur de x dans G (i.e. est l'ensemble des éléments de G commutant à x), on a $O_x = G/Z_x$, et donc $|O_x| = \frac{|G|}{|Z_x|}$. On est donc ramené au problème de déterminer l'ensemble des matrices inversibles commutant à une matrice donnée. Or, si $A \in \mathbf{M}_2(\mathbf{F})$, l'ensemble des matrices $M \in \mathbf{M}_2(\mathbf{F})$ qui commutent à A sont :

- $\mathbf{M}_2(\mathbf{F})$ tout entier, si A est la matrice d'une homothétie,

• le sous-espace vectoriel engendré par $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et A , si A n'est pas la matrice d'une homothétie⁽⁸⁾.

On déduit de cette discussion les résultats suivants :

- Tout commute à A_x et donc il y a un seul élément dans a_x .
- Les matrices commutant à B_x sont de la forme $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$, avec $a, b \in \mathbf{F}$, et une telle matrice est inversible si et seulement si $a \neq 0$. Le centralisateur de B_x est donc de cardinal $q(q-1)$ et le nombre d'éléments de b_x est $\frac{(q^2-1)(q^2-q)}{q(q-1)} = q^2 - 1$.
- si $x \neq y$, les matrices commutant à $B_{x,y}$ sont de la forme $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, avec $a, b \in \mathbf{F}$, et une telle matrice est inversible si et seulement si $a \neq 0$ et $b \neq 0$. Le centralisateur de $B_{x,y}$ est donc de cardinal $(q-1)^2$ et le nombre d'éléments de $b_{x,y}$ est $\frac{(q^2-1)(q^2-q)}{(q-1)^2} = q(q+1)$.
- si $z \in \mathbf{K} - \mathbf{F}$, les matrices commutant à C_z sont de la forme $C_{z'}$, avec $z' \in \mathbf{K}$, et une telle matrice est inversible si et seulement si $z' \neq 0$. Le centralisateur de C_z est donc de cardinal $q^2 - 1$ et le nombre d'éléments de c_z est $\frac{(q^2-1)(q^2-q)}{q^2-1} = q(q-1)$.

Ceci permet de conclure.

4. La table des caractères de $\mathbf{GL}_2(\mathbf{F})$

On suppose dorénavant que q n'est pas une puissance de 2 (il y a des petites modifications (cf. note 4) à faire dans le cas $p = 2$; nous les laissons en exercice). Les caractères de G se paramètrent de manière parallèle aux classes de conjugaison de G . Il y a aussi 4 types paramétrés par les ensembles E_I^* , E_{II}^* , E_{III}^* et E_{IV}^* de la figure 3.

	description	cardinal
E_I^*	$\widehat{\mathbf{F}}^*$	$q - 1$
E_{II}^*	$\widehat{\mathbf{F}}^*$	$q - 1$
E_{III}^*	$\{(\eta_1, \eta_2) \in \widehat{\mathbf{F}}^* \times \widehat{\mathbf{F}}^*, \eta_1 \neq \eta_2\} / (\eta_1, \eta_2) \sim (\eta_2, \eta_1)$	$\frac{1}{2}(q-1)(q-2)$
E_{IV}^*	$\{\eta \in \widehat{\mathbf{K}}^*, \eta^* \neq \eta\} / \eta^* \sim \eta$	$\frac{1}{2}q(q-1)$

FIG. 3. Paramètres des caractères de $\mathbf{GL}_2(\mathbf{F})$

Dans le tableau ci-dessus, on a noté η^* le caractère linéaire $z \mapsto \eta(\bar{z})$, si $\eta \in \widehat{\mathbf{K}}^*$.

Théorème B.3.2. — *La table des caractères de $\mathbf{GL}_2(\mathbf{F})$ est celle donnée par la figure 4.*

⁽⁸⁾Un calcul brutal montre que l'espace des solutions est de dimension 2 si A n'est pas une homothétie, et comme I et A commutent à A , et sont linéairement indépendantes, cela permet de conclure. Cet argument est raisonnable en dimension 2, mais pour déterminer le commutant d'une matrice en dimension quelconque, il vaut mieux utiliser le point de vue du n° 5.2 du § 5.

type	I	II	III	IV
nombre	$q - 1$	$q - 1$	$\frac{(q-1)(q-2)}{2}$	$\frac{q(q-1)}{2}$
caractère	$\alpha_\eta, \eta \in E_I^*$	$\beta_\eta, \eta \in E_{II}^*$	$\beta_{\eta_1, \eta_2}, (\eta_1, \eta_2) \in E_{III}^*$	$\gamma_\eta, \eta \in E_{IV}^*$
dimension	1	q	$q + 1$	$q - 1$
a_x	$\eta(x)^2$	$q\eta(x)^2$	$(q + 1)\eta_1\eta_2(x)$	$(q - 1)\eta(x)$
b_x	$\eta(x)^2$	0	$\eta_1\eta_2(x)$	$-\eta(x)$
$b_{x,y}$	$\eta(xy)$	$\eta(xy)$	$\eta_1(x)\eta_2(y) + \eta_1(y)\eta_2(x)$	0
c_z	$\eta(z\bar{z})$	$-\eta(z\bar{z})$	0	$-(\eta(z) + \eta^*(z))$

FIG. 4. Table des caractères de $\mathbf{GL}_2(\mathbf{F})$

Démonstration. — On note X l'ensemble des fonctions centrales apparaissant dans le tableau (et donc de la forme $\alpha_\eta, \beta_\eta, \beta_{\eta_1, \eta_2}$ ou γ_η). Comme il y a autant de paramètres pour les éléments de X que pour les classes de conjugaison de G , il suffit de vérifier que les χ correspondant à deux paramètres distincts sont différents (lemme B.3.6) et que $\chi \in \text{Irr}(G)$, si $\chi \in X$. Pour vérifier ce deuxième point, il suffit, d'après l'ex. I.2.18, de vérifier que

- $\chi(1) \geq 0$ (ce qui est évident),
- $\langle \chi, \chi \rangle = 1$ (ce qui fait l'objet du lemme B.3.5),
- $\chi \in R_{\mathbf{Z}}(G)$ (ce qui fait l'objet du lemme B.3.4).

Remarque B.3.3. — (i) Remplacer le couple (η_1, η_2) par (η_2, η_1) ne change rien dans la colonne du type III, ce qui montre que l'on peut effectivement passer au quotient par la relation d'équivalence $(\eta_1, \eta_2) \sim (\eta_2, \eta_1)$.

(ii) De même, remplacer η par η^* ne change rien dans la colonne du type IV (car $\eta^*(x) = \eta(x)$, si $x \in \mathbf{F}^*$), ce qui montre que l'on peut effectivement passer au quotient par la relation d'équivalence $\eta \sim \eta^*$.

5. Démonstrations

Lemme B.3.4. — (i) Si $\eta \in \widehat{\mathbf{F}^*}$, alors $\alpha_\eta = \eta \circ \det$.

(ii) Si $\eta \in \widehat{\mathbf{F}^*}$, alors β_η est le caractère de $\text{St} \otimes (\eta \circ \det)$.

(iii) Si $(\eta_1, \eta_2) \in E_{III}$, alors β_{η_1, η_2} est le caractère de $\text{Ind}_B^G \eta_1 \otimes \eta_2$.

(iv) Si $\eta \in E_{IV}$, alors $\gamma_\eta = \text{Ind}_{ZU}^G \eta \otimes \psi - \text{Ind}_C^G \eta$.

Démonstration. — Le (i) est immédiat. Pour démontrer le reste, nous noterons (e_1, e_2) la base canonique de \mathbf{F}^2 .

• Pour démontrer le (ii), il faut commencer par calculer le caractère χ_{St} de St . Pour ce faire, on commence par calculer le caractère $\chi_{\mathbf{P}^1(\mathbf{F})}$ de la représentation de permutation

$V_{\mathbf{P}^1(\mathbf{F})}$, et on est ramené (alinéa 2.3 du § I.1) à calculer le nombre de points fixes de l'action de $g \in G$ sur $\mathbf{P}^1(\mathbf{F})$. Comme les points fixes sont exactement les droites propres de \mathbf{F}^2 sous l'action de g , on voit que le nombre de points fixes de A_x est $q + 1$, celui de B_x est 1 (la droite $\mathbf{F}e_1$), celui de $B_{x,y}$ est 2 (les droites $\mathbf{F}e_1$ et $\mathbf{F}e_2$) et celui de C_z est 0. Comme $V_{\mathbf{P}^1(\mathbf{F})} = \text{St} \oplus \mathbf{1}$, on a $\chi_{\text{St}} = \chi_{\mathbf{P}^1(\mathbf{F})} - 1$, et donc $\chi_{\text{St}}(A_x) = q$, $\chi_{\text{St}}(B_x) = 0$, $\chi_{\text{St}}(B_{x,y}) = 1$ et $\chi_{\text{St}}(C_z) = -1$. Le (ii) s'en déduit en tordant par le caractère $\eta \circ \det$.

• Passons au (iii). Soit $\chi = \text{Ind}_B^G \eta_1 \otimes \eta_2$, avec $\eta_1, \eta_2 \in \widehat{\mathbf{F}^*}$. Pour calculer χ , on part de la formule du th. I.3.10

$$\chi(g) = \frac{1}{|B|} \sum_{s \in G, sgs^{-1} \in B} \eta_1 \otimes \eta_2(sgs^{-1}).$$

– Comme A_x est dans le centre de G qui est inclus dans B , on a $sA_x s^{-1} \in B$ et $\eta_1 \otimes \eta_2(sA_x s^{-1}) = \eta_1 \eta_2(x)$, quel que soit $s \in G$. On en déduit que

$$\chi(A_x) = \frac{|G|}{|B|} \eta_1 \eta_2(x) = (q + 1) \eta_1 \eta_2(x).$$

– On a $sB_x s^{-1} \in B$ si et seulement si $sB_x s^{-1}(e_1) \in \mathbf{F}e_1$, et donc si et seulement si $B_x(s^{-1}(e_1)) \in \mathbf{F}s^{-1}(e_1)$. Comme $\mathbf{F}e_1$ est la seule droite propre de B_x , cela équivaut à $s^{-1}(e_1) \in \mathbf{F}e_1$ et donc à $s \in B$. Comme $\eta_1 \otimes \eta_2(sB_x s^{-1}) = \eta_1 \eta_2(x)$, si $s \in B$, on obtient

$$\chi(B_x) = \frac{1}{|B|} \sum_{s \in B} \eta_1 \eta_2(x) = \eta_1 \eta_2(x).$$

– On montre de même que $sB_{x,y} s^{-1} \in B$, si et seulement si $B_{x,y}(s^{-1}(e_1)) \in \mathbf{F}s^{-1}(e_1)$, et donc si et seulement si $s^{-1}(e_1) \in \mathbf{F}e_1$ ou $s^{-1}(e_1) \in \mathbf{F}e_2$. Le premier cas équivaut à $s \in B$ comme ci-dessus, et on a $\eta_1 \otimes \eta_2(sB_{x,y} s^{-1}) = \eta_1(x) \eta_2(y)$; le second équivaut à $s \in Bw$, où $w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, et on a $\eta_1 \otimes \eta_2(sB_{x,y} s^{-1}) = \eta_1(y) \eta_2(x)$. On obtient donc

$$\chi(B_{x,y}) = \frac{1}{|B|} \left(\sum_{s \in B} \eta_1(x) \eta_2(y) + \sum_{s \in Bw} \eta_1(y) \eta_2(x) \right) = \eta_1(x) \eta_2(y) + \eta_1(y) \eta_2(x).$$

– Si $z \in \mathbf{K}^* - \mathbf{F}^*$, il n'existe aucun $s \in G$ tel que $sC_z s^{-1} \in B$, et donc $\chi(C_z) = 0$.

• Finalement, venons-en au (iv). Soit $\eta \in \widehat{\mathbf{K}^*}$, soient χ_1 et χ_2 les caractères de $\text{Ind}_{ZU}^G \eta \otimes \psi$ et $\text{Ind}_C^G \eta$, et soit $\chi = \chi_1 - \chi_2$. D'après le th. I.3.10,

$$\chi_1(g) = \frac{1}{|ZU|} \sum_{s \in G, sgs^{-1} \in ZU} \eta \otimes \psi(sgs^{-1}) \quad \text{et} \quad \chi_2(g) = \frac{1}{|C|} \sum_{s \in G, sgs^{-1} \in C} \eta(sgs^{-1}).$$

– Comme $A_x \in ZU$ et $A_x \in C$, on a, comme ci-dessus,

$$\chi_1(A_x) = \frac{|G|}{|ZU|} \eta(x) = (q^2 - 1) \eta(x) \quad \text{et} \quad \chi_2(A_x) = \frac{|G|}{|C|} \eta(x) = (q^2 - q) \eta(x),$$

et donc $\chi(A_x) = (q - 1) \eta(x)$.

- Comme aucun élément de ZU n'est conjugué à $B_{x,y}$ ou C_z , on a $\chi_1(B_{x,y}) = 0$ et $\chi_1(C_z) = 0$.
- Comme aucun élément de C n'est conjugué à $B_{x,y}$ ou B_x , on a $\chi_2(B_{x,y}) = 0$ et $\chi_2(B_x) = 0$.
- On a $ZU \subset B$. Or on a vu plus haut que $sB_x s^{-1} \in B$ implique $s \in B$. De plus, comme B_x a une seule valeur propre, il en est de même de $sB_x s^{-1}$, et donc $sB_x s^{-1} \in ZU$ si et seulement si $s \in B$. Maintenant, on a

$$\begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix} \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix}^{-1} = \begin{pmatrix} \alpha x & \alpha + \beta x \\ 0 & \delta x \end{pmatrix} \begin{pmatrix} \alpha^{-1} & -\beta \alpha^{-1} \delta^{-1} \\ 0 & \delta^{-1} \end{pmatrix} = \begin{pmatrix} x & \alpha \delta^{-1} \\ 0 & x \end{pmatrix}.$$

On en déduit que

$$\chi_1(B_x) = \frac{1}{|ZU|} \sum_{\alpha, \delta \in \mathbf{F}^*, \beta \in \mathbf{F}} \eta(x) \psi(\alpha \delta^{-1}) = \frac{\eta(x)}{q-1} \sum_{\alpha, \delta \in \mathbf{F}^*} \psi(\alpha \delta^{-1}).$$

En faisant le changement de variable $\alpha' = \alpha \delta^{-1}$, $\delta' = \delta$, et en utilisant le fait que $\sum_{\alpha' \in \mathbf{F}} \psi(\alpha') = 0$ puisque l'on a supposé ψ non trivial, on obtient finalement

$$\chi_1(B_x) = \eta(x)(-\psi(0)) = -\eta(x) \quad \text{et} \quad \chi(B_x) = -\eta(x).$$

– Si $sC_z s^{-1} \in C$, alors $sC_z s^{-1}$ est égal à C_z ou $C_{\bar{z}}$ (le polynôme caractéristique de C_z est égal à $X^2 - (z + \bar{z})X + z\bar{z}$, et donc celui de $sC_z s^{-1}$ aussi). Si $sC_z s^{-1} = C_z$, alors s est dans le centralisateur de C_z qui, comme on l'a vu (dém. de la prop. B.3.1), est égal à C ; si $sC_z s^{-1} = C_{\bar{z}}$, alors $s \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} C_{\bar{z}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} s^{-1} = C_{\bar{z}}$, et donc $s \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ est dans le centralisateur de $C_{\bar{z}}$, et $s \in C \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. On en déduit que

$$\chi_2(C_z) = \frac{1}{|C|} \left(\sum_{s \in C} \eta(z) + \sum_{s \in C} \eta(\bar{z}) \right) = \eta(z) + \eta^*(z) \quad \text{et} \quad \chi(C_z) = -(\eta(z) + \eta^*(z)).$$

Ceci permet de conclure

Lemme B.3.5. — Si $\chi \in X$, alors $\langle \chi, \chi \rangle = 1$.

Démonstration. — Si ϕ est une fonction centrale sur G , il résulte de la table des classes de conjugaison, que

$$\begin{aligned} |G| \langle \phi, \phi \rangle &= \sum_{x \in \mathbf{F}^*} |\phi(a_x)|^2 + (q^2 - 1) \sum_{x \in \mathbf{F}^*} |\phi(b_x)|^2 \\ &\quad + q(q+1) \sum_{\substack{(x,y) \in (\mathbf{F}^*)^2, \\ \text{mod } (x,y) \mapsto (y,x)}} |\phi(b_{x,y})|^2 + q(q-1) \sum_{\substack{z \in \mathbf{K}^* - \mathbf{F}^* \\ \text{mod } z \mapsto \bar{z}}} |\phi(c_z)|^2 \end{aligned}$$

Il nous faut vérifier que $|G| \langle \phi, \phi \rangle = |G| = (q^2 - 1)(q^2 - q)$, si ϕ est un des caractères apparaissant dans la table 4.

- Pour un caractère du type $\eta \circ \det$, il n'y a rien à faire puisque l'on a affaire au caractère d'une représentation de dimension 1 qui est automatiquement irréductible.

• On remarque que, si ϕ est le caractère de la représentation $\text{St} \otimes (\eta \circ \det)$, les contributions des classes $b_{x,y}$ et c_z sont les mêmes que pour $\eta \circ \det$. Il suffit donc de vérifier que $\sum_{x \in \mathbf{F}^*} |\phi(a_x)|^2 + (q^2 - 1) \sum_{x \in \mathbf{F}^*} |\phi(b_x)|^2$ a la même valeur que dans le cas $\phi = \eta \circ \det$, ce qui est immédiat, les deux sommes valant $(q - 1)q^2$.

• Dans le cas où $\phi = \beta_{\eta_1, \eta_2}$, avec $\eta_1, \eta_2 \in \widehat{\mathbf{F}^*}$ et $\eta_1 \neq \eta_2$, la somme à évaluer devient

$$(q - 1)(q + 1)^2 + (q - 1)(q^2 - 1) + q(q + 1) \sum_{\substack{(x,y) \in (\mathbf{F}^*)^2, \ x \neq y \\ \text{mod } (x,y) \mapsto (y,x)}} |\eta_1(x)\eta_2(y) + \eta_2(x)\eta_1(y)|^2.$$

En remarquant que $\bar{\eta}(a) = \eta(a)^{-1} = \eta(a^{-1})$, on peut écrire $|\eta_1(x)\eta_2(y) + \eta_2(x)\eta_1(y)|^2$ sous la forme

$$2 + \eta_1(x)\eta_2(y)\overline{\eta_2(x)\eta_1(y)} + \overline{\eta_1(x)\eta_2(y)}\eta_2(x)\eta_1(y) = 2 + (\eta_1/\eta_2)(x/y) + (\eta_2/\eta_1)(y/x).$$

Ceci permet de mettre $\sum |\eta_1(x)\eta_2(y) + \eta_2(x)\eta_1(y)|^2$ sous la forme

$$(q - 1)(q - 2) + \sum_{(x,y) \in (\mathbf{F}^*)^2, \ x \neq y} (\eta_1/\eta_2)(x/y) = (q - 1)(q - 2) + \sum_{(x,y) \in (\mathbf{F}^*)^2} (\eta_1/\eta_2)(x/y) - (q - 1).$$

Maintenant, comme $\eta_1 \neq \eta_2$, on a $\sum_{(x,y) \in (\mathbf{F}^*)^2} (\eta_1/\eta_2)(x/y) = 0$. La somme à évaluer est donc, finalement, égale à

$$(q - 1)(q + 1)^2 + (q - 1)(q^2 - 1) + q(q + 1)(q - 1)(q - 3) = (q^2 - 1)(q + 1 + q - 1 + q(q - 3)) = (q^2 - 1)(q^2 - q),$$

ce que l'on voulait.

• Dans le cas où $\phi = \gamma_\eta$, avec $\eta \in \widehat{\mathbf{K}^*}$ et $\eta \neq \eta^*$, la somme à évaluer est

$$(q - 1)(q - 1)^2 + (q - 1)(q^2 - 1) + q(q - 1) \sum_{z \in \mathbf{K}^* - \mathbf{F}^*, \text{ mod } z \mapsto \bar{z}} |\eta(z) + \eta^*(z)|^2.$$

Les mêmes calculs que ci-dessus (en utilisant le fait que $\eta^*(z) = \eta(\bar{z})$) permettent de mettre $\sum_{z \in \mathbf{K}^* - \mathbf{F}^*, \text{ mod } z \mapsto \bar{z}} |\eta(z) + \eta^*(z)|^2$ sous la forme

$$q(q - 1) + \sum_{z \in \mathbf{K}^* - \mathbf{F}^*} (\eta/\eta^*)(z) = q(q - 1) + \sum_{z \in \mathbf{K}^*} (\eta/\eta^*)(z) - (q - 1),$$

et comme ci-dessus $\sum_{z \in \mathbf{K}^*} (\eta/\eta^*)(z) = 0$ puisque $\eta \neq \eta^*$. La somme à évaluer est donc, finalement, égale à

$$(q - 1)(q - 1)^2 + (q - 1)(q^2 - 1) + q(q - 1)(q - 1)^2 = (q - 1)^2(q - 1 + q + 1 + q(q - 1)) \\ = (q - 1)^2(q^2 + q) = (q^2 - 1)(q^2 - q),$$

ce que l'on voulait.

Lemme B.3.6. — Si χ et χ' sont deux éléments de X tels que $\chi = \chi'$, alors les paramètres de χ et χ' sont égaux.

Démonstration. — Si $\chi = \chi'$, on a en particulier $\chi(1) = \chi'(1)$, ce qui implique que χ et χ' sont de même type.

- Dans les cas des types I et II, il suffit de regarder la valeur en les $b_{x,y}$, pour prouver que les paramètres de χ et χ' sont les mêmes.

- Pour traiter le cas du type III, associons à un couple (δ_1, δ_2) d'éléments $\widehat{\mathbf{F}^*}$, le caractère linéaire $\delta_1 \otimes \delta_2$ de $\mathbf{F}^* \times \mathbf{F}^*$ défini par $(\delta_1 \otimes \delta_2)(x, y) = \delta_1(x)\delta_2(y)$. Si (η_1, η_2) et (η'_1, η'_2) sont deux éléments de E_{III} tels que $\beta_{\eta_1, \eta_2} = \beta_{\eta'_1, \eta'_2}$, on voit, en regardant ce que cela signifie sur les $b_{x,y}$ et les b_x , que les caractères linéaires $\eta_1 \otimes \eta_2$, $\eta_2 \otimes \eta_1$, $\eta'_1 \otimes \eta'_2$ et $\eta'_2 \otimes \eta'_1$ sont liés par la relation

$$\eta_1 \otimes \eta_2 + \eta_2 \otimes \eta_1 = \eta'_1 \otimes \eta'_2 + \eta'_2 \otimes \eta'_1.$$

En utilisant l'orthogonalité des caractères, cela prouve que $\eta_1 \otimes \eta_2$ est égal à un des deux caractères $\eta'_1 \otimes \eta'_2$ ou $\eta'_2 \otimes \eta'_1$, et donc que les éléments (η_1, η_2) et (η'_1, η'_2) de E_{III} sont égaux.

- Le cas du type IV se traite de la même manière. Si η_1 et η_2 sont deux éléments de E_{IV} tels que $\gamma_{\eta_1} = \gamma_{\eta_2}$, alors en regardant ce que cela signifie sur les c_z et les b_x , on voit que l'on a $\eta_1 + \eta_1^* = \eta_2 + \eta_2^*$ sur \mathbf{K}^* . On en déduit, en utilisant l'orthogonalité des caractères, que η_1 est égal à η_2 ou à η_2^* , et donc que η_1 et η_2 sont égaux dans E_{IV} .

Ceci permet de conclure.

ANNEXE C

LE PROBLÈME DES NOMBRES CONGRUENTS

Ce chapitre est une introduction à la conjecture de Birch et Swinnerton-Dyer (un des problèmes à un million de dollar), à travers le problème des nombres congruents qui est probablement le plus vieux problème non résolu à ce jour.

C.1. Introduction

Définition C.1.1. — Un entier D , sans facteur carré (divisible par le carré d'aucun nombre premier), est *congruent*, s'il existe un triangle rectangle de cotés rationnels dont l'aire est D ; autrement dit, si et seulement s'il existe $a, b, c \in \mathbf{Q}$ avec $a^2 + b^2 = c^2$ et $D = \frac{ab}{2}$.

Pour étudier les nombres congruents, on peut commencer par étudier l'ensemble des triangles rectangles à côtés rationnels, c'est-à-dire résoudre l'équation $a^2 + b^2 = c^2$ en nombres rationnels. On pose $u = \frac{a}{c}$ et $v = \frac{b}{c}$, et on est ramené à trouver les points rationnels sur le cercle $u^2 + v^2 = 1$ avec $u > 0$ et $v > 0$. Pour cela, on note t la pente de la droite joignant (u, v) à $(-1, 0)$, dont l'équation est donc $v = t(u + 1)$; on a $t \in \mathbf{Q}$ et $(u, v) = (\frac{1-t^2}{t^2+1}, \frac{2t}{t^2+1})$. En conclusion, $a, b, c \in \mathbf{Q}$ sont les côtés d'un triangle rectangle si et seulement s'il existe $t \in \mathbf{Q}$, $0 < t < 1$, tel que $a = \frac{1-t^2}{t^2+1}c$ et $b = \frac{2t}{t^2+1}c$. En posant $x = -t$ et $y = \frac{t^2+1}{c}$, ce qui précède permet presque⁽¹⁾ de démontrer le résultat suivant.

Proposition C.1.2. — Si D est un entier positif sans facteur carré, alors les conditions suivantes sont équivalentes :

⁽¹⁾La relation $D = \frac{ab}{2}$ devient $D = \frac{x^3-x}{y^2}$, et la condition $0 < t < 1$ équivaut à $-1 < x < 0$. On a donc démontré que D est congruent si et seulement si l'équation $Dy^2 = x^3 - x$ a une solution dans \mathbf{Q}^2 avec $-1 < x < 0$. La courbe $C_D(\mathbf{R}) = \{(x, y) \in \mathbf{R}^2, Dy^2 = x^3 - x\}$ a deux composantes connexes : un ovale dans la région $-1 \leq x \leq 0$, et une courbe avec une direction asymptotique verticale dans la région $x \geq 1$. L'application qui, à $P = (x, y) \in C_D(\mathbf{R})$, associe $P' = (x', y')$, intersection de la droite $(P, (-1, 0))$ avec C_D , échange les deux composantes connexes comme le montre un petit dessin (ou un calcul explicite), et envoie $C_D(\mathbf{Q})$ dans lui-même comme il est expliqué au § C.2 (ou comme le montre un calcul explicite). Ceci permet de montrer que l'existence d'une solution dans \mathbf{Q}^2 avec $-1 < x < 0$ est équivalente à celle d'une solution dans \mathbf{Q}^2 avec $x > 1$. On en déduit la proposition.

- (i) D est congruent
- (ii) L'équation $Dy^2 = x^3 - x$ a une solution dans \mathbf{Q}^2 avec $y \neq 0$.

Déterminer si un entier est congruent ou pas, est un problème très ancien et très difficile. On a par exemple le résultat suivant « conjecturé » par Fibonacci (1175-1240).

Théorème C.1.3 (Fermat (1601-1665)). — 1 n'est pas un nombre congruent.

C'est une des nombreuses utilisations que Fermat a trouvées pour sa méthode de « la descente infinie ». Remarquons que si a, b, c sont des entiers non nuls vérifiant $a^4 - b^4 = c^4$, et si $x = \frac{a^2}{b^2}$, $y = \frac{ac^2}{b^3}$, alors $y = x^3 - x$. Le fait que 1 n'est pas congruent implique donc le théorème de Fermat⁽²⁾ pour l'exposant 4.

Exemple C.1.4 (Zagier). — L'entier 157 est congruent, mais le triangle (a, b, c) le plus simple d'aire 157 est

$$a = \frac{6803298487826435051217540}{411340519227716149383203}, \quad b = \frac{411340519227716149383203}{21666555693714761309610},$$

$$c = \frac{224403517704336969924557513090674863160948472041}{8912332268928859588025535178967163570016480830}.$$

Cet exemple montre que la chasse aux triangles rectangles à côtés rationnels d'aire D risque d'être un peu acrobatique... Le résultat suivant de Tunnell (1983) n'en est que plus remarquable.

Théorème C.1.5. — Soit D un entier impair sans facteur carré. Si D est congruent, alors

$$|\{x, y, z \in \mathbf{Z}, 2x^2 + y^2 + 8z^2 = D\}| = 2 \cdot |\{x, y, z \in \mathbf{Z}, 2x^2 + y^2 + 32z^2 = D\}|. \quad (*)$$

Réciproquement, si D vérifie $(*)$, et si (une forme faible de) la conjecture de Birch et Swinnerton-Dyer est vraie, alors D est congruent.

Il y a un résultat similaire pour D pair. Comme il est très facile de décider si D vérifie ou non $(*)$, cela fournit un critère effectif permettant de décider qu'un nombre donné est non congruent, ou (sous Birch et Swinnerton-Dyer) congruent, et ce, sans exhiber de triangle rectangle d'aire D . Un entier congru à 5 ou 7 modulo 8 vérifie $(*)$ car les deux ensembles sont vides, mais on ne sait pas montrer que cela implique que D est congruent...

Comme le lecteur le constatera, la démonstration de ce théorème emprunte des chemins très détournés (ce qui en fait le charme); on peut légitimement se demander si une preuve plus directe ne serait pas possible, maintenant qu'on connaît la réponse.

⁽²⁾Il semble que Fermat se soit légèrement laissé emporté par son enthousiasme devant cette découverte...

C.2. Arithmétique des courbes elliptiques

Si C est une conique, on peut étudier l'ensemble $C(\mathbf{Q})$, des points de C à coordonnées rationnelles, comme on l'a fait pour le cercle. On trouve un point $P \in C(\mathbf{Q})$ sur la conique et on paramètre les points de la conique par la pente d'une droite variable passant par P . Cette stratégie ne marche plus pour une courbe C donnée par une équation de degré 3 (comme la courbe C_D d'équation $Dy^2 = x^3 - x$) car, si on coupe par une droite passant par un point de $C(\mathbf{Q})$, et qu'on élimine y entre les deux équations, on obtient une équation de degré 3 en x dont on sait seulement qu'une des solutions est rationnelle; les deux autres vivent donc, en général, dans une extension quadratique de \mathbf{Q} , mais pas dans \mathbf{Q} . Par contre, si on prend une droite passant par deux points rationnels de C ou tangente à un point rationnel de C , alors on obtient une équation dont deux des solutions (ou une solution double) sont rationnelles; comme la somme des racines est aussi rationnelle, cela montre que cette droite recoupe C en un point rationnel.

Une *courbe elliptique* E sur un corps K est une courbe d'équation $y^2 = P(x)$, avec $P \in K[X]$, de degré 3, sans racine double⁽³⁾. On note $E(K)$ l'ensemble des solutions dans K^2 de $y^2 = P(x)$, et $\bar{E}(K) = E(K) \cup \{\infty\}$, avec la convention qu'une droite passe par ∞ si et seulement si elle est verticale⁽⁴⁾. On munit $\bar{E}(K)$ d'une loi de composition $+$ qui en fait un groupe commutatif⁽⁵⁾ avec ∞ comme élément neutre et $P + Q + R = \infty$ si et seulement si (P, Q, R) sont alignés (avec les conventions évidentes si deux ou trois des points sont confondus; en particulier, P est d'ordre 2 si et seulement si ∞ appartient à la tangente à E en P , c'est-à-dire si et seulement si $y = 0$; de même, P est d'ordre 3 si et seulement si la tangente en P à E a un contact d'ordre 3).

⁽³⁾ Ceci se traduit par la non nullité du discriminant $\Delta(P)$ de P : si $P(x) = ax^3 + bx^2 + cx + d$, on a

$$\Delta(P) = \begin{vmatrix} a & 0 & 3a & 0 & 0 \\ b & a & 2b & 3a & 0 \\ c & b & c & 2b & 3a \\ d & c & 0 & c & 2b \\ 0 & d & 0 & 0 & c \end{vmatrix}.$$

La matrice ci-dessus est celle de l'application $(U, V) \mapsto UP + VP'$, où U est de degré ≤ 1 , V de degré ≤ 2 . La nullité de $\Delta(P)$ est donc équivalente à l'existence de (U, V) avec $UP = -VP'$ et U de degré ≤ 1 , V de degré ≤ 2 , ce qui est possible si et seulement si P et P' ne sont pas premiers entre eux.

⁽⁴⁾ Cette définition de $\bar{E}(K)$ est parfaitement artificielle. Une définition naturelle demande de travailler dans le plan projectif \mathbf{P}^2 , espace des droites de l'espace vectoriel de dimension 3. Celui-ci peut être vu comme la réunion du plan affine et d'une droite (projective) à l'infini dont les points correspondent aux directions de droites du plan affine; notre ∞ est le point de cette droite à l'infini correspondant à la direction verticale.

⁽⁵⁾ L'associativité n'est pas une évidence. Elle peut se vérifier par un calcul explicite assez pénible (mais on peut demander l'aide d'un ordinateur...). Une solution plus élégante consiste, si K est un sous-corps de \mathbf{C} , à passer par les fonctions elliptiques (cf. § VI.3), ce qui permet de montrer que $\bar{E}(\mathbf{C})$ est isomorphe, en tant que groupe, à \mathbf{C}/Λ , où Λ est un réseau de \mathbf{C} . (Dans le cas général, il y a une jolie démonstration passant par la géométrie projective.)

Théorème C.2.1. — Si E est une courbe elliptique sur \mathbf{Q} , le groupe $\overline{E}(\mathbf{Q})$ est engendré par un nombre fini d'éléments; il est donc isomorphe à $\overline{E}(\mathbf{Q})_{\text{tors}} \oplus \mathbf{Z}^{r(E)}$, où $\overline{E}(\mathbf{Q})_{\text{tors}}$, sous-groupe des points d'ordre fini⁽⁶⁾, est un groupe fini, et $r(E) \in \mathbf{N}$.

Ce résultat, conjecturé par Poincaré vers 1900, a été démontré par Mordell en 1922 en adaptant la méthode de la descente infinie de Fermat; c'est un cas particulier du célèbre théorème de Mordell-Weil. Le groupe $\overline{E}(\mathbf{Q})_{\text{tors}}$ se calcule très facilement; par contre la détermination du rang $r(E)$ et des générateurs de $\mathbf{Z}^{r(E)}$ est très délicate. À ce jour, il n'y a pas d'algorithme⁽⁷⁾ dont on peut prouver qu'il va permettre de les déterminer, ce qui ne nous arrange pas en ce qui concerne le problème des nombres congruents, mais la conjecture de Birch et Swinnerton-Dyer, dont il sera question plus loin, fournirait un tel algorithme si elle était démontrée.

Exemple C.2.2. — On note C_D la courbe elliptique d'équation $Dy^2 = x^3 - x$. Alors $Q_1 = (-1, 0)$, $Q_2 = (0, 0)$ et $Q_3 = (1, 0)$ sont d'ordre 2, et $\overline{C}_D(\mathbf{Q})_{\text{tors}} = \{\infty, Q_1, Q_2, Q_3\}$. En conséquence, D est congruent si et seulement si $r(C_D) \geq 1$.

C.3. L'heuristique de Birch et Swinnerton-Dyer

Si p est un nombre premier, $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ est un corps. Si $r = \frac{a}{b} \in \mathbf{Q}$ et p ne divise pas b , on peut voir r comme un élément de \mathbf{F}_p en réduisant a et b modulo p (i.e. en prenant le quotient des images de a et b dans \mathbf{F}_p , ce qui ne dépend pas des choix de a et b). En particulier, si E est une courbe elliptique sur \mathbf{Q} d'équation $y^2 = P(x)$, on peut aussi considérer E comme une courbe elliptique sur \mathbf{F}_p pour tous les *bons* nombres premiers (ceux ne divisant ni les dénominateurs des coefficients de P , ni le numérateur de son discriminant (note 3)).

⁽⁶⁾Si K est un sous-corps de \mathbf{C} , et si $\overline{E}(\mathbf{C}) \cong \mathbf{C}/\Lambda$, alors le sous-groupe des points de n -torsion de $\overline{E}(K)$ s'identifie à un sous-groupe de $\frac{1}{n}\Lambda/\Lambda \cong (\mathbf{Z}/n\mathbf{Z})^2$; en particulier il est de cardinal $\leq n^2$.

⁽⁷⁾Lors du congrès international des mathématiciens de 1900, Hilbert a énoncé une série de 23 problèmes, dont le 10-ième était de produire un algorithme permettant de décider si une équation polynomiale (en plusieurs variables), à coefficients entiers, a, ou non, des solutions en nombres entiers. Ce problème fut finalement résolu par Matiyasevich en 1970, qui prouva qu'un tel algorithme ne peut pas exister, au grand soulagement des arithméticiens qui voyaient d'un mauvais œil l'idée qu'un ordinateur puisse les mettre au chômage. En poussant plus loin les méthodes de Matiyasevich, on peut construire des polynômes explicites pour lesquels on peut décider arbitrairement de l'existence ou de la non existence de solutions en nombres entiers, sans rajouter de contradiction dans les mathématiques... C'est un peu ennuyeux, car cela veut dire qu'on n'est jamais sûr que le problème auquel on s'attaque n'est pas de ce type. Le théorème de Matiyasevich n'exclut pas, a priori, l'existence d'un algorithme pour décider si un polynôme (en plusieurs variables) a des solutions rationnelles ou pas. Les courbes elliptiques fournissent le premier test non trivial dans cette direction, la conjecture de Birch et Swinnerton-Dyer fournissant un tel algorithme (cf. cor. C.7.2), si elle est vraie...

Si E est une courbe elliptique⁽⁸⁾ sur \mathbf{F}_p , on a trivialement, $|\overline{E}(\mathbf{F}_p)| \leq 2p + 1$, mais on dispose du résultat plus précis suivant.

Théorème C.3.1 (Hasse, 1933). — Si E est une courbe elliptique sur \mathbf{F}_p , et si $a_p = p + 1 - |\overline{E}(\mathbf{F}_p)|$, alors $|a_p| \leq 2\sqrt{p}$

L'idée de Birch et Swinnerton-Dyer (1960-1965), est que, si $r(E) \geq 1$, alors il devrait y avoir en moyenne plus de points dans $E(\mathbf{F}_p)$ que si $r(E) = 0$, à cause de la réduction modulo p des éléments de $E(\mathbf{Q})$. Comme ce nombre de points est à peu près p d'après le théorème de Hasse, le produit $\prod_p \frac{p}{|\overline{E}(\mathbf{F}_p)|}$ devrait avoir des chance de diverger (d'être nul), si $r(E) \geq 1$, et de converger, si $r(E) = 0$. Comme le produit n'est pas convergent au sens usuel, nous allons devoir passer par les fonctions holomorphes pour donner corps à cette heuristique.

C.4. Fonction L d'une courbe elliptique

Soit E une courbe elliptique sur \mathbf{Q} . Si p est un bon nombre premier, soit a_p l'entier défini par $a_p = 1 + p - |\overline{E}(\mathbf{F}_p)|$. On définit la⁽⁹⁾ fonction $L(E, s)$, et des entier a_n , pour $n \in \mathbf{N} - \{0\}$ par

$$L(E, s) = \prod_{p \text{ bon}} \frac{1}{1 - a_p p^{-s} + p^{1-2s}} = \sum_{n=1}^{+\infty} a_n n^{-s}.$$

Il est facile de voir que le produit converge pour $\operatorname{Re}(s) > 2$, et même $\operatorname{Re}(s) > \frac{3}{2}$ si on utilise la majoration $|a_p| \leq 2\sqrt{p}$ de Hasse, et définit une fonction holomorphe sur ce demi-plan.

Théorème C.4.1. — La fonction $L(E, s)$ admet un prolongement analytique à \mathbf{C} tout entier.

Ce résultat a été conjecturé par Hasse vers 1935 ; c'est un cas particulier de la conjecture de Hasse-Weil. Le premier résultat dans sa direction est celui, dû à Weil, de la famille⁽¹⁰⁾ des courbes C_D . Shimura (1958), inspiré par des travaux d'Eichler (1954), a démontré de

⁽⁸⁾On peut aussi considérer des courbes elliptiques sur $\mathbf{Z}/D\mathbf{Z}$, où D n'est pas premier, ou sur \mathbf{F}_q . Ceci est utilisé pour factoriser des grands nombres, prouver la primalité de grands nombres (de plus de 1000 chiffres), ou pour fabriquer des signatures plus sûres, à taille égale, que celles fournies par \mathbf{F}_q^* .

⁽⁹⁾Cette fonction n'est pas celle qui est habituellement considérée ; elle en diffère par la multiplication par des facteurs en les mauvais p , mais comme ceux-ci ne s'annulent pas en $s = 1$, cela ne change rien en ce qui concerne la conjecture de Birch et Swinnerton-Dyer. La bonne fonction $L(E, s)$ a une équation fonctionnelle plus sympathique que celle considérée dans cet article : il existe $\varepsilon \in \{\pm 1\}$ et $N \in \mathbf{N}$ tel que, si $\Lambda(E, s) = \frac{\Gamma(s)}{(2\pi)^s} N^{s/2} L(E, s)$, alors $\Lambda(E, 2-s) = \varepsilon \Lambda(E, s)$; en particulier, si $\varepsilon = -1$, alors $r_\infty(E)$ est impair et $L(E, 1) = 0$.

⁽¹⁰⁾Dans ce cas, on définit $\delta : \mathbf{Z}[i] \rightarrow \{0, 1, i, -1, -i\}$ par

$$\begin{cases} \delta(\omega) = 0 & \text{si } \omega \text{ est divisible par } 1+i \text{ dans } \mathbf{Z}[i], \\ \omega\delta(\omega) - 1 \text{ est divisible par } (1+i)^3 & \text{sinon.} \end{cases}$$

nombreux cas de cette conjecture en utilisant la théorie des formes modulaires dont il sera question plus loin. Le pas le plus important a été accompli par Wiles en 1994, dans sa quête de la démonstration du théorème de Fermat, qui a démontré cette conjecture dans le cas où E est d'équation $y^2 = P(x)$ et P a toutes ses racines dans \mathbf{Q} . Le cas général a finalement été résolu par Breuil, Conrad, Diamond et Taylor en 1999.

La quantité $\prod_{p \text{ bon}} \frac{p}{|\overline{E}(\mathbf{F}_p)|}$ apparaissant dans l'heuristique de Birch et Swinnerton-Dyer est, au moins formellement, égale à $L(E, 1)$, et leur heuristique devient :

Conjecture C.4.2 (Birch et Swinnerton-Dyer (forme faible))

« $r(E) \geq 1$ » si et seulement si « $L(E, 1) = 0$ ».

On peut préciser cet énoncé⁽¹¹⁾. Notons $r_\infty(E)$ l'ordre du zéro en $s = 1$ de $L(E, s)$. La conjecture de Birch et Swinnerton-Dyer prend alors la forme suivante.

Conjecture C.4.3. — (Birch et Swinnerton-Dyer) On a l'égalité $r(E) = r_\infty(E)$.

C'est sous cette forme que le problème vaut un million de dollar. Il y a en fait une forme plus précise⁽¹²⁾ de cette conjecture (donnant une formule pour $\lim_{s \rightarrow 1} (s - 1)^{-r(E)} L(E, s)$), et plus générale (\mathbf{Q} peut être remplacé par une extension finie, ou même par des corps de caractéristique p , extensions finies du corps $\mathbf{F}_p(T)$).

On a alors

$$L(C_1, s) = \frac{1}{4} \sum_{\omega \in \mathbf{Z}[i] - \{0\}} \frac{\omega \delta(\omega)}{|\omega|^{2s}} = \sum_{n=1}^{+\infty} a_n n^{-s}.$$

Le cas D général se déduit facilement du cas $D = 1$: on a $L(C_D, s) = \sum_{n=1}^{+\infty} \chi_D(n) a_n n^{-s}$, où $\chi_D : \mathbf{Z} \rightarrow \{-1, 0, 1\}$ est le symbole de Legendre modulo D . Ce symbole de Legendre est caractérisé par les propriétés suivantes : $\chi_D(n + 4D) = \chi_D(n)$, $\chi_D(n) = 0$ si $(D, n) \neq 1$, et

$$\chi_D(nm) = \chi_D(n)\chi_D(m), \quad \chi_D(p) = \begin{cases} 1 & \text{si } D \text{ est un carré dans } \mathbf{F}_p^*, \\ -1 & \text{si } D \text{ n'est pas un carré dans } \mathbf{F}_p^*. \end{cases}$$

L'existence de χ_D est une conséquence de la loi de réciprocité quadratique conjecturée par Euler en 1783 et démontrée par Gauss en 1801.

⁽¹¹⁾En fait, Birch et Swinnerton-Dyer avaient été plus optimistes et avaient conjecturé que

$$\prod_{p \text{ bon}, p \leq x} \frac{p}{|\overline{E}(\mathbf{F}_p)|} \sim C(\log x)^{-r(E)}.$$

Goldfeld (1982) a prouvé que, si c'est le cas, alors $r_\infty(E) = r(E)$, la fonction $L(E, s)$ vérifie l'hypothèse de Riemann (i.e. elle ne s'annule pas pour $\operatorname{Re}(s) > 1$), mais, de manière surprenante, que l'on a $C = \frac{L(E, 1)}{\sqrt{2}}$ au lieu de $C = L(E, 1)$, si $r(E) = 0$.

⁽¹²⁾Celle-ci prend la forme $\lim_{s \rightarrow 1} (s - 1)^{-r(E)} L(E, s) = |\operatorname{III}(E)| \cdot R_\infty(E) \cdot \Omega_\infty(E) \cdot \prod_{p \text{ mauvais}} c_p$, où c_p est un nombre rationnel explicite, $\Omega_\infty(E)$ est la période réelle de E (donnée par $\Omega_\infty(E) = 2 \int_\alpha^{+\infty} \frac{dx}{\sqrt{P(x)}}$, si E est d'équation $y^2 = P(x)$ et α est la plus grande racine réelle de P), $R_\infty(E)$ est un « régulateur » mesurant la taille des générateurs de $\overline{E}(\mathbf{Q})$, et $\operatorname{III}(E)$, le groupe de Tate-Shafarevich de E , est un groupe mystérieux, conjecturalement fini.

Les résultats sont peu nombreux ; ce sont les suivants.

- Coates et Wiles (1977) ont démontré que, si $E = C_D$ (ou si E est une courbe elliptique sur \mathbf{Q} à multiplication complexe⁽¹³⁾), alors « $L(E, 1) \neq 0$ » \Rightarrow « $r(E) = 0$ ».
- Gross et Zagier (1983) ont donné une formule explicite⁽¹⁴⁾ pour $L'(E, 1)$ en termes de certains points rationnels sur E , dits « de Heegner », et qui sont construits de manière purement analytique (ce sont ces points qui permettent d'amuser la galerie en exhibant des triangles rectangles à côtés rationnels avec un nombre astronomique de chiffres). Comme conséquence, ils obtiennent l'implication : « $r_\infty(E) = 1$ » \Rightarrow « $r(E) \geq 1$ ».
- Kolyvagin (1989) a démontré, en utilisant ces points de Heegner, l'implication suivante « $r_\infty(E) \leq 1$ » \Rightarrow « $r(E) = r_\infty(E)$ ».

C'est tout ! On est dans la situation paradoxale où plus il est censé y avoir de points rationnels ($r_\infty(E) \geq 2$), moins on sait en construire... Mentionnons quand-même, qu'en général, le rang $r(E)$ est égal à 0 ou 1, mais il y a tout lieu de croire que $r(E)$ peut prendre des valeurs arbitrairement grandes. Le record actuel est détenu par N. Elkies (2006) avec une courbe vérifiant $r(E) \geq 28$.

C.5. La stratégie de Tunnell

La théorème de Coates-Wiles mentionné ci-dessus fournit un critère pour que D ne soit pas congruent : il suffit que $L(C_D, 1) \neq 0$. Réciproquement, si la conjecture de Birch et Swinnerton-Dyer est vraie (même sous sa forme faible), alors la nullité de $L(C_D, 1)$ implique que D est congruent. C'est le point de départ de la démonstration du théorème de Tunnell. Le problème est donc de calculer $L(C_D, 1)$ et de décider si ce nombre est nul ou pas. Il y a deux problèmes sérieux qui se posent : le produit définissant $L(C_D, 1)$ converge beaucoup trop lentement (s'il converge..., cf. note 11) pour qu'on puisse l'utiliser pour le calcul de $L(C_D, 1)$, et de toute façon, il est impossible de prouver qu'un nombre réel est nul en le calculant de manière approchée, sauf si on sait par ailleurs qu'il s'agit d'un entier. La solution que Tunnell apporte à ces deux problèmes est particulièrement élégante.

On note $\mathcal{H} = \{z \in \mathbf{C}, \operatorname{Im}(z) > 0\}$ le demi-plan de Poincaré. On pose $q = e^{2i\pi z}$, et on définit $\Theta : \mathcal{H} \rightarrow \mathbf{C}$ par

$$\Theta(z) = \sum_{n \in \mathbf{Z}} q^{n^2}.$$

⁽¹³⁾ Si $(x, y) \in C_D(\mathbf{C})$, alors $(-x, iy) \in C_D(\mathbf{C})$. Si Λ est le réseau de \mathbf{C} correspondant à $\overline{C_D}(\mathbf{C})$ (cf. note 5), la remarque précédente se traduit par le fait que $i\Lambda = \Lambda$. On dit qu'une courbe elliptique E définie sur un sous-corps de \mathbf{C} a de la multiplication complexe si, Λ étant le réseau de \mathbf{C} qui lui correspond, il existe $\tau \in \mathbf{C} - \mathbf{R}$ tel que $\tau\Lambda \subset \Lambda$ (c'est donc le cas de C_D , avec $\tau = i$) ; un tel τ est alors racine d'un polynôme unitaire de degré 2 à coefficients dans \mathbf{Z} .

⁽¹⁴⁾ La démonstration de cette formule occupe une centaine de pages...

Le résultat que démontre Tunnell est alors le suivant.

Théorème C.5.1. — (Tunnell) Soit $\Omega = \int_1^{+\infty} \frac{dx}{\sqrt{x^3-x}}$, et soit $\sum_{n=0}^{+\infty} b_n q^n$ le développement de

$$\Theta(z) \cdot \Theta(2z) \cdot (2\Theta(32z) - \Theta(8z)),$$

alors, si D est impair (il y a une formule similaire pour les entiers pairs) sans facteur carré,

$$L(C_D, 1) = \frac{\Omega}{16\sqrt{D}} \cdot b_D^2.$$

Comme b_D est la différence des deux termes apparaissant dans la condition (*) du th. C.1.5, cela explique comment ledit théorème peut se déduire du théorème de Coates-Wiles. La démonstration du théorème C.5.1 repose sur la théorie des formes modulaires dont il est question au § suivant.

C.6. Formes modulaires

Si f est holomorphe sur \mathcal{H} et vérifie $f(z+1) = f(z)$, alors f a un développement de Fourier (q -développement)

$$f(z) = \sum_{n \in \mathbf{Z}} a_n q^n, \quad \text{avec } q = e^{2i\pi z}.$$

On dit que f est à croissance lente à l'infini si $a_n = 0$ pour tout $n < 0$ et s'il existe $C \in \mathbf{R}$ tel que $a_n = O(n^C)$.

Si N est un entier, on note $\Gamma_0(N)$ le sous-groupe de $\mathbf{SL}_2(\mathbf{Z})$ des $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec c divisible par N . On note $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Définition C.6.1. — Si $k \in \frac{1}{2}\mathbf{N}$ et $j : \Gamma_0(N) \rightarrow \{\text{racines de l'unité}\}$ vérifie $j(T) = 1$, l'espace $M_k(\Gamma_0(N), j)$ des formes modulaires de poids k et type j pour $\Gamma_0(N)$ est l'espace des fonctions f , holomorphes sur \mathcal{H} , vérifiant

$$f\left(\frac{az+b}{cz+d}\right) = j\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) (cz+d)^k f(z), \quad \text{quels que soient } z \in \mathcal{H} \text{ et } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N),$$

et qui sont à croissance lente à l'infini.

Remarque C.6.2. — (i) Il n'est pas du tout clair que de telles formes existent ; et de fait, il faut choisir correctement la fonction j pour $M_k(\Gamma_0(N), j)$ soit non nul.

(ii) $M_k(\Gamma_0(N), j)$ est un \mathbf{C} -espace vectoriel de dimension finie $\leq 1 + \frac{Nk}{12} \prod_{p|N} (1 + \frac{1}{p})$.

(iii) Les formes modulaires ont un don d'ubiquité assez remarquable. On les rencontre en théorie des nombres, en combinatoire ou en physique théorique, bien que ce soient des objets définis de manière purement analytique.

Pour un aperçu de la théorie des formes modulaires pour $\mathbf{SL}_2(\mathbf{Z})$, le lecteur est invité à se reporter aux exercices VII.6.1-VII.6.11 du chap. VII.

- La fonction $\theta(z) = \Theta(\frac{z}{2})$ est étudiée dans l'ex. VII.6.6, et les résultats de cet exercice montrent que Θ est une forme modulaire de poids $\frac{1}{2}$ pour $\Gamma_0(4)$ et un j un peu compliqué.

- Si k est un entier pair ≥ 3 , la série d'Eisenstein $G_k(z) = \frac{(k-1)!}{2 \cdot (2i\pi)^k} \sum'_{m,n} \frac{1}{(mz+n)^k}$ fait l'objet des ex. VII.6.4 et VII.6.5 ; elle appartient à $M_k(\mathbf{SL}_2(\mathbf{Z}), 1)$, et son q -développement est donné par $G_k = \frac{(k-1)!\zeta(k)}{(2i\pi)^k} + \sum_{n=1}^{+\infty} \sigma_{k-1}(n)q^n$, où $\sigma_t(n) = \sum_{d|n, d \geq 1} d^t$, si $t \in \mathbf{C}$ et $n \in \mathbf{N} - \{0\}$.

- La fonction $G_2 = \frac{\zeta(2)}{(2i\pi)^2} + \sum_{n=1}^{+\infty} \sigma_1(n)q^n$ est étudiée dans l'ex. VII.6.7. Elle n'est pas modulaire, mais presque, et $4G_2(4z) - G_2(z)$ est modulaire.

Comme échauffement pour le théorème de Tunnell, mentionnons (cf. ex. VII.6.9) l'identité de Jacobi (1829) : $4G_2(4z) - G_2(z) = \frac{3\zeta(2)}{(2i\pi)^2} \Theta^4$, qui se démontre en constatant que les deux membres appartiennent à $M_2(\Gamma_0(4), 1)$ qui est de dimension 2, et que la différence est divisible par q^2 . On en tire, en comparant les q -développements, une forme effective du théorème des 4 carrés de Lagrange (1770).

$$|\{(a, b, c, d) \in \mathbf{Z}^4, a^2 + b^2 + c^2 + d^2 = n\}| = 8 \sum_{d|n, 4 \nmid d} d.$$

C.7. Courbes elliptiques et formes modulaires

Théorème C.7.1. — Si E est une courbe elliptique définie sur \mathbf{Q} et $L(E, s) = \sum_{n=1}^{+\infty} a_n n^{-s}$, alors $f = \sum_{n=1}^{+\infty} a_n q^n \in M_2(N_E, 1)$, où N_E est un entier explicite ne dépendant que des p mauvais.

Autrement dit, une courbe elliptique définie sur \mathbf{Q} est modulaire. Ce résultat, conjecturé de manière vague par Taniyama en 1955, et précisé par Weil en 1966 suite aux travaux de Shimura sus-mentionnés, est celui que démontrent Wiles⁽¹⁵⁾ et Breuil-Conrad-Diamond-Taylor. Le prolongement analytique de $L(E, s)$ s'en déduit en utilisant la formule (cf. ex. VII.6.3)

$$L(E, s) = \frac{(2\pi)^s}{\Gamma(s)} \int_0^{+\infty} f(iy) y^s \frac{dy}{y},$$

⁽¹⁵⁾ Si $a^p + b^p = c^p$ est un contre-exemple au théorème de Fermat, on peut considérer la courbe elliptique introduite par Hellegouarch et Frey, d'équation $y^2 = x(x - a^p)(x + b^p)$. Wiles montre que cette courbe est modulaire, ce qui est en contradiction avec la « conjecture ε » de Serre (1984) démontrée par Ribet (1988). La « conjecture ε » décrit les congruences que l'on peut attendre entre formes modulaires. Dans le cas qui nous intéresse, cette conjecture prédit une congruence modulo p entre le q -développement de la forme modulaire attachée à la courbe elliptique $y^2 = x(x - a^p)(x + b^p)$, et celui de $g \in M_2(\Gamma_0(2), 1)$, ce qui n'est pas possible car cet espace est de dimension 1, et la divisibilité par p du terme constant du q -développement de g entraîne celle de tous les termes du q -développement.

ce qui permet d'utiliser les propriétés analytiques de f pour étudier $L(E, s)$. C'est un cas particulier de la philosophie de Langlands sur les fonctions L arithmétiques (elle devraient provenir de *formes automorphes*, généralisations des formes modulaires, et donc avoir des tas de propriétés mirifiques). Dans le cas de la courbe C_D , la forme modulaire que l'on obtient est une combinaison linéaire de fonctions thêta.

La modularité d'une courbe elliptique E en fournit une description analytique en termes du demi-plan de Poincaré. Ceci est à la base de la construction des points de Heegner (ce sont les images⁽¹⁶⁾ des points $\tau \in \mathcal{H}$ solutions d'une équation du second degré à coefficients dans \mathbf{Q}) qui, comme nous l'avons déjà mentionné, jouent un rôle essentiel dans la démonstration du résultat de Kolyvagin ($r(E) = r_\infty(E)$ si $r_\infty(E) \leq 1$) ; ce résultat n'est donc devenu valable pour toutes les courbes elliptiques sur \mathbf{Q} que depuis les travaux de Wiles et Breuil-Conrad-Diamond-Taylor.

Une autre application de la modularité des courbes elliptiques est le résultat suivant qui permet, modulo un calcul numérique, de déterminer la valeur de $L(E, 1)$, ce qui fournit, modulo la conjecture de Birch et Swinnerton-Dyer (sous sa forme faible), un algorithme pour décider de l'existence de solutions en nombres rationnels pour une équation $y^2 = P(x)$, avec P de degré 3.

Corollaire C.7.2. — (Manin-Drinfeld, 1973) *Si E est d'équation $y^2 = P(x)$, et si α est la plus grande racine réelle de P , alors*

$$\left(\int_{\alpha}^{+\infty} \frac{dx}{\sqrt{P(x)}} \right)^{-1} L(E, 1)$$

est un nombre rationnel de dénominateur explicite.

Le point de départ de la démonstration du théorème C.5.1 est un théorème de Waldspurger (1979). Si $f = \sum_{n=1}^{+\infty} a_n q^n \in M_{2k}(\Gamma_0(N), 1)$, k entier, si D est sans facteur carré et premier à N , et si χ_D est le caractère de Legendre (cf. note 10), on peut montrer que $f \otimes \chi_D$, défini par $f \otimes \chi_D = \sum \chi_D(n) a_n q^n$, est un élément de $M_{2k}(\Gamma_0(ND^2), 1)$. Le théorème de Waldspurger dit, de manière vague, que les $L(f \otimes \chi_D, k)$ sont, quand D varie, les carrés de coefficients de Fourier de formes modulaires de poids $k + \frac{1}{2}$ pour $\Gamma_0(N')$, avec N' explicite. « Il n'y a plus qu'à » exhiber une base de l'espace de ces formes modulaires et calculer quelques coefficients pour obtenir une identité valable pour tout D .

⁽¹⁶⁾La théorie de la multiplication complexe (note 13) permet de déterminer le corps de définition de ces points ; ce sont, d'après un théorème de Schneider (1937), les seuls éléments de \mathcal{H} , algébriques sur \mathbf{Q} , qui fournissent des points algébriques de E .

ANNEXE D

INTRODUCTION AU PROGRAMME DE LANGLANDS

Un titre plus honnête pour ce chapitre, qui contient beaucoup d'énoncés magnifiques, mais peu de démonstrations⁽¹⁾, serait « Introduction à l'existence du programme de Langlands », une vraie introduction au programme de Langlands pouvant difficilement se faire avant le M2.

On peut faire remonter les thèmes menant au programme de Langlands à la *loi de réciprocité quadratique* conjecturée par Euler (1783) et démontrée par Gauss (1801). Le problème est, étant donné un nombre premier ℓ ou, plus généralement, un entier $d \in \mathbf{Z}$ divisible par le carré d'aucun nombre premier, de décrire l'ensemble des nombres premiers p tels que d soit un carré dans \mathbf{F}_p . De manière équivalente, il s'agit de déterminer le nombre de zéros dans \mathbf{F}_p du polynôme $X^2 - d$. La surprise est que la réponse ne dépend que de la classe de p modulo D , avec $D = d$ si $d \equiv 1 \pmod{4}$, et modulo $D = 4d$ si $d \equiv 2, 3 \pmod{4}$. Plus précisément, il existe un caractère de Dirichlet χ_D de conducteur D et à valeurs dans $\{\pm 1\}$ tel que le nombre de zéros de $X^2 - d$ dans \mathbf{F}_p soit égal à $1 + \chi_D(p)$.

Du point de vue qui va nous intéresser dans ce chapitre, cette loi de réciprocité peut s'encoder dans une identité multiplicative entre fonctions L. On note P_D le polynôme $X^2 - X + \frac{1-d}{4}$, si $d \equiv 1 \pmod{4}$, et $X^2 - d$, si $d \equiv 2, 3 \pmod{4}$. Dans tous les cas, ce polynôme est à coefficients entiers, son discriminant est D , et ses racines sont $\frac{1 \pm \sqrt{d}}{2}$ ou $\pm \sqrt{d}$ suivant la congruence de d modulo 4. On note A_D l'anneau $\mathbf{Z}[X]/(P_D)$. On a donc $A_D = \mathbf{Z}[\frac{1+\sqrt{d}}{2}]$, si $d \equiv 1 \pmod{4}$, et $A_D = \mathbf{Z}[\sqrt{d}]$, si $d \equiv 2, 3 \pmod{4}$. On définit la fonction $\zeta_{A_D}(s)$ par la formule $\zeta_{A_D}(s) = \sum_{\mathfrak{n}} \frac{1}{|A_D/\mathfrak{n}|^s}$, où \mathfrak{n} décrit l'ensemble des idéaux de A_D tels que l'anneau A_D/\mathfrak{n} soit un anneau fini.⁽²⁾ Un peu de théorie algébrique des nombres

⁽¹⁾Les seules démonstrations, outre celles esquissées en notes de bas de page, se trouvent dans le § D.2 où l'on étend aux adèles la transformée de Fourier du chap. IV et la transformée de Mellin du § VII.2, qui en est l'analogue multiplicatif. Comme le lecteur le constatera, la plupart des difficultés sont concentrées sur les nombres réels, les nombres p -adiques se comportant plutôt comme des groupes finis.

⁽²⁾On remarquera que si on remplace A_D par \mathbf{Z} dans la définition précédente, on tombe sur la fonction zêta de Riemann.

permet de montrer que la série de Dirichlet $\zeta_{A_D}(s)$ converge pour $\operatorname{Re}(s) > 1$, et que la loi de réciprocité quadratique est équivalente à la factorisation $\zeta_{A_D}(s) = \zeta(s)L(\chi_D, s)$.

Par exemple, si $d = -1$ (et donc $D = -4$ et $A_D = \mathbf{Z}[i]$), le caractère χ_D est donné par $\chi_D(n) = 1$ si $n \equiv 1 \pmod{4}$, et $\chi_D(n) = -1$ si $n \equiv 3 \pmod{4}$. On en déduit la relation⁽³⁾

$$\frac{1}{4} \sum_{\substack{(n,m) \in \mathbf{Z}^2 \\ (n,m) \neq (0,0)}} \frac{1}{(n^2 + m^2)^s} = \zeta(s)L(\chi_4, s) = \frac{1}{1 - 2^{-s}} \prod_{p \equiv 1 \pmod{4}} \frac{1}{(1 - p^{-s})^2} \prod_{p \equiv 3 \pmod{4}} \frac{1}{1 - p^{-2s}},$$

et en identifiant les termes en p^{-s} des deux côtés, on retrouve le résultat de Fermat selon lequel *un nombre premier impair est somme de deux carrés si et seulement si il est de la forme $4n + 1$* .

La définition de ζ_{A_D} ci-dessus se généralise à tout anneau $A = \mathbf{Z}[X_1, \dots, X_d]/(P_1, \dots, P_r)$, où P_1, \dots, P_r sont des polynômes en X_1, \dots, X_d à coefficients dans \mathbf{Z} , ce qui permet d'associer à un tel anneau une *fonction zêta de Hasse-Weil* $\zeta_A(s) = \sum_n \frac{1}{|A/\mathfrak{n}|^s}$ comme ci-dessus. Comme un anneau fini de cardinal n est de manière unique un produit d'anneaux de cardinaux $p^{v_p(n)}$ (lemme des restes chinois), la fonction $\zeta_A(s)$ admet une décomposition en produit de facteurs d'Euler $\zeta_A(s) = \prod_{p \in \mathcal{P}} \zeta_{A,p}(s)$, et A. Weil a conjecturé en 1949 que $\zeta_{A,p}(s)$ est une fonction rationnelle de p^{-s} ce qui fut démontré par B. Dwork en 1959. Weil a aussi conjecturé que cette fonction rationnelle a une factorisation sous une forme indépendante de p , et ne dépendant que de la géométrie de l'espace des solutions $P_1(z_1, \dots, z_d) = \dots = P_r(z_1, \dots, z_d)$ dans \mathbf{C}^d , ce qui fut démontré par A. Grothendieck⁽⁴⁾ comme aboutissement d'un énorme programme ayant totalement révolutionné la géométrie algébrique. À la factorisation des facteurs d'Euler correspond une factorisation de la fonction ζ_A en fonctions L, et l'un des buts principaux du programme de Langlands est de comprendre chacune de ces fonctions L en vue, en particulier, de démontrer la *conjecture de Hasse-Weil* selon laquelle les fonctions zêta de Hasse-Weil possèdent un prolongement méromorphe à tout le plan complexe.

L'exemple le plus célèbre est probablement $A = \mathbf{Z}[X, Y]/(Y^2 - X^3 - \alpha X^2 - \beta X - \gamma)$, avec $\alpha, \beta, \gamma \in \mathbf{Z}$. Dans ce cas, A. Wiles (1994, dans le cas « semi-stable ») et C. Breuil, B. Conrad, F. Diamond et R. Taylor (1999, dans le cas général) ont démontré, ce qui avait été conjecturé de manière vague par Y. Taniyama en 1956, et précisé par A. Weil en 1966, qu'il existe une forme modulaire primitive f (cf. n° 4.1 du § D.1) telle que $\zeta_A(s) = \frac{\zeta(s-1)}{L(f,s)}$, à multiplication près par un nombre fini de facteurs d'Euler innocents. Ceci a permis à A. Wiles, en utilisant des résultats antérieurs de K. Ribet (1987), d'en déduire la non

⁽³⁾L'anneau $\mathbf{Z}[i]$ est principal, et si $n + mi \in \mathbf{Z}[i]$ est non nul, l'anneau $\mathbf{Z}[i]/(n + mi)$ est fini de cardinal $n^2 + m^2$; le facteur $\frac{1}{4}$ s'explique par le fait que $\alpha, -\alpha, i\alpha$ et $-i\alpha$ engendrent le même idéal, si $\alpha = n + mi \in \mathbf{Z}[i] - \{0\}$.

⁽⁴⁾Cela lui a valu une médaille Fields en 1966; P. Deligne en a obtenu une en 1978 pour avoir démontré la dernière des conjectures de Weil, *l'hypothèse de Riemann sur les corps finis*, selon laquelle les zéros et les pôles de $\zeta_{A,p}(s)$ se répartissent sur un nombre fini de droites verticales explicites.

existence d'un A de la forme ci-dessus avec $X^3 + \alpha X^2 + \beta X + \gamma = X(X - a^p)(X + b^p)$, où $a^p + b^p = c^p$ est un contreexemple au théorème de Fermat, et donc de démontrer le théorème de Fermat.

D.1. La conjecture d'Artin

1. Le groupe $\mathcal{G}_{\mathbf{Q}}$

Rappelons que $x \in \mathbf{C}$ est *algébrique* s'il existe $P \in \mathbf{Q}[X]$, non nul, tel que $P(x) = 0$. De manière équivalente, $x \in \mathbf{C}$ est algébrique si et seulement si la sous- \mathbf{Q} -algèbre $\mathbf{Q}[x]$ de \mathbf{C} engendrée par x est de dimension finie sur \mathbf{Q} ; c'est alors un sous-corps⁽⁵⁾ de \mathbf{C} . Ceci permet de montrer que, si x et y sont algébriques, alors $x + y$ et xy sont algébriques⁽⁶⁾. On en déduit que l'ensemble $\overline{\mathbf{Q}}$ des nombres algébriques est un sous-corps de \mathbf{C} .

On note $\mathcal{G}_{\mathbf{Q}}$ l'ensemble des *automorphismes de corps* de $\overline{\mathbf{Q}}$, c'est-à-dire l'ensemble des permutations σ de $\overline{\mathbf{Q}}$ telles que l'on ait

$$\sigma(x + y) = \sigma(x) + \sigma(y) \text{ et } \sigma(xy) = \sigma(x)\sigma(y), \quad \text{quels que soient } x, y \in \overline{\mathbf{Q}}.$$

Muni de la composition, $\mathcal{G}_{\mathbf{Q}}$ est un sous-groupe du groupe des permutations de $\overline{\mathbf{Q}}$; c'est même un sous-groupe du groupe des automorphismes \mathbf{Q} -linéaires⁽⁷⁾ de $\overline{\mathbf{Q}}$.

Le groupe $\mathcal{G}_{\mathbf{Q}}$ est un groupe gigantesque et extrêmement mystérieux, mais dont la compréhension serait cruciale pour beaucoup de problèmes. Le simple fait qu'il existe est déjà très utile.

- Le seul élément de $\mathcal{G}_{\mathbf{Q}}$ que l'on sache décrire est la conjugaison complexe, que nous noterons frob_{∞} . On a $(\text{frob}_{\infty})^2 = 1$, et Artin (1924) a démontré que tout élément d'ordre fini de $\mathcal{G}_{\mathbf{Q}}$ est conjugué à 1 ou frob_{∞} , et donc est d'ordre 1 ou 2.

- Si $P \in \mathbf{Q}[X]$ est irréductible de degré n , et si $\alpha_1, \dots, \alpha_n \in \overline{\mathbf{Q}}$ sont les racines de P , alors quel que soient $i, j \in \{1, \dots, n\}$, il existe⁽⁸⁾ $\sigma_{i,j} \in \mathcal{G}_{\mathbf{Q}}$, avec $\sigma_{i,j}(\alpha_i) = \alpha_j$, ce qui montre qu'il y a bien d'autres éléments que frob_{∞} dans $\mathcal{G}_{\mathbf{Q}}$.

- Si H est un sous-groupe de $\mathcal{G}_{\mathbf{Q}}$, alors l'ensemble $\overline{\mathbf{Q}}^H$ des éléments de $\overline{\mathbf{Q}}$ fixes par H est un sous-corps de $\overline{\mathbf{Q}}$. Réciproquement, si K est un sous-corps de $\overline{\mathbf{Q}}$, l'ensemble \mathcal{G}_K des

⁽⁵⁾Si $z \in \mathbf{Q}[x] - \{0\}$, la multiplication par x sur $\mathbf{Q}[x]$ est injective et \mathbf{Q} -linéaire; comme on est en dimension finie, elle est surjective, et donc x a un inverse.

⁽⁶⁾En effet, $x + y$ et xy sont tous deux éléments de $\mathbf{Q}[x, y]$, qui est engendré par les $x^i y^j$, avec $i \leq n - 1$ et $j \leq m - 1$, si les x^i , avec $i \leq n - 1$ (resp. les y^j , avec $j \leq m - 1$), engendrent $\mathbf{Q}[x]$ (resp. $\mathbf{Q}[y]$). On en déduit que $\mathbf{Q}[x + y]$ et $\mathbf{Q}[xy]$ sont de dimensions finies sur \mathbf{Q} puisqu'inclus dans $\mathbf{Q}[x, y]$.

⁽⁷⁾On commence par montrer que $\sigma(1) = 1$, puis que $\sigma(n) = n$, quel que soit $n \in \mathbf{Z}$, et finalement que σ induit l'identité sur \mathbf{Q} .

⁽⁸⁾C'est une conséquence de la théorie de Galois (cf. cours d'Y. Laszlo "Théorie de Galois" en seconde année).

$\sigma \in \mathcal{G}_{\mathbf{Q}}$ fixant tous les éléments de K est un sous-groupe fermé⁽⁹⁾ de $\mathcal{G}_{\mathbf{Q}}$. On obtient de la sorte⁽¹⁰⁾ une bijection entre les *corps de nombres* (i.e. les sous-corps de $\overline{\mathbf{Q}}$ dont la dimension sur \mathbf{Q} est finie), et les sous-groupes fermés d'indice⁽¹¹⁾ fini de $\mathcal{G}_{\mathbf{Q}}$, et on a $\overline{\mathbf{Q}}^{\mathcal{G}_K} = K$, si K est un corps de nombres et $\mathcal{G}_{\overline{\mathbf{Q}}^H} = H$, si H est un sous-groupe fermé d'indice fini de $\mathcal{G}_{\mathbf{Q}}$.

• On conjecture que $\mathcal{G}_{\mathbf{Q}}$ est tellement compliqué que tout groupe fini G est un quotient⁽¹²⁾ de $\mathcal{G}_{\mathbf{Q}}$ (problème inverse de Galois). On sait montrer⁽¹³⁾ que c'est vrai pour tout groupe d'ordre impair (I. Shafarevich, 1954), pour S_n et A_n quel que soit n , pour le monstre... ; je ne pense pas qu'on sache le démontrer pour $\mathbf{SL}_3(\mathbf{F}_9)$ par exemple (\mathbf{F}_9 est le corps à 9 éléments), bien qu'il soit beaucoup plus petit que le monstre.

2. Représentations de $\mathcal{G}_{\mathbf{Q}}$

Comme nous l'avons mentionné au chap. I, on n'a de prise sur $\mathcal{G}_{\mathbf{Q}}$ qu'à travers ses représentations. Nous ne nous intéresserons qu'aux représentations complexes ; celles-ci jouent un rôle fondamental mais sont, pour des raisons topologiques⁽¹⁴⁾, un peu trop rigides pour beaucoup d'applications, et on est aussi amené à considérer des représentations de

⁽⁹⁾ On met sur $\mathcal{G}_{\mathbf{Q}}$ la *topologie de Krull* qui, par définition, est la plus faible rendant continue l'action de $\mathcal{G}_{\mathbf{Q}}$ sur $\overline{\mathbf{Q}}$ muni de la topologie discrète. Ceci signifie que, si $\sigma_0 \in \mathcal{G}_{\mathbf{Q}}$, et si $\alpha \in \overline{\mathbf{Q}}$, l'ensemble $U_{\alpha}(\sigma_0) = \{\sigma \in \mathcal{G}_{\mathbf{Q}}, \sigma(\alpha) = \sigma_0(\alpha)\}$ est un ouvert de $\mathcal{G}_{\mathbf{Q}}$, et les $U_{\alpha}(\sigma_0)$, pour $\sigma_0 \in \mathcal{G}_{\mathbf{Q}}$ et $\alpha \in \overline{\mathbf{Q}}$ forment une base d'ouverts de $\mathcal{G}_{\mathbf{Q}}$. En particulier, tout ouvert de $\mathcal{G}_{\mathbf{Q}}$ contenant 1 contient $U_{\alpha}(1)$, pour un certain $\alpha \in \overline{\mathbf{Q}}$, et donc contient un sous-groupe d'indice fini.

⁽¹⁰⁾ C'est encore une conséquence de la théorie de Galois.

⁽¹¹⁾ Rappelons que, si $H \subset G$ sont des groupes, l'indice $[G : H]$ de H dans G est le cardinal de G/H .

⁽¹²⁾ Si c'est vrai, alors $G \times \cdots \times G$ est aussi un quotient de $\mathcal{G}_{\mathbf{Q}}$ et donc G est un quotient de $\mathcal{G}_{\mathbf{Q}}$ d'une infinité de manières différentes.

⁽¹³⁾ Pour montrer un énoncé de ce type, on peut procéder de la manière suivante. Si $P \in \mathbf{Q}[X]$, si $\alpha \in \overline{\mathbf{Q}}$ vérifie $P(\alpha) = 0$, et si $\sigma \in \mathcal{G}_{\mathbf{Q}}$, alors $0 = \sigma(P(\alpha)) = P(\sigma(\alpha))$ (car σ fixe les coefficients de P). On obtient de la sorte un morphisme de groupes de $\mathcal{G}_{\mathbf{Q}}$ dans le groupe des permutations des racines de P . La description de l'image Gal_P (qui, par construction, est un quotient de $\mathcal{G}_{\mathbf{Q}}$) fait l'objet de la théorie de Galois. Si P est irréductible de degré n , alors Gal_P est un sous-groupe de S_n , en général égal à S_n ; il faut pas mal d'astuce pour construire des polynômes irréductibles P tels que Gal_P ne soit pas un groupe symétrique. Par exemple, la construction d'un polynôme P tel que Gal_P soit le monstre utilise un mélange de géométrie complexe, de topologie algébrique, de théorie des groupes finis, de géométrie algébrique et de théorie des nombres.

⁽¹⁴⁾ Si $\rho : \mathcal{G}_{\mathbf{Q}} \rightarrow \mathbf{GL}_n(\mathbf{C})$ est un morphisme continu de groupes, on a $\rho(1) = 1$, et la continuité de ρ implique (cf. note 9) qu'il existe un sous-groupe H d'indice fini de $\mathcal{G}_{\mathbf{Q}}$ tel que, si $g \in H$, alors $\rho(g) - 1$ a toutes ses coordonnées $a_{i,j}$ vérifiant $|a_{i,j}| \leq \frac{1}{2n}$. En appliquant ceci à g^k , on en déduit que $|\text{Tr}(\rho(g)^k) - n| \leq \frac{1}{2}$ quel que soit $k \in \mathbf{Z}$, et il n'est pas très difficile d'en conclure que ceci implique que toutes les valeurs propres de $\rho(g)$ sont égales à 1, puis que $\rho(g) = 1$. En conclusion, le noyau de ρ contient un sous-groupe d'indice fini, et donc $\rho(\mathcal{G}_{\mathbf{Q}})$ est un groupe fini. En particulier, chaque représentation complexe de $\mathcal{G}_{\mathbf{Q}}$ ne fournit de renseignements que sur une toute petite partie de $\mathcal{G}_{\mathbf{Q}}$. La situation est nettement plus favorable sur \mathbf{Q}_p , et la définition naturelle des fonctions L apparaissant dans la factorisation de la fonction zêta de Hasse-Weil ζ_A de $A = \mathbf{Z}[X_1, \dots, X_d]/(P_1, \dots, P_r)$ utilise des représentations de $\mathcal{G}_{\mathbf{Q}}$ que Grothendieck a associées à la variété d'équation $P_1 = \cdots = P_r = 0$.

$\mathcal{G}_{\mathbf{Q}}$ sur d'autres corps (et mêmes sur des anneaux) que \mathbf{C} , comme le corps \mathbf{Q}_p des nombres p -adiques. Comme il est expliqué dans la note 14, si V est une représentation complexe de $\mathcal{G}_{\mathbf{Q}}$, alors $\mathcal{G}_{\mathbf{Q}}$ agit à travers un groupe fini, ce qui permet d'utiliser les résultats du chap. I. De plus, tous les calculs à effectuer pour expliciter les objets apparaissant dans la conjecture d'Artin du n° 3 se font dans le corps de nombres K , fixé par le noyau de ρ_V , ce qui explique qu'ils soient faisables bien qu'on ne maîtrise pas du tout $\mathcal{G}_{\mathbf{Q}}$.

Si $P \in \mathbf{Q}[X]$, alors $\mathcal{G}_{\mathbf{Q}}$ permute les racines de P , ce qui nous fournit une représentation de permutation V_P de $\mathcal{G}_{\mathbf{Q}}$. On peut montrer que toute \mathbf{C} -représentation irréductible de $\mathcal{G}_{\mathbf{Q}}$ apparaît comme une composante irréductible d'une représentation V_P , mais essayer de comprendre $\mathcal{G}_{\mathbf{Q}}$ en énumérant les éléments P de $\mathbf{Q}[X]$, et en décomposant les V_P en composantes irréductibles est à peu près aussi efficace que d'espérer tomber sur une démonstration de l'hypothèse de Riemann en faisant la liste de toutes les propositions logiques. Il y a quand même deux exemples de représentations intéressantes de $\mathcal{G}_{\mathbf{Q}}$ que l'on peut obtenir de cette manière.

- *Racines carrées.* Si $d \in \mathbf{Q}^*$ n'est pas un carré dans \mathbf{Q} , alors $\sqrt{d} \notin \mathbf{Q}$, et si $\sigma \in \mathcal{G}_{\mathbf{Q}}$, il existe $\eta_d(\sigma) \in \{\pm 1\}$ tel que $\sigma(\sqrt{d}) = \eta_d(\sigma)\sqrt{d}$, et il est facile de vérifier que $\sigma \mapsto \eta_d(\sigma)$ est un caractère linéaire de $\mathcal{G}_{\mathbf{Q}}$.

- *Racines de l'unité.* Soit D un entier ≥ 3 , et soit $\alpha = e^{2i\pi/D}$. Alors $\alpha^D = 1$ et $\alpha^d \neq 1$ pour tout diviseur strict d de D . Maintenant, si $\sigma \in \mathcal{G}_{\mathbf{Q}}$, alors $\sigma(\alpha)^D = \sigma(\alpha^D) = 1$ et $\sigma(\alpha)^d = \sigma(\alpha^d) \neq 1$, pour tout diviseur strict d de D . On en déduit l'existence de $\chi_{\text{cycl},D}(\sigma) \in (\mathbf{Z}/D\mathbf{Z})^*$ tel que $\sigma(\alpha) = e^{2i\pi\chi_{\text{cycl},D}(\sigma)/D} = \alpha^{\chi_{\text{cycl},D}(\sigma)}$. On a alors $\sigma(\alpha^n) = \alpha^{n\chi_{\text{cycl},D}(\sigma)}$, quel que soit $n \in \mathbf{Z}$; on en déduit que $\chi_{\text{cycl},D} : \mathcal{G}_{\mathbf{Q}} \rightarrow (\mathbf{Z}/D\mathbf{Z})^*$ est un morphisme de groupes⁽¹⁵⁾. Ceci permet d'associer une représentation continue ρ_{χ} de $\mathcal{G}_{\mathbf{Q}}$, de dimension 1, à tout caractère de Dirichlet primitif : si χ est un tel caractère, et si D est son conducteur, alors $\rho_{\chi} = \chi \circ \chi_{\text{cycl},D}$ est un morphisme de groupes de $\mathcal{G}_{\mathbf{Q}}$ dans \mathbf{C}^* .

Théorème D.1.1. — (Kronecker-Weber) *Toute représentation de dimension 1 de $\mathcal{G}_{\mathbf{Q}}$ est de la forme ρ_{χ} , où χ est un caractère de Dirichlet primitif.*

Ce théorème a été énoncé par Kronecker (1853), mais il a fallu attendre 1886 pour une démonstration (presque) juste (par Weber).

Au vu du théorème de Kronecker-Weber, on peut se demander ce qui se passe si on considère les représentations de dimension 1 de \mathcal{G}_K , où K est un corps de nombres, ou bien si on considère les représentations de dimension n de $\mathcal{G}_{\mathbf{Q}}$, avec n fixé, ou encore si on mélange les deux problèmes et on considère les représentations de dimension n de \mathcal{G}_K , où K est un corps de nombres et n est fixé. La description des représentations de dimension 1 de \mathcal{G}_K est ce qu'on appelle la *théorie du corps de classes* (qui a occupé les arithméticiens pendant une bonne trentaine d'années au début du 20-ième siècle).

La dimension $n \geq 2$ a longtemps été considérée comme sans espoir.

⁽¹⁵⁾C'est le caractère cyclotomique.

3. Fonctions L d'Artin

Soit \mathcal{P} l'ensemble des nombres premiers. Si $p \in \mathcal{P}$, on note $|\cdot|_p$ la norme p -adique sur \mathbf{Q} , et on choisit⁽¹⁶⁾ une extension de $|\cdot|_p$ en une norme sur $\overline{\mathbf{Q}}$. On note alors $\mathcal{G}_{\mathbf{Q}_p}$ l'ensemble des $\sigma \in \mathcal{G}_{\mathbf{Q}}$ qui sont des isométries pour $|\cdot|_p$. On note I_p le sous-groupe de $\mathcal{G}_{\mathbf{Q}_p}$ des σ tels que $\sigma(e^{2i\pi/D}) = e^{2i\pi/D}$, quel que soit D premier à p . On dit qu'un élément σ de $\mathcal{G}_{\mathbf{Q}_p}$ est un frobenius en p , si $\sigma(e^{2i\pi/D}) = e^{2i\pi p/D}$, quel que soit D premier à p . Par définition de I_p , deux frobenius en p ne diffèrent que par un élément de I_p . On choisit un frobenius frob_p en p .

Soit V une représentation de $\mathcal{G}_{\mathbf{Q}}$ de dimension n . Si p est un nombre premier, on note V^{I_p} le sous-espace de V fixe par I_p . On démontre⁽¹⁷⁾ que l'on a $V^{I_p} = V$ pour presque tout⁽¹⁸⁾ p . L'action de frob_p sur V^{I_p} ne dépend alors pas du choix de frob_p , et $E_p(V, T) = \det(1 - T\rho_{V^{I_p}}(\text{frob}_p))$ ne dépend ni du choix de frob_p , ni de celui de l'extension de $|\cdot|_p$ à $\overline{\mathbf{Q}}$. C'est un polynôme de degré $\dim V^{I_p}$, et donc $E_p(V, T)$ est un polynôme de degré n pour presque tout p . On définit alors la fonction L d'Artin $L(V, s)$ attachée à V par le produit eulérien

$$L(V, s) = \prod_{p \in \mathcal{P}} E_p(V, p^{-s})^{-1}.$$

Exemple D.1.2. — (i) Si $V = \mathbf{1}$ est la représentation triviale, alors $V^{I_p} = V$ et $E_p(V, T) = 1 - T$ quel que soit p . On a donc $L(\mathbf{1}, s) = \prod_{p \in \mathcal{P}} (1 - p^{-s})^{-1} = \zeta(s)$.

(ii) Plus généralement, si χ est un caractère de Dirichlet primitif, on a $L(\rho_\chi, s) = L(\chi, s)$.

(iii) Soit $d \in \mathbf{Z}$ divisible par le carré d'aucun nombre premier, et soient D et χ_D le caractère de Dirichlet modulo D à valeurs dans $\{\pm 1\}$ apparaissant dans l'introduction. La loi de réciprocité quadratique de l'introduction peut se reformuler comme une égalité $L(\eta_d, s) = L(\chi_D, s)$ entre une fonction L d'Artin et une fonction L de Dirichlet.

(iv) Si V_1 et V_2 sont deux représentations de $\mathcal{G}_{\mathbf{Q}}$, et si $V = V_1 \oplus V_2$, alors $L(V, s) = L(V_1, s)L(V_2, s)$. Par exemple, la représentation $V = \text{Ind}_{\mathcal{G}_{\mathbf{Q}[\sqrt{d}]}}^{\mathcal{G}_{\mathbf{Q}}} \mathbf{1}$, se décompose sous la forme $V = \mathbf{1} \oplus \eta_d$, et l'identité $\zeta_{A_D}(s) = \zeta(s)L(\chi_D, s)$ de l'introduction est une traduction de l'identité $L(V, s) = L(\mathbf{1}, s)L(\eta_d, s)$.

Le *degré d'une fonction L d'Artin* est la dimension de la représentation de $\mathcal{G}_{\mathbf{Q}}$ sous-jacente; c'est aussi le degré de presque tous les facteurs d'Euler. D'après le théorème de Kronecker-Weber, les fonctions L de Dirichlet (en incluant la fonction zêta) décrivent l'ensemble des fonctions L d'Artin de degré 1. Il résulte des th. VII.3.4 et VII.4.4, que :

⁽¹⁶⁾Démontrer qu'il en existe demande un peu de travail; cela fait partie de la théorie algébrique des nombres. On démontre aussi que si on a deux extensions $|\cdot|_{p,1}$ et $|\cdot|_{p,2}$ de $|\cdot|_p$ à $\overline{\mathbf{Q}}$, alors il existe $\sigma \in \mathcal{G}_{\mathbf{Q}}$ tel que $|x|_{p,2} = |\sigma(x)|_{p,1}$, quel que soit $x \in \overline{\mathbf{Q}}$, ce qui explique que ce qui suit ne dépend pas du choix de cette extension.

⁽¹⁷⁾C'est encore un résultat de théorie algébrique des nombres.

⁽¹⁸⁾L'expression « pour presque tout p » signifie « à l'exception d'un nombre fini de p ».

Théorème D.1.3. — Une fonction L d'Artin de degré 1 possède un prolongement méromorphe à \mathbf{C} tout entier, holomorphe en dehors d'un pôle simple en $s = 1$, si $L = \zeta$.

Conjecture D.1.4. — (Artin, 1923) Si V est irréductible de dimension ≥ 2 , alors $L(V, s)$ a un prolongement holomorphe à \mathbf{C} tout entier.

Cette conjecture est très loin d'être démontrée, mais il y a eu récemment des progrès spectaculaires en dimension 2.

La théorie du corps de classes permet (cf. th. D.1.8 et D.1.9) de prouver que, si K est un corps de nombres, et si χ est un caractère linéaire de \mathcal{G}_K , alors $L(\text{Ind}_{\mathcal{G}_K}^{\mathcal{G}_{\mathbf{Q}}} \chi, s)$ est une fonction méromorphe sur \mathbf{C} , holomorphe en dehors d'un pôle simple en $s = 1$, si $\chi = 1$.

Artin lui-même a démontré, en utilisant ce résultat et la théorie des représentations des groupes finis (plus précisément le th. I.3.15 dont c'était la motivation principale⁽¹⁹⁾) qu'une puissance de $L(V, s)$ a un prolongement méromorphe à \mathbf{C} tout entier. Si on utilise le théorème de Brauer (th. I.3.16) au lieu du théorème d'Artin, on peut montrer (et c'était la motivation principale de Brauer) que $L(V, s)$ a un prolongement méromorphe à \mathbf{C} tout entier.

R. Langlands (1967) a proposé un gigantesque programme dont un des buts est la démonstration de la conjecture d'Artin.

4. Fonctions L de degré 2

4.1. Représentations impaires et formes modulaires

Soit D un entier, et soit $\Gamma_0(D)$ l'ensemble des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ à coefficients dans \mathbf{Z} , vérifiant $ad - bc = 1$ et $c \equiv 0$ modulo D . C'est le sous-groupe de $\mathbf{SL}_2(\mathbf{Z})$, image inverse du groupe des matrices triangulaires supérieures dans $\mathbf{SL}_2(\mathbf{Z}/D\mathbf{Z})$.

Si $k \geq 1$ est un entier, et si χ est un caractère de Dirichlet modulo D , une forme quasi-modulaire f de poids k , niveau⁽²⁰⁾ D et caractère χ est une fonction holomorphe sur \mathcal{H} , telle que $f\left(\frac{az+b}{cz+d}\right) = \chi(d)(cz+d)^k f(z)$ quel que soit $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(D)$. Une forme quasi-modulaire est en particulier périodique de période 1, et donc a un développement

⁽¹⁹⁾ Soit H le noyau de ρ_V , et soit $G = \mathcal{G}_{\mathbf{Q}}/H$; c'est un groupe fini, et V peut être considérée comme une représentation de G . D'après le th. I.3.15, il existe $d_V \in \mathbf{N}$, une famille (C_i, χ_i, n_i) , pour $i \in I$, où C_i est un sous-groupe (cyclique) de G , χ_i est un caractère linéaire de C_i , et $n_i \in \mathbf{Z}$, telle que $\chi_V = \sum_{i \in I} n_i \text{Ind}_{C_i}^G \chi_i$. Soit alors K_i le corps de nombres tel que \mathcal{G}_{K_i} soit l'image réciproque de C_i dans $\mathcal{G}_{\mathbf{Q}}$. On peut considérer χ_i comme un caractère linéaire de \mathcal{G}_{K_i} , et on déduit du (iv) de l'exemple D.1.2, la relation

$$L(V, s)^{d_V} = \prod_{i \in I} L(\text{Ind}_{\mathcal{G}_{K_i}}^{\mathcal{G}_{\mathbf{Q}}} \chi_i, s)^{n_i}.$$

⁽²⁰⁾ Si $D = 1$, on a $\Gamma_0(D) = \mathbf{SL}_2(\mathbf{Z})$ et $\chi = 1$. Comme $\mathbf{SL}_2(\mathbf{Z})$ est engendré par $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ et $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, la condition de quasi-modularité est équivalente aux conditions $f(z+1) = f(z)$ et $z^{-k}f(-1/z) = f(z)$ utilisée dans les exercices du § VII.5.

de Fourier (*q-développement*) du type

$$f(z) = \sum_{n \in \mathbf{Z}} a_n(f) q^n, \quad \text{avec } q = e^{2i\pi z}.$$

On dit que f est une *forme modulaire parabolique* (ou cuspidale) de poids k , niveau D et caractère χ si elle est quasi-modulaire, et si elle est à *décroissance rapide à l'infini*, ce qui signifie que $\text{Im}(z)^N (cz + d)^{-k} f\left(\frac{az+b}{cz+d}\right)$ est bornée sur le demi-plan $\text{Im}(z) \geq 1$, quels que soient⁽²¹⁾ $N \in \mathbf{N}$ et $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbf{Z})$. On note $S_k(D, \chi)$ l'espace de ces formes.

Si $f \in S_k(D, \chi)$, on a en particulier $a_n(f) = 0$, si $n \leq 0$ à cause de la condition de décroissance à l'infini, et on démontre comme dans l'ex. VII.6.3 que $a_n(f) = O(n^{k/2})$, ce qui permet d'associer à f une fonction L définie par

$$L(f, s) = \frac{(2\pi)^s}{\Gamma(s)} \int_0^{+\infty} f(iy) y^s \frac{dy}{y} = \sum_{n=1}^{+\infty} \frac{a_n(f)}{n^s}.$$

Un petit calcul montre que $f \mapsto f|_k w$, où $f|_k w$ est définie par $(f|_k w)(z) = z^{-k} f(-1/Dz)$, envoie $S_k(D, \chi)$ dans $S_k(D, \bar{\chi})$. Ceci permet, en coupant, comme dans les ex. VII.6.3, VII.6.6, l'intégrale ci-dessus en \sqrt{D} , de démontrer que $L(f, s)$ a un prolongement holomorphe à \mathbf{C} tout entier et vérifie une équation fonctionnelle reliant s et $k - s$. On dit que f est *primitive*, si $L(f, s)$ admet

- un développement en produit de facteurs d'Euler

$$L(f, s) = \prod_{p|D} \frac{1}{1 - a_p(f)p^{-s}} \prod_{p \nmid D} \frac{1}{1 - a_p(f)p^{-s} + \chi(p)p^{k-1-2s}},$$

- une équation fonctionnelle du type $\Lambda(f, k - s) = w \Lambda(f^*, s)$, où $f^*(z) = \overline{f(-\bar{z})}$, w est un nombre complexe de module 1, et

$$\Lambda(f, s) = \frac{\Gamma(s)}{(2\pi)^s} D^{s/2} L(f, s).$$

L'existence de formes modulaires primitives relève du miracle, mais Atkin et Lehner ont démontré que les $f\left(\frac{az+b}{d}\right)$, où f décrit les formes primitives de poids k , et a, b, d les entiers ($a \neq 0$ et $d \neq 0$) forment une famille génératrice de l'espace vectoriel engendré par les formes paraboliques de poids k de tout niveau.

Théorème D.1.5. — Si $\rho : \mathcal{G}_{\mathbf{Q}} \rightarrow \mathbf{GL}_2(\mathbf{C})$ est une représentation irréductible impaire⁽²²⁾, alors il existe une forme modulaire primitive f , de poids 1, telle que $L(f, s) = L(\rho, s)$. En particulier, la conjecture d'Artin est vérifiée pour une telle représentation.

⁽²¹⁾ Comme $f\left(\frac{az+b}{cz+d}\right) = \chi(d)(cz + d)^k f(z)$, si $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(D)$, il suffit de vérifier ceci pour un système de représentants de $\Gamma_0(D) \backslash \mathbf{SL}_2(\mathbf{Z})$, qui est un ensemble fini.

⁽²²⁾ Cela signifie que $\rho(\text{frob}_{\infty}) \neq \pm 1$ et donc que $\det \rho(\text{frob}_{\infty}) = -1$, ce qui est, de loin, le cas le plus fréquent. Si $\det \rho(\text{frob}_{\infty}) = 1$, la représentation ρ est dite *paire*.

Ce théorème est tout récent ; il date de janvier 2007. Le cas où l'image⁽²³⁾ de ρ dans $\mathbf{PGL}_2(\mathbf{C})$ est un groupe diédral remonte aux années 30, et suit des travaux de Artin et Hecke. L'un des premiers succès du programme de Langlands a été de démontrer un tel résultat dans le cas où l'image n'est pas A_5 (Langlands (1980) et J. Tunnell (1981)). Ceci couvre en particulier le cas où l'image de ρ dans $\mathbf{GL}_2(\mathbf{C})$ est le groupe⁽²⁴⁾ $\mathbf{GL}_2(\mathbf{F}_3)$, qui a servi de point de départ à Wiles pour démontrer le théorème de Fermat. Les méthodes de Wiles ont été transformées en une machine très efficace, par des mathématiciens de toute la planète, pour passer d'un point à un autre dans l'ensemble des représentations de dimension 2 de $\mathcal{G}_{\mathbf{Q}}$. La touche finale vient d'être apportée⁽²⁵⁾ par M. Kisin (australien en poste à Chicago), C. Khare (indien en poste à Los Angeles) et J-P. Wintenberger (Strasbourg). Un des points amusants de la démonstration est qu'elle passe par une récurrence sur l'ensemble des nombres premiers.

4.2. Représentations paires et formes de Maass

Les formes de Maass sont des objets beaucoup plus récents que les formes modulaires puisqu'elles n'ont été introduites, par H. Maass, qu'en 1949.

Une *forme de Maass* f de niveau D , caractère χ et valeur propre $\lambda \in \mathbf{C}$ est une fonction \mathcal{C}^∞ sur \mathcal{H} (vu comme un ouvert de \mathbf{R}^2) et vérifiant les conditions suivantes :

- (i) $f\left(\frac{az+b}{cz+d}\right) = \chi(d)f(z)$, quel que soit $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(D)$;
- (ii) $\Delta f = \lambda f$, où $\Delta = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$ est le laplacien hyperbolique.
- (iii) f est à décroissance rapide à l'infini.

Une telle forme a un développement de Fourier du type⁽²⁶⁾

$$f(z) = \sum_{n \in \mathbf{Z} - \{0\}} a_n(f) \sqrt{y} K_\nu(2\pi|n|y) e^{2i\pi nx},$$

⁽²³⁾L'image de ρ dans $\mathbf{GL}_2(\mathbf{C})$ peut être arbitrairement grande car on peut toujours tordre une représentation par un caractère linéaire attaché à un caractère de Dirichlet. Par contre, si on regarde l'image de ρ dans le quotient $\mathbf{PGL}_2(\mathbf{C})$ de $\mathbf{GL}_2(\mathbf{C})$ par le sous-groupe des homothéties, on tue ce degré de liberté, et l'image de ρ est soit de type diédral (groupe des symétries d'un polygone régulier), soit isomorphe à A_4 , S_4 ou A_5 . Le cas le plus difficile est celui de A_5 car ce groupe est simple contrairement aux autres groupes de la liste qui se déviennent en une suite de groupes commutatifs.

⁽²⁴⁾Les représentations du type IV dans la figure 4 de l'annexe B sont de dimension 2 pour $\mathbf{GL}_2(\mathbf{F}_3)$.

⁽²⁵⁾Ils obtiennent ce résultat en démontrant une conjecture plus fine et plus générale de J-P. Serre(1986) qui a servi de fil conducteur pour tous les résultats du sujet postérieurs à ceux de Langlands et Tunnell, dont la démonstration du théorème de Fermat.

⁽²⁶⁾L'équation aux dérivées partielles satisfaite par f se traduit en une équation différentielle du second ordre pour les coefficients de Fourier, et une seule des solutions (à multiplication près par une constante) n'explose pas à l'infini.

où $\frac{1}{4} - \nu^2 = \lambda$ et $K_\nu(y) = \frac{1}{2} \int_0^{+\infty} e^{-y(t+t^{-1})/2} t^\nu \frac{dt}{t}$ est une *fonction de Bessel* dont une des vertus est de vérifier, si $\operatorname{Re}(s) > |\operatorname{Re}(\nu)|$,

$$\int_0^{+\infty} K_\nu(y) y^s \frac{dy}{y} = 2^{s-2} \Gamma\left(\frac{s+\nu}{2}\right) \Gamma\left(\frac{s-\nu}{2}\right).$$

On dit que f est paire si $f(-\bar{z}) = f(z)$ et impaire si $f(-\bar{z}) = -f(z)$; toute forme de Maass peut s'écrire comme somme d'une forme paire et d'une forme impaire. On peut attacher une fonction L à une forme de Maass f par la formule $L(f, s) = \sum_{n=1}^{+\infty} \frac{a_n(f)}{n^s}$, et on a

$$\begin{aligned} \int_0^{+\infty} f(iy) y^{s-1/2} \frac{dy}{y} &= \pi^{-s} \Gamma\left(\frac{s+\nu}{2}\right) \Gamma\left(\frac{s-\nu}{2}\right) L(f, s), \quad \text{si } f \text{ est paire,} \\ \int_0^{+\infty} \frac{1}{4i\pi} \frac{\partial f}{\partial x}(iy) y^{s+1/2} \frac{dy}{y} &= \pi^{-s} \Gamma\left(\frac{s+1+\nu}{2}\right) \Gamma\left(\frac{s+1-\nu}{2}\right) L(f, s), \quad \text{si } f \text{ est impaire.} \end{aligned}$$

On dit qu'une forme de Maass f est primitive, si sa fonction L possède :

- un développement en produit de facteurs d'Euler

$$L(f, s) = \prod_{p|D} \frac{1}{1 - a_p(f)p^{-s}} \prod_{p \nmid D} \frac{1}{1 - a_p(f)p^{-s} + \chi(p)p^{k-1-2s}},$$

- une équation fonctionnelle du type $\Lambda(f, k-s) = w \Lambda(f^*, s)$, où $f^*(z) = \overline{f(-\bar{z})}$, w est un nombre complexe de module 1, $c = 0$ ou 1 suivant que f est paire ou impaire, et

$$\Lambda(f, s) = \pi^{-s} \Gamma\left(\frac{s+c+\nu}{2}\right) \Gamma\left(\frac{s+c-\nu}{2}\right) D^{s/2} L(f, s).$$

Conjecture D.1.6. — Si $\rho : \mathcal{G}_{\mathbf{Q}} \rightarrow \mathbf{GL}_2(\mathbf{C})$ est une représentation irréductible paire, il existe une forme de Maass primitive f , de valeur propre $\frac{1}{4}$, telle que $L(f, s) = L(\rho, s)$.

Cet énoncé impliquerait la conjecture d'Artin pour une telle représentation. On sait le démontrer, grâce aux travaux de Langlands et J. Tunnell sus-mentionnés, si l'image de ρ dans $\mathbf{PGL}_2(\mathbf{C})$ n'est pas A_5 .

5. La théorie du corps de classes

Soit F un corps de nombres⁽²⁷⁾. L'ensemble \mathcal{O}_F des $x \in K$ annulés par un polynôme unitaire $P \in \mathbf{Z}[X]$ (dépendant de x) est un sous-anneau de F appelé *anneau des entiers*. Par exemple, si $d \in \mathbf{Z}$ n'est divisible par le carré d'aucun nombre premier, l'anneau des entiers $\mathbf{Q}(\sqrt{d})$ est l'anneau A_D de l'introduction.

Un tel anneau n'est pas forcément principal, mais tout idéal \mathfrak{n} non nul de \mathcal{O}_F peut s'écrire de manière unique sous la forme⁽²⁸⁾ $\mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_r^{n_r}$, où les \mathfrak{p}_i sont des idéaux premiers

⁽²⁷⁾Tous les énoncés qui suivent sont des résultats de base de théorie algébrique des nombres, ce qui ne signifie pas, loin de là, qu'ils sont faciles à établir.

⁽²⁸⁾Un anneau ayant cette propriété est un *anneau de Dedekind*.

distincts⁽²⁹⁾. Deux idéaux \mathfrak{a} , \mathfrak{b} de \mathcal{O}_F sont premiers entre eux (ce qui signifie que $\mathfrak{a} + \mathfrak{b} = \mathcal{O}_F$), si et seulement si il n'y a pas d'idéal premier de \mathcal{O}_F apparaissant à la fois dans la décomposition de \mathfrak{a} et \mathfrak{b} en idéaux premiers.

On peut généraliser les notions de caractère de Dirichlet et de fonctions L de Dirichlet aux corps de nombres. Si \mathfrak{f} est un idéal de \mathcal{O}_F , on note $I_{\mathfrak{f}}$ l'ensemble des idéaux de \mathcal{O}_F premiers à \mathfrak{f} . Un *caractère de Hecke* χ (d'ordre fini) modulo \mathfrak{f} est une fonction multiplicative $\chi : I_{\mathfrak{f}} \rightarrow \mathbf{C}^*$, telle que, si $\alpha \in \mathcal{O}_F$ est positif⁽³⁰⁾, et si $\alpha - 1 \in \mathfrak{f}$, alors $\chi(\mathfrak{n}) = 1$, si \mathfrak{n} est l'idéal principal engendré par α . La *fonction L de Hecke* attachée à un caractère de Hecke χ (d'ordre fini) modulo \mathfrak{f} est alors définie par

$$L(\chi, s) = \sum_{\mathfrak{n} \in I_{\mathfrak{f}}} \frac{\chi(\mathfrak{n})}{N(\mathfrak{n})^s}, \quad \text{où } N(\mathfrak{n}) = |\mathcal{O}_F/\mathfrak{n}|.$$

Comme χ est multiplicatif et comme $\mathfrak{n} \mapsto N(\mathfrak{n})$ est aussi multiplicatif, les mêmes arguments que pour les fonctions L de Dirichlet montrent que l'on a une décomposition en produit de facteurs d'Euler

$$L(\chi, s) = \prod_{\mathfrak{p} \in \mathcal{P}_F} \frac{1}{1 - \chi(\mathfrak{p})N(\mathfrak{p})^{-s}} = \prod_{p \in \mathcal{P}} \left(\prod_{\mathfrak{p}|(p)} \frac{1}{1 - \chi(\mathfrak{p})N(\mathfrak{p})^{-s}} \right),$$

et comme $\prod_{\mathfrak{p}|(p)} N(\mathfrak{p}) = p^{[F:\mathbf{Q}]}$ pour presque tout p , le facteur $\prod_{\mathfrak{p}|(p)} (1 - \chi(\mathfrak{p})N(\mathfrak{p})^{-s})$ est un polynôme de degré $[F:\mathbf{Q}]$ en p^{-s} , sauf pour un nombre fini de p .

Ce qui précède s'applique en particulier au caractère trivial envoyant tout idéal \mathfrak{n} de \mathcal{O}_F sur 1. La fonction L de Hecke associée $\zeta_F(s) = \sum_{\mathfrak{n}} \frac{1}{N(\mathfrak{n})^s}$ avait été considérée longtemps auparavant par R. Dedekind; c'est la *fonction zêta de Dedekind de F* .

Exemple D.1.7. — (i) Si $F = \mathbf{Q}$, on retombe sur la fonction zêta de Riemann.

(ii) Si $F = \mathbf{Q}[\sqrt{d}]$, on retombe sur la fonction ζ_{A_D} de l'introduction.

Théorème D.1.8. — (Hecke, 1920) *Si χ est un caractère de Hecke primitif, alors $L(\chi, s)$ admet un prolongement méromorphe à \mathbf{C} tout entier, holomorphe en dehors d'un pôle simple en $s = 1$ si χ est le caractère trivial.*

La situation est donc exactement la même pour un corps de nombres quelconque que pour \mathbf{Q} . La démonstration du théorème de Hecke repose sur les mêmes idées que celles utilisées dans l'ex. VII.6.6 pour démontrer l'existence d'un prolongement méromorphe pour

⁽²⁹⁾ Si I et J sont deux idéaux d'un anneau A , le produit IJ de I et J est l'idéal de A engendré par les éléments de la forme ab , avec $a \in I$ et $b \in J$. Par exemple, si I et J sont principaux engendrés par α et β , alors IJ est l'idéal principal engendré par $\alpha\beta$. Par ailleurs, un idéal \mathfrak{p} est premier si la condition $ab \in \mathfrak{p}$ implique $a \in \mathfrak{p}$ ou $b \in \mathfrak{p}$.

⁽³⁰⁾ Cela signifie que $\sigma(\alpha) > 0$ si $\sigma \in \mathcal{G}_{\mathbf{Q}}$ et $\sigma(\alpha)$ est réel; il n'y a pas de condition si $\sigma(\alpha)$ n'est pas un nombre réel.

la fonction zêta de Riemann. La présence d'unités dans \mathcal{O}_F rend toutefois les arguments un peu plus délicats.

Le théorème principal de la théorie du corps de classes est alors le suivant.

Théorème D.1.9. — *Si ρ est une représentation de dimension 1 de \mathcal{G}_F , alors il existe un caractère de Hecke d'ordre fini χ de F tel que $L(\text{Ind}_{\mathcal{G}_F}^{\mathcal{G}_{\mathbf{Q}}} \rho, s) = L(\chi, s)$.*

Le théorème de Kronecker-Weber est déjà un théorème profond, mais sa démonstration est grandement facilitée par l'existence des racines de l'unité qui donne une construction systématique des *extensions abéliennes de \mathbf{Q}* (ce sont les corps de nombres fixés par le noyau d'un caractère linéaire de $\mathcal{G}_{\mathbf{Q}}$). Dans le cas d'un corps de nombres général, on ne connaît pas de tel procédé pour construire ces extensions⁽³¹⁾, ce qui explique, en grande partie, qu'on ait mis tant de temps à démontrer le théorème ci-dessus (il a déjà fallu beaucoup de temps pour arriver à concevoir son énoncé). L'histoire a commencé dans la fin des années 1880 avec L. Kronecker et H. Weber, s'est continuée avec Hilbert⁽³²⁾ autour de 1900, et le théorème ci-dessus date de la fin des années 1920, avec des contributions essentielles de P. Furtwängler, T. Takagi, E. Artin et H. Hasse⁽³³⁾. Grâce aux efforts de tous ces mathématiciens et de ceux des trois décennies suivantes, on a maintenant une théorie puissante, parfaitement bien comprise, aux énoncés compacts, et qui peut très bien s'utiliser comme une boîte noire. Elle fait régulièrement l'objet de cours fondamentaux de troisième cycle.

D.2. Le théorème de Kronecker-Weber revisité

C. Chevalley s'est demandé s'il existait un groupe naturel, construit directement à partir de \mathbf{Q} (resp. de F , où F est un corps de nombres), dont les représentations de dimension 1 seraient exactement les caractères de Dirichlet (resp. les caractères de Hecke d'ordre fini de F) primitifs. Cela l'a mené à la construction (1940) du groupe des idèles⁽³⁴⁾ esquissée ci-dessous. Nous allons aller à l'envers de l'ordre historique (les adèles sont nés, de la main d'Artin et Whaples, 5 ans plus tard que les idèles (sous le nom fort peu poétique de “valuation vectors”)), mais ce sont toujours les idées simples qui viennent en dernier⁽³⁵⁾. Les adèles sont construits en considérant simultanément tous les complétés de \mathbf{Q} (resp. d'un

⁽³¹⁾C'est un des derniers problèmes (le “liebster Jugendtraum de Kronecker”) énoncés par D. Hilbert en 1900 qu'on ne sait toujours pas résoudre.

⁽³²⁾3 des problèmes de Hilbert (les 9-ième, 11-ième et 12-ième) sont reliés à la question.

⁽³³⁾L'école allemande a joué un rôle absolument primordial dans le sujet, et tous les articles de l'époque (y compris ceux du mathématicien japonais T. Takagi) sont en allemand.

⁽³⁴⁾Le mot “idèle” vient du mot “idéal” (dans tous les sens du terme...); il a donné naissance 20 ans plus tard au mot “adèle”, et J. Roubaud a fait d'Adèle et Idèle des jumelles dans “La belle Hortense”.

⁽³⁵⁾Comme le dit A. Weil : “ Pour franchir toutes ces étapes, un bon étudiant ne met guère plus de jours à présent qu'il n'y a fallu d'années à l'époque...”

corps de nombres F), et J. Tate a montré dans sa thèse (1950), sous la direction d'E. Artin, comment l'analyse de Fourier sur les adèles permet de démontrer naturellement toutes les propriétés analytiques des fonctions L de Dirichlet (resp. de Hecke) rencontrées au chap. VII (resp. mentionnées au n° 5 du § D.1).

1. Adèles

1.1. Le théorème d'Ostrowski

Si K est un corps, une *norme* sur K est une application $x \mapsto |x|$ de K dans \mathbf{R}_+ vérifiant les trois propriétés suivantes :

$$(i) \quad |x| = 0 \Leftrightarrow x = 0, \quad (ii) \quad |xy| = |x||y|, \quad (iii) \quad |x + y| \leq |x| + |y|.$$

Si K est un corps muni d'une norme $|\cdot|$, et x, y sont deux éléments de K , on pose $d(x, y) = |x - y|$. Les propriétés (i) et (iii) des normes assurent que d est une distance sur K et donc définit une topologie sur K . Deux normes sur un corps K sont dites *équivalentes* si elles définissent la même topologie. Une norme est dite *triviale* si elle induit la topologie discrète sur K (on a alors $|x| = 1$ quel que soit $x \neq 0$).

En dehors de la norme usuelle que nous noterons $|\cdot|_\infty$, on peut définir, pour chaque nombre premier p , une norme p -adique $|\cdot|_p$ sur \mathbf{Q} . Rappelons que celle-ci est définie (cf. alinéa 3.4 du § 1) par la formule $|\frac{a}{b}|_p = p^{-v_p(a) + v_p(b)}$, si $a, b \in \mathbf{Z} - \{0\}$, et $v_p(n)$ est le plus grand entier v tel que p^v divise n .

Théorème D.2.1. — (Ostrowski, 1918) *Une norme non triviale sur \mathbf{Q} est équivalente à la norme usuelle $|\cdot|_\infty$ ou à la norme p -adique $|\cdot|_p$ pour un unique nombre premier p .*

Démonstration. — Commençons par supposer qu'il existe $k \in \mathbf{N}$ tel que $\|k\| > 1$. Comme $\|1\| = 1$, l'inégalité triangulaire implique $\|k\| \leq k$ et il existe $\alpha \in]0, 1]$ tel que l'on ait $\|k\| = k^\alpha$. Soit $m \in \mathbf{N}$. On peut écrire m en base k sous la forme $m = \sum_{i=0}^n a_i k^i$ avec $a_i \in \{0, 1, \dots, k-1\}$ et $a_n \neq 0$ de telle sorte que l'on a $m \geq k^n$. Comme $\|a_i\| \leq a_i \leq k-1$ et $\|k^i\| = \|k\|^i$, on obtient la majoration

$$\|m\| \leq (k-1) \sum_{i=1}^n k^{i\alpha} = \frac{k-1}{k^\alpha - 1} (k^{(n+1)\alpha} - 1) \leq \frac{k^\alpha(k-1)}{k^\alpha - 1} k^{n\alpha} \leq C m^\alpha,$$

où $C = \frac{k^\alpha(k-1)}{k^\alpha - 1}$ est indépendant de m . On peut appliquer cette inégalité à m^n , ce qui nous donne $\|m\|^n \leq C m^{n\alpha}$ et, prenant la racine n -ième de cette égalité et passant à la limite, l'inégalité $\|m\| \leq m^\alpha$. Par symétrie, on en déduit le fait que si $\|m\| > 1$, alors cette inégalité est une égalité.

Maintenant, si $m \in \mathbf{N} - \{0\}$ est quelconque, il existe $n \in \mathbf{N}$ tel que l'on ait $\|k^n m\| > 1$. On a alors $\|m\| = \|k^n\|^{-1} \|k^n m\| = k^{-\alpha n} (k^n m)^\alpha = m^\alpha$, ce qui montre que l'on a égalité quel que soit $m \in \mathbf{N}$ puis, utilisant la multiplicativité de la norme et le fait que $\|-1\| = 1$, que $\|x\| = |x|_\infty^\alpha$ quel que soit $x \in \mathbf{Q}$. On en déduit que s'il existe $k \in \mathbf{N}$ tel que $\|k\| > 1$, alors $\|\cdot\|$ est équivalente à la norme usuelle.

Dans le cas contraire, on a $\|\ell\| \leq 1$ pour tout nombre premier ℓ . Comme on a supposé $\|\cdot\|$ non triviale, il existe au moins un nombre premier p tel que $\|p\| < 1$. Si il en existe un autre q , alors quel que soit $n \in \mathbf{N}$, on peut, d'après le théorème de Bezout, trouver $u_n, v_n \in \mathbf{Z}$ tels que l'on ait $u_n p^n + v_n q^n = 1$. On obtient donc

$$1 = \|1\| = \|u_n p^n + v_n q^n\| \leq \|u_n\| \cdot \|p\|^n + \|v_n\| \cdot \|q\|^n \leq \|p\|^n + \|q\|^n,$$

ce qui est impossible pour n assez grand. Il existe donc un et un seul nombre premier p tel que $\|p\| < 1$ et $\|\cdot\|$ est équivalente à la norme p -adique. Ceci termine la démonstration.

Le résultat suivant est une conséquence immédiate de l'unicité de la décomposition d'un entier en produit de facteurs premiers, mais est très important ; c'est ce qui justifie la normalisation utilisée pour $|\cdot|_p$.

Théorème D.2.2. — (Formule du produit) *Si $x \in \mathbf{Q}^*$, alors*

$$|x|_\infty \cdot \prod_{p \text{ premier}} |x|_p = 1.$$

1.2. L'anneau des adèles de \mathbf{Q}

On renvoie à l'alinéa 3.4 du § 1 pour ce qui concerne la construction des nombres p -adiques et leurs propriétés élémentaires.

Soit $\mathcal{V} = \{\infty\} \cup \mathcal{P}$. Les éléments de \mathcal{V} sont les *places* de \mathbf{Q} . La place ∞ est la *place à l'infini* et les $p \in \mathcal{P}$ sont les *places finies*. Si $v \in \mathcal{V}$, on note \mathbf{Q}_v le complété de \mathbf{Q} correspondant ; on a donc $\mathbf{Q}_\infty = \mathbf{R}$.

D'après le théorème d'Ostrowski, les complétés de \mathbf{Q} sont exactement les \mathbf{Q}_v , pour $v \in \mathcal{V}$. Il est donc naturel⁽³⁶⁾ d'essayer de considérer tous ces complétés ensemble. On peut plonger \mathbf{Q} diagonalement dans $\prod_{v \in \mathcal{V}} \mathbf{Q}_v$, c'est-à-dire en envoyant $a \in \mathbf{Q}$ sur $(x_v)_{v \in \mathcal{V}}$, avec $x_v = a$, quel que soit $v \in \mathcal{V}$. On a alors $x_p \in \mathbf{Z}_p$ sauf si p divise le dénominateur de a , ce qui montre que l'image de \mathbf{Q} est incluse dans le sous-anneau de $\prod_{v \in \mathcal{V}} \mathbf{Q}_v$ des $x = (x_\infty, \dots, x_p, \dots)$, avec $x_p \in \mathbf{Z}_p$ pour presque tout p . Cet anneau est l'*anneau des adèles* \mathbf{A} de \mathbf{Q} ; c'est le *produit restreint* des \mathbf{Q}_v relativement aux \mathbf{Z}_p , pour $p \in \mathcal{P}$. On note souvent $(x_\infty, x^{[\infty]})$ un élément x de \mathbf{A} , et $x^{[\infty]}$ est la partie finie (i.e. en dehors de ∞) de x . Si $v \in \mathcal{V}$, on peut identifier \mathbf{Q}_v au sous-anneau de \mathbf{A} des x dont toutes les composantes sont nulles sauf la composante en v .

On munit \mathbf{A} de la *topologie du produit restreint* obtenue en prenant comme base d'ouverts les $\prod_{v \in S} U_v \times \prod_{p \notin S} \mathbf{Z}_p$, où S décrit les parties finies de \mathcal{V} contenant ∞ , et U_v , pour $v \in S$, décrit les ouverts de \mathbf{Q}_v . Les propositions D.2.3 et D.2.4 ci-dessous montrent que \mathbf{Q} est discret dans \mathbf{A} , mais n'est pas loin d'être dense.

Proposition D.2.3. — (i) *Tout élément x de \mathbf{A} peut s'écrire de manière unique sous la forme $x = \alpha + y$, avec $\alpha \in \mathbf{Q}$ et $y \in [0, 1[\times \widehat{\mathbf{Z}}$, où $\widehat{\mathbf{Z}} = \prod_{p \in \mathcal{P}} \mathbf{Z}_p$.*

(ii) *\mathbf{Q} est discret dans \mathbf{A} et \mathbf{A}/\mathbf{Q} est isomorphe à $(\mathbf{R}/\mathbf{Z}) \times \widehat{\mathbf{Z}}$, et est compact.*

Démonstration. — (i) Commençons par démontrer qu'une telle écriture, si elle existe, est unique. Si $\alpha_1 + y_1 = \alpha_2 + y_2$, avec $\alpha_1, \alpha_2 \in \mathbf{Q}$ et $y_1, y_2 \in [0, 1[\times \widehat{\mathbf{Z}}$, alors en particulier, on a $\alpha_2 - \alpha_1 \in \mathbf{Z}_p$ quel que soit $p \in \mathcal{P}$, et donc $\alpha_2 - \alpha_1 \in \mathbf{Z}$. Mais alors $\alpha_1 + y_{1,\infty} = \alpha_2 + y_{2,\infty}$ implique, puisque $y_{1,\infty}, y_{2,\infty} \in [0, 1[$, que $\alpha_2 = \alpha_1$, et donc aussi que $y_2 = y_1$. D'où l'unicité.

⁽³⁶⁾C'est naturel, mais il s'est écoulé plus de 50 ans entre la construction des nombres p -adiques par K. Hensel et les constructions qui vont suivre.

Passons à l'existence. Soit S un sous-ensemble fini de \mathcal{P} tel que $x_p \in \mathbf{Z}_p$, si $p \notin S$. Comme $\mathbf{Z}[\frac{1}{p}]$ est dense dans \mathbf{Q}_p d'après le th. 3.4.2, il existe, pour tout $p \in S$, un élément α_p de $\mathbf{Z}[\frac{1}{p}]$ tel que $x_p - \alpha_p \in \mathbf{Z}_p$. Mais alors $x' = x - \sum_{p \in S} \alpha_p \in \mathbf{R} \times \widehat{\mathbf{Z}}$, et il suffit de poser $\alpha = [x'_\infty] + \sum_{p \in S} \alpha_p$ et $y = x' - [x'_\infty]$ (où $[x'_\infty]$ est la partie entière de la composante x'_∞ en ∞ de x') pour obtenir une décomposition de x sous la forme souhaitée. Ceci termine la démonstration du (i).

Le fait que \mathbf{Q} est discret suit de ce que, si $x \in \mathbf{A}$, le voisinage ouvert $x + (\frac{1}{2}, \frac{1}{2}[\times \widehat{\mathbf{Z}}])$ de x contient au plus un élément de \mathbf{Q} , vu que la différence de deux de ses éléments est dans $] -1, 1[\times \widehat{\mathbf{Z}}$ qui ne contient que 0 comme élément de \mathbf{Q} . L'isomorphisme $\mathbf{A}/\mathbf{Q} \cong (\mathbf{R}/\mathbf{Z}) \times \widehat{\mathbf{Z}}$ est une conséquence immédiate du (i), et la compacité de \mathbf{A}/\mathbf{Q} suit alors du fait qu'un produit dénombrable de compacts est compact.

Si $S = \{p_1, \dots, p_s\}$ est un sous-ensemble fini de \mathcal{P} , on note $\mathbf{Z}[S^{-1}]$ le sous-anneau $\mathbf{Z}[\frac{1}{p_1}, \dots, \frac{1}{p_s}]$ de \mathbf{Q} obtenu en rendant p_1, \dots, p_s inversibles.

Proposition D.2.4. — Si $S \subset \mathcal{P}$ est fini, alors $\mathbf{Z}[S^{-1}]$ est dense dans $\prod_{p \in S} \mathbf{Q}_p$.

Démonstration. — Comme les p_i^n sont premiers entre eux deux à deux, il existe, d'après le théorème des restes chinois, $a_{i,n} \in \mathbf{Z}$ congru à 1 modulo p_i^n , et à 0 modulo p_j^n , pour tout $j \in S - \{i\}$. Si maintenant $y_i \in \mathbf{Q}_{p_i}$, on peut trouver $x_{i,n} \in \mathbf{Z}[\frac{1}{p_i}]$ tel que $|x_{i,n} - y_i| \leq p_i^{-n}$ car $\mathbf{Z}[\frac{1}{p_i}]$ est dense dans \mathbf{Q}_{p_i} d'après le th. 3.4.2. En posant $x_n = \sum_{j=1}^s a_{j,n} x_{j,n}$, cela fournit une suite d'éléments de $\mathbf{Z}[S^{-1}]$ vérifiant

$$\begin{aligned} |x_n - y_i|_{p_i} &= |a_{i,n} x_{i,n} - y_i + \sum_{j \neq i} a_{j,n} x_{j,n}|_{p_i} = |a_{i,n}(x_{i,n} - y_i) + (a_{i,n} - 1)y_i + \sum_{j \neq i} a_{j,n} x_{j,n}|_{p_i} \\ &\leq \sup(|a_{i,n}|_{p_i} |x_{i,n} - y_i|_{p_i}, |(a_{i,n} - 1)|_{p_i} |y_i|_{p_i}, \sup_{j \neq i} |a_{j,n}|_{p_i} |x_{j,n}|_{p_i}). \end{aligned}$$

Maintenant, on a $|x_{j,n}|_{p_i} \leq 1$, si $j \neq i$, et comme $|x_{i,n} - y_i|_{p_i}$, $|(a_{i,n} - 1)|_{p_i}$ et $|a_{j,n}|_{p_i}$, si $j \neq i$, sont tous $\leq p_i^{-n}$, on voit que x_n tend vers (y_1, \dots, y_s) dans $\prod_{p \in S} \mathbf{Q}_p$. Ceci permet de conclure.

1.3. Le groupe des idèles de \mathbf{Q}

Le groupe des idèles \mathbf{A}^* de \mathbf{Q} est le groupe des éléments inversibles de l'anneau \mathbf{A} ; c'est donc l'ensemble des $(x_v)_{v \in \mathcal{V}}$, avec $x_v \in \mathbf{Q}_v^*$, et $x_p \in \mathbf{Z}_p^*$ pour presque tout p . Autrement dit, \mathbf{A}^* est le produit restreint des \mathbf{Q}_v^* relativement aux \mathbf{Z}_p^* , pour $p \in \mathcal{P}$, et on le munit de la topologie du produit restreint dont une base d'ouverts est constituée des $\prod_{v \in S} U_v \times \prod_{p \notin S} \mathbf{Z}_p^*$, où S décrit les parties finies de \mathcal{V} contenant ∞ , et U_v les ouverts de \mathbf{Q}_v^* , si $v \in S$.

Si $\beta \in \mathbf{Q}^*$, on a $\beta \in \mathbf{Z}_p^*$ sauf si p divise le dénominateur ou le numérateur de β , ce qui prouve que l'élément (β, β, \dots) de $\prod_{v \in \mathcal{V}} \mathbf{Q}_v$ appartient à \mathbf{A}^* , et permet d'identifier \mathbf{Q}^* à un sous-groupe de \mathbf{A}^* . Si $v \in \mathcal{V}$, on peut identifier \mathbf{Q}_v^* au sous-groupe de \mathbf{A} des x dont toutes les composantes sont égales à 1 sauf la composante en v .

Si $x = (x_v)_{v \in \mathcal{V}} \in \mathbf{A}^*$, on a $|x_p|_p = 1$ pour presque tout p , ce qui permet de définir $|x|_{\mathbf{A}}$ par $|x|_{\mathbf{A}} = \prod_{v \in \mathcal{V}} |x_v|_v$. Il résulte de la formule du produit (th. D.2.2) que $|\beta|_{\mathbf{A}} = 1$, quel que soit $\beta \in \mathbf{Q}^*$.

Lemme D.2.5. — (i) Tout élément x de \mathbf{A}^* peut s'écrire de manière unique sous la forme $x = \beta b_\infty b^{[\infty]}$, avec $\beta \in \mathbf{Q}^*$, $b_\infty \in \mathbf{R}_+^*$, et $b^{[\infty]} \in \widehat{\mathbf{Z}}^*$, où $\widehat{\mathbf{Z}}^* = \prod_{p \in \mathcal{P}} \mathbf{Z}_p^*$.

(ii) $\mathbf{A}^*/\mathbf{Q}^* \cong \mathbf{R}_+^* \times \widehat{\mathbf{Z}}^*$.

Démonstration. — On peut et doit poser $\beta = \text{sign}(x_\infty) \prod_{p \in \mathcal{P}} p^{v_p(x_p)}$, et alors $b = \beta^{-1}x$ vérifie les conditions $b_\infty \in \mathbf{R}_+$ et $b^{[\infty]} \in \prod_{p \in \mathcal{P}} \mathbf{Z}_p^*$. On en déduit le (i), et le (ii) en est une conséquence immédiate.

2. La formule de Poisson adélique

2.1. Transformée de Fourier sur \mathbf{Q}_p

Si $p \in \mathcal{P}$, on note $\mathcal{S}(\mathbf{Q}_p)$ l'espace (de Schwartz) des fonctions à valeurs dans \mathbf{C} , localement constantes, à support compact dans \mathbf{Q}_p . Soit $\phi \in \mathcal{S}(\mathbf{Q}_p)$. Comme ϕ est à support compact (et donc borné), il existe $m \in \mathbf{Z}$ tel que $\phi(x) = 0$ si $|x|_p > p^{-m}$; autrement dit, il existe $m \in \mathbf{Z}$ tel que $\phi(x) = 0$ si $x \notin p^m \mathbf{Z}_p$. Par ailleurs, comme ϕ est localement constante, il existe pour tout $x \in p^m \mathbf{Z}_p$, un entier $n_x \in \mathbf{N}$ tel que ϕ soit constante sur $B(x, p^{-n_x}) = x + p^{n_x} \mathbf{Z}_p$. Comme $p^m \mathbf{Z}_p$ est compact, on peut extraire un recouvrement fini du recouvrement de $p^m \mathbf{Z}_p$ par les $x + p^{n_x} \mathbf{Z}_p$. En notant n le minimum des n_x intervenant dans le sous-recouvrement fini, on voit qu'il existe $n \in \mathbf{N}$ tel que ϕ soit constante sur $x + p^n \mathbf{Z}_p$, quel que soit $x \in p^m \mathbf{Z}_p$. En notant $\mathbf{1}_{a+p^n \mathbf{Z}_p}$ la fonction caractéristique de $a + p^n \mathbf{Z}_p$, et en utilisant la densité de $\mathbf{Z}[\frac{1}{p}]$ dans \mathbf{Q}_p , on en déduit le résultat suivant.

Proposition D.2.6. — $\mathcal{S}(\mathbf{Q}_p)$ est le \mathbf{C} -espace vectoriel engendré par les $\mathbf{1}_{a+p^n \mathbf{Z}_p}$, pour $a \in \mathbf{Z}[\frac{1}{p}]$ et $n \in \mathbf{N}$. De plus, les relations entre les $\mathbf{1}_{a+p^n \mathbf{Z}_p}$ sont engendrées par les relations suivantes :

- $\mathbf{1}_{a+p^n \mathbf{Z}_p} = \mathbf{1}_{b+p^n \mathbf{Z}_p}$, si $b \in a + p^n \mathbf{Z}$,
- $\mathbf{1}_{a+p^n \mathbf{Z}_p} = \sum_{i=0}^{p-1} \mathbf{1}_{a+ip^n+p^{n+1} \mathbf{Z}_p}$,

la seconde traduisant le fait qu'il y a p possibilités pour le reste de la division par p^{n+1} , si on connaît le reste de la division par p^n .

Si $\phi \in \mathcal{S}(\mathbf{Q}_p)$, on définit $\int_{\mathbf{Q}_p} \phi(x) dx$ par linéarité, en posant $\int_{\mathbf{Q}_p} \mathbf{1}_{a+p^n \mathbf{Z}_p}(x) dx = p^{-n}$, quels que soient $a \in \mathbf{Z}[\frac{1}{p}]$ et $n \in \mathbf{N}$, ce qui revient à demander que \mathbf{Z}_p ait mesure 1 et que l'intégration soit invariante par translation. Que ce soit possible suit du fait que les seules relations entre les $\mathbf{1}_{a+p^n \mathbf{Z}_p}$ sont celles ci-dessus.

Si $x \in \mathbf{Q}_p$, il existe $\xi \in \mathbf{Z}[\frac{1}{p}]$ tel que $x - \xi \in \mathbf{Z}_p$ (c'est une conséquence de la densité de $\mathbf{Z}[\frac{1}{p}]$ dans \mathbf{Q}_p), et ξ est bien déterminé à addition près d'un élément de $\mathbf{Z}[\frac{1}{p}] \cap \mathbf{Z}_p = \mathbf{Z}$. On en déduit que $e^{2i\pi \xi}$ ne dépend que de x , et pas du choix de ξ , ce qui nous autorise à le noter $e^{2i\pi x}$. Il est alors immédiat que $x \mapsto e^{2i\pi x}$ est un caractère linéaire de \mathbf{Q}_p ($e^{2i\pi(x+y)} = e^{2i\pi x} e^{2i\pi y}$), localement constant (et même constant sur $a + \mathbf{Z}_p$, quel que soit $a \in \mathbf{Q}_p$).

On définit alors la *transformée de Fourier* $\hat{\phi} = \mathcal{F}_p \phi$ de $\phi \in \mathcal{S}(\mathbf{Q}_p)$ par la formule

$$\hat{\phi}(y) = \int_{\mathbf{Q}_p} \phi(x) e^{2i\pi xy} dx,$$

ce qui a un sens car $x \mapsto \phi(x)e^{2i\pi xy}$ a même support que ϕ et est localement constante comme produit de deux fonctions localement constantes.

Exemple D.2.7. — (i) La transformée de Fourier de $\mathbf{1}_{\mathbf{Z}_p}$ est $\mathbf{1}_{\mathbf{Z}_p}$. En effet, si $y \in \mathbf{Z}_p$, on a $e^{2i\pi xy} = 1$ quel que soit $x \in \mathbf{Z}_p$, et donc

$$\int_{\mathbf{Q}_p} \mathbf{1}_{\mathbf{Z}_p}(x) e^{2i\pi xy} dx = \int_{\mathbf{Q}_p} \mathbf{1}_{\mathbf{Z}_p}(x) dx = 1.$$

Par contre, si $v_p(y) = -n < 0$, le caractère linéaire $x \mapsto e^{2i\pi xy}$ de \mathbf{Z}_p est non trivial et constant modulo $p^n \mathbf{Z}_p$; on a donc

$$\int_{\mathbf{Q}_p} \mathbf{1}_{\mathbf{Z}_p}(x) e^{2i\pi xy} dx = \int_{\mathbf{Q}_p} \sum_{a \in \mathbf{Z}_p/p^n \mathbf{Z}_p} e^{2i\pi ya} \mathbf{1}_{a+p^n \mathbf{Z}_p}(x) dx = p^{-n} \sum_{a \in \mathbf{Z}_p/p^n \mathbf{Z}_p} e^{2i\pi ya} = 0,$$

d'après l'orthogonalité des caractères linéaires d'un groupe fini, puisque $a \mapsto e^{2i\pi ya}$ est un caractère linéaire non trivial de $\mathbf{Z}_p/p^n \mathbf{Z}_p \cong \mathbf{Z}/p^n \mathbf{Z}$.

(ii) Les mêmes calculs montrent que, si ϕ est constante modulo $p^n \mathbf{Z}_p$, et à support dans $p^{-m} \mathbf{Z}_p$, alors $\hat{\phi}$ est constante modulo $p^m \mathbf{Z}_p$, et à support dans $p^{-n} \mathbf{Z}_p$.

(iii) Soit $\chi : (\mathbf{Z}/p^n \mathbf{Z})^* \rightarrow \mathbf{C}^*$ un caractère de Dirichlet de conducteur p^n , avec $n \geq 1$. En utilisant l'isomorphisme $\mathbf{Z}_p/p^n \mathbf{Z}_p \cong \mathbf{Z}/p^n \mathbf{Z}$, on peut associer à χ une fonction ϕ_χ à support dans \mathbf{Z}_p (et même dans \mathbf{Z}_p^*), constante modulo $p^n \mathbf{Z}_p$, en posant $\phi_\chi(x) = 0$ si $x \in p \mathbf{Z}_p$, et $\phi_\chi(x) = \chi(\bar{x})$, si $x \in \mathbf{Z}_p^*$ et \bar{x} est l'image de x modulo $p^n \mathbf{Z}_p$. Alors la transformée de Fourier de ϕ_χ est donnée par la formule

$$\hat{\phi}_\chi(y) = \frac{G(\chi)}{p^n} \phi_{\bar{\chi}}(p^n y), \quad \text{où } G(\chi) = \sum_{a \in (\mathbf{Z}/p^n \mathbf{Z})^*} \chi(a) e^{2i\pi \frac{a}{p^n}}$$

est la somme de Gauss introduite au n° 2 du § VII.4. En effet, ϕ_χ étant constante modulo $p^n \mathbf{Z}_p$, la fonction $\hat{\phi}_\chi$ est nulle en dehors de $p^{-n} \mathbf{Z}_p$, et si $v_p(y) \geq -n$, la formule à vérifier est équivalente à celle du lemme VII.4.3.

La transformée de Fourier sur \mathbf{Q}_p vérifie les mêmes propriétés que sur \mathbf{R} en ce qui concerne les dilatations, translations... De manière précise, on a le résultat suivant.

Proposition D.2.8. — (i) Si $a \in \mathbf{Q}_p^*$ et $b, c \in \mathbf{Q}_p$, alors

$$\begin{aligned} \mathcal{F}_p(\phi(ax))(y) &= |a|_p^{-1} \hat{\phi}(a^{-1}y), \\ \mathcal{F}_p(\phi(x+b))(y) &= e^{-2i\pi by} \hat{\phi}(y), \quad \text{et} \quad \mathcal{F}_p(e^{2i\pi cx} \phi(x))(y) = \hat{\phi}(y+c). \end{aligned}$$

(ii) $(\mathcal{F}_p \circ \mathcal{F}_p \phi)(x) = \phi(-x)$ (formule d'inversion de Fourier).

Démonstration. — Si $a \in \mathbf{Q}_p^*$ et $b, c \in \mathbf{Q}_p$, soit $u_{a,b,c} : \mathcal{S}(\mathbf{Q}_p) \rightarrow \mathcal{S}(\mathbf{Q}_p)$ l'application définie par $(u_{a,b,c} \phi)(x) = e^{2i\pi cx} \phi(ax+b)$. Un changement de variable immédiat dans l'intégrale définissant $\mathcal{F}_p \circ u_{a,b,c}$ montre que

$$\mathcal{F}_p \circ u_{a,b,c} = |a|_p^{-1} e^{-2i\pi \frac{bc}{a}} u_{a^{-1}, a^{-1}c, -a^{-1}b} \circ \mathcal{F}_p.$$

On en déduit le (i). En appliquant le (i) deux fois, on montre que $\mathcal{F}_p \circ \mathcal{F}_p \circ u_{a,b,c} = u_{a,-b,-c} \circ \mathcal{F}_p \circ \mathcal{F}_p$. Soit Y le sous-espace de $\mathcal{S}(\mathbf{Q}_p)$ des ϕ tels que $\mathcal{F}_p \circ \mathcal{F}_p \phi = u_{-1,0,0} \phi$. Comme $u_{a,-b,-c} \circ u_{-1,0,0} = u_{-1,0,0} \circ u_{a,b,c}$, on voit que Y est stable par les $u_{a,b,c}$. Par ailleurs, Y contient $\mathbf{1}_{\mathbf{Z}_p}$, et $u_{p^{-n},-p^{-n}a,0} \mathbf{1}_{\mathbf{Z}_p} = \mathbf{1}_{a+p^n \mathbf{Z}_p}$; on en déduit que Y contient les $\mathbf{1}_{a+p^n \mathbf{Z}_p}$ pour $a \in \mathbf{Q}_p$ et $n \in \mathbf{Z}$, et donc $Y = \mathcal{S}(\mathbf{Q}_p)$. Ceci permet de conclure.

2.2. Transformée de Fourier adélique

Soit $\mathcal{S}(\mathbf{A})$ l'espace de Schwartz de \mathbf{A} . C'est l'ensemble des combinaisons linéaires des fonctions de la forme $\phi(x) = \prod_{v \in S} \phi_v(x_v) \prod_{p \notin S} \mathbf{1}_{\mathbf{Z}_p}(x_p)$, où S décrit les parties finies de \mathcal{V} contenant ∞ , et $\phi_v \in \mathcal{S}(\mathbf{Q}_v)$, si $v \in S$. Une telle fonction sera aussi notée $\otimes_{v \in \mathcal{V}} \phi_v$, étant sous-entendu que $\phi_p = \mathbf{1}_{\mathbf{Z}_p}$ pour presque tout $p \in \mathcal{P}$.

Comme les $\mathbf{1}_{a+p^n \mathbf{Z}_p}$ engendrent $\mathcal{S}(\mathbf{Q}_p)$, on voit que $\mathcal{S}(\mathbf{A})$ est aussi engendré par des fonctions de la forme

$$\phi_\infty(x_\infty) \prod_{p \in S} \mathbf{1}_{a_p+p^{n_p} \mathbf{Z}_p}(x_p) \prod_{p \notin S} \mathbf{1}_{\mathbf{Z}_p}(x_p) = \phi_\infty(x_\infty) \mathbf{1}_{a+N\hat{\mathbf{Z}}}(x^{|\infty|}),$$

où $N = \prod_{p \in S} p^{n_p}$ et a est un élément de $\mathbf{Z}[S^{-1}]$ tel que $a - a_p \in \mathbf{Z}_p$, quel que soit $p \in S$ (l'existence d'un tel a est garantie par le lemme D.2.4). De plus, les relations entre ces générateurs sont engendrées par les relations (en notant plus simplement $\phi_\infty \otimes \mathbf{1}_{a+N\hat{\mathbf{Z}}}$ l'élément de $\mathcal{S}(\mathbf{A})$ apparaissant dans le membre de droite ci-dessus) :

- $\phi_\infty \otimes \mathbf{1}_{a+N\hat{\mathbf{Z}}} = \phi_\infty \otimes \mathbf{1}_{b+N\hat{\mathbf{Z}}}$, si $a - b \in N\mathbf{Z}$,
- $\phi_\infty \otimes \mathbf{1}_{a+N\hat{\mathbf{Z}}} = \sum_{i=0}^{M-1} \phi_\infty \otimes \mathbf{1}_{a+iN+NM\hat{\mathbf{Z}}}$, si $a \in \mathbf{Q}$, $M, N \in \mathbf{N} - \{0\}$.

Si $\phi \in \mathcal{S}(\mathbf{A})$, on définit $\int_{\mathbf{A}} \phi dx$ par linéarité en imposant que $\int_{\mathbf{A}} \phi_\infty \otimes \mathbf{1}_{a+N\hat{\mathbf{Z}}} dx = \frac{1}{N} \int_{\mathbf{R}} \phi_\infty(x_\infty) dx_\infty$, ce qui revient à demander que la mesure de $[0, 1[\times \hat{\mathbf{Z}}$ soit 1, et que l'intégration sur \mathbf{A} soit invariante par translation. On remarquera que si $\phi = \otimes_{v \in \mathcal{V}} \phi_v$, alors $\int_{\mathbf{A}} \phi dx = \prod_{v \in \mathcal{V}} \int_{\mathbf{Q}_v} \phi_v(x_v) dx_v$, et presque tous les termes du produit valent 1.

Si $x = (x_v)_{v \in \mathcal{V}} \in \mathbf{A}$, on définit $e^{2i\pi x}$ par $e^{2i\pi x} = e^{-2i\pi x_\infty} \prod_{p \in \mathcal{P}} e^{2i\pi x_p}$, et presque tous les termes du produit sont égaux à 1. Il est immédiat sur la formule que l'on a $e^{2i\pi(x+y)} = e^{2i\pi x} e^{2i\pi y}$, si $x, y \in \mathbf{A}$. De plus, on vérifie sans peine que $e^{2i\pi(x+\nu)} = e^{2i\pi x}$, si $\nu \in \mathbf{Z}$ ou si ν est de la forme $\frac{1}{p^n}$, avec $p \in \mathcal{P}$ et $n \in \mathbf{N}$. Comme tout élément de \mathbf{Q} peut s'écrire comme une somme d'éléments de ce type (décomposition en éléments simples), on en déduit que $x \mapsto e^{2i\pi x}$ est périodique de période \mathbf{Q} .

On définit la *transformée de Fourier adélique* $\hat{\phi} = \mathcal{F}_{\mathbf{A}} \phi$ de $\phi \in \mathcal{S}(\mathbf{A})$ par la formule $\hat{\phi}(y) = \int_{\mathbf{A}} \phi(x) e^{2i\pi xy} dx$. Il est clair que si $\phi = \otimes_{v \in \mathcal{V}} \phi_v$, alors $\mathcal{F}_{\mathbf{A}} \phi = \otimes_{v \in \mathcal{V}} \mathcal{F}_v \phi_v$, (ce qui a un sens puisque, pour presque tout p , on a $\mathcal{F}_p \phi_p = \mathbf{1}_{\mathbf{Z}_p}$).

On déduit des formules locales (i.e. sur \mathbf{R} et sur \mathbf{Q}_p) d'inversion de Fourier et des formules locales pour les dilatations, que

$$(\mathcal{F}_{\mathbf{A}} \circ \mathcal{F}_{\mathbf{A}} \phi)(y) = \phi(-y), \quad \text{et} \quad \mathcal{F}_{\mathbf{A}}(\phi(bx))(y) = |b|_{\mathbf{A}}^{-1} \hat{\phi}(b^{-1}y), \quad \text{si } b \in \mathbf{A}^*.$$

Théorème D.2.9. — (Formule de Poisson) Soit $\phi \in \mathcal{S}(\mathbf{A})$. Alors

- (i) $\sum_{\alpha \in \mathbf{Q}} \phi(\alpha) = \sum_{\alpha \in \mathbf{Q}} \hat{\phi}(\alpha)$.

(ii) Plus généralement, si $b \in \mathbf{A}^*$, alors $\sum_{\alpha \in \mathbf{Q}} \phi(\alpha b) = |b|_{\mathbf{A}}^{-1} \sum_{\alpha \in \mathbf{Q}} \hat{\phi}(\alpha b^{-1})$.

Démonstration. — Par linéarité, on peut supposer que $\phi = \phi_{\infty} \otimes \mathbf{1}_{a+N\hat{\mathbf{Z}}}$, avec $a \in \mathbf{Q}$, et $N \in \mathbf{N}$. Dans ce cas, $\hat{\phi}(y) = N^{-1} \hat{\phi}(y_{\infty}) e^{2i\pi a y} \mathbf{1}_{N^{-1}\hat{\mathbf{Z}}}$, et on est ramené à prouver l'identité

$$\sum_{\alpha \in a+N\mathbf{Z}} \phi_{\infty}(\alpha) = \frac{1}{N} \sum_{\alpha \in N^{-1}\mathbf{Z}} e^{2i\pi a\alpha} \hat{\phi}_{\infty}(\alpha),$$

ce qui suit de la formule de Poisson classique (th. IV.3.18) appliquée à la fonction $f : \mathbf{R} \rightarrow \mathbf{C}$, définie par $f(x) = \phi_{\infty}(Nx + a)$, dont la transformée de Fourier est $\hat{f}(y) = \frac{1}{N} e^{2i\pi \frac{ay}{N}} \hat{\phi}_{\infty}(\frac{y}{N})$.

Le (ii) suit de ce que la transformée de Fourier de ϕ_b définie par $\phi_b(x) = \phi(bx)$ est $\hat{\phi}_b(y) = |b|_{\mathbf{A}}^{-1} \hat{\phi}(b^{-1}y)$.

3. Transformée de Mellin adélique et fonctions L

3.1. Intégration sur \mathbf{Q}_p^*

Soit $\mathcal{L}_0(\mathbf{Q}_p^*)$ l'ensemble des $\phi : \mathbf{Q}_p^* \rightarrow \mathbf{C}$, localement constantes et à support compact dans \mathbf{Q}_p^* (autrement dit, $\mathcal{L}_0(\mathbf{Q}_p^*)$ est le sous-espace de $\mathcal{S}(\mathbf{Q}_p)$ des ϕ vérifiant $\phi(0) = 0$). Si $\phi \in \mathcal{L}_0(\mathbf{Q}_p^*)$, on définit $\int_{\mathbf{Q}_p^*} \phi(x) d^*x$ par la formule

$$\int_{\mathbf{Q}_p^*} \phi(x) d^*x = \frac{p}{p-1} \int_{\mathbf{Q}_p} \phi(x) |x|_p^{-1} dx,$$

ce qui revient à demander que la mesure de \mathbf{Z}_p^* soit 1 et que l'intégration sur \mathbf{Q}_p^* soit invariante par $x \mapsto bx$, si $b \in \mathbf{Q}_p^*$.

Soit $\mathcal{L}(\mathbf{Q}_p^*)$ l'ensemble des $\phi : \mathbf{Q}_p^* \rightarrow \mathbf{C}$, à support compact dans \mathbf{Q}_p , dont la restriction à $p^n \mathbf{Z}_p^* = \{x, |x|_p = p^{-n}\}$ appartient, quel que soit $n \in \mathbf{Z}$, à $\mathcal{L}_0(\mathbf{Q}_p^*)$, et qui sont *sommables*, ce qui signifie que $\sum_{n \in \mathbf{Z}} \int_{p^n \mathbf{Z}_p^*} |\phi(x)| d^*x < +\infty$. Si $\phi \in \mathcal{L}(\mathbf{Q}_p^*)$, on définit $\int_{\mathbf{Q}_p^*} \phi(x) d^*x$ par la formule

$$\int_{\mathbf{Q}_p^*} \phi(x) d^*x = \sum_{n \in \mathbf{Z}} \int_{p^n \mathbf{Z}_p^*} \phi(x) d^*x.$$

3.2. Intégration sur le groupe des idèles

Soit $\mathcal{L}(\mathbf{R}^*)$ l'ensemble des $\phi : \mathbf{R}^* \rightarrow \mathbf{C}$ à décroissance rapide à l'infini et telles que $\int_{\mathbf{R}^*} |\phi(t)| d^*t < +\infty$, où l'on a posé $d^*t = \frac{dt}{|t|}$. La mesure d^*t est invariante par le changement de variable $t \mapsto bt$, si $b \in \mathbf{R}^*$.

Soit $\mathcal{L}_0(\mathbf{A}^*)$ l'espace engendré par les fonctions de la forme $\phi(x) = \prod_{v \in \mathcal{V}} \phi_v(x_v)$, où $\phi_{\infty} \in \mathcal{L}(\mathbf{R}^*)$, $\phi_p \in \mathcal{L}_0(\mathbf{Q}_p^*)$ et $\phi_p = \mathbf{1}_{\mathbf{Z}_p^*}$ pour presque tout p . On notera une fonction de ce type sous la forme $\otimes_{v \in \mathcal{V}} \phi_v$, ce qui sous-entend que $\phi_p = \mathbf{1}_{\mathbf{Z}_p^*}$ pour presque tout p . Si $\phi \in \mathcal{L}_0(\mathbf{A}^*)$, on définit $\int_{\mathbf{A}^*} \phi(x) d^*x$ par linéarité en imposant que

$$\int_{\mathbf{A}^*} \otimes_{v \in \mathcal{V}} \phi_v d^*x = \prod_{v \in \mathcal{V}} \left(\int_{\mathbf{Q}_v^*} \phi_v(x_v) d^*x_v \right),$$

où presque tous les termes du produit sont égaux à 1. Cela revient à demander que la mesure de $[1, a] \times \hat{\mathbf{Z}}^*$ soit $\log a$ et que d^*x soit invariante par $x \mapsto bx$, si $b \in \mathbf{A}^*$.

On note $\mathcal{L}(\mathbf{A}^*)$ l'espace des fonctions dont la restriction à $\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*$ appartient à $\mathcal{L}_0(\mathbf{A}^*)$, quel que soit $\beta \in \mathbf{Q}^*$, et qui sont *sommables*, i.e. $\sum_{\beta \in \mathbf{Q}^*} \int_{\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*} |\phi(x)| d^*x < +\infty$. Si $\phi \in \mathcal{L}(\mathbf{A}^*)$, on définit $\int_{\mathbf{A}^*} \phi(x) d^*x$ par la formule

$$\int_{\mathbf{A}^*} \phi(x) d^*x = \sum_{\beta \in \mathbf{Q}^*} \int_{\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*} \phi(x) d^*x.$$

Proposition D.2.10. — Soit $\phi_v \in \mathcal{L}(\mathbf{Q}_v^*)$, pour $v \in \mathcal{V}$, vérifiant :

- $\phi_p = 1$ sur \mathbf{Z}_p^* pour presque tout p ,
- $\prod_{v \in \mathcal{V}} \left(\int_{\mathbf{Q}_v^*} |\phi_v(x_v)| d^*x_v \right) < +\infty$.

Alors $\phi : \mathbf{A}^* \rightarrow \mathbf{C}$ définie par $\phi(x) = \prod_{v \in \mathcal{V}} \phi_v(x_v)$ est un élément de $\mathcal{L}(\mathbf{A}^*)$, et $\int_{\mathbf{A}^*} \phi(x) d^*x = \prod_{v \in \mathcal{V}} \left(\int_{\mathbf{Q}_v^*} \phi_v(x_v) d^*x_v \right)$.

Démonstration. — Si $\beta \in \mathbf{Q}^*$, la restriction de ϕ à $\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*$ est le produit de $\mathbf{1}_{\mathbf{R}_{\text{sign}(\beta)}^*} \phi_\infty$ et des $\mathbf{1}_{p^{v_p(\beta)}\mathbf{Z}_p^*} \phi_p$, et comme $v_p(\beta) = 0$ pour presque tout p , et $\phi_p = 1$ sur \mathbf{Z}_p^* pour presque tout p , cette restriction est visiblement élément de $\mathcal{L}_0(\mathbf{A}^*)$.

Le reste se démontre comme la prop. VII.3.1. Soit $\mathcal{P}(X)$ l'ensemble des nombres premiers $\leq X$, $I(X)$ l'ensemble des $\beta \in \mathbf{Q}^*$, positifs ou négatifs, dont la décomposition en facteurs premiers ne fait intervenir que des éléments de $\mathcal{P}(X)$ et, si $k \in \mathbf{N}$, soit $I(X, k)$ l'ensemble fini des $\beta \in I(X)$ tels que $-k \leq v_p(\beta) \leq k$, si $p \in \mathcal{P}(X)$. En notant $C_p(-k, k)$ la couronne $\{x \in \mathbf{Q}_p, p^{-k} \leq |x_p| \leq p^k\}$, on obtient

$$\sum_{\beta \in I(X, k)} \int_{\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*} |\phi(x)| d^*x = \int_{\mathbf{R}^* \times \prod_{p \leq X} C_p(-k, k) \times \prod_{p > X} \mathbf{Z}_p^*} |\phi(x)| d^*x.$$

Or $\phi(x) = \prod_{v \in \mathcal{V}} \phi_v(x_v)$, ce qui permet de mettre le membre de droite sous la forme

$$\left(\int_{\mathbf{R}^*} |\phi_\infty(x_\infty)| d^*x_\infty \right) \prod_{p \leq X} \left(\int_{C_p(-k, k)} |\phi_p(x_p)| d^*x_p \right),$$

si X est assez grand pour que $\phi_p = 1$ sur \mathbf{Z}_p^* , pour tout $p > X$. En faisant tendre k vers $+\infty$, on en déduit

$$\sum_{\beta \in I(X)} \int_{\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*} |\phi(x)| d^*x = \left(\int_{\mathbf{R}^*} |\phi_\infty(x_\infty)| d^*x_\infty \right) \prod_{p \leq X} \left(\int_{\mathbf{Q}_p^*} |\phi_p(x_p)| d^*x_p \right),$$

et en faisant tendre X vers $+\infty$, on obtient

$$\sum_{\beta \in \mathbf{Q}^*} \int_{\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*} |\phi(x)| d^*x = \prod_{v \in \mathcal{V}} \left(\int_{\mathbf{Q}_v^*} |\phi_v(x_v)| d^*x_v \right) < +\infty,$$

et donc ϕ est sommable. En particulier la série $\sum_{\beta \in \mathbf{Q}^*} \int_{\beta\mathbf{R}_+^*\hat{\mathbf{Z}}^*} \phi(x) d^*x$ est absolument convergente, ce qui permet, en reprenant les calculs ci-dessus sans les valeurs absolues, de démontrer la formule $\int_{\mathbf{A}^*} \phi(x) d^*x = \prod_{v \in \mathcal{V}} \left(\int_{\mathbf{Q}_v^*} \phi_v(x_v) d^*x_v \right)$, que l'on cherchait à obtenir.

3.3. Transformée de Mellin sur \mathbf{Q}_p

Lemme D.2.11. — Si χ est un caractère linéaire continu de \mathbf{Q}_p^* , il existe $n \in \mathbf{N} - \{0\}$ tel que $\chi = 1$ sur $1 + p^n\mathbf{Z}_p$.

Démonstration. — Comme χ est continu, il existe $n > 0$ tel que $|\chi(x) - 1| \leq 1/2$, si $|x - 1|_p \leq p^{-n}$. Or $B(1, p^{-n}) = 1 + p^n\mathbf{Z}_p$, est un sous-groupe d'indice fini de \mathbf{Z}_p^* (c'est l'image réciproque de $1 \in (\mathbf{Z}/p^n\mathbf{Z})^*$ par la projection naturelle $\mathbf{Z}_p^* \rightarrow (\mathbf{Z}_p/p^n\mathbf{Z}_p)^* = (\mathbf{Z}/p^n\mathbf{Z})^*$). Comme χ est un morphisme de groupes,

l'image de $1 + p^n \mathbf{Z}_p$ est un sous-groupe de \mathbf{C}^* inclus dans le disque $D(1, 1/2)$; on en déduit que cette image est réduite à $\{1\}$, ce qui permet de conclure.

On dit que χ est *non ramifié* si $\chi = 1$ sur \mathbf{Z}_p^* . Si χ est ramifié, on note $n(\chi)$ le plus petit entier n tel que $\chi = 1$ sur $1 + p^n \mathbf{Z}_p$; alors $p^{n(\chi)}$ est le *conducteur* de χ .

Si G est un groupe, un caractère linéaire $\chi : G \rightarrow \mathbf{C}^*$ est *unitaire* si $|\chi(g)| = 1$, quel que soit $g \in G$. On a alors $\chi(g^{-1}) = \overline{\chi}(g)$, pour tout $g \in G$.

Proposition D.2.12. — Soit $\chi : \mathbf{Q}_p^* \rightarrow \mathbf{C}^*$ un caractère linéaire unitaire continu.

(i) Si $\phi \in \mathcal{S}(\mathbf{Q}_p)$, alors $x \mapsto \phi(x)\chi(x)|x|_p^s \in \mathcal{L}(\mathbf{Q}_p^*)$, si $\operatorname{Re}(s) > 0$.

Si $\operatorname{Re}(s) > 0$, soit $M_p(\phi, \chi, s) = \int_{\mathbf{Q}_p^*} \chi(x)|x|_p^s d^*x$ la transformée de Mellin de ϕ .

(ii) $M_p(\mathbf{1}_{\mathbf{Z}_p}, \chi, s) = \frac{1}{1 - \chi(p)p^{-s}}$, si χ est non ramifié, et $M_p(\mathbf{1}_{\mathbf{Z}_p}, \chi, s) = 0$, si χ est ramifié.

(iii) Dans le cas général, $M_p(\phi, \chi, s)$ est un polynôme en p^s et p^{-s} si χ est ramifié, et de la forme $\frac{\phi(0)}{1 - \chi(p)p^{-s}} + R(s)$, où $R(s)$ est un polynôme en p^s et p^{-s} , si χ est non ramifié.

Démonstration. — On écrit ϕ sous la forme $\phi(0)\mathbf{1}_{\mathbf{Z}_p} + \psi$, où $\psi \in \mathcal{L}_0(\mathbf{Q}_p^*)$. Il existe alors $n \in \mathbf{N}$ tel que $\psi\chi$ soit constant modulo $p^n \mathbf{Z}_p$, ce qui permet de l'écrire sous la forme $\sum_{a \in \mathbf{I}} \lambda_a \mathbf{1}_{a+p^n \mathbf{Z}_p}$. On a alors $M_p(\psi, \chi, s) = \sum_{a \in \mathbf{I}} \lambda_a \frac{p^{1-n}}{p-1} |a|_p^{s-1}$, et comme $|a|_p \in p^{\mathbf{Z}}$, cela montre que $M_p(\psi, \chi, s)$ est un polynôme en p^s et p^{-s} .

Maintenant, on a $\int_{p^n \mathbf{Z}_p^*} |x|_p^s d^*x = p^{-n \operatorname{Re}(s)}$. On en déduit que $x \mapsto \mathbf{1}_{\mathbf{Z}_p}(x)\chi(x)|x|_p^s$ est sommable si $\operatorname{Re}(s) > 0$. De plus, on a

$$\begin{aligned} M_p(\mathbf{1}_{\mathbf{Z}_p}, \chi, s) &= \sum_{n=0}^{+\infty} \int_{p^n \mathbf{Z}_p^*} \chi(x)|x|_p^s d^*x = \sum_{n=0}^{+\infty} \int_{\mathbf{Z}_p^*} \chi(p^n x) |p^n x|_p^s d^*x \\ &= \sum_{n=0}^{+\infty} \chi(p)^n p^{-ns} \int_{\mathbf{Z}_p^*} \chi(x) d^*x = \frac{1}{1 - \chi(p)p^{-s}} \int_{\mathbf{Z}_p^*} \chi(x) d^*x, \end{aligned}$$

et le résultat suit de ce que $\int_{\mathbf{Z}_p^*} \chi(x) d^*x = 1$ si $\chi = 1$ sur \mathbf{Z}_p^* et est nul si χ n'est pas trivial sur \mathbf{Z}_p^* , puisqu'il se factorise alors à travers le groupe fini $(\mathbf{Z}/p^{n(\chi)}\mathbf{Z})^*$, et que la somme des valeurs d'un caractère linéaire non trivial d'un groupe fini est nulle (orthogonalité des caractères linéaires).

3.4. La transformée de Mellin adélique

Proposition D.2.13. — Soit χ un caractère continu de \mathbf{A}^* , et si $v \in \mathcal{V}$, soit χ_v la restriction de χ à \mathbf{Q}_v^* . Alors, pour presque tout $p \in \mathcal{P}$, on a $\chi_p = 1$ sur \mathbf{Z}_p^* , et, si $x \in \mathbf{A}^*$, $\chi(x) = \prod_{v \in \mathcal{V}} \chi_v(x_v)$.

Démonstration. — Comme χ est continu sur \mathbf{A}^* , il est continu sur $\widehat{\mathbf{Z}}^* = \prod_{p \in \mathcal{P}} \mathbf{Z}_p^*$, et par définition de la topologie produit, il existe $S \subset \mathcal{P}$ fini et, si $p \in S$, $n_p \in \mathbf{N} - \{0\}$, tels que $|\chi(x) - 1| \leq 1/2$, si $x \in U = \prod_{p \in S} (1 + p^{n_p} \mathbf{Z}_p) \prod_{p \notin S} \mathbf{Z}_p^*$. Or U est un sous-groupe de $\widehat{\mathbf{Z}}^*$, et donc son image par χ est un sous-groupe de \mathbf{C}^* inclus dans le disque $D(1, 1/2)$; on en déduit que cette image est réduite à $\{1\}$. En particulier, on a $\chi_p = 1$ sur \mathbf{Z}_p^* pour tout $p \notin S$.

Maintenant, si $x \in \mathbf{A}^*$, on a $\chi_v(x_v) = 1$ pour presque tout v , ce qui fait que $\chi'(x) = \prod_{v \in \mathcal{V}} \chi_v(x_v)$ est bien défini, et que χ' est un caractère linéaire continu de \mathbf{A}^* . On conclut en remarquant que χ' et χ coïncident sur le sous-groupe des idéles dont presque toutes les composantes sont égales à 1, et en utilisant la densité de ce sous-groupe dans \mathbf{A}^* . (Si U est un ouvert de \mathbf{A}^* , il contient un ouvert de la

forme $\prod_{v \in S} U_v \times \prod_{p \notin S} \mathbf{Z}_p^*$, et donc contient des éléments dont toutes les composantes en dehors de S sont égales à 1.)

Si $\phi \in \mathcal{L}(\mathbf{R}^*)$, et si χ est un caractère linéaire unitaire continu de \mathbf{R}^* , la *transformée de Mellin* $M_\infty(\phi, \chi, s)$ de ϕ est la fonction $s \mapsto \int_{\mathbf{R}^*} \phi(t) \chi(t) |t|^s d^*t$, qui est holomorphe dans le demi-plan $\text{Re}(s) > 0$, comme le montre le th. V.2.18.

Proposition D.2.14. — Soit χ un caractère unitaire continu de \mathbf{A}^* et, si $v \in \mathcal{V}$, soit χ_v la restriction de χ à \mathbf{Q}_v^* .

(i) Si $\phi \in \mathcal{S}(\mathbf{A})$, alors $x \mapsto \phi(x) \chi(x) |x|_{\mathbf{A}}^s \in \mathcal{L}(\mathbf{A}^*)$, si $\text{Re}(s) > 1$.

(ii) La transformée de Mellin $M_{\mathbf{A}}(\phi, \chi, s) = \int_{\mathbf{A}^*} \phi(x) \chi(x) |x|_{\mathbf{A}}^s d^*x$ de ϕ est holomorphe sur le demi-plan $\text{Re}(s) > 1$.

(iii) Si $\phi = \otimes_{v \in \mathcal{V}} \phi_v$, et si $\text{Re}(s) > 1$, alors $M_{\mathbf{A}}(\phi, \chi, s) = \prod_{v \in \mathcal{V}} M_v(\phi_v, \chi_v, s)$.

Démonstration. — Commençons par supposer que $\phi = \otimes_{v \in \mathcal{V}} \phi_v$, où $\phi_v \in \mathcal{S}(\mathbf{Q}_v)$, et $\phi_p = \mathbf{1}_{\mathbf{Z}_p}$, pour presque tout p . Soit $\psi_v = \phi_v \chi_v |x_v|^s$. Il résulte de la prop. D.2.12 que, si $\text{Re}(s) > 0$, alors $\psi_v \in \mathcal{L}(\mathbf{Q}_v)$ quel que soit $v \in \mathcal{V}$, et que $\int_{\mathbf{Q}_p^*} |\psi_p| d^*x_p = \frac{1}{1-p^{-\text{Re}(s)}}$ pour presque tout p . Comme le produit des $\frac{1}{1-p^{-\text{Re}(s)}}$ est convergent pour $\text{Re}(s) > 1$, et comme $\psi_p = 1$ sur \mathbf{Z}_p^* pour presque tout p , d'après la prop. D.2.13, on est dans les conditions d'application de la prop. D.2.10. On en déduit que $\prod_{v \in \mathcal{V}} \psi_v = \phi(x) \chi(x) |x|_{\mathbf{A}}^s \in \mathcal{L}(\mathbf{A}^*)$, si $\text{Re}(s) > 1$, et que

$$M_{\mathbf{A}}(\phi, \chi, s) = \prod_{v \in \mathcal{V}} \int_{\mathbf{Q}_v^*} \psi_v(x_v) d^*x_v = \prod_{v \in \mathcal{V}} M_v(\phi_v, \chi_v, s).$$

Comme chacune des $M_v(\phi_v, \chi_v, s)$ est holomorphe sur $\text{Re}(s) > 0$, et comme le produit est absolument convergent sur tout demi-plan $\text{Re}(s) > c$, si $c > 1$, on en déduit (th. V.2.15) l'holomorphie de $M_{\mathbf{A}}(\phi, \chi, s)$ sur le demi-plan $\text{Re}(s) > 1$. Ceci termine la démonstration dans le cas où ϕ est un produit. Les (i) et (ii) dans le cas général s'en déduisent par linéarité. Ceci permet de conclure.

3.5. Le théorème de Tate

Soit χ un caractère unitaire continu de \mathbf{A}^* , et soit χ_v , si $v \in \mathcal{V}$, la restriction de χ à \mathbf{Q}_v^* . Soit $S(\chi)$, l'ensemble des p tels que χ_p soit ramifié. D'après la prop. D.2.13 ci-dessus, cet ensemble est fini. On définit le *conducteur* $D(\chi)$ de χ par la formule $D(\chi) = \prod_{p \in S(\chi)} p^{n(\chi_p)}$.

Par ailleurs, la restriction de χ_∞ à \mathbf{R}_+^* est unitaire continue; il existe donc $t(\chi) \in \mathbf{R}$ tel que $\chi_\infty(x) = x^{it(\chi)}$, si $x \in \mathbf{R}_+^*$.

Théorème D.2.15. — (Tate, 1950) Si $\phi \in \mathcal{S}(\mathbf{A})$, et si $\chi : \mathbf{A}^* \rightarrow \mathbf{C}^*$ est un caractère linéaire unitaire continu, trivial sur \mathbf{Q}^* , alors $M_{\mathbf{A}}(\phi, \chi, s)$ admet un prolongement méromorphe à \mathbf{C} tout entier, holomorphe en dehors de pôles simples en $s = -it(\chi)$ et $s = 1 - it(\chi)$, si $D(\chi) = 1$, de résidus respectifs $-\phi(0)$ et $\hat{\phi}(0)$, et vérifie l'équation fonctionnelle $M_{\mathbf{A}}(\phi, \chi, s) = M_{\mathbf{A}}(\hat{\phi}, \bar{\chi}, 1 - s)$.

Démonstration. — Si $\text{Re}(s) > 1$, on a

$$M_{\mathbf{A}}(\phi, \chi, s) = \int_{\mathbf{A}^*} \phi(x) \chi(x) |x|_{\mathbf{A}}^s d^*x = \sum_{\beta \in \mathbf{Q}^*} \int_{\beta \cdot \mathbf{R}_+^* \cdot \hat{\mathbf{Z}}^*} \phi(x) \chi(x) |x|_{\mathbf{A}}^s d^*x$$

Or d^*x et $\chi(x)|x|_{\mathbf{A}}^s$ sont invariants par $x \mapsto \beta x$, ce qui permet de réécrire l'expression précédente sous la forme

$$\sum_{\beta \in \mathbf{Q}^*} \int_0^{+\infty} \int_{\hat{\mathbf{Z}}^*} \phi(\beta x) \chi(x) |x|_{\mathbf{A}}^s d^*x = \int_0^{+\infty} \int_{\hat{\mathbf{Z}}^*} \left(\sum_{\beta \in \mathbf{Q}^*} \phi(\beta x) \right) \chi(x) |x|_{\mathbf{A}}^s d^*x,$$

l'échange des signes \sum et \int étant justifié par la sommabilité de $\phi(x)\chi(x)|x|_{\mathbf{A}}^s$ sur \mathbf{A}^* . Maintenant, comme $\mathbf{Q}^* = \mathbf{Q} - \{0\}$, la formule de Poisson adélique (th. D.2.9) permet d'écrire $\sum_{\beta \in \mathbf{Q}^*} \phi(\beta x)$ sous la forme

$$-\phi(0) + \sum_{\beta \in \mathbf{Q}} \phi(\beta x) = -\phi(0) + |x|_{\mathbf{A}}^{-1} \sum_{\beta \in \mathbf{Q}} \hat{\phi}(\beta x^{-1}) = \hat{\phi}(0) |x|_{\mathbf{A}}^{-1} - \phi(0) + |x|_{\mathbf{A}}^{-1} \sum_{\beta \in \mathbf{Q}^*} \hat{\phi}(\beta x^{-1}).$$

On peut alors couper l'intégrale $\int_0^{+\infty}$ en $\int_0^1 + \int_1^{+\infty}$, remplacer $\sum_{\beta \in \mathbf{Q}^*} \phi(\beta x)$ par l'expression ci-dessus dans \int_0^1 , faire le changement de variable $x \mapsto x^{-1}$ dans \int_0^1 , et utiliser l'identité $\chi(x^{-1}) = \bar{\chi}(x)$ car χ est unitaire. Comme

$$\int_0^1 \int_{\hat{\mathbf{Z}}^*} |x|_{\mathbf{A}}^u \chi(x) d^*x = \left(\int_0^1 x_{\infty}^{u+it(\chi)} \frac{dx_{\infty}}{x_{\infty}} \right) \left(\prod_{p \in \mathcal{P}} \int_{\mathbf{Z}_p^*} \chi_p(x_p) d^*x_p \right) = \begin{cases} 0 & \text{si } D(\chi) \neq 1, \\ \frac{1}{u+it(\chi)} & \text{si } D(\chi) = 1, \end{cases}$$

(si $p|D(\chi)$, alors $\int_{\mathbf{Z}_p^*} \chi_p(x_p) d^*x_p = 0$), on obtient

$$M_{\mathbf{A}}(\phi, \chi, s) = \int_1^{+\infty} \int_{\hat{\mathbf{Z}}^*} \sum_{\beta \in \mathbf{Q}^*} (\phi(\beta x) \chi(x) |x|_{\mathbf{A}}^s + \hat{\phi}(\beta x) \bar{\chi}(x) |x|_{\mathbf{A}}^{1-s}) d^*x,$$

si $D(\chi) \neq 1$, auquel il faut rajouter $-\frac{\phi(0)}{s+it(\chi)} - \frac{\hat{\phi}(0)}{1-s-it(\chi)}$, si $D(\chi) = 1$. On en déduit le théorème car $\sum_{\beta \in \mathbf{Q}^*} (\phi(\beta x) \chi(x) |x|_{\mathbf{A}}^s + \hat{\phi}(\beta x) \bar{\chi}(x) |x|_{\mathbf{A}}^{1-s})$ est sommable pour toute valeur de s , et définit une fonction holomorphe de s sur \mathbf{C} tout entier, et la formule $\hat{\phi}(x) = \phi(-x)$ montre que ce qu'on obtient ne change pas si on remplace simultanément ϕ par $\hat{\phi}$, χ par $\bar{\chi}$ et s par $1-s$ (on a $t(\bar{\chi}) = -t(\chi)$).

4. Application aux fonctions L de Dirichlet

Le th. D.2.15 permet de redémontrer l'existence de prolongements analytiques et d'équations fonctionnelles pour les fonctions L de Dirichlet. En ce qui concerne les fonctions L de Dirichlet, le gain n'est pas flagrant, mais les calculs adéliques ne sont pas plus compliqués pour les fonctions L de Hecke, et dans ce cas, le gain par rapport à la méthode originelle de Hecke devient très appréciable.

4.1. *La fonction zêta de Riemann.* — Le résultat suivant a déjà été démontré, et la démonstration proposée ci-dessous n'est qu'une traduction adélique de celle de l'ex. VII.6.6.

Proposition D.2.16. — Soit $\xi(s) = \frac{\Gamma(s/2)}{\pi^{s/2}} \zeta(s)$. Alors $\xi(s)$ a un prolongement méromorphe à \mathbf{C} , holomorphe en dehors de pôles simples en $s = 0$ et $s = 1$, de résidus respectifs -1 et 1 , et vérifie l'équation fonctionnelle $\xi(1-s) = \xi(s)$.

Démonstration. — Soit $\phi = \otimes_v \phi_v \in \mathcal{S}(\mathbf{A})$, avec $\phi_\infty(t) = e^{-\pi t^2}$, et $\phi_p = \mathbf{1}_{\mathbf{Z}_p}$, quel que soit $p \in \mathbf{Z}_p$. On a $\hat{\phi}_v = \phi_v$, quel que soit v , et donc $\hat{\phi} = \phi$. De plus $\phi(0) = \hat{\phi}(0) = 1$. Il résulte donc du th. D.2.15 que $M_{\mathbf{A}}(\phi, \mathbf{1}, s)$ admet un prolongement méromorphe à \mathbf{C} , holomorphe en dehors de pôles simples en $s = 0$ et $s = 1$, de résidus respectifs -1 et 1 , et vérifie l'équation fonctionnelle $M_{\mathbf{A}}(\phi, \mathbf{1}, 1-s) = M_{\mathbf{A}}(\phi, \mathbf{1}, s)$.

Par ailleurs, on a $M_{\mathbf{A}}(\phi, \mathbf{1}, s) = \prod_{v \in \mathcal{V}} M_v(\phi_v, \mathbf{1}, s)$, si $\operatorname{Re}(s) > 1$, d'après le (iii) de la prop. D.2.14. Or $M_p(\phi_p, \mathbf{1}, s) = \frac{1}{1-p^{-s}}$, si $p \in \mathcal{P}$, d'après le (ii) de la prop. D.2.12, et donc $\prod_{p \in \mathcal{P}} M_p(\phi_p, \mathbf{1}, s) = \zeta(s)$. Finalement,

$$M_\infty(\phi_\infty, \mathbf{1}, s) = M_\infty(e^{-\pi t^2}, \mathbf{1}, s) = \int_{\mathbf{R}^*} e^{-\pi t^2} t^s \frac{dt}{|t|} = \int_0^{+\infty} e^{-\pi u} u^{s/2} \frac{du}{u} = \frac{\Gamma(s/2)}{\pi^{s/2}}.$$

On en déduit que $\xi(s) = M_{\mathbf{A}}(\phi, \mathbf{1}, s)$, ce qui permet de conclure.

4.2. Fonctions L de caractères de \mathbf{A}^*

Soit χ un caractère linéaire unitaire continu de \mathbf{A}^* , trivial sur \mathbf{Q}^* , dont le conducteur $D(\chi)$ est différent de 1, et soit χ_v la restriction de χ à \mathbf{Q}_v^* , si $v \in \mathcal{V}$.

Comme on a supposé que χ est unitaire, on a soit $\chi_\infty(x_\infty) = |x_\infty|^{it(\chi)}$, soit $\chi_\infty(x_\infty) = \operatorname{sign}(x_\infty)|x_\infty|^{it(\chi)}$. Dans le premier cas, on dit que χ est *pair*, dans le second, qu'il est *impair*. On pose

$$w_\infty(\chi) = \begin{cases} 1 & \text{si } \chi \text{ pair,} \\ -i & \text{si } \chi \text{ impair,} \end{cases} \quad c(\chi) = \begin{cases} 0 & \text{si } \chi \text{ pair,} \\ 1 & \text{si } \chi \text{ impair.} \end{cases}$$

On définit $\varepsilon(\chi, s)$ par la formule $\varepsilon(\chi, s) = \prod_v \varepsilon_v(\chi, s)$, avec

$$\varepsilon_v(\chi, s) = \begin{cases} w_\infty(\chi) & \text{si } v = \infty, \\ w_p(\chi) p^{-n(\chi_p)s} & \text{si } p \in \mathcal{P}, \end{cases}, \quad \text{où } w_p(\chi) = \begin{cases} 1 & \text{si } p \nmid D(\chi), \\ \chi_p(p^{n(\chi_p)}) G(\bar{\chi}_p) & \text{si } p \mid D(\chi). \end{cases}$$

On a donc $\varepsilon(\chi, s) = (\prod_v w_v(\chi)) D(\chi)^{-s}$. (La somme de Gauss $G(\bar{\chi}_p)$ est la somme de Gauss du caractère de Dirichlet de conducteur $p^{n(\chi_p)}$ obtenu en utilisant l'isomorphisme $(\mathbf{Z}_p/p^{n(\chi_p)}\mathbf{Z}_p)^* \cong (\mathbf{Z}/p^{n(\chi_p)}\mathbf{Z})^*$.)

Finalement, on définit la fonction L de χ par $L(\chi, s) = \prod_{p \nmid D(\chi)} \frac{1}{1-\chi_p(p)p^{-s}}$, si $\operatorname{Re}(s) > 1$.

Proposition D.2.17. — La fonction $\Lambda(\chi, s) = \frac{\Gamma((s+it(\chi)+c(\chi))/2)}{\pi^{(s+it(\chi)+c(\chi))/2}} L(\chi, s)$ admet un prolongement holomorphe à \mathbf{C} , et vérifie l'équation fonctionnelle $\Lambda(\chi, s) = \varepsilon(\chi, s) \Lambda(\bar{\chi}, 1-s)$.

Démonstration. — Soit $\phi = \otimes_v \phi_v$, on l'on a posé

- $\phi_\infty(t) = e^{-\pi t^2}$, si χ est pair, et $\phi_\infty(t) = t e^{-\pi t^2}$, si χ est impair,
- $\phi_p = \mathbf{1}_{\mathbf{Z}_p}$, si $p \nmid D(\chi)$,
- $\phi_p = \mathbf{1}_{\mathbf{Z}_p^*} \bar{\chi}_p$, si $p \mid D(\chi)$.

Nous allons calculer les facteurs $M_v(\phi_v, \chi_v, s)$ et $M_v(\hat{\phi}_v, \bar{\chi}_v, 1-s)$, pour tout v .

• On a $\hat{\phi}_\infty = w_\infty(\chi) \phi_\infty$. En effet, si χ est pair, ceci est équivalent au fait que la transformée de Fourier de $e^{-\pi t^2}$ est $e^{-\pi x^2}$. Si χ est impair, ceci équivaut au fait que la transformée de Fourier de $t e^{-\pi t^2}$ est $-i x e^{-\pi x^2}$, ce qui peut se vérifier en remarquant que $-2\pi t e^{-\pi t^2}$ est la dérivée de $e^{-\pi t^2}$ et donc sa dérivée est $2i\pi x$ fois la transformée de Fourier de $e^{-\pi t^2}$. On en déduit que

$$M_\infty(\phi_\infty, \chi_\infty, s) = \frac{\Gamma((s+it(\chi)+c(\chi))/2)}{\pi^{(s+it(\chi)+c(\chi))/2}} \quad \text{et} \quad M_\infty(\hat{\phi}_\infty, \bar{\chi}_\infty, s) = w_\infty(\chi) \frac{\Gamma((s+it(\chi)+c(\chi))/2)}{\pi^{(s+it(\chi)+c(\chi))/2}},$$

par un petit calcul analogue à celui que l'on a fait pour déterminer $M_\infty(e^{-\pi t^2}, \mathbf{1}, s)$.

- Si $p \nmid D(\chi)$, on a $\hat{\phi}_p = \mathbf{1}_{\mathbf{Z}_p}$, et donc, d'après le (ii) de la prop. D.2.12, si $\text{Re}(s) > 0$,

$$M_p(\phi_p, \chi_p, s) = \frac{1}{1 - \chi_p(p)p^{-s}} \quad \text{et} \quad M_p(\hat{\phi}_p, \bar{\chi}_p, s) = \frac{1}{1 - \bar{\chi}_p(p)p^{-s}}.$$

• Si $p \mid D(\chi)$, on $M_p(\phi_p, \chi_p, s) = 1$, comme le montre un calcul immédiat. Par ailleurs, on a $\hat{\phi}_p(x) = \frac{G(\bar{\chi}_p)}{p^{n(\chi_p)}} \chi_p(p^{n(\chi_p)} x) \mathbf{1}_{p^{-n(\chi_p)} \mathbf{Z}_p^*}(x)$, d'après le (iii) de l'ex. D.2.7, et donc

$$\begin{aligned} M_p(\hat{\phi}_p, \bar{\chi}_p, 1-s) &= \frac{G(\bar{\chi}_p)}{p^{n(\chi_p)}} \int_{p^{-n(\chi_p)} \mathbf{Z}_p^*} \chi_p(p^{n(\chi_p)} x) \bar{\chi}_p(x) |x|_p^{1-s} d^*x \\ &= \frac{G(\bar{\chi}_p)}{p^{n(\chi_p)}} \chi_p(p^{n(\chi_p)}) p^{n(\chi_p)(1-s)} \int_{p^{-n(\chi_p)} \mathbf{Z}_p^*} d^*x = w_p(\chi) p^{-n(\chi_p)(s)}. \end{aligned}$$

On a donc, si $\text{Re}(s) > 1$,

$$M_{\mathbf{A}}(\phi, \chi, s) = \prod_{v \in \mathcal{V}} M_v(\phi_v, \chi_v, s) = \frac{\Gamma((s+it(\chi)+c(\chi))/2)}{\pi^{(s+it(\chi)+c(\chi))/2}} \prod_{p \nmid D(\chi)} \frac{1}{1 - \chi_p(p)p^{-s}} = \Lambda(\chi, s).$$

De même, si $\text{Re}(s) < 0$, alors $M_{\mathbf{A}}(\hat{\phi}, \bar{\chi}, 1-s) = \prod_{v \in \mathcal{V}} M_v(\hat{\phi}_v, \bar{\chi}_v, 1-s)$ est aussi égal à

$$w_\infty(\chi) \frac{\Gamma((1-s+it(\chi)+c(\chi))/2)}{\pi^{(1-s+it(\chi)+c(\chi))/2}} \left(\prod_{p \mid D(\chi)} w_p(\chi) p^{-n(\chi_p)(s)} \right) \prod_{p \nmid D(\chi)} \frac{1}{1 - \bar{\chi}_p(p)p^{-(1-s)}},$$

c'est-à-dire à $\varepsilon(\chi, s) \Lambda(\bar{\chi}, 1-s)$. Le résultat suit donc du th. D.2.15, en utilisant le fait que $\phi(0) = \hat{\phi}(0) = 0$ car $\phi_p(0) = \hat{\phi}_p(0) = 0$, si $p \mid D(\chi)$.

4.3. Caractères de Dirichlet et caractères linéaires continus des idèles

A un caractère de Dirichlet χ de conducteur D , on peut associer, grâce au lemme D.2.5, un caractère linéaire $\chi_{\mathbf{A}}$ de \mathbf{A}^* , continu, d'ordre fini (et donc unitaire), par la formule $\chi_{\mathbf{A}}(x) = \chi(\pi_D(b^{[\infty]}))^{-1}$, si $x = \beta b_{\infty} b^{[\infty]}$, où $\beta \in \mathbf{Q}^*$, $b_{\infty} \in \mathbf{R}_+^*$, $b^{[\infty]} \in \widehat{\mathbf{Z}}^*$, et où l'on note $\pi_D : \widehat{\mathbf{Z}}^* \rightarrow (\mathbf{Z}/D\mathbf{Z})^*$ la projection naturelle⁽³⁷⁾. On remarquera que $\chi_{\mathbf{A}}$ est trivial sur \mathbf{Q}^* par construction.

Lemme D.2.18. — $L(\chi_{\mathbf{A}}, s) = L(\chi, s)$.

Démonstration. — On note χ_p la restriction de $\chi_{\mathbf{A}}$ à \mathbf{Q}_p^* , et il s'agit de vérifier que si $p \nmid D$, alors $\chi(p) = \chi_p(p)$. Soit x (resp. y) l'idèle dont toutes les composantes sont 1 (resp. p^{-1}) sauf la composante en p qui est égale à p (resp. 1). On a alors $\chi_p(p) = \chi_{\mathbf{A}}(x)$ par définition, et comme $x = py$, on a $\chi_p(p) = \chi(\pi_D(y^{[\infty]}))^{-1}$. Or $y^{[\infty]}$ a toutes ses composantes en les ℓ divisant D égales à p^{-1} ; on en déduit que $\pi_D(y^{[\infty]}) = p^{-1}$, et donc que $\chi_p(p) = \chi(p^{-1})^{-1} = \chi(p)$, ce qui permet de conclure.

On déduit de ce lemme, et de la prop. D.2.17, une démonstration adélique de l'existence d'un prolongement analytique et d'une équation fonctionnelle pour $L(\chi, s)$. Par ailleurs, ce lemme permet aussi de reformuler le théorème de Kronecker-Weber sous la forme suivante.

Théorème D.2.19. — Si ρ est une représentation de dimension 1 de $\mathcal{G}_{\mathbf{Q}}$, il existe un caractère linéaire unitaire continu $\chi(\rho)$ de $\mathbf{A}^* = \mathbf{GL}_1(\mathbf{A})$, trivial sur $\mathbf{Q}^* = \mathbf{GL}_1(\mathbf{Q})$, tel que $L(\rho, s) = L(\chi(\rho), s)$.

D.3. Le programme de Langlands

Le programme de Langlands consiste à remplacer 1 par n dans l'énoncé du th. D.2.19 ci-dessus. (C'est plus facile à dire qu'à faire...)

1. Représentations automorphes

Notons G le groupe \mathbf{GL}_n , et donc $G(\mathbf{A})$, $G(\mathbf{Q})$, $G(\mathbf{R})$, $G(\mathbf{Q}_p)$ et $G(\mathbf{Z}_p)$ désignent respectivement les groupes $\mathbf{GL}_n(\mathbf{A})$, $\mathbf{GL}_n(\mathbf{Q})$, $\mathbf{GL}_n(\mathbf{R})$, $\mathbf{GL}_n(\mathbf{Q}_p)$ et $\mathbf{GL}_n(\mathbf{Z}_p)$. Alors $G(\mathbf{A})$ est le produit restreint des $G(\mathbf{Q}_v)$, pour $v \in \mathcal{V}$, relativement aux $G(\mathbf{Z}_p)$, pour $p \in \mathcal{P}$; autrement dit, un élément x de $G(\mathbf{A})$ est de la forme $(x_v)_{v \in \mathcal{V}}$, avec $x_v \in G(\mathbf{Q}_v)$, et $x_p \in G(\mathbf{Z}_p)$ pour presque tout p . On écrit aussi x sous la forme $(x_{\infty}, x^{[\infty]})$, où $x^{[\infty]} = (x_p)_{p \in \mathcal{P}}$ est la partie de x en dehors de ∞ . Comme d'habitude, $G(\mathbf{Q}_v)$ s'identifie à un sous-groupe de $G(\mathbf{A})$, si $v \in \mathcal{V}$.

⁽³⁷⁾On envoie \mathbf{Z}_p^* sur 1, si $p \nmid D$, et \mathbf{Z}_p^* sur $(\mathbf{Z}_p/p^{v_p(D)}\mathbf{Z}_p)^* = (\mathbf{Z}/p^{v_p(D)}\mathbf{Z})^*$, si $p \mid D$, et on utilise l'isomorphisme $(\mathbf{Z}/D\mathbf{Z})^* = \prod_{p \mid D} (\mathbf{Z}/p^{v_p(D)}\mathbf{Z})^*$ fourni par le théorème des restes chinois.

Soit $\mathcal{A}_0(G(\mathbf{Q}) \backslash G(\mathbf{A}))$ le \mathbf{C} -espace vectoriel des *formes automorphes*⁽³⁸⁾ cuspidales pour $G(\mathbf{A})$, c'est-à-dire des fonctions $\phi : G(\mathbf{A}) \rightarrow \mathbf{C}$ vérifiant :

- (i) $\phi(\gamma x) = \phi(x)$ quels que soient $\gamma \in G(\mathbf{Q})$ et $x \in G(\mathbf{A})$,
- (ii) il existe K_ϕ d'indice fini dans $\prod_{p \in \mathcal{P}} G(\mathbf{Z}_p)$ tel que $\phi(xh) = \phi(x)$ quels que soient $x \in G(\mathbf{A})$ et $h \in K_\phi$, et les $\phi(xh)$, pour⁽³⁹⁾ $h \in \mathbf{O}_n(\mathbf{R})$, engendrent un espace de dimension finie,
- (iii) il existe χ caractère linéaire unitaire continu de \mathbf{A}^* , trivial sur \mathbf{Q}^* , tel que $\phi(A_z x) = \phi(x A_z) = \chi(z) \phi(x)$, si A_z est la matrice de l'homothétie de rapport $z \in \mathbf{A}^*$,
- (iv) Si $x^{|\infty|}$ est fixé, $\phi(x_\infty, x^{|\infty|})$ est une fonction de classe \mathcal{C}^∞ des coordonnées de x_∞ , vecteur propre de tous les opérateurs différentiels commutant à l'action de $G(\mathbf{R})$.
- (v) ϕ est à décroissance rapide à l'infini (en un sens que nous ne précisons pas).

On fait agir $g \in G(\mathbf{A})$ sur les fonctions $\phi : G(\mathbf{A}) \rightarrow \mathbf{C}$ par translation à droite sur la variable, c'est-à-dire que l'on pose $g \cdot \phi(x) = \phi(xg)$. Alors $\mathcal{A}_0(G(\mathbf{Q}) \backslash G(\mathbf{A}))$ n'est pas stable par l'action de $G(\mathbf{A})$, mais presque⁽⁴⁰⁾, et nous allons un peu tricher en prétendant qu'il l'est (cf. note 40 pour un énoncé correct), et donc que $\mathcal{A}_0(G(\mathbf{Q}) \backslash G(\mathbf{A}))$ est une représentation de $G(\mathbf{A})$. Alors $\mathcal{A}_0(G(\mathbf{Q}) \backslash G(\mathbf{A}))$ se décompose en somme directe de représentations irréductibles (ses composantes irréductibles sont les *représentations automorphes cuspidales*). Le théorème ci-dessous représente un travail certain (dû, pour $n = 2$, à H. Jacquet et R. Langlands(1969), et pour $n \geq 3$, à R. Godement et H. Jacquet(1972)). La démonstration de l'existence et du prolongement analytique des fonctions L automorphes a fortement été inspirée par la méthode introduite par J. Tate pour étudier les fonctions L de Dirichlet et de Hecke.

Théorème D.3.1. — *Soit Π une représentation automorphe cuspidale de degré n . Alors*

⁽³⁸⁾L'automorphie traduit la condition (i) qui est la plus subtile des conditions (i)-(v) ci-dessus, et la cuspidalité correspond à la condition (v).

⁽³⁹⁾Le groupe $\mathbf{O}_n(\mathbf{R})$ est le groupe des isométries de \mathbf{R}^n ; c'est un sous-groupe compact de $G(\mathbf{R})$, et il est maximal pour cette propriété, de la même manière que $G(\mathbf{Z}_p)$ est un sous-groupe compact de $G(\mathbf{Q}_p)$, et est maximal pour cette propriété.

⁽⁴⁰⁾Il est stable par $G(\mathbf{A}^{|\infty|})$, par $\mathbf{O}_n(\mathbf{R})$, mais pas par $G(\mathbf{R})$ à cause de la condition (iii) imposant que les $\phi(xh)$, pour $h \in \mathbf{O}_n(\mathbf{R})$, engendrent un espace de dimension finie. Par contre, il est stable par l'action infinitésimale de $G(\mathbf{R})$, c'est-à-dire par les opérateurs différentiels ∂_A , pour $A \in \mathbf{M}_n(\mathbf{R})$, avec $\partial_A \phi(x) = \lim_{t \rightarrow 0} t^{-1}(\phi(xe^{tA}) - \phi(x))$, et $e^{tA} = \sum_{n=0}^{+\infty} \frac{t^n A^n}{n!} \in G(\mathbf{R})$ est l'exponentielle de la matrice tA . Il y a des conditions de compatibilité évidentes que doivent satisfaire les actions de $\mathbf{O}_n(\mathbf{R})$ et $\mathbf{M}_n(\mathbf{R})$ car $e^{tA} \in \mathbf{O}_n(\mathbf{R})$ si (et seulement si) A est antisymétrique. Dans ce qui suit, nous commettrons l'abus d'appeler représentation de $G(\mathbf{R})$ (resp. de $G(\mathbf{A})$) un \mathbf{C} -espace vectoriel muni d'actions de $\mathbf{O}_n(\mathbf{R})$ et $\mathbf{M}_n(\mathbf{R})$ vérifiant les relations de compatibilité évoquées ci-dessus.

(i) Π admet une factorisation⁽⁴¹⁾ sous la forme $\Pi = \otimes'_{v \in \mathcal{V}} \Pi_v$, où Π_v est une représentation irréductible (de dimension infinie) de $G(\mathbf{Q}_v)$.

(ii) Π admet une fonction L se factorisant sous la forme⁽⁴²⁾ $L(\Pi, s) = \prod_{v \in \mathcal{V}} L(\Pi_v, s)$, où $L(\Pi_\infty, s)$ est un produit de fonctions Γ qui ne dépend que de Π_∞ , et $L(\Pi_p, s) = \frac{1}{E_p(p^{-s})}$, où $E_p(X)$ est un polynôme de degré $\leq n$ et de degré n pour presque tout p , dont le terme constant est 1, et qui ne dépend que de Π_p .

(iii) $L(\Pi, s)$ admet un prolongement holomorphe à tout le plan complexe, et admet une équation fonctionnelle du type $L(\Pi, s) = \varepsilon(s)L(\Pi^\vee, 1-s)$, où Π^\vee est une autre représentation automorphe cuspidale (contragrédiente de Π), et $\varepsilon(s)$ est de la forme $A \cdot B^s$, avec $A \in \mathbf{C}^*$ et $B \in \mathbf{R}_+^*$.

Conjecture D.3.2. — (Langlands, 1968) Si ρ est une représentation irréductible de dimension n de $\mathcal{G}_{\mathbf{Q}}$, il existe une représentation automorphe cuspidale de degré n telle que $L(\rho, s) = L(\Pi(\rho), s)$.

Au vu du théorème ci-dessus, cette conjecture implique la conjecture d'Artin, mais va en fait bien au-delà. Ce n'est qu'un petit bout de l'édifice dont Langlands conjecture l'existence.

Nous montrerons comment transformer une forme modulaire primitive ou une forme de Maass primitive en une représentation automorphe cuspidale de $\mathbf{GL}_2(\mathbf{A})$ dans le n° 2, ce qui permet de voir les représentations automorphes cuspidales comme une vaste généralisation des formes modulaires (ou de Maass) primitives. Notons qu'en degré ≥ 3 , il ne semble pas forcément y avoir d'équivalents des formes modulaires ou de Maass, et le recours au langage des représentations devient difficilement contournable.

⁽⁴¹⁾Le groupe $G_{\mathbf{A}}$ est essentiellement un produit. Or on a vu, dans le cas des groupes finis, que si $G = G_1 \times G_2$, alors toute représentation irréductible de G se factorise sous la forme $V_1 \otimes V_2$ où V_i est une représentation irréductible de G_i , si $i = 1, 2$. Il est donc naturel de penser qu'une représentation irréductible de $G_{\mathbf{A}}$ va aussi admettre une factorisation, et c'est ce que prétend le (i) du th. D.3.1. D'un autre côté, il n'est pas très raisonnable de faire le produit tensoriel d'une infinité d'espaces vectoriels, et le produit tensoriel du théorème est un *produit tensoriel restreint*. (On a déjà vu des exemples de cette notion : $\mathcal{S}(\mathbf{A})$ est le produit tensoriel restreint des $\mathcal{S}(\mathbf{Q}_v)$, pour $v \in \mathcal{V}$, relativement aux fonctions $\mathbf{1}_{\mathbf{Z}_p}$, pour $p \in \mathcal{P}$; de même, $\mathcal{L}_0(\mathbf{A}^*)$ est le produit tensoriel restreint de $\mathcal{L}(\mathbf{R}^*)$ et des $\mathcal{L}_0(\mathbf{Q}_p^*)$, pour $p \in \mathcal{P}$, relativement aux fonctions $\mathbf{1}_{\mathbf{Z}_p^*}$, pour $p \in \mathcal{P}$.) De manière précise, pour tout p tel que le polynôme E_p du (ii) du théorème soit de degré n , la représentation Π_p possède un vecteur privilégié x_p (bien déterminé à multiplication près par un élément de \mathbf{C}^* , mais cette indétermination est sans importance) qui est fixe par $G(\mathbf{Z}_p)$. Alors le produit tensoriel restreint $\otimes_{v \in \mathcal{V}} \Pi_v$ relativement aux vecteurs x_p est engendré par des éléments de la forme $\otimes_{v \in \mathcal{V}} y_v$, avec $y_p = x_p$ pour presque tout p . On remarquera que $g \in G_{\mathbf{A}}$ agit naturellement sur un élément de ce type par la formule $g \cdot (\otimes_{v \in \mathcal{V}} y_v) = \otimes_{v \in \mathcal{V}} (g_v \cdot y_v)$, comme dans le cas du produit tensoriel de représentations de deux groupes, le point étant que $g_p \in G(\mathbf{Z}_p)$ pour presque tout p et donc que $g_p \cdot y_p = x_p$, pour presque tout p .

⁽⁴²⁾Ces fonctions L automorphes sont de vastes généralisations des fonctions L de Dirichlet.

2. Des formes modulaires aux représentations automorphes

2.1. Préliminaires

Si $D \in \mathbf{N}$, soit K^D le sous-groupe de $\mathbf{GL}_2(\widehat{\mathbf{Z}}) = \prod_{p \in \mathcal{P}} \mathbf{GL}_2(\mathbf{Z}_p)$ des $h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, avec $c \in D\widehat{\mathbf{Z}}$. Si χ est un caractère de Dirichlet modulo D , on fabrique un caractère linéaire $\tilde{\chi}$ de K^D , en posant $\tilde{\chi}(h) = \chi(\bar{d})$, où l'on a noté \bar{d} l'image de d dans $(\widehat{\mathbf{Z}}/D\widehat{\mathbf{Z}})^* = (\mathbf{Z}/D\mathbf{Z})^*$. On a alors $\tilde{\chi}(\gamma) = \chi(d)$, si $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(D)$.

Proposition D.3.3. — (i) Tout élément de $\mathbf{GL}_2(\mathbf{A})$ peut s'écrire sous la forme $\gamma g_\infty \kappa$, avec $\gamma \in \mathbf{GL}_2(\mathbf{Q})$, $g_\infty \in \mathbf{GL}_2(\mathbf{R})^+ = \{g \in \mathbf{GL}_2(\mathbf{R}), \det g > 0\}$, et $\kappa \in K^D$.

(ii) Cette écriture est unique à multiplication près de γ à droite par $\alpha \in \Gamma_0(D)$ et de g_∞ et κ à gauche par α^{-1} .

Démonstration. — Cf. alinéa 2.3.

2.2. La forme automorphe associée à une forme modulaire

Soit $f \in S_k(D, \chi)$. On peut utiliser la décomposition de $\mathbf{GL}_2(\mathbf{A})$ ci-dessus pour attacher à f une fonction ϕ_f sur $\mathbf{GL}_2(\mathbf{A})$, grâce à la formule,

$$\phi_f(x) = \tilde{\chi}(h)^{-1} \left(\frac{(a_\infty d_\infty - b_\infty c_\infty)^{1/2}}{c_\infty i + d_\infty} \right)^k f\left(\frac{a_\infty i + b_\infty}{c_\infty i + d_\infty}\right), \quad \text{si } x = \gamma g_\infty h \text{ et } g_\infty = \begin{pmatrix} a_\infty & b_\infty \\ c_\infty & d_\infty \end{pmatrix}.$$

Le (ii) de la prop. D.3.3 et le fait que $(cz + d)^{-k} f\left(\frac{az+b}{cz+d}\right) = \chi(d)f(z)$, si $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(D)$, montrent que ceci ne dépend pas de la décomposition de x choisie. Il n'est pas très difficile de vérifier que ϕ_f est une forme automorphe cuspidale.

On procède de même avec une forme de Maass f , de caractère χ et valeur propre λ pour $\Gamma_0(D)$ en définissant ϕ_f par la formule $\phi_f(x) = \tilde{\chi}(h)^{-1} f(g_\infty)$.

Dans les deux cas, on note Π_f le sous-espace de $\mathcal{A}_0(\mathbf{GL}_2(\mathbf{Q}) \backslash \mathbf{GL}_2(\mathbf{A}))$ engendré par ϕ_f sous l'action de $\mathbf{GL}_2(\mathbf{A})$. Le théorème ci-dessous peut être considéré comme le point de départ du programme de Langlands.

Théorème D.3.4. — (i) Si f est primitive, alors Π_f est une représentation automorphe cuspidale de degré 2, et donc admet une factorisation sous la forme $\Pi_f = \otimes_{v \in \mathcal{V}} \Pi_{f,v}$.

(ii) Si f est de niveau D , et si p ne divise pas D , la représentation $\Pi_{f,p}$ est complètement décrite par le facteur d'Euler $E_p(f, s)$ en p de la fonction $L(f, s)$, et $L(\Pi_{f,p}, s) = E_p(f, s)$.

Pour préciser le (ii), factorisons le facteur d'Euler $E_p(f, s)$ sous la forme

$$\begin{cases} E_p(f, s + \frac{k-1}{2}) = (1 - \alpha_p p^{-s})(1 - \beta_p p^{-s}) & \text{si } f \text{ est une forme modulaire de poids } k, \\ E_p(f, s) = (1 - \alpha_p p^{-s})(1 - \beta_p p^{-s}) & \text{si } f \text{ est une forme de Maass.} \end{cases}$$

Ceci permet de définir deux caractères μ_1, μ_2 de \mathbf{Q}_p^* par la formule $\mu_1(x) = \alpha_p^{-v_p(x)}$ et $\mu_2(x) = \beta_p^{-v_p(x)}$. La représentation $\Pi_{f,p}$ est alors la représentation $I(\mu_1, \mu_2)$ obtenue en

faisant agir $\mathbf{GL}_2(\mathbf{Q}_p)$, par translation à droite sur la variable, sur l'espace des fonctions ϕ localement constantes sur $\mathbf{GL}_2(\mathbf{Q}_p)$ telles que

$$\phi\left(\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} x\right) = \mu_1(a)\mu_2(d) \left|\frac{a}{d}\right|^{1/2} \phi(x), \quad \text{si } a, d \in \mathbf{Q}_p^*, b \in \mathbf{Q}_p, \text{ et } x \in \mathbf{GL}_2(\mathbf{Q}_p).$$

La construction ci-dessus n'est pas sans rappeler la construction des représentation induites pour les groupes finis (cf. n° 2 du § B.3) et, si on note B le sous-groupe de Borel de $\mathbf{GL}_2(\mathbf{Q}_p)$, (i.e. le sous-groupe des matrices triangulaires supérieures), alors $I(\mu_1, \mu_2)$ est l'induite localement constante à G du caractère linéaire $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mapsto \mu_1(a)\mu_2(d) \left|\frac{a}{d}\right|^{1/2}$ de B.

La représentation $\Pi_{f,p}$, pour p divisant D est plus subtile à décrire. Sa description rentre dans le cadre de la *correspondance de Langlands locale* (cf. n° 3).

2.3. La décomposition de $\mathbf{GL}_2(\mathbf{A})$

Lemme D.3.5. — Si K est un corps, et si $A \in \mathbf{SL}_2(K)$, il existe $t, x, y, z \in K$ tels que l'on ait $A = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}$.

Démonstration. — On part de la formule $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix} \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+xy & x+z+xyz \\ y & 1+yz \end{pmatrix}$. On en déduit que si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(K)$ vérifie $c \neq 0$, alors A est le produit des 3 matrices ci-dessus, avec $y = c$, $x = c^{-1}(a-1)$ et $z = c^{-1}(d-1)$ et on peut prendre $t = 0$. Si $c = 0$, il suffit de multiplier A par une matrice de la forme $\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}$, pour obtenir une nouvelle matrice qui peut s'écrire comme produit de 3 matrices comme ci-dessus. Ceci permet de conclure.

Lemme D.3.6. — Si S est un sous-ensemble fini de \mathcal{P} , alors $\mathbf{SL}_2(\mathbf{Z}[S^{-1}])$ est dense dans $\prod_{p \in S} \mathbf{SL}_2(\mathbf{Q}_p)$.

Démonstration. — Soit $A = (A_p)_{p \in S} \in \prod_{p \in S} \mathbf{SL}_2(\mathbf{Q}_p)$. D'après le lemme D.3.5, on peut écrire A_p sous la forme $A_p = \begin{pmatrix} 1 & 0 \\ t_p & 1 \end{pmatrix} \begin{pmatrix} 1 & x_p \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y_p & 1 \end{pmatrix} \begin{pmatrix} 1 & z_p \\ 0 & 1 \end{pmatrix}$, avec $t_p, x_p, y_p, z_p \in \mathbf{Q}_p$. Comme $\mathbf{Z}[S^{-1}]$ est dense dans $\prod_{p \in S} \mathbf{Q}_p$, on peut trouver des suites $(t'_n)_{n \in \mathbf{N}}$, $(x'_n)_{n \in \mathbf{N}}$, $(y'_n)_{n \in \mathbf{N}}$ et $(z'_n)_{n \in \mathbf{N}}$ d'éléments de $\mathbf{Z}[S^{-1}]$ tendant respectivement vers t_p , x_p , y_p et z_p , pour tout $p \in S$. Si on pose $A'_n = \begin{pmatrix} 1 & 0 \\ t'_n & 1 \end{pmatrix} \begin{pmatrix} 1 & x'_n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y'_n & 1 \end{pmatrix} \begin{pmatrix} 1 & z'_n \\ 0 & 1 \end{pmatrix}$, alors $(A'_n)_{n \in \mathbf{N}}$ est une suite d'éléments de $\mathbf{SL}_2(\mathbf{Z}[S^{-1}])$ tendant vers A dans $\prod_{p \in S} \mathbf{SL}_2(\mathbf{Q}_p)$. Ceci permet de conclure.

Venons-en à la démonstration de la prop. D.3.3. Soit $g \in \mathbf{GL}_2(\mathbf{A})$. Comme $\det g \in \mathbf{A}^*$, on peut écrire $\det g$ de manière unique sous la forme βb , où $b = (b_\infty, b^{|\infty|})$, avec $\beta \in \mathbf{Q}^*$, $b_\infty \in \mathbf{R}_+^*$ et $b^{|\infty|} \in \prod_{p \in \mathcal{P}} \mathbf{Z}_p^*$. Mais alors $h = \begin{pmatrix} \beta^{-1} & 0 \\ 0 & 1 \end{pmatrix} g \begin{pmatrix} b^{-1} & 0 \\ 0 & 1 \end{pmatrix} \in \mathbf{SL}_2(\mathbf{A})$. Soit S le sous-ensemble de \mathcal{P} constitués des p divisant D et des p tels que $h_p \notin \mathbf{SL}_2(\mathbf{Z}_p)$. D'après le lemme D.3.6, on peut trouver $\gamma_0 \in \mathbf{SL}_2(\mathbf{Z}[S^{-1}])$ aussi proche que l'on veut de $(h_p)_{p \in S}$; en particulier, on peut trouver $\gamma_0 \in \mathbf{SL}_2(\mathbf{Z}[S^{-1}])$ tel que $\gamma_0^{-1} h_p = \begin{pmatrix} a_p & b_p \\ c_p & d_p \end{pmatrix} \in \mathbf{SL}_2(\mathbf{Z}_p)$ quel que soit $p \in S$, et vérifie $v_p(c_p) \geq v_p(D)$, quel que soit p divisant D. Alors $\gamma = \begin{pmatrix} \beta & 0 \\ 0 & 1 \end{pmatrix} \gamma_0$, $g_\infty = \gamma_0^{-1} h_\infty \begin{pmatrix} b_\infty & 0 \\ 0 & 1 \end{pmatrix}$ et $\kappa = \gamma_0^{-1} h^{|\infty|} \begin{pmatrix} b^{|\infty|} & 0 \\ 0 & 1 \end{pmatrix}$ fournissent une décomposition de g ayant les propriétés voulues. Ceci démontre le (i).

Le (ii) suit juste du fait qu'un élément $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ dans l'intersection de $\mathbf{GL}_2(\mathbf{Q})$ et de $\mathbf{GL}_2(\mathbf{R})^+ K^D$ est à coefficients dans \mathbf{Z} ainsi que son inverse, de déterminant positif, et vérifie $v_p(c) \geq v_p(D)$ quel que soit $p \in \mathcal{P}$. Autrement dit, c'est un élément de $\Gamma_0(D)$.

3. Quelques autres aspects du programme de Langlands

Le programme de Langlands a plusieurs autres facettes. En particulier, on peut remplacer \mathbf{Q} par un corps de nombres F dans tout ce qu'on a fait ; l'anneau des adèles de F étant construit de la même manière, à partir de tous les complétés F_v de F .

- *La correspondance de Langlands locale.* Il s'agit d'une correspondance entre les représentations irréductibles de dimension n de $\mathcal{G}_{\mathbf{Q}_p}$ et certaines représentations irréductibles (les représentations cuspidales) de $\mathbf{GL}_n(\mathbf{Q}_p)$. Cette correspondance a finalement été établie, en toute généralité⁽⁴³⁾, par M. Harris (Paris 7) et R. Taylor (Harvard) en 1999, et simplifiée par G. Henniart (Orsay), toujours en 1999. Ceci fournit une description indirecte fort utile du groupe $\mathcal{G}_{\mathbf{Q}_p}$.

- *La correspondance de Langlands pour les corps de fonctions.* Fixons un nombre premier p . On peut considérer, au lieu de \mathbf{Q} et des corps de nombres, les extensions finies K du corps $\mathbf{F}_p(X)$ des fractions rationnelles à coefficients dans \mathbf{F}_p . L'anneau des adèles de K est défini à partir des complétés de $\mathbf{F}_p(X)$ pour les différentes normes que l'on peut mettre sur $\mathbf{F}_p(X)$; une différence avec le cas des corps de nombres est que ces complétés se ressemblent beaucoup : ils sont tous de la forme $\mathbf{F}_q((T))$, corps des fractions de l'anneau des séries formelles à coefficients dans \mathbf{F}_q , où \mathbf{F}_q est le corps fini à q éléments et q est une puissance de p . Dans ce cadre, la correspondance de Langlands a été établie par L. Lafforgue (I.H.E.S., Bures sur Yvette) en 1999, ce qui lui a valu la médaille Fields (2002). La démonstration utilise, entre autres, de puissants outils de géométrie algébrique introduits par A. Grothendieck dans les années 60 ; le cas $n = 2$ avait été établi par V. Drinfeld (médaillé Fields en 1990) dans les années 70, et la démonstration de L. Lafforgue est une vaste généralisation de celle de V. Drinfeld.

- *La fonctorialité de Langlands.* Du côté galoisien, on dispose de tas de constructions donnant de nouvelles représentations à partir de représentations connues. Par exemple, si $F \subset K$ sont deux corps de nombres, et si ρ est une représentation de \mathcal{G}_F (resp. de \mathcal{G}_K), on peut considérer la restriction (resp. l'induction) de ρ à \mathcal{G}_K (resp. à \mathcal{G}_F). Toutes ces constructions devraient avoir des analogues du côté automorphe, mais on est encore bien loin de comprendre vraiment ce qui se passe.

- *La correspondance de Langlands géométrique* (qui semble beaucoup intéresser l'armée américaine). Elle ne fait pas partie du programme initial de Langlands : c'est une extension (due à l'école russe autour de V. Drinfeld et A. Beilinson (maintenant à Chicago tous les deux)) dont la motivation vient de la physique mathématique en lien avec la théorie des cordes. Il s'agit d'une correspondance pour les corps de fonctions où on remplace le corps fini \mathbf{F}_p par le corps \mathbf{C} des nombres complexes.

⁽⁴³⁾De la même manière que l'on peut considérer un corps de nombres au lieu de \mathbf{Q} , on peut considérer un complété F_v d'un corps de nombres au lieu de \mathbf{Q}_p .

Index terminologique

- adhérence, 55
- anneau
 - intègre, 44
 - noethérien, 45
 - principal, 44
- bande de Moebius, 54, 68
- base
 - hilbertienne, 142–144, 146, 148, 191–193
 - orthonormale, 141
- boule
 - fermée, 50
 - ouverte, 50
- bouteille de Klein, 54
- caractère
 - conducteur, 265, 339, 340
 - cyclotomique, 323
 - Dirichlet, 115, 264–266, 268, 280, 289, 319, 323–325, 327, 335
 - Hecke, 329, 330
 - irréductible, 111, 294, 305
 - linéaire, 103, 115, 123, 184, 185, 188, 191, 193, 294, 297, 299, 335
 - orthogonalité, 104, 111, 113, 335, 339
 - primitif, 265
 - représentation, 103–105, 107, 111, 125–127
 - table, 117, 128, 302
 - unitaire continu, 181, 339, 340, 342, 344, 345
- catégorie, 12
- centralisateur, 21
- centre, 22, 293
- classe
 - d'équivalence, 13
 - de conjugaison, 21, 293, 300
 - formule des, 32
- coefficient de Fourier, 144, 149, 190, 192, 193
- compacité, 28, 57–60, 63, 77, 83, 132, 134–136, 146, 148, 152, 171, 175, 181, 209–211, 214–216, 218, 228, 229, 239, 242, 252, 253, 271, 272, 332–334, 337, 345
- complétion, 25, 26, 29, 71, 136, 204
- complétude, 26, 29, 69, 131–133, 136, 139, 141, 146, 149, 151, 163, 164, 167, 202, 204
- congruence, 15
- conjugaison, 21
- connexité, 65–67, 142, 209, 210, 226, 231, 236, 239, 242
 - composante connexe, 66, 239
 - par arcs, 67
- constante γ d'Euler, 253, 263, 275
- continuité, 51
 - uniforme, 52
- convergence
 - en moyenne, 159
 - en moyenne quadratique, 144, 164
 - normale, 81, 132
 - simple, 73, 136
 - simple p.p., 155
 - uniforme, 73, 82, 134
- coupures de Dedekind, 25
- courbe de Peano, 82
- critère de Cauchy, 9, 68, 252
 - uniforme, 74
- cycle, 33
- cylindre, 54, 68
- dénombrable, 9, 132, 138, 142, 153, 154
- demi-plan de Poincaré, 211, 226, 237, 270
- densité, 27, 55, 57, 71–73, 77, 78, 82, 84, 85, 134–138, 142, 143, 146–149, 154, 155, 164, 166, 167, 171, 186, 188, 189, 192, 198, 200–202, 204, 237, 332, 333, 348
- diagramme de Young, 297, 298
- distance, 50
 - équivalence, 50
 - p -adique, 26
 - triviale, 51
 - ultramétrique, 50
- distribution, 137, 166, 186, 196, 201, 257
- dollar (million de), 264, 309, 314
- domaine fondamental, 21, 193, 194
- endomorphisme, 38
 - diagonalisable, 39
 - trace, 39
- ensemble triadique de Cantor, 81, 84

- équivalence
 - classe, 13
 - quotient par une relation, 14
 - relation, 13
- espace
 - caractéristique, 40
 - de Banach, 131–134, 136–139, 141, 145, 148, 163, 166
 - de Hilbert, 140–142, 145, 148, 163, 166, 202
 - métrique, 50
 - métrisable, 50
 - propre, 39
- espace fonctionnel
 - \mathcal{C} , \mathcal{C}_b ou \mathcal{C}_c , 134, 135, 142, 186
 - \mathcal{C}^k ou \mathcal{C}^∞ , 136, 137
 - \mathcal{L}^1 ou L^1 , 55, 158, 159, 161–169, 171–173, 184–187, 196, 198, 200, 201, 208, 256
 - \mathcal{L}^2 ou L^2 , 55, 144, 148, 149, 163, 164, 166, 197–202, 208
 - L^p , 166
 - Schwartz \mathcal{S} , 194, 334–337, 339, 340, 342
 - Sobolev H^k , 136, 142, 201, 202
- facteur d’Euler, 259, 320, 324, 326, 328, 329, 347
- fermé, 48
 - de Zariski, 49
- fonction
 - centrale, 103, 110–112, 115, 126, 303, 305
 - continue, 51
 - de Bessel, 328
 - de carré sommable, 131, 137, 163, 164, 199–202, 208
 - elliptique, 243
 - η de Dedekind, 295
 - Γ d’Euler, 181, 183, 206, 219, 243, 253, 260, 264, 346
 - holomorphe, 2, 135, 207–210, 213, 214, 216–222, 226, 228, 231–242, 250, 252, 253, 257, 259, 260, 265, 267, 268, 270–273, 279, 281, 282, 284, 285, 289, 325, 326, 329, 340–342, 346
 - φ indicatrice d’Euler, 15
 - lipschitzienne, 52
 - méromorphe, 220, 235–237, 243, 250, 253, 259, 261–263, 268, 272, 279, 281, 282, 325, 329, 340, 342
 - mesurable, 155, 156, 159, 165, 168, 170, 174, 178–181, 219
 - sommable, 158–160, 162, 166, 172, 173, 181, 184–186, 196, 197, 200–202, 219, 240, 256, 280, 337–339, 341
 - τ de Ramanujan, 269
 - θ de Jacobi, 272, 273
 - uniformément continue, 52
 - \wp de Weierstrass, 244
- fonction L, 2, 319, 320
 - Artin, 324
 - automorphe, 346
 - Dirichlet, 251, 264, 265, 267, 277, 281, 324, 329, 331
 - forme modulaire, 271, 326, 328
 - Hecke, 329, 331
 - idéles, 342
- fonction zêta
 - de Dedekind, 329
 - de Hasse-Weil, 320, 322
 - de Riemann, 2, 249, 259, 260, 263, 264, 267, 272, 276–278, 281, 289–292
- forme
 - alternée, 212
 - automorphe, 345, 347
 - bilinéaire, 122, 123
 - différentielle, 212, 226
 - de Jordan, 38, 40, 43
 - linéaire, 18, 122, 137, 139, 145, 147–149, 152, 166, 212
 - de Maass, 327, 328, 346, 347
 - modulaire, 4, 100, 269–274, 295, 320, 326, 346, 347
 - parabolique, 326
 - primitive, 326, 346
 - quadratique, 4
 - sesquilinéaire, 163
- formule
 - $\frac{k}{12}$, 270, 271, 274
 - sommatoire d’Abel, 251, 252, 291
 - de Burnside, 114
 - de Cauchy, 29, 213, 214, 216, 219, 225, 229, 234, 238
 - des compléments, 254
 - des équerres, 297, 298
 - explicite, 282, 290, 291

- de réciprocity de Frobenius, 126, 127
- de Gauss, 183, 253
- d'inversion de Fourier, 115, 181, 198, 200, 201, 335, 336
- d'inversion de Moebius, 269
- de Jacobi, 274
- de Poisson, 194, 334, 337, 341
- du produit, 332, 333
- des résidus, 225, 232, 240, 242, 243, 262, 267, 272, 279, 282, 287
- de Stirling, 183, 206, 254, 255, 281
- de Stokes, 226
- de Taylor, 19, 135
- $\zeta(2) = \frac{\pi^2}{6}$, 2, 149, 190, 243
- groupe
 - abélien, 29
 - alterné, 36
 - cyclique, 30
 - distingué, 22
 - le monstre, 99, 322
 - orthogonal, 21
 - p -groupe, 37, 293, 294
 - p -Sylow, 37
 - simple, 23, 99
 - sous-groupe, 30
 - sporadique, 99
 - symétrique, 33, 295
 - symplectique, 21
 - unitaire, 21
- homéomorphisme, 51
- homothétie, 38
- idéal, 18
 - maximal, 44
 - premier, 44
 - principal, 44
- inégalité
 - Cauchy, 215
 - Cauchy-Schwarz, 140
 - triangulaire, 50
 - ultramétrique, 26
- intégrale
 - de Lebesgue, 151, 152, 156
 - de Riemann, 151, 152, 174
- intérieur, 55
- Jordan
 - bloc, 40
 - forme, 40
- limite
 - inférieure, 63
 - projective, 204
 - simple, 73
 - supérieure, 63
 - uniforme, 74
- mesurabilité, 155
- mesure
 - de Haar, 157
 - de Lebesgue, 156, 157
 - extérieure d'un ensemble, 153, 176
 - nulle (ensemble de), 153
- module, 40
 - de torsion, 41
 - de type fini, 41
 - engendré, 41
- nombre
 - algébrique, 10, 154, 318
 - complexe, 26
 - congruent, 309
 - duaux, 19
 - entier, 23
 - de Liouville, 155
 - p -adique, 26
 - premier, 7, 275–277, 327
 - réel, 25
 - rationnel, 24, 154
 - transcendant, 10, 154
- norme
 - équivalence, 76
 - espace vectoriel, 75
 - opérateur, 75
 - p -adique, 26
- orbite, 20, 293
- ordre
 - d'un élément, 31
 - d'un groupe, 37
- ouvert, 48
 - base, 49
- p -adique

- entiers, 27
- nombres, 26, 28, 323, 332
- norme, 72, 324, 331, 332
- transformée de Fourier, 335, 336
- transformée de Mellin, 339
- partie finie de Hadamard, 257
- partition
 - d'un ensemble, 13
 - d'un entier, 34, 295
- permutation, 20, 33
 - signature, 35
 - support, 33
- polynôme
 - caractéristique, 39
 - irréductible, 41
 - de Legendre, 148
 - minimal, 39
- produit scalaire, 140
- produit tensoriel, 112, 121, 123, 124, 191, 346
- projectif
 - droite, 20, 299
 - espace, 20
 - plan, 54, 311
- prolongement analytique, 209, 249–251, 255–257, 259, 266, 268, 271, 272, 320, 325, 326, 329, 340, 342, 344–346
- propre
 - espace, 39
 - valeur, 39
 - vecteur, 39
- propriété universelle, 11, 71
- réduction
 - des endomorphismes, 38
 - modulo D , 15
- série
 - de Dirichlet, 249–252, 255, 256, 258, 259, 264, 280, 289, 290, 319
 - d'Eisenstein, 271, 272, 317
 - entière, 203, 205–208, 210, 214, 215
 - formelle, 203, 204, 349
 - de Fourier, 142, 149, 181, 191, 195, 272, 326, 327
 - de Laurent, 235
 - de Taylor, 205, 214, 215
- sommabilité
 - fonction, 158
 - série, 132
- somme directe
 - d'espaces vectoriels, 12
 - de groupes, 12
- spectre, 39
- stabilisateur, 21
- suite
 - convergente, 56
 - de Cauchy, 68
 - extraite, 56
 - valeur d'adhérence, 58
- tableau de Young, 297, 298
- tipi de Cantor, 84
- topologie, 27, 48, 331
 - algébrique, 225, 322
 - discrète, 49, 54, 322, 331
 - espace topologique, 49
 - grossière, 49, 54
 - induite, 53
 - Krull, 322
 - produit, 53, 77, 339
 - produit restreint, 332, 333
 - quotient, 53
 - séparée, 54, 55, 57, 162, 163, 166
 - totalement discontinue, 27, 66
 - Zariski, 49, 54
- tore, 54
- transformée de Fourier, 181
 - adélique, 336
 - dans $\mathcal{S}(\mathbf{R})$, 194–196, 337, 343
 - dans $L^1(\mathbf{R}^m)$, 184, 186, 196, 198, 200, 201, 208
 - dans $L^2(\mathbf{R}^m)$, 198, 200, 201, 208
 - groupes finis, 115
 - p -adique, 335, 336
- transformée de Mellin
 - adélique, 340
 - p -adique, 339
 - sur \mathbf{R} , 256, 340
- transposition, 34
- valuation p -adique, 8
- voisinage, 49
 - base, 49

Conjectures et théorèmes

- application ouverte, 221
- Artin (conjecture d'), 325, 326, 328, 346
- Artin (théorème d'), 127, 325
- Bézout (théorème de), 7, 46
- Baire (lemme de), 70, 71, 79, 85, 137, 138, 154
- Banach-Steinaus (théorème de), 137, 138
- Banach-Tarski (paradoxe de), 156
- Bateman-Horn (conjecture de), 276, 277
- Beilinson (conjecture de), 2
- Birch et Swinnerton-Dyer (conjecture de), 309
- Bloch-Kato (conjecture de), 2
- Borel-Cantelli (théorème de), 154, 178
- Borel-Lebesgue (théorème de), 59
- Brauer (théorème de), 128, 299, 325
- Cayley-Hamilton (théorème de), 39
- Coates-Wiles (théorème de), 315, 316
- continuité d'une intégrale dépendant d'un paramètre, 181, 199, 239
- convergence dominée, 151, 160, 162, 165, 169, 181, 182, 200, 214, 229, 258, 290
- convergence dominée pour les séries, 161
- convergence monotone, 159, 161, 169, 178, 180
- dérivation sous le signe somme, 182, 183, 187
- Deligne (conjecture de), 2
- Fatou (lemme de), 159
- fermés emboîtés, 70, 71
- Fermat
 - grand théorème de, 100, 260, 271, 321, 327
 - nombres premiers de la forme $4n + 1$, 278, 320
 - non congruence de 1, 310
 - petit théorème de, 32
- Fischer-Riesz (théorème de), 164
- Fubini (sur $\mathbf{N} \times \mathbf{X}$), 161
- Fubini (théorème de), 151, 167, 169, 196, 199, 219
- Gauss (lemme de), 7, 45
- graphe fermé, 139
- GRH (conjecture), 268
- Hahn-Banach (théorème de), 139
- Hasse (théorème de), 313
- Hasse-Weil (conjecture de), 320
- Hecke (théorème de), 329
- hypothèse de Lindelöf, 292
- hypothèse de Riemann, 16, 264, 268, 290–292, 323
- image ouverte, 138
- inversion globale, 222
- inversion locale pour les fonctions holomorphes, 220–222
- Jordan (théorème de), 238
- Kronecker-Weber (théorème de), 323, 324, 330, 344
- Lagrange
 - l'ordre d'un sous-groupe divise celui du groupe, 32
 - théorème des 4 carrés, 273, 317
- Landau (théorème de), 250, 289
- Liouville (théorème de), 215
- loi de réciprocité quadratique, 17, 313, 319, 324
- Mertens (conjecture de), 291
- Mordell-Weil (théorème de), 312
- Morera (théorème de), 232
- nombres premiers, 275–277, 280, 288, 290
- Ostrowski (théorème d'), 331, 332
- Picard (théorème de), 237
- point fixe, 70
- problème d'Erdős, 277
- problème de Bâle, 2
- problème de Kakeya, 154
- problème des nombres congruents, 309
- problème inverse de Galois, 322
- problèmes de Hilbert, 312, 330
- programme de Langlands, 251, 319, 320, 325, 327, 344, 347, 349
- progression arithmétique, 115, 275, 277, 279, 290

projection sur un convexe fermé, 146
Pythagore (théorème de), 140

Ramanujan (conjecture de), 269
Ramanujan-Petersson (conjecture de), 269
représentation conforme, 226
restes chinois, 16
Riemann-Lebesgue (théorème de), 185
Riesz (théorème de), 77, 140, 145, 166
Rouché (théorème de), 243

Serre (conjecture de), 317, 327
Stone-Weierstrass (théorème de), 134, 142
structure des groupes abéliens finis, 31, 117
structure des modules de torsion sur un anneau
 principal, 38, 47
Sylow (théorème de), 37

Taniyama-Weil (conjecture de), 320
théorème fondamental de l'algèbre, 26, 208, 210,
 215, 242
théorème fondamental de l'analyse, 79, 160, 226
théorème fondamental de l'arithmétique, 8

valeurs intermédiaires, 66

Wedderburn (théorème de), 32
Weil (conjectures de), 269, 320

zéros isolés, 209, 236

Index des noms propres

- Apéry R., 154
 Artin E., 127, 321, 325, 327, 330, *see* conjecture, théorème, fonction L
 Baire R., 1, 138, *see* théorème
 Banach S., 1, 78, 131, 138, 139, 156, *see* espace, théorème (Banach-Steinaus, Hahn-Banach, image ouverte), paradoxe de Banach-Tarski
 Beilinson A., 2, 349
 Besicovitch A., 154
 Birch B., 312
 Bloch S., 2
 Bombieri F., 292
 Borchers R., 100
 Bourgain J., 277
 Brauer R., 128, 325
 Breuil C., 314, 320
 Burnside W., 1
 Cantor G., 5, 25, 26
 Carleson L., 149
 Cauchy A., 1, 19, 37, 138, 209, *see* critère, formule, formule des résidus
 Chevalley C., 99, 330
 Coates J., 315
 Conrad B., 314, 320
 Dedekind R., 23, 25, 26, 329
 Deligne P., 2, 269, 320
 Diamond F., 314, 320
 Dirichlet G., 2, 115, 249, 277, *see* caractère, fonction L, théorème de la progression arithmétique, série
 Drinfeld V., 318, 349
 Dwork B., 320
 Elkies N., 315
 Erdős P., 277
 Euler L., 2, 17, 205, 209, 217, 222, 275, 313, *see* constante d'Euler, facteur d'Euler, fonction Γ
 de Fermat P., 310
 de Fermat P., 100, 260, 271, 278, 320
 Fibonacci L., 310
 Fields (médaille), 100, 144, 155, 166, 269, 277, 320, 349
 Fischer E., 131
 Fontaine J.-M., 5, 29
 Fréchet M., 1
 Friedlander J., 278
 Frobenius F., 1, 111, 126
 Furtwängler P., 330
 Gauss C.-F., 6, 17, 313
 Godement R., 4, 345
 Goldfeld D., 314
 Goldston D., 277
 Gowers T., 144, 277
 Green B., 277
 Griess R., 100
 Gross B., 315
 Grothendieck A., 54, 270, 320, 322, 349
 Hadamard J., 2, 276
 Hahn H., 1, 131, 139
 Harris M., 349
 Hasse H., 313, 330
 Hausdorff F., 48, 251, 292
 Hecke E., 327, 329, 341, *see* caractère, fonction L
 Henniart G., 349
 Hensel K., 26, 332
 Hilbert D., 1, 131, *see* espace, conjecture, 312
 Iwaniec H., 278, 292
 Jacobi C., 273, *see* fonction θ , formule, théorème des 4 carrés
 Jacquet H., 345
 Kato K., 2
 Katz N., 154
 Khare K., 327
 Kisin M., 327
 Kolmogorov A., 149
 Kolyvagin V., 315

- Kronecker L., 4, 31, 323, 330
 Lafforgue L., 349
 Langlands R., 318, 325, 327, 328, 345, 346
 Lebesgue H., 1, 174, *see* intégrale, théorème (Borel-Lebesgue, Riemann-Lebesgue, convergence dominée)
 Lindenstrauss J., 145
 Maass H., 327
 Manin Y., 318
 Matiyasevich Y., 312
 McKay J., 100
 Mertens F., 291
 Mordell L., 269, 312
 de Morgan A., 5
 Odlysko A., 292
 Ostrowski A., 331
 Peano G., 23, 82
 Plancherel M., 1
 Poincaré H., 1, 51, 214, 312
 Ramanujan S., 269
 Ribet K., 317, 320
 de Riele H., 292
 Riemann B., 2, 264, 276, *see* conjecture, fonction zêta, intégrale, théorème
 Riesz F., 1, 77, 131, 140, 145, 163, 166
 Rivoal T., 154
 Roth K., 155, 277
 Rutherford E., 100
 Schur I., 1, 109, 298
 Schwartz L., 166
 Serre J-P., 4, 54, 317, 327
 Shafarevich I., 314, 322
 Shimura G., 313, 317
 Solovay R., 156
 Steinhaus H., 1, 131
 Swinnerton-Dyer H., 312
 Sylow L., 37
 Takagi T., 330
 Taniyama Y., 317, 320
 Tao T., 154, 277
 Tarski A., 156
 Tate J., 29, 314, 331, 345
 Taylor R., 314, 320, 349
 Tunnell J., 310, 315, 327, 328
 Tzafriri L., 145
 de la Vallée Poussin C., 2, 276, 277
 Waldspurger J.-L., 318
 Weber H., 330
 Weierstrass K., 1, 23, 78, 214
 Weil A., 4, 54, 313, 317, 320, 330, *see* conjecture, fonction zêta
 Wiles A., 100, 271, 314, 315, 320, 327
 Wintenberger J-P., 327
 Yildirim C., 277
 Zagier D., 310, 315
 Zudilin W., 154