

MASTER DE SCIENCES ET TECHNOLOGIES  
MENTION PHYSIQUE FONDAMENTALE ET APPLICATIONS



# Symétries en Physique

Module MU4PY215

Samuel WALLON



version du 12 mars 2020



# Table des matières

<b>I</b>	<b>Des symétries géométriques à la théorie des groupes</b>	<b>3</b>
1	Les symétries : qu'ès aquò ?	5
<b>2</b>	<b>Symétries géométriques</b>	<b>11</b>
2.1	Quelques rappels de géométrie affine . . . . .	11
2.1.1	Espace affine . . . . .	11
2.1.2	Applications affines . . . . .	12
2.2	Isométries vectorielles . . . . .	14
2.2.1	Linéarité des isométries vectorielles . . . . .	14
2.2.2	Décomposition en produit de réflexions . . . . .	15
2.2.3	Isométries vectorielles en dimensions 0, 1 et 2 . . . . .	15
2.2.4	Isométries vectorielles de $\mathbb{R}^n$ . . . . .	16
2.3	Structure du groupe $\text{Isom}(\mathbb{R}^n)$ . . . . .	19
2.3.1	Décomposition en produit de réflexions . . . . .	19
2.3.2	Structure de produit semi-direct de $\text{Isom}(\mathbb{R}^n)$ . . . . .	19
2.4	Isométries affines . . . . .	19
2.4.1	Les isométries propres (ou déplacements) . . . . .	19
2.4.2	Les isométries impropres . . . . .	20
2.5	Classification des isométries affines pour les petites dimensions . . . . .	22
2.5.1	Dimensions 0 et 1 . . . . .	22
2.5.2	Dimension 2 . . . . .	23
2.5.3	Dimension 3 . . . . .	25
2.6	Quelques résultats sur la composition des isométries . . . . .	26
2.6.1	Composition de deux rotations de même centre (2D) ou d'axes concourants (3D)	26
2.6.2	Composition de deux réflexions . . . . .	26
<b>3</b>	<b>Symétries et topologie</b>	<b>27</b>
3.1	Introduction . . . . .	27
3.2	Caractéristique d'Euler-Descartes . . . . .	28
3.2.1	Surface bidimensionnelles de genre nul . . . . .	28
3.2.2	Les polyèdres réguliers convexes . . . . .	30
3.2.3	Caractéristique d'Euler-Descartes de quelques objets élémentaires . . . . .	31
3.2.4	Surfaces bidimensionnelles orientables de genre quelconque . . . . .	32
3.2.5	Surfaces bidimensionnelles non orientables de genre quelconque . . . . .	35
3.2.6	Quelques éléments de géométrie projective . . . . .	36

3.2.7	Classification des surfaces fermées . . . . .	40
3.3	Les polyèdres réguliers non convexes . . . . .	41
3.4	Caractéristique d'Euler-Poincaré . . . . .	41
3.4.1	Polytopes . . . . .	42
3.4.2	Simplexes . . . . .	43
<b>4</b>	<b>Quelques éléments de théorie des groupes</b>	<b>45</b>
4.1	Groupe . . . . .	45
4.1.1	Définitions . . . . .	45
4.1.2	Exemples de groupes . . . . .	46
4.1.3	Groupes finis et table de Cayley . . . . .	47
4.2	Morphismes de groupe, sous-groupe . . . . .	48
4.2.1	Morphisme . . . . .	48
4.2.2	Sous-groupe . . . . .	49
4.2.3	Quelques sous-groupes particuliers . . . . .	50
4.2.4	Exemples de morphismes et sous-groupes associés . . . . .	52
4.3	Générateurs d'un groupe, groupe cyclique . . . . .	52
4.4	Classes, sous-groupes distingués, groupes quotient . . . . .	54
4.4.1	Automorphismes intérieurs . . . . .	54
4.4.2	Classes de conjugaison d'un groupe . . . . .	55
4.4.3	Classes par rapport à un sous-groupe . . . . .	55
4.4.4	Sous-groupe distingué . . . . .	57
4.4.5	Groupe quotient . . . . .	58
4.4.6	Exemples . . . . .	59
4.5	Suite exacte* . . . . .	59
4.5.1	Exemples . . . . .	59
4.5.2	Suites exactes courtes . . . . .	60
4.5.3	Extension de groupes . . . . .	60
4.6	Groupe opérant sur un ensemble . . . . .	61
4.7	Produit semi-direct . . . . .	63
4.7.1	Exemples . . . . .	64
4.8	Le groupe symétrique $\mathcal{S}_n$ . . . . .	65
4.8.1	Notions élémentaires sur les permutations . . . . .	66
4.8.2	Décomposition en produits de cycles . . . . .	67
4.8.3	Décomposition en produits de transpositions . . . . .	69
4.8.4	Signature . . . . .	69
4.8.5	Théorème de Cayley . . . . .	71
4.9	Présentation d'un groupe . . . . .	73
4.9.1	Groupe libre . . . . .	73
4.9.2	Exemples . . . . .	73
4.9.3	Présentation . . . . .	73

<b>II</b>	<b>Symétries discrètes et cristallographie</b>	<b>77</b>
<b>5</b>	<b>Introduction à la cristallographie</b>	<b>79</b>
5.1	L'état cristallin . . . . .	79
5.1.1	Les cristaux parfaits . . . . .	79
5.1.2	Les défauts dans les cristaux . . . . .	80
5.2	Qu'est-ce qu'un cristal du point de vue mathématique? . . . . .	81
5.2.1	Réseau . . . . .	81
5.2.2	Maille primitive . . . . .	82
5.2.3	Quelques choix possibles de maille primitive . . . . .	83
<b>6</b>	<b>Groupe ponctuel à deux dimensions</b>	<b>85</b>
6.1	Définitions et propriétés principales . . . . .	85
6.2	Les 10 groupes ponctuels bidimensionnels . . . . .	87
6.3	Les 5 types de réseau . . . . .	90
6.3.1	Groupes $C_1, C_2$ : les réseaux parallélogrammes . . . . .	90
6.3.2	Groupes $C_4, D_4$ : les réseaux carrés . . . . .	91
6.3.3	Groupes $C_3, D_3, C_6, D_6$ : réseaux hexagonaux . . . . .	92
6.3.4	Groupes $D_1, D_2$ : réseaux rectangulaires ou orthorhombiques . . . . .	94
<b>7</b>	<b>Classification des 17 groupes du papier peint</b>	<b>97</b>
7.1	Les 17 groupes du papier peint et leur nomenclature . . . . .	97
7.2	Des groupes ponctuels aux groupes du papier peint : le problème de l'extension . . . . .	98
7.3	La théorie des groupes en marche : construction explicite des isométries du réseau . . . . .	99
7.4	Mise en oeuvre : des 10 groupes ponctuels aux 17 groupes du papier peint . . . . .	100
7.4.1	$G = 0 = D_{1,p}$ . . . . .	100
7.4.2	$G = 0 = D_{1,c}$ . . . . .	101
7.4.3	$G = 0 = D_{3,\ell}$ ou $D_{3,s}$ . . . . .	102
7.4.4	$G = 0 = D_6$ . . . . .	103
7.4.5	$G = 0 = D_4$ . . . . .	104
7.4.6	$G = 0 = D_2$ . . . . .	105
7.5	Quelques compléments sur les centres et axes de symétries des 17 pavages . . . . .	107
<b>8</b>	<b>Quelques éléments sur la cristallographie à 3 dimensions</b>	<b>109</b>
8.1	Les 32 groupes ponctuels . . . . .	109
8.1.1	Notation de Schönflies . . . . .	109
8.1.2	Notation d'Hermann- Mauguin . . . . .	110
8.2	Des 32 groupes ponctuels aux 230 groupes d'espace . . . . .	111
8.2.1	Les 7 systèmes cristallins . . . . .	111
8.2.2	Les 7 systèmes réticulaires . . . . .	112
8.2.3	Les 6 familles cristallines . . . . .	113
8.2.4	Les 14 réseaux de Bravais . . . . .	114

<b>III</b>	<b>Symétries dynamiques en physique classique</b>	<b>115</b>
<b>9</b>	<b>Mécanique analytique</b>	<b>117</b>
9.1	Principe de moindre action . . . . .	117
9.1.1	Equations d'Euler-Lagrange . . . . .	117
9.1.2	Principe de relativité galiléen : . . . . .	118
9.1.3	Fonction de lagrange pour un système de points matériels . . . . .	119
9.2	Intégrales du mouvement . . . . .	119
9.3	Equations canoniques . . . . .	121
9.3.1	Equations de Hamilton . . . . .	121
9.3.2	Fonction de Routh . . . . .	122
9.3.3	Crochets de Poisson . . . . .	123
9.4	L'action en fonction des coordonnées . . . . .	124
9.4.1	Variation de l'action pour deux trajectoires réelles voisines . . . . .	124
9.4.2	Théorème de Noether . . . . .	127
9.4.3	Principe de Maupertuis . . . . .	128
9.5	Transformations canoniques . . . . .	131
9.5.1	Fonction génératrice d'une transformation canonique . . . . .	131
9.5.2	Condition en terme de crochets de Poisson . . . . .	132
9.5.3	Théorème de Liouville . . . . .	133
<b>10</b>	<b>Théorie classique des champs</b>	<b>135</b>
10.1	Introduction . . . . .	135
10.2	Formulation lagrangienne . . . . .	135
10.3	Equations d'Euler-Lagrange . . . . .	136
10.4	Les symétries globales . . . . .	136
10.4.1	Seconde variation de l'action . . . . .	136
10.4.2	Théorème de Noether . . . . .	138
10.4.3	Applications . . . . .	138
<b>11</b>	<b>Symétries en Mécanique Quantique</b>	<b>143</b>
11.1	Introduction . . . . .	143
11.2	Description des états . . . . .	144
11.3	Lois de symétrie . . . . .	144
11.4	Théorème de Wigner . . . . .	146
11.5	Conséquence du théorème de Wigner . . . . .	147
<b>IV</b>	<b>Symétries continues</b>	<b>149</b>
<b>12</b>	<b>Groupes et algèbres de Lie</b>	<b>151</b>
12.1	Généralités . . . . .	151
12.1.1	Groupe continu, groupe de Lie, algèbre de Lie . . . . .	151
12.1.2	Générateurs . . . . .	152
12.1.3	Application exponentielle . . . . .	152
12.1.4	Algèbre de Lie . . . . .	153

12.1.5	Quelques définitions utiles . . . . .	154
12.2	Quelques résultats sur les groupes et algèbres de Lie . . . . .	156
12.2.1	Lien entre groupe de Lie et algèbre de Lie correspondante . . . . .	156
12.2.2	Structure des algèbres de Lie et groupes de Lie ; caractérisation de Cartan . . . . .	158
12.2.3	Algèbre de Lie compacte . . . . .	162
12.2.4	Classification . . . . .	168
<b>13</b>	<b>Représentations</b>	<b>173</b>
13.1	Généralités . . . . .	173
13.1.1	Définition, premiers concepts . . . . .	173
13.1.2	Représentations réductibles et irréductibles . . . . .	176
13.1.3	Lemme de Schur . . . . .	180
13.2	Produit tensoriel de représentations . . . . .	182
13.2.1	Produit tensoriel de représentations . . . . .	182
13.2.2	Décomposition et coefficients de Clebsch-Gordan . . . . .	183
13.3	Représentation régulière . . . . .	186
13.3.1	Généralités . . . . .	186
13.3.2	Représentation régulière d'un groupe fini . . . . .	187
13.4	Représentations des groupes finis . . . . .	187
13.4.1	Orthogonalité et complétude . . . . .	187
13.4.2	Applications . . . . .	191
13.5	Représentation des algèbres et groupes de Lie . . . . .	193
13.5.1	Représentation unitaires d'un groupe de Lie . . . . .	193
13.5.2	Lien entre représentation des algèbres et groupes de Lie . . . . .	194
13.5.3	Exemples importants de représentations . . . . .	195
13.5.4	Normalisation . . . . .	198
<b>14</b>	<b>Les Groupes <math>SO(3)</math> et <math>SU(2)</math></b>	<b>199</b>
14.1	Le groupe des rotations à 3 dimensions . . . . .	199
14.1.1	Définition . . . . .	199
14.1.2	Forme explicite . . . . .	199
14.1.3	Topologie . . . . .	200
14.1.4	Algèbre de Lie de $SO(3)$ . . . . .	200
14.2	Groupe $SU(2)$ : définition et paramétrisation . . . . .	200
14.2.1	Définition . . . . .	200
14.2.2	Caractérisation . . . . .	201
14.2.3	Topologie . . . . .	203
14.3	Morphisme du groupe $SU(2)$ sur le groupe $SO(3)$ . . . . .	206
14.3.1	Morphisme . . . . .	206
14.3.2	Caractérisation . . . . .	207
14.3.3	Formules d'inversion . . . . .	208
14.3.4	Eléments conjugués . . . . .	208
14.3.5	Forme infinitésimale . . . . .	209
14.4	Représentations de $SU(2)$ . . . . .	209
14.4.1	Générateurs dans une représentation quelconque . . . . .	209

14.4.2	Exemples de représentation de $SU(2)$ . . . . .	211
14.5	Représentations unitaires irréductibles de $SU(2)$ . . . . .	213
14.5.1	Spectre . . . . .	214
14.5.2	Vecteurs propres de $J^2$ et $J_3$ et construction des sous-espaces invariants $\mathcal{E}^{(j)}$	216
14.6	Angles d'Euler . . . . .	217
14.6.1	Définition et construction . . . . .	217
14.6.2	Expression explicite . . . . .	219
14.6.3	Propriétés de symétrie de la paramétrisation d'Euler . . . . .	219
14.7	Définition des matrices de rotation . . . . .	219
14.8	Transformation des états et des observables sous une rotation . . . . .	221
14.8.1	Etats et observables . . . . .	221
14.8.2	Notion d'opérateur tensoriel . . . . .	222
14.8.3	Exemples . . . . .	222
14.9	Représentation à une phase près de $SO(3)$ . . . . .	225
<b>15</b>	<b>Représentations spinorielles de <math>SU(2)</math></b>	<b>227</b>
15.1	Spineurs . . . . .	227
15.1.1	Action contravariante . . . . .	227
15.1.2	Métrique . . . . .	227
15.2	Calcul tensoriel . . . . .	229
15.2.1	Spineur covariant . . . . .	229
15.2.2	Action d'une rotation sur un spineur covariant . . . . .	230
15.2.3	Equivalence des représentations $U$ et $\bar{U}$ . . . . .	231
15.2.4	Norme sur les spineurs . . . . .	231
15.3	Tenseurs spinoriels . . . . .	232
15.3.1	Tenseurs de rang quelconque . . . . .	232
15.3.2	Tenseurs antisymétriques . . . . .	233
15.3.3	Norme sur les tenseurs spinoriels . . . . .	233
15.3.4	Opérations sur les tenseurs spinoriels . . . . .	233
15.4	Représentations irréductibles de $SU(2)$ . . . . .	234
15.4.1	Lien avec les tenseurs spinoriels . . . . .	234
15.4.2	Lien avec les représentations $D^j$ . . . . .	236
15.4.3	Retour sur le morphisme $SU(2) \rightarrow SO(3)$ . . . . .	237
<b>16</b>	<b>Notion de groupe de recouvrement</b>	<b>239</b>
16.1	Homotopie . . . . .	239
16.1.1	Chemins . . . . .	239
16.1.2	Groupe d'homotopie (ou groupe fondamental) . . . . .	240
16.1.3	Topologie des groupes $SU(2)$ et $SO(3)$ . . . . .	241
16.2	Groupe de recouvrement . . . . .	242
<b>V</b>	<b>Symétries relativistes</b>	<b>245</b>
<b>17</b>	<b>Le groupe de Lorentz</b>	<b>247</b>
17.1	Intervalle d'espace-temps . . . . .	247



17.2	Groupe de Poincaré et groupe de Lorentz . . . . .	249
17.2.1	Caractérisation de $\mathcal{P}$ . . . . .	249
17.2.2	Structure du groupe de Lorentz . . . . .	251
17.2.3	Vecteurs contravariants et covariants . . . . .	253
17.2.4	Transformation d'un vecteur covariant . . . . .	254
17.2.5	Calcul tensoriel . . . . .	255
17.3	Algèbre de Lie du groupe de Lorentz restreint $L_+^\uparrow$ . . . . .	256
17.3.1	Forme non covariante de générateurs . . . . .	256
17.3.2	Algèbre de Lie . . . . .	258
17.3.3	Transformation de Lorentz pure . . . . .	259
17.3.4	Rotation . . . . .	261
17.3.5	Forme covariante de l'algèbre de Lie de $L_+^\uparrow$ . . . . .	262
17.3.6	Représentation sur les fonctions scalaires (représentation appelée régulière) . . . . .	264
17.3.7	Forme réelle compacte . . . . .	265
<b>18</b>	<b>Représentations spinorielles du groupe de Lorentz</b>	<b>269</b>
18.1	Introduction . . . . .	269
18.2	Homomorphismes du groupe $SL(2, \mathbb{C})$ sur le groupe $L_+^\uparrow$ . . . . .	269
18.2.1	Préliminaires . . . . .	269
18.2.2	Théorème . . . . .	270
18.3	Représentations spinorielles de $L_+^\uparrow$ . . . . .	272
18.3.1	Identification des représentations spinorielles . . . . .	272
18.3.2	Représentations de $SL(2, \mathbb{C})$ . . . . .	274
18.4	Spineurs quadridimensionnels . . . . .	275
18.4.1	Spineurs contravariants et covariants . . . . .	275
18.4.2	Produit scalaire invariant . . . . .	276
18.4.3	Spineurs d'ordre supérieur . . . . .	276
18.4.4	Représentation irréductibles de $SL(2, \mathbb{C})$ . . . . .	277
18.4.5	Application : lien entre spineurs mixtes et quadrivecteurs . . . . .	278
18.5	Inversion d'espace et bispineurs . . . . .	279
<b>19</b>	<b>Le groupe de Poincaré</b>	<b>283</b>
19.1	Structure . . . . .	283
19.2	Représentations . . . . .	284
<b>20</b>	<b>L'équation de Klein-Gordon</b>	<b>285</b>
20.1	Construction . . . . .	285
20.2	Solutions . . . . .	286
<b>21</b>	<b>L'équation de Dirac</b>	<b>287</b>
21.1	L'équation de Dirac dans la représentation spinorielle . . . . .	287
21.2	Matrices $\gamma$ . . . . .	289
21.3	Propriétés de l'équation de Dirac . . . . .	292
21.3.1	Equation conjuguée . . . . .	292
21.3.2	Le courant . . . . .	293
21.3.3	L'Hamiltonien . . . . .	293

21.3.4	La représentation standard (ou représentation de Dirac)	294
21.3.5	Le couplage avec le champ électromagnétique	295
21.3.6	L'invariance relativiste	295
21.3.7	Les constantes du mouvement	298
21.3.8	Lois de transformations de quantités bilinéaires sous $\mathbf{L}_+^\uparrow$	300
21.4	Les solutions ondes planes	304
21.5	Spin et hélicité	309
21.5.1	quadrivecteur de Pauli Lubanski	309
21.5.2	Spin et hélicité	310
21.6	La conjugaison de charge	313
21.7	Lagrangien et propagateur	315
<b>VI</b>	<b>Les symétries en action</b>	<b>317</b>
<b>22</b>	<b>Théories de Yang-Mills</b>	<b>319</b>
22.1	Théorie de jauge abélienne	319
22.2	Vade-mecum de théorie des groupes	322
22.2.1	Algèbre de Lie	322
22.2.2	Représentations	322
22.2.3	Compacité, simplicité et semi-simplicité	327
22.3	Théories de jauge non abéliennes	328
22.3.1	Principe de jauge et couplage matière-champ de jauge	328
22.3.2	Dynamique du champ de jauge	330
	<b>Appendices</b>	<b>334</b>
<b>A</b>	<b>Abréviations</b>	<b>337</b>
<b>B</b>	<b>Quelques éléments de topologie</b>	<b>339</b>
B.1	Les concepts de base de la topologie	339
B.1.1	Topologie générale	339
B.1.2	Topologie des espaces métriques	341
B.1.3	Topologie des espaces vectoriels normés	343
B.2	Limite, continuité	344
B.2.1	Limite	344
B.2.2	Continuité	344
B.3	Espace métrique complet	346
B.3.1	Suites de Cauchy	346
B.3.2	Espace complet	346
B.4	Compacité	347
B.4.1	Compacité en topologie générale	347
B.4.2	Compacité locale	348
B.4.3	Quelques résultats concernant les espaces vectoriels normés	349
B.4.4	Compactifié d'Alexandroff	350

B.5	Connexité . . . . .	350
<b>C</b>	<b>Petite introduction aux variétés</b>	<b>353</b>
C.1	Variété topologique . . . . .	353
C.2	Variété différentielle . . . . .	354
C.3	Immersion, plongement, submersion . . . . .	355
C.3.1	Immersion . . . . .	355
C.3.2	Plongement . . . . .	355
C.3.3	Submersion . . . . .	356
C.4	Connexion affine . . . . .	356
C.5	Application exponentielle . . . . .	356
	<b>References</b>	<b>359</b>
	<b>Figures</b>	<b>359</b>
	<b>Tables</b>	<b>360</b>
	<b>Bibliographie</b>	<b>363</b>



Les chapitres ou parties dont le titre est suivi du symbole \* peuvent être sautés en première lecture.



## Première partie

# Des symétries géométriques à la théorie des groupes





# Chapitre 1

## Les symétries : qu'ès aquò ?

La notion de symétrie nous paraît au premier abord familière, presque intuitive. Elle est multiforme, aussi bien présente dans la nature que dans toute activité humaine (arts, sports, sciences...). Avant de considérer quelques exemples de symétries, il semble important de fixer quelques termes de vocabulaire, afin préciser le champ d'étude. En particulier, il nous faut distinguer la notion de transformation et de symétrie.

### **Transformations :**

On demandera à toutes les transformations qui agissent sur un système d'associer à chaque point de l'espace considéré un et un seul point du même espace. La transformation sera donc inversible. Ainsi, on exclura les projections sur un sous-espace, qui ne sont pas inversibles.

### **Symétries :**

Les symétries d'un système seront quand à elles les transformations qui laissent invariantes certaines propriétés de ce système. Les propriétés considérées pourront être très variées, qu'elles concernent la géométrie de ce système (dimensions, angles, volume, etc.) ou ses propriétés physiques (dynamique, structure interne, etc.).

Ces transformations et ces symétries peuvent être composées, de sorte que la structure mathématique naturelle sous-jacente est la notion de *groupe*, nous y reviendrons en détail au chap. 4.

Selon le système considéré, et selon l'échelle des phénomènes étudiés, l'espace sur lequel les transformations ou symétries précédentes vont agir avoir une structure mathématique très variable.

Le premier espace naturel est bien sûr l'espace géométrique à trois dimensions dans lequel nous vivons, et que nous étudierons en détail au chap. 2. Les symétries correspondantes font intimement partie de notre environnement. Présentes dans la nature sous toutes ses formes, elles ont également été exploitées depuis l'aube de l'humanité, dans toutes les formes d'art.

Citons l'artisanat, avec de multiples (mais cependant en nombre fini..., nous y reviendrons dans le chap. 7) façons de tisser des étoffes avec un motif régulier, l'architecture, avec les célèbres pyramides de Gizeh (fig. 1.1), ou encore en musique.



FIGURE 1.1 – Les pyramides de Gizeh, 2500 ans avant J-C.

Chez les grecs, le concept de symétrie dans l'espace géométrique permet d'aboutir à la fameuse classification des polyèdres réguliers convexes, dits solides de Platon, illustrés dans la fig. 1.2. Cette classification est intimement liée aux propriétés topologiques de l'espace, que

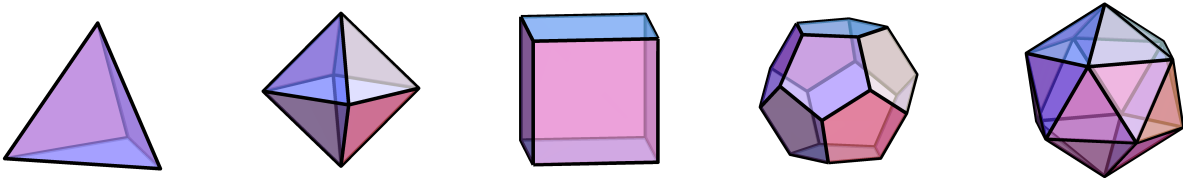


FIGURE 1.2 – Les cinq solides de Platon, seuls polyèdres réguliers convexes. De gauche à droite, le tétraèdre, l'octaèdre, le cube ou hexaèdre, le dodécaèdre et l'icosaèdre.

nous examinerons en détail dans le chap. 3. Notons également que la notion de centre de gravité d'un solide, introduite par Archimède, est également intimement reliée à la notion de symétrie.

Historiquement, le premier domaine où le concept de symétrie a été utilisé de manière systématique en physique est la cristallographie. Il s'agit alors de symétries dans l'espace du cristal, qui n'est pas continu. Les cristaux réalisent ici de façon tangible un système physique possédant un ensemble de symétries discrètes. Contrairement aux transformations qui laissent invariante un boule de billard sphérique par exemple, en nombre infini, l'ensemble des transformations laissant invariant un cristal est en nombre *fini*. Ces exemples nous amènent ainsi à faire une distinction entre *symétrie continue* (le *groupe* sous-jacent, dont nous formaliserons le concept plus loin, dépend d'un ou plusieurs paramètres que l'on peut faire varier continûment) et *symétrie discrète* (absence de paramètre). Le concept de symétrie continue joue un rôle essentiel en physique théorique. En effet, comme nous le montrerons plus loin, un théorème due à Emmy Noether permet d'affirmer qu'à toute symétrie continue d'un système correspond une grandeur conservée, en mécanique (voir chap. 9 comme en théorie des champs (voir chap. 10).

Quittons un instant les symétries dans l'espace physique à trois dimensions pour nous intéresser au temps. Imaginons que nous filmons la trajectoire suivie par une pierre qui a été lancée (voir fig. 1.3). Le fait de regarder le film dans le sens habituel ou de le regarder en le projetant à l'envers ne produit aucune différence sensible pour le spectateur, qui sera bien incapable de déterminer le sens suivant lequel la pierre a été réellement lancée. Ceci est une illustration de l'invariance par renversement du temps, qui constitue un autre exemple de symétrie discrète.



FIGURE 1.3 – Mouvement d’une pierre. A gauche : dans le sens du déroulement temporel réel du mouvement. A droite : film inversé.

Vérifiée par de nombreux systèmes mécaniques (à condition qu’ils ne soient pas plongés dans des champs magnétiques extérieurs, qui brisent l’invariance par renversement du temps, comme on le voit aisément sur l’expression de la force de Lorentz), elle semble être antagoniste avec la flèche du temps, qui peut se formuler en termes d’augmentation d’entropie et de perte d’information dans le cadre de la thermodynamique.

Une autre symétrie discrète joue un rôle essentiel, en physique, en chimie et en biologie. Il s’agit de la symétrie miroir, illustrée par le cas d’une espèce de nautilus qui existe dans ses deux chiralités dans la fig. 1.4. De nombreuses molécules sont chirales, i.e. non superposables



FIGURE 1.4 – Deux exemplaires de coquilles polies de Nautilus pompilius, symétriques l’une de l’autre dans un miroir.

à leur image dans un miroir (on parle des deux énantiomères d’une même molécule). Deux énantiomères ont les mêmes propriétés physiques (par exemple le même spectre en RMN). Mais ils n’ont pas les mêmes propriétés lorsqu’elles interagissent avec d’autres molécules ou objets chiraux. Ainsi en biologie, les acides aminés qui sont les Legos des protéines, ou encore les sucres, sont chiraux.

Le premier exemple historique de résolution chirale (séparation d’énantiomères), qui constitua l’un de ses tout premiers travaux de recherche, fut celui accompli par Louis Pasteur [1]. En 1848, il étudia la nature des dépôts apparaissant dans les tonneaux d’un vin des Vosges après fermentation. Ce dépôt était un sel d’un acide appelé racémique (de *racemus*, grappe de raisin en latin). Il avait été étudié par Eihard Mitscherlich : en étudiant le sel sodico-ammonique de cet acide racémique, il avait montré que celui-ci était parfaitement identique au sel de l’acide tartrique, présent dans de nombreuses plantes et en particulier l’un des principaux acide du vin. Cependant, le sel de l’acide tartrique était optiquement actif, dextrogyre (le plan de polarisation de la lumière, lorsque l’on fait face à la lumière, tourne vers la droite, i.e. dans le sens inverse trigonométrique, par opposition au pouvoir rotatoire lévogyre), alors que l’acide

racémique ne l'était pas.

Pasteur parvint à isoler les deux énantiomères de l'acide tartrique, mélange racémique de deux énantiomères, par tri manuel au microscope des cristaux énantiomorphes, constatant que ces deux formes cristallines étaient dissymétriques, donc discernable, contrairement aux l'observations de Mitscherlich. Ensuite, par dissolution séparée dans l'eau des deux types de cristaux, avec des concentrations identiques, il put alors montrer qu'une des deux solutions avait un pouvoir rotatoire dextrogyre, donc identique à celui de l'acide tartrique, tandis que l'autre avait un pouvoir lévogyre, avec un angle de déviation exactement opposé. Ceci expliquait l'absence de pouvoir rotatoire de la solution d'acide racémique<sup>1</sup>.

En physique, la distinction entre champ électrique et champ magnétique est bien connue : tandis que l'image d'un système de charge dans un miroir produit un champ électrique qui est lui-même miroir du champ électrique produit par le système de charges initial, le champ magnétique produit par l'image miroir d'un système de charge en mouvement est l'*antisymétrique* du champ magnétique miroir produit par le mouvement des charges sans application du miroir. Le champ électrique est donc un vecteur, tandis que le champ magnétique est un pseudo vecteur.

Associée à une rotation dans le plan du miroir, une autre symétrie lui est intimement reliée, la parité. Cette symétrie a en particulier joué un rôle essentiel dans le développement des théorie moderne de physique des particules élémentaires, dans le cadre quantique et relativiste. En particulier, l'électrodynamique et l'interaction forte respectent la parité, tandis que l'interaction faible viole (de façon maximale...) la parité.

Enfin, la symétrie par conjugaison de charge, qui fait passer l'électron au positron, constitue un autre exemple de symétrie discrète, qui joue elle aussi un rôle essentiel en physique corpusculaire.

En théorie quantique des champs relativistes, cadre conceptuel fondamental de toutes les théories modernes de la physique des particules, ces trois symétries discrètes, prises ensemble, jouent un rôle important. En effet, on peut démontrer que toute théorie des champs quantique relativiste locale est invariante sous le produit de ces trois symétries.

Enfin, un résultat important des théories des champs relativistes est que la symétrie statistique d'un ensemble de particules est relié au spin de ces particules. Un ensemble particules de spin entier (bosons) possède une fonction d'onde symétrique sous l'échange de deux particules de cet ensemble, tandis qu'un ensemble de particules de spin demi-entier (fermions) possède une fonction d'onde antisymétrique sous l'échange de deux particules de cet ensemble. Ce résultat joue un rôle central dans la quantification des théories des champs relativistes.

Plusieurs des exemples précédents concernent des symétries qui n'agissent pas sur l'espace-temps dans lequel baigne le système. On parle alors de *symétries internes*, lorsqu'elles agissent sur l'objet physique considéré (exemple : champs magnétiques, symétrie de permutation d'un ensemble de particules, symétrie d'isospin entre proton et neutron, ou entre quark *u* et *d*, etc.).

---

1. Le terme « racémique » est maintenant utilisé pour caractériser une solution constitué d'un mélange équimoléculaire de deux composés énantiomères.

La dernière distinction importante concerne la différence entre *symétrie globale* et *symétrie locale*. Dans le premier cas, la symétrie concerne tous les points de l'espace-temps considéré. C'est la cas de toutes les symétries envisagées précédemment. Dans le second cas, les paramètres de la transformation peuvent dépendre de la position dans l'espace-temps. On parlera alors de symétrie de jauge. Un tel type de symétrie a d'abord été découvert en électromagnétisme classique, puis en relativité générale. Le concept de symétrie de jauge a et continue de jouer un rôle majeur dans la construction des théories modernes des interactions fondamentales. Elevée au rang de principe, elle permet en effet de construire la dynamique des champs d'une théorie en interaction, comme on le verra au chap. 22.

Nous venons d'examiner de multiples situations dans lesquelles un système possède une certaine symétrie, en accord avec les symétries des équations décrivant ce système. Il existe cependant beaucoup de cas où une telle symétrie attendue n'est pas présente. On parle alors de brisure de symétrie.

Une premier type de brisure de symétrie survient lorsque les équations qui régissent le comportement d'un système ne sont plus invariantes par les transformations considérées, en raison de la présence d'un terme supplémentaire dans ces équations. Nous sommes soumis à une telle *brisure explicite de symétrie* depuis notre naissance. En effet, à la surface de la Terre, l'isotropie spatiale est brisée par l'existence de la pesanteur dirigée vers le centre de la Terre, et brise donc l'invariance haut/bas.

Une second type de brisure de symétrie, plus intrigant, survient lorsque les lois décrivant le système sont invariantes sous la symétrie, mais que la réalisation particulière du système observé ne l'est pas. Ainsi, en moyennant sur toute les réalisations possible du système, la symétrie considérée est restaurée. Un premier exemple familier nous est fourni par le jeu de pile ou face. Une pièce de monnaie est invariante par symétrie dans son plan, de même que les équations de la mécanique qui régissent son mouvement. Pourtant, la probabilité que la pièce reste sur sa tranche est extrêmement faible, et le résultat obtenu (pile ou face) brise la symétrie... De la même façon, la formation des planètes du système solaire, à partir d'un nuage de matière isotrope constitue une autre illustration de brisure dite *spontanée* de symétrie. Une nébuleuse avec une distribution de matière a priori isotrope s'effondre par contraction gravitationnelle. La matière acquiert alors une distribution dite de disque protoplanétaire, ayant la forme d'un disque représenté par des lignes concentriques de lumière vive dans lesquelles on peut observer des trous, visible sur la fig. 1.5. Au sein de ce disque, les poussières et autres particules finissent par s'entrechoquer et s'agglutiner en attirant toute la matière autour. Ces corps grossissent jusqu'à former des planètes (bandes noires sur la fig. 1.5). Ici, bien que l'agitation des poussières soit chaotique, le moment cinétique total du nuage de matière n'a aucune raison de s'annuler. Lors de la contraction, c'est l'existence de ce moment cinétique non nul qui va briser la symétrie et expliquer l'évolution vers un disque, par combinaison de la force d'attraction gravitationnelle (qui pointe en moyenne vers le centre de gravité du nuage de poussières) et de la force centrifuge (due à l'existence d'un moment cinétique total), plutôt que vers une distribution sphérique de matière.

De façon similaire, considérons le phénomène de cristallisation. A haute température, la matière possède toutes les symétries des équations décrivant le mouvement des atomes ou molécules, cet état initial étant dans une phase désordonnée liquide, gazeuse ou solide. Le groupe de symétrie est simplement l'ensemble des isométries à trois dimensions (la distance typique

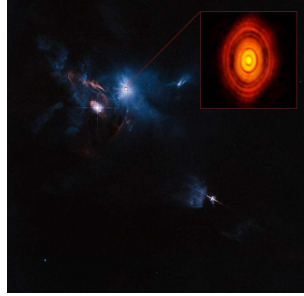


FIGURE 1.5 – Un disque proto-planétaire autour de l'étoile HL Tauri, située à 450 années-lumière de la Terre.

entre constituants est typiquement imposée par la température). Lors de la cristallisation, une brisure spontanée de symétrie a lieu, le cristal n'étant symétrique que sous un sous-groupe d'espace fini du groupe des isométries à trois dimensions. La direction spécifique de cristallisation est liée à la présence d'impuretés, et à l'existence de la pesanteur<sup>2</sup>.

En physique des hautes énergies, le mécanisme de Brout-Englert-Higgs-Hagen-Guralnik-Kibble, responsable de la masse des bosons de jauge  $W^\pm$  et  $Z^0$  et de la masse de quarks et des leptons, est un mécanisme de brisure spontanée de symétrie, en l'occurrence de la symétrie de jauge électrofaible. La récente découverte du boson BEH<sup>3</sup> au LHC (par les collaborations ATLAS et CMS) a confirmé la justesse de ce mécanisme.

---

2. dont on peut s'abstraire dans les expériences en micro-gravité, à coût élevé (vols paraboliques) ou très élevé (Station Spatiale Internationale)...

3. Acronyme pour Brout-Englert-Higgs. On l'appelle également couramment le boson de Higgs. Cela ne rend pas grâce au fait que ce mécanisme a été proposé simultanément par Higgs, par Brout et Englert et par Guralnik, Hagen et Kibble, en 1964...

# Chapitre 2

## Symétries géométriques

*Chapitre en cours d'écriture...*

### 2.1 Quelques rappels de géométrie affine

#### 2.1.1 Espace affine

### 2.1.2 Applications affines

**Définition 2.1 :** *Application affine*

Soient  $\mathcal{E}$  et  $\mathcal{F}$  deux espaces affines sur  $\mathbb{K}$ , et  $E$  et  $F$  leur espace directeur respectifs. On dit que  $f$  est une *application affine* s'il existe une application linéaire  $\varphi : E \rightarrow F$  telle que

$$\forall M \in \mathcal{E}, \forall u \in E, f(M + u) = f(M) + \varphi(u). \quad (2.1)$$

Il est facile de montrer que  $\varphi$  est unique. Dans la suite, on notera  $\vec{f}$  cette application linéaire.

**Proposition 2.2 :**

Soient  $E$  et  $F$  deux espaces vectoriels considérés comme espaces affines. Alors si  $\ell$  est une application affine,  $\vec{\ell} = \ell$ .



**Matrice d'une application affine****Proposition 2.3 :**

Une application  $f$  de  $\mathcal{E}$  dans  $\mathcal{F}$  est affine si et seulement si elle peut s'écrire, en coordonnées dans les repères  $\mathcal{R} = (O, e_1, \dots, e_m)$  et  $\mathcal{S} = (\Omega, f_1, \dots, f_n)$  respectivement des espaces  $\mathcal{E}$  dans  $\mathcal{F}$

$$f(x_1, \dots, x_m) = \left( \sum a_{1j} x_j + b_1, \dots, \sum a_{nj} x_j + b_n \right). \quad (2.2)$$

On déduit immédiatement de ce résultat l'écriture matricielle

$$\begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \\ 1 \end{pmatrix} = \begin{pmatrix} & & & & b_1 \\ & & & & \cdot \\ & A & & & \cdot \\ & & & & \cdot \\ & & & & b_n \\ 0 & \cdot & \cdot & \cdot & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \\ 1 \end{pmatrix} \quad (2.3)$$

où  $A$  est la matrice de l'application linéaire  $\vec{f}$ .

## 2.2 Isométries vectorielles

### 2.2.1 Linéarité des isométries vectorielles

**Théorème 2.4 :**

Les isométries vectorielles sont des applications linéaires.

### 2.2.2 Décomposition en produit de réflexions

**Théorème 2.5 :**

Soit  $E$  un espace vectoriel euclidien de dimension  $n$ , et  $\varphi$  une isométrie vectorielle. Alors on peut décomposer  $\varphi$  en produit de  $s$  réflexions vectorielles, avec  $s \leq n$ .

### 2.2.3 Isométries vectorielles en dimensions 0, 1 et 2



**Les isométries vectorielles directes en dimension 3**

**Les isométries vectorielles indirectes en dimension 3**

## 2.3 Structure du groupe Isom( $\mathbb{R}^n$ )

### 2.3.1 Décomposition en produit de réflexions

**Théorème 2.8 :**

Soit  $\mathcal{E}$  un espace affine euclidien de dimension  $n$ , et  $f$  une isométrie affine. Alors on peut décomposer  $f$  en produit de  $s$  réflexions vectorielles, avec  $s \leq n + 1$ .

### 2.3.2 Structure de produit semi-direct de Isom( $\mathbb{R}^n$ )

## 2.4 Isométries affines

### 2.4.1 Les isométries propres (ou déplacements)

### 2.4.2 Les isométries impropres



**Inversions**

Symétrie miroir, parité

## 2.5 Classification des isométries affines pour les petites dimensions

### 2.5.1 Dimensions 0 et 1

## 2.5.2 Dimension 2

**Lemme 2.9 :**

**Lemme 2.10 :**

Si  $r$  est une rotation autour de l'origine d'angle  $\theta$ , et  $v \in \mathbb{R}^2$ , alors  $\tau_v \circ r$  est une rotation autour de  $(1 - r)^{-1}(v)$  d'angle  $\theta$ .

**Lemme 2.11 :**

Soit  $f$  une réflexion autour d'un axe  $\ell$  passant par l'origine  $O$ , et  $v \in \mathbb{R}^2$ .

Soit  $g = \tau_v \circ f$ .

Alors  $g$  est une réflexion si et seulement si  $v$  est perpendiculaire à  $\ell$ .

◇ Lorsque  $v \perp \ell$ , alors l'axe de la réflexion  $g$  est  $\ell + \frac{1}{2}v$ .

◇ Lorsque  $v$  n'est pas perpendiculaire à  $\ell$ , alors  $g$  est un glissement et  $g^2$  est une translation de vecteur  $v + f(v)$ , *i.e.*  $g$  est la composition de la réflexion  $f$  suivie d'une translation de vecteur  $\frac{1}{2}(v + f(v))$ .

**Les isométries propres (ou déplacements)**

translations et rotations

**Les isométries impropres**

glissements

### 2.5.3 Dimension 3

Les isométries propres (ou déplacements)

translations, rotations, vissages

**Les isométries impropres**

- ◇ Composées d'une rotation et d'une réflexion par rapport à un plan orthogonal à l'axe de rotation
- ◇ glissements (ou réflexions glissées)

**2.6 Quelques résultats sur la composition des isométries****2.6.1 Composition de deux rotations de même centre (2D) ou d'axes concourants (3D)**

rotation

**2.6.2 Composition de deux réflexions**

Réflexions d'axes (2D) ou de plan (3D) concourants

rotations

Réflexions d'axes (2D) ou de plan (3D) parallèles

translation

# Chapitre 3

## Symétries et topologie

### 3.1 Introduction

Le fait qu'à trois dimensions il existe cinq polyèdres réguliers convexes, illustrés fig. 1.2, est connu depuis l'époque de Platon<sup>1</sup>.

Une démonstration du fait qu'il existe au maximum cinq polyèdres convexes est donnée par Euclide (proposition 18 du livre 13), basé sur des arguments géométriques. Nous allons la présenter.

a) Pour commencer, examinons les sommets d'un polyèdre. Chacun d'eux doit être relié par une arête à au moins trois sommets pour être considéré comme un sommet, puisque sinon ce sommet est en fait un point intérieur à une arête.

b) Ensuite, considérons les faces adjacentes à un sommet quelconque. Si l'on fait la somme, parmi ces faces adjacentes, des angles entre côtés adjacents à ce sommet, cette somme doit être strictement inférieure à  $360^\circ$ , puisque si cette somme est égale à  $360^\circ$ , ce n'est pas un sommet mais un point intérieur à une face.

c) Par définition d'un polyèdre régulier, les angles de tous les sommets de chacune de ses faces sont identiques. D'après a), chaque sommet doit donc contribuer pour un angle strictement inférieur à  $360^\circ/3 = 120^\circ$ .

d) Les polygones réguliers qui constituent les faces des polyèdres réguliers doivent avoir un nombre de côtés inférieur strictement à six. En effet, les polygones réguliers à six côtés ou plus ont des angles au sommet de  $120^\circ$  ou plus (voir ci-dessous), ce qui est interdit par c). Les faces des polyèdres réguliers sont donc soit des triangles, soit des carrés, soit des pentagones.

- ◇ Faces triangulaires : les angles au sommet d'un triangle équilatéral sont de  $60^\circ$ . De a) on déduit que chaque sommet est un point de rencontre de 3 (tétraèdre), 4 (octaèdre), ou 5 (icosaèdre) triangles.
- ◇ Faces carrées : les angles au sommet d'un carré étant de  $90^\circ$ , la seule possibilité d'après a) est que chaque sommet soit commun à 3 faces : c'est le cube.

---

1. L'existence de plusieurs modèles de solides, parmi lesquels certains solides « platoniciens » apparaissent, est attestée chez les peuples néolithiques d'Ecosse, mais l'identification de la régularité des fameux cinq solides n'est pas avérée

- ◇ Faces pentagonales : les angles au sommet d'un pentagone sont de  $108^\circ$ , et l'on déduit de a) que la seule possibilité est que chaque sommet soit commun à 3 faces : c'est le dodécaèdre.

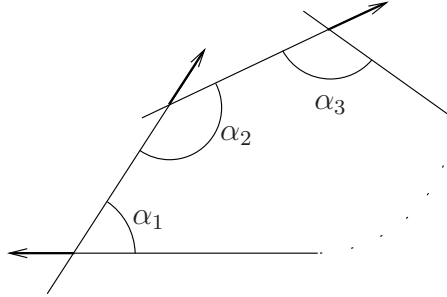


FIGURE 3.1 – Angles aux sommets d'un polygone.

Rappelons le lien entre le nombre de côtés  $n$  d'un polygone régulier et angle au sommet  $\alpha$ . Considérons tout d'abord un polygone quelconque, et un vecteur donné porté par l'un des côtés, pointant par exemple vers l'extérieur, voir fig. 3.1. Si l'on fait tourner ce vecteur pour qu'il ait la même position relative par rapport au côté suivant (dans le sens inverse trigonométrique dans l'illustration de la fig. 3.1), et que l'on répète le processus jusqu'à avoir fait le tour du polygone, ce vecteur a tourné d'un angle  $2\pi$ . L'angle entre le vecteur et le côté suivant étant de  $\pi - \alpha_i$ , nous avons donc

$$n\pi - \sum_{i=1}^n \alpha_i = 2\pi, \quad (3.1)$$

*i.e.*  $\sum \alpha_i = (n - 2)\pi$ . Pour un polygone régulier, ceci conduit à  $\alpha = \frac{n-2}{n}\pi$ . La condition  $\alpha < 2\pi/3$  conduit à  $n < 6$ .

---

La classification précédente peut s'obtenir avec des outils plus modernes, basés sur la topologie algébrique. Cette branche de la topologie cherche à associer des invariants algébriques aux différentes structures topologiques associées. Cet outil joue en particulier un rôle important en théorie des groupes et de leurs représentations. Nous allons maintenant présenter quelques outils de ce domaine.

## 3.2 Caractéristique d'Euler-Descartes

### 3.2.1 Surface bidimensionnelles de genre nul

C'est le premier invariant topologique qui a été introduit. Il est basé sur l'observation, que nous allons démontrer, du fait que pour les polyèdres convexes

$$\chi = S - A + F = 2, \quad (3.2)$$

où  $S$  est le nombre de sommets,  $A$  le nombre d'arêtes et  $F$  le nombre de faces



Une preuve du résultat précédent a été obtenue par Cauchy, que nous allons maintenant présenter. Considérons un polyèdre quelconque,

◇ Etape 1 :

On enlève une face du polyèdre. Par homéomorphisme (*i.e.* bijection continue), on peut écartier vers l'extérieur les côtés de cette face manquante. On déforme le polyèdre en l'aplatissant, ce qui conduit à un graphe plan dont :

- les nœuds sont les sommets
- les arcs sont les arêtes déformées.

Avant de poursuivre, deux remarques importantes doivent être faites :

- Les nombres  $S$ ,  $A$ ,  $F$  n'ont pas changé, à condition de compter l'extérieur du graphe comme une face enlevée, et donc  $S - A + F$  est inchangée dans cette étape.
- Le fait de déformer les lignes par homéomorphisme ne modifie aucune des quantités  $S$ ,  $A$ ,  $F$ .

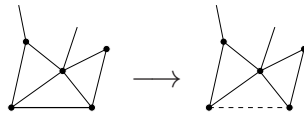
◇ Etape 2 :

Pour toute face ayant plus de 3 côtés : on trace une diagonale. On obtient ainsi, par triangulation, un nouveau graphe, dans lequel à la fois le nombre de face et d'arêtes augmente d'une unité. La quantité  $S - A + F$  demeure inchangée dans cette étape.

◇ Etape 3 :

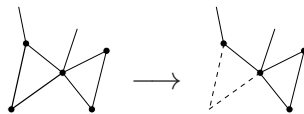
On répète les deux opérations ci-dessous, qui conservent  $S - A + F$  :

- On supprime un à un tous les triangles qui comportent exactement un seul côté aux frontières extérieures du graphe. Par exemple,



Alors  $A \rightarrow A - 1$ ,  $F \rightarrow F - 1$  et  $S$  est constant, donc  $S - A + F$  est constant.

- On supprime un à un tous les triangles qui comportent exactement deux arêtes aux frontières extérieures du graphe. Par exemple,



Alors  $A \rightarrow A - 2$ ,  $F \rightarrow F - 1$  et  $S \rightarrow S - 1$ , donc  $S - A + F$  est constant.

On obtient finalement un triangle, pour lequel  $S = 3$ ,  $A = 3$  et  $F = 2$  (ne pas oublier la face extérieure!), et donc  $S - A + F = 2$ , qui a été conservée tout au long du processus, d'où le résultat cherché.

Ces différentes étapes sont illustrées pour le cas d'un parallélépipède, dans la fig. 3.2.

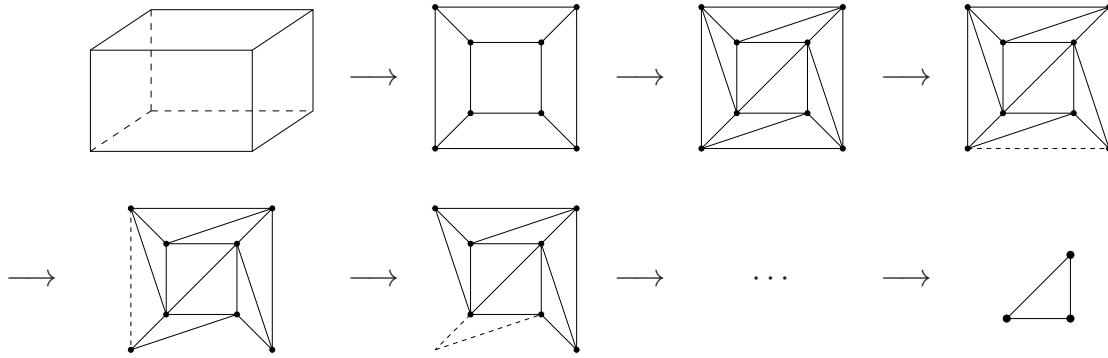
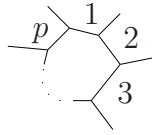


FIGURE 3.2 – Les étapes de la réduction du graphe d'un parallélépipède.

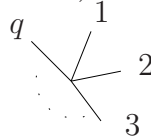
### 3.2.2 Les polyèdres réguliers convexes

Le résultat précédent permet d'obtenir la classification des solides de Platon. En effet, introduisons le symbole de Schläfli  $\{p, q\}$ , défini par :

- ◇  $p$  = le nombre de côtés de chaque face (qui est aussi le nombre de sommets sur chacune des faces)



- ◇  $q$  = le nombre de faces qui se rencontrent à chaque sommet (qui est aussi le nombre d'arêtes se rencontrant à chaque sommet).



Chacune des arêtes joint deux sommets, et possède deux faces qui lui sont adjacentes. En parcourant ainsi deux fois l'ensemble des sommets, on peut donc écrire

$$pF = 2A = qS. \quad (3.3)$$

Combinées à la caractéristique d'Euler-Descartes (3.2), ces deux égalités conduisent donc à la relation

$$\frac{2A}{q} - A + \frac{2A}{p} = 2 \quad (3.4)$$

et donc

$$\frac{1}{q} + \frac{1}{p} = \frac{1}{2} + \frac{1}{A}. \quad (3.5)$$

Le nombre d'arêtes  $A$  étant strictement positif, on a donc  $\frac{1}{q} + \frac{1}{p} > \frac{1}{2}$ . Les paramètres  $p$  et  $q$  devant être au moins égaux à 3, on en déduit que les seules solutions possibles pour  $\{p, q\}$  sont :

- {3, 3} tétraèdre
- {4, 3} cube
- {3, 4} octaèdre
- {5, 3} dodécaèdre
- {3, 5} icosaèdre .

On en déduit donc finalement les différents paramètres caractérisant ces cinq solides de Platon, qui sont donnés dans la table 3.1.

p	q	S	A	F	
3	3	4	6	4	tétraèdre
4	3	8	12	6	cube
3	4	6	12	8	octaèdre
5	3	20	30	12	dodécaèdre
3	5	12	30	20	icosaèdre

TABLE 3.1 – Les différents polyèdres connexes réguliers et leurs paramètres.

Il est intéressant de noter la dualité qui existe entre ces polyèdres : deux polyèdres sont duaux si leurs indices  $p$  et  $q$  sont échangés, autrement dit, on met les faces de l'un en correspondance avec les sommets de l'autre. Il suffit pour cela de relier les centres des faces adjacentes d'un polyèdre pour obtenir son polyèdre dual. Le tétraèdre est ainsi auto dual, tandis que le cube est dual de l'octaèdre, et le dodécaèdre et dual de l'icosaèdre.

### 3.2.3 Caractéristique d'Euler-Descartes de quelques objets élémentaires





	S	A	F	$\chi$	
	2	1	0	1	segment
	1	1	0	0	cercle $S^1$
	1	1	1	1	disque
	1	1	2	2	sphère $S^2$

TABLE 3.2 – Caractéristique d'Euler-Descartes de quelques objets élémentaires.

La caractéristique d'Euler-Descartes est facile à calculer pour les objets simples. Les valeurs pour le segment, le cercle  $S^1$ , le disque et la sphère  $S^2$  sont donnés dans la table 3.2. Dans le

cas de la sphère, le résultat peut également s'obtenir en ajoutant la caractéristique d'Euler-Descartes de deux disques et en enlevant la contribution de l'arête et du sommet commun, *i.e.* de  $\mathbb{S}^1$  qui a une contribution nulle, soit  $1 + 1 - 0 = 2$ .

### 3.2.4 Surfaces bidimensionnelles orientables de genre quelconque

L'expression de la caractéristique d'Euler-Descartes (3.2) est correcte à condition de pouvoir réduire le graphe à un triangle. Mais la situation se complique si le graphe est non planaire. A l'ordre le bas dans le nombre de sommets  $S$ , ceci arrive pour  $S = 4$  dans ce cas d'un graphe homéomorphe à un carré, dont les deux diagonales passent l'une au-dessus de l'autre, comme illustré dans la fig. 3.3.

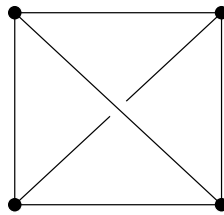


FIGURE 3.3 – Diagramme non planaire d'ordre minimal.

Le comptage du nombre de faces est malaisé. Il est facilité en utilisant l'astuce suivante : on remplace chaque arête par une double ligne. Chaque face se trouve ainsi entourée par une ligne simple, et il suffit de compter le nombre de ces lignes simples pour obtenir  $F$  (y compris la face extérieure). Ceci est illustré pour notre diagramme d'ordre minimal dans la fig. 3.4. On constate

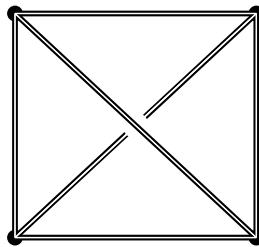


FIGURE 3.4 – Le diagramme non planaire d'ordre minimal, où chaque arête est remplacée par une double ligne afin de faciliter le comptage des faces.

sur cette représentation que seules deux lignes subsistent : l'une correspond au parcours de la ligne interne, tandis que la seconde correspond à faire le tour de la face externe. Il y a donc exactement  $F = 2$  faces. Comme  $S = 4$  et  $A = 6$ , on en déduit que  $\chi = 0$ .

Ce résultat surprenant est directement relié à la topologie de la surface bidimensionnelle qui porte notre graphe. Dans le cas d'un polyèdre usuel, le graphe est porté par une surface bidimensionnelle homéomorphe à la sphère  $\mathbb{S}^2$ . Il est en effet clair que le parallélépipède initial est homéomorphe à  $\mathbb{S}^2$ . Mais dans le cas de notre graphe de la fig. 3.3, il n'est pas possible de dessiner ce graphe sur une sphère. Un tel graphe peut en revanche être dessiné sur un tore à un trou  $\mathbb{T}^2$ , comme illustré dans la fig. 3.5. Sur cette figure, on a utilisé la même convention

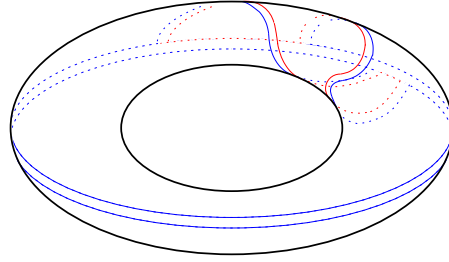


FIGURE 3.5 – Le diagramme non planaire d'ordre minimal, dessiné sur un tore à un trou.

des doubles lignes, ce qui permet bien de vérifier que deux zones de la surface du tore sont délimitées (entourées par les lignes bleues et rouges).

Si l'on revient aux polyèdres, il est facile d'en imaginer qui soient homéomorphes à un tore  $\mathbb{T}^2$ . Un exemple est fourni par la fig. 3.6. Le comptage du nombre de sommets, d'arêtes et de faces donne  $S = 6 \times 4 = 24$ ,  $A = 6 \times 4 + 6 \times 4 = 48$  et  $F = 6 \times 4$ , ce qui conduit bien à une caractéristique  $\chi = 0$ .

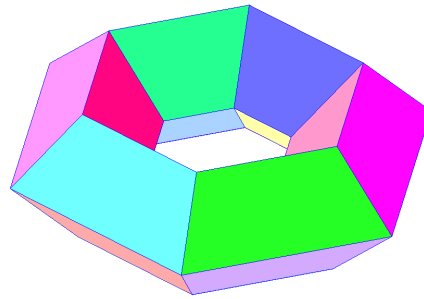


FIGURE 3.6 – Un exemple de polyèdre à un trou.

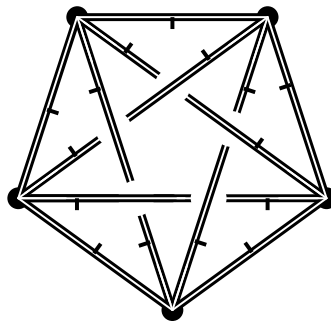


FIGURE 3.7 – Diagramme non planaire d'ordre minimal (où chaque arête est remplacée par une double ligne afin de faciliter le comptage des faces) pouvant se dessiner sur un tore à deux trous. Pour faciliter le comptage, l'une des deux lignes interne porte un symbole. La troisième ligne court autour du graph. Il y a donc 3 faces.

Il est maintenant aisé de généraliser le résultat précédent. Le diagramme non planaire qui suit par ordre de complexité le cas du carré avec diagonales non concourantes est un pentagone

avec diagonales non courantes, comme illustré dans la fig. 3.7. On a dans ce cas  $S = 5$ ,  $A = 10$ ,  $F = 3$ , de sorte que  $\chi = -2$ . Ce graphe peut être dessiné sur un tore à deux trous, même s'il n'est pas tout à fait évident de s'en convaincre...

### Domaine fondamental

Etant donné l'action d'un groupe  $G$  sur un espace topologique  $X$  (voir chap. 4) par homéomorphisme, un domaine fondamental pour cette action est un sous-ensemble de  $X$  qui contient exactement un représentant de chacune des orbites.

*Exemples :*

- ◇ Si  $X = \mathbb{R}^2$  et  $G$  le réseau  $\mathbb{Z}^2$  agissant sur  $X$  par translation, l'espace quotient  $X/G$  est simplement le tore  $\mathbb{T}^2$ . Un domaine fondamental est donné par  $[0, 1[ \times [0, 1[$ . Il est illustré fig. 3.8 et il est facile de se convaincre que l'on obtient bien le tore  $\mathbb{T}^2$  par homéomorphisme (il suffit d'identifier les lignes bleues et rouges sur la figure : on obtient un cylindre à la première étape, puis un tore).
- ◇ Dans le cadre de la cristallographie en dimension  $d$ , le domaine fondamental pour l'action du groupe des translations du réseau (isomorphe au réseau,  $\mathbb{Z}$ -module constitué par les combinaisons linéaires à coefficients dans  $\mathbb{Z}$  de  $d$  vecteurs de bases) sur  $\mathbb{R}^d$  sera appelé maille primitive.

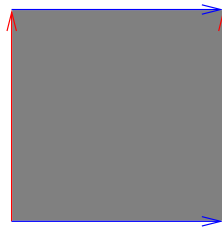


FIGURE 3.8 – Le domaine fondamental du tore.

### Décomposition en somme connexe d'une surface

Un outil permet de calculer facilement la caractéristique d'Euler-Poincaré. Partant de deux surfaces connexes (on parle mathématiquement plus généralement de *variétés* connexes de dimensions  $n$ , au sens où il faut  $n$  paramètres pour se repérer localement au voisinage d'un point de cette variété), la *somme connexe* de ces deux variétés s'obtient en retirant à chacune de ces deux variétés un petit voisinage d'un point formé d'une boule ouverte, et en recollant les deux variétés ainsi obtenues le long des deux sphères  $\mathbb{S}^{n-1}$  ainsi apparues. Le résultat ainsi obtenu est une variété de dimension  $n$ , bien définie à homéomorphisme près, et connexe (exception faite du cas où les deux variétés initiales sont homéomorphes à des droites réelles : on obtient dans ce cas particulier à nouveau deux droites).

Considérons l'exemple du tore à deux trous. On peut le voir comme la somme connexe de deux tores à un trou.

Pour calculer la caractéristique du tore à deux trous, il suffit de faire la somme de la caractéristique de chacun des deux tores, dont on a enlevé la caractéristique d'un disque (soit donc  $0 - 1 = -1$  pour chaque tore percé), à laquelle on ajoute la caractéristique de  $\mathbb{S}^1$  (qui sert à recoller), qui vaut 0, soit au total  $\chi = -2$ . On peut facilement vérifier ce résultat en considérant

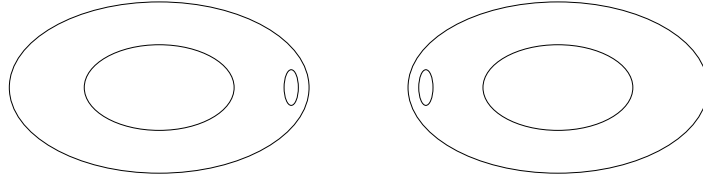


FIGURE 3.9 – Recollement de deux tores à un trou.

l'exemple du polyèdre à un trou de la fig. 3.6. Le recollement de ce polyèdres, par exemple au niveau de la face verte au premier plan de la figure, avec un second polyèdre identique, conduit comme attendu à

$$\chi = 0 + 0 - 1 - 1 = -2 \quad (3.6)$$

où l'on a utilisé le fait que chaque face enlevée conduit à soustraire une contribution  $4 - 4 + 1 = 1$  (*i.e.* topologiquement la contribution d'un disque), tandis que le recollement conduit à l'identification de 4 sommets et 4 arêtes, ce qui ne change pas  $\chi$ .

De façon générale, en utilisant la même méthode, on peut établir la relation suivante :

$$\chi = 2 - 2g \quad (3.7)$$

où  $g$  est le genre de la surface. Cela signifie que la surface est homéomorphe à un tore à  $g$  trous. De façon équivalente, c'est le nombre de tores (à un trou)  $\mathbb{T}^2$  apparaissant dans la décomposition en somme connexe de la surface  $\mathcal{S}$  :

$$\mathcal{S} = \mathbb{T}^2 \# \mathbb{T}^2 \cdots \# \mathbb{T}^2. \quad (3.8)$$

### 3.2.5 Surfaces bidimensionnelles non orientables de genre quelconque

Dans le cas des surfaces non orientables, les choses se compliquent. La surface non orientable

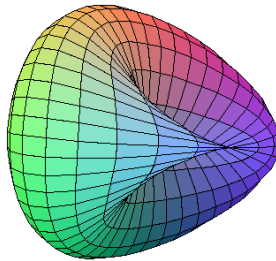


FIGURE 3.10 – Le bonnet croisé.

fermée la plus simple qui conduit à une caractéristique impaire (égale ici à 1) est le bonnet croisé, illustré dans la fig. 3.10. C'est le plan projectif réel  $P^2(\mathbb{R})$ , comme on va le justifier dans la



FIGURE 3.11 – Deux choix possibles de domaine fondamental pour le bonnet croisé : un disque (à gauche) ou un carré (à droite).

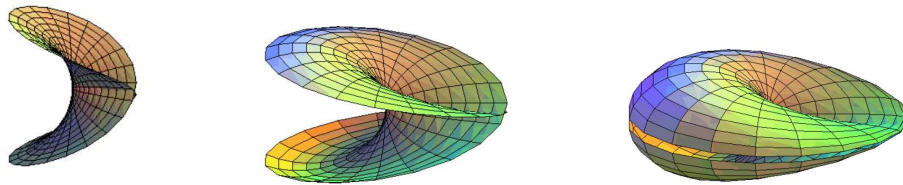


FIGURE 3.12 – Le bonnet croisé construit par recollement du bord d'un disque, en accord avec la représentation du plan projectif réel  $P^2(\mathbb{R})$ .

paragraphe qui suit. Il s'obtient par identification des points opposés du bord du disque unité du plan réel, suivant la fig. 3.11. Cette construction du bonnet croisé est illustrée sur la fig. 3.12.

On montre de façon générale que pour toute surface fermée connexe non orientable,

$$\chi = 2 - k \quad (3.9)$$

où  $k$  est le genre de la surface non orientable. Cela signifie que  $k$  est le nombre de plan projectif réels apparaissant dans la décomposition en somme connexe de la surface.

Nous allons maintenant présenter quelques éléments de géométrie projective<sup>2</sup>, en nous concentrant sur les aspects topologiques.

### 3.2.6 Quelques éléments de géométrie projective

Considérons un corps commutatif  $\mathbb{K}$  (en pratique,  $\mathbb{R}$  ou  $\mathbb{C}$ , mais on peut également considérer un corps fini).

En géométrie affine, il est naturel de considérer l'ensemble des directions de cet espace (donc l'ensemble des droites vectorielles). En particulier, en géométrie affine du plan réel, si l'on considère deux droites affines quelconques, soit elles se coupent en un point unique, soit elles sont parallèles. Si l'on ajoute un point à l'infini, alors toute paire de droite se coupera en un point unique. Voilà le point de départ de la géométrie projective. C'est en particulier le cadre naturel du dessin en perspective.

**Définition 3.1 :** *Espace projectif*

2. De nombreuses références existent sur le sujet, parmi lesquelles [2] et [3]. Pour un avant goût de niveau M1 de mathématiques ou agrégation, voir [4, 5].



Soit  $E$  un  $\mathbb{K}$ -espace vectoriel. L'espace projectif déduit de  $E$ , et noté  $P(E)$ , est l'ensemble des droites vectorielles de  $E$ .

De façon équivalente, si l'on définit la relation de colinéarité  $\mathcal{R}$  sur l'espace  $E \setminus \{0\}$  par  $x\mathcal{R}y \Leftrightarrow \exists \lambda \in \mathbb{K}^*$  tel que  $y = \lambda x$ , alors  $P(E) \cong (E \setminus \{0\})/\mathcal{R}$  : l'espace projectif est simplement l'espace des classes d'équivalence des vecteurs non nuls de  $E$  pour la relation de colinéarité.

Si  $E$  est de dimension finie, alors on dit que  $P(E)$  est de dimension finie, et sa dimension est alors  $\dim(E) - 1$ .

Si  $E = \mathbb{K}^{n+1}$ , on note alors  $P(\mathbb{K}^{n+1}) = P^n(\mathbb{K}) = \mathbb{K}P^n$ .

*Exemples :*

◇ Si  $\dim E = 0$ , *i.e.* dans le cas trivial  $E = \{0\}$ , il n'y a pas de droite vectorielle, et  $P(\{0\}) = \emptyset$ .

◇ Si  $\dim E = 1$ ,  $E$  contient une seule droite vectorielle, et  $P(E)$  est un singleton.

◇ Si  $\dim E = 2$ ,  $P(E)$  est une *droite projective*.

– L'exemple le plus simple est de considérer  $\mathbb{K} = \mathbb{R}$ . Alors  $P^1(\mathbb{R}) \cong \mathbb{S}^1/(x \sim -x) \cong \mathbb{S}^1$ .

La première égalité vient du fait qu'une droite vectorielle est donnée par un angle, défini au signe près. La seconde égalité se montre en considérant la projection stéréographique de la sphère  $\mathbb{S}^1$  sur la droite  $\mathbb{R}$  complétée par un point à l'infini  $\infty$ , illustrée dans la fig. 3.13.

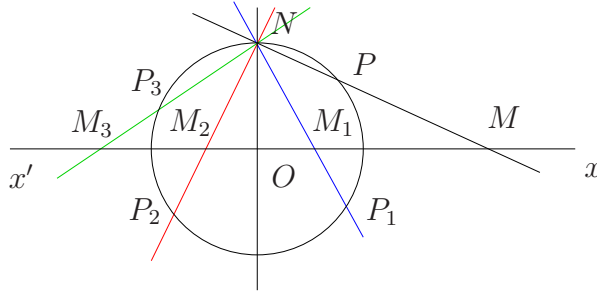


FIGURE 3.13 – La projection stéréographique de  $P^1(\mathbb{R})$  sur  $\mathbb{S}^1$ .

On considère pour cela le cercle unité  $\mathbb{S}^1$  et le *pôle*  $N(0, 1)$ . Soit l'application  $f : P^1(\mathbb{R}) \rightarrow \mathbb{S}^1$  définie de la façon suivante :

La droite passant la  $N$  et coupant l'axe  $x'x \simeq \mathbb{R}$  en  $M$  coupe  $\mathbb{S}^1$  en un point unique  $P$ . Le pôle  $N$  lui-même est en bijection avec le point à l'infini.

Un calcul simple montre que

$$f(x) = \left( \frac{2x}{x^2 + 1}, \frac{x^2 - 1}{x^2 + 1} \right) \text{ et } f(\infty) = (0, 1). \quad (3.10)$$

La bijection réciproque s'écrit

$$f^{-1}(\alpha, \beta) = \sqrt{\frac{1 + \beta}{1 - \beta}} \text{ et } f^{-1}(N) = \infty. \quad (3.11)$$

- Lorsque  $\mathbb{K} = \mathbb{C}$ ,  $P^1(\mathbb{C}) \cong \mathbb{C} \cup \{\infty\} \cong \mathbb{R}^2 \cup \{\infty\} \cong \mathbb{S}^2$ . Ce dernier isomorphisme s'établit cette fois en considérant la projection stéréographique sur la sphère  $\mathbb{S}^2$  (appelée sphère de Riemann dans ce contexte), comme illustré sur la fig. 3.14.

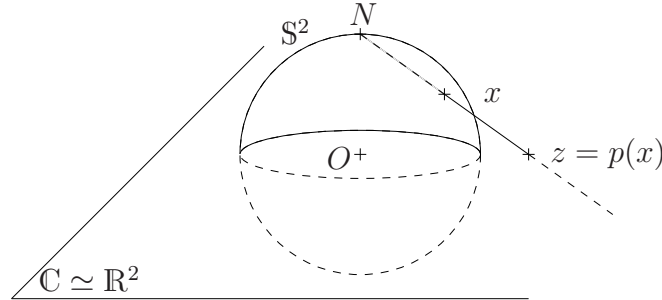


FIGURE 3.14 – La projection stéréographique de  $P^1(\mathbb{C})$  sur  $\mathbb{S}^2$ .

- ◇ Si  $\dim E = 3$ , on parle alors de *plan projectif*. Le plan projectif réel  $\mathbb{R}P^2$  (ou encore  $P^2(\mathbb{R})$ ) est donc l'ensemble des droites réelles, dans l'espace de dimension 3.

### Carte affine

La visualisation d'un espace projectif  $P^n(\mathbb{K})$  est grandement facilitée par la notion de carte affine.

Commençons par deux remarques très simples :

- ◇ Si  $E$  est un  $\mathbb{K}$ -espace vectoriel de dimension  $n + 1 > 0$  et  $F$  un sous-espace vectoriel de  $E$ , alors toute droite vectorielle de  $E$  passant par un élément de  $F$  est entièrement contenue dans  $F$ . Ainsi, l'espace  $F \setminus \{0\}$  est stable sous la relation d'équivalence  $\mathcal{R}$ .
- ◇ De plus, sous la projection (*i.e.* le morphisme quotient qui à un vecteur non nul de  $E$  associe sa direction)  $\pi : E \setminus \{0\} \rightarrow P(E)$ , l'espace projectif  $P(F)$  est isomorphe à  $\pi(F \setminus \{0\}) \subseteq P(E)$ , ce qui implique l'inclusion (intuitivement évidente...)  $P(F) \subseteq P(E)$ .

### Définition 3.2 : Sous-espace projectif

Un sous-ensemble  $V \subseteq P(E)$  est appelé *sous-espace projectif* s'il existe un sous-espace vectoriel  $F$  de  $E$  tel que  $V = P(F)$ . En particulier, la dimension de  $V$  vérifie  $\dim V = \dim F - 1$ .

Ceci permet donc d'établir une bijection naturelle entre les sous-espaces vectoriels de  $E$  et l'ensemble des sous-espaces projectifs de  $P(E)$ .

*Exemples :*

- ◇ Si  $\dim V = 0$ ,  $V$  est un point, correspondant à une droite de  $E$ .
- ◇ Si  $\dim V = 1$ ,  $V$  est une droite, correspondant à un plan de  $E$ .
- ◇ Si  $\dim V = \dim P(E) - 1$ ,  $V$  est appelé hyperplan projectif ; il correspond à un hyperplan de  $E$ .

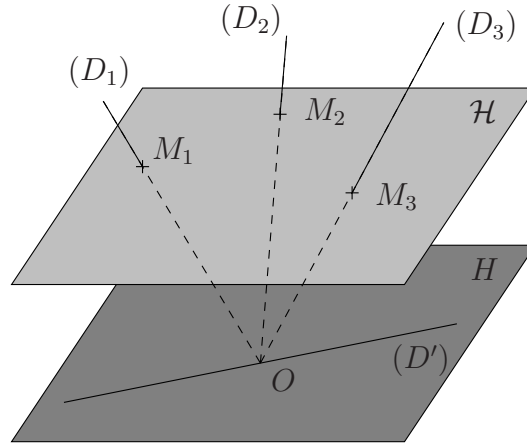


FIGURE 3.15 – L'espace projectif  $P(E)$  vu comme l'union d'un hyperplan affine  $\mathcal{H}$  et de l'espace projectif  $P(H)$  des droites vectorielles  $(D')$  de cet hyperplan  $H$ .

Considérons les sous-espaces projectifs de  $P^2(\mathbb{R})$  :

- ◇ Les points de  $P^2(\mathbb{R})$  sont les projetés des droites vectorielles de  $\mathbb{R}^3$ .
- ◇ Les droites de  $P^2(\mathbb{R})$  sont les projetés des plans vectoriels de  $\mathbb{R}^3$ , que l'on peut aussi appeler hyperplan projectifs en accord avec la définition ci-dessus.

Considérons maintenant un hyperplan  $H$  de  $E$ . On peut munir le  $\mathbb{K}$ -espace vectoriel  $E$  d'une structure de  $\mathbb{K}$ -espace affine. Soit dans ce  $\mathbb{K}$ -espace affine  $E$  un hyperplan affine  $\mathcal{H}$ , de direction  $H$  et ne contenant pas l'origine  $O$ . L'hyperplan  $\mathcal{H}$  est donc distinct de  $H$ .

Nous pouvons maintenant étudier la structure de l'espace projectif  $P(E)$ , comme illustré dans la fig. 3.15. Le but est d'identifier dans  $P(E)$  une partie affine. Soit  $D$  une droite vectorielle de  $E$ . Il y a alors deux possibilités :

- ◇ Soit  $D$  n'est pas contenue dans  $H$ . Alors  $E = H \oplus D$  et la droite  $D$  intersecte  $\mathcal{H}$  en un point unique.
- ◇ Soit  $D \subseteq H$  et la droite  $D$  ne rencontre pas  $\mathcal{H}$ .

Il y a donc une bijection entre  $\mathcal{H}$  et  $P(E) \setminus P(H)$ . Dans ce raisonnement, l'hyperplan  $\mathcal{H}$  n'est bien entendu pas unique. Si l'on choisit un autre hyperplan  $\mathcal{H}'$ , le théorème de Thalès permet de construire une bijection entre  $\mathcal{H}$  et  $\mathcal{H}'$ , qui préserve la structure affine. Donc on peut conclure qu'un espace projectif contient un espace affine.

**Proposition 3.3 :** Si  $H$  est un hyperplan de  $E$ , le sous-espace  $P(E) \setminus P(H)$  est naturellement muni d'une structure d'espace affine  $\mathcal{H}$ , de direction  $H$ . On peut donc écrire

$$P(E) = \mathcal{H} \cup P(H). \quad (3.12)$$

On dira que  $\mathcal{H}$  est la *partie affine* de  $H$  et on appellera *carte affine* de  $P(E)$  l'application  $P(E) \setminus P(H) \longrightarrow \mathcal{H}$ .

Connaissant un espace affine  $\mathcal{E}$ , de dimension  $n$  et de direction  $E$ , il est possible de faire le chemin inverse.

On peut tout d'abord démontrer qu'il est possible de construire un espace vectoriel  $\hat{E}$  (appelé *espace universel* de  $E$ ) tel que  $E$  soit un hyperplan vectoriel de  $\hat{E}$  (voir par exemple [6]). L'idée est d'ajouter une dimension  $n + 1$ . On peut alors toujours choisir un système de coordonnées pour que  $z = 0$  corresponde à  $E$  (vu comme un espace affine) et  $z = 1$  corresponde à  $\mathcal{E}$ .

On a alors une bijection entre  $\mathcal{E}$  et  $P(\hat{E}) \setminus P(E)$ , et donc

$$P(\hat{E}) = \mathcal{E} \cup P(E). \quad (3.13)$$

L'espace  $P(\hat{E})$  ainsi construit est appelé la *complétion projective* de l'espace affine  $\mathcal{E}$ . Prenons quelques exemples pour rendre les choses plus explicites.

*Exemples :*

- ◇ La droite projective se décompose suivant  $P^1(\mathbb{K}) = \mathbb{K} \cup \{\infty\}$  : à la droite affine  $\mathbb{K}$  on a ajouté un point à l'infini.
- ◇  $P^2(\mathbb{K}) = \mathbb{K}^2 \cup P^1(\mathbb{R})$ , et en particulier  $P^2(\mathbb{R}) = \mathbb{R}^2 \cup P^1(\mathbb{R})$ . En effet, le plan projectif est constitué de l'ensemble des points du plan affine  $\mathbb{R}^2$  (qui codent les directions des droites vectorielles non incluses dans le plan affine  $\mathbb{R}^2$ ), auquel on ajoute toutes les directions des droites vectorielles de  $\mathbb{R}^2$ ; cette seconde partie de  $P^2(\mathbb{R})$  constitue l'ensemble des points à l'infini.
- ◇ De façon générale, on obtient par récurrence immédiate que  $P^n(\mathbb{K}) = \mathbb{K}^n \cup \mathbb{K}^{n-1} \cup \dots \cup \mathbb{K} \cup \{\infty\}$ .

### Caractéristique d'Euler-Descartes du plan projectif

Il nous reste maintenant à justifier que  $\chi(P^2(\mathbb{R})) = 1$ . Pour cela, il suffit d'utiliser la décomposition  $P^2(\mathbb{R}) = \mathbb{R}^2 \cup P^1(\mathbb{R})$ , qui se simplifie dans le cas particulier de  $\mathbb{K} = \mathbb{R}$ , puisque  $P^1(\mathbb{R}) \cong \mathbb{S}^1/(x \sim -x)$ .

- ◇ D'une part, en utilisant la projection stéréographique 3.14, mais restreinte à  $\mathbb{C} \cong \mathbb{R}^2 \longrightarrow \mathbb{S}^2 \setminus \{N\}$  (donc sans le point à l'infini en bijection avec le pôle  $N$ ), on établit un homéomorphisme entre  $\mathbb{R}^2$  et le disque ouvert (il est facile de se convaincre de la sphère  $\mathbb{S}^2$  privée du pôle  $N$  est homéomorphe à  $D_1(0)$ , disque unité ouvert centré en 0).
- ◇ D'autre part,  $P^1(\mathbb{R})$  est homéomorphe au cercle  $\mathbb{S}^1$  dont on a identifié les points opposés.

On en déduit donc que  $P^2(\mathbb{R}) \cong D_1(0) \cup \mathbb{S}^1/(x \sim -x)$ , qui est homéomorphe au disque unité dont on a identifié les points opposés du bord, suivant la fig. 3.11. Un tel disque a une caractéristique  $\chi = 1$ .

### 3.2.7 Classification des surfaces fermées

Nous allons simplement énoncer le théorème de classification des surfaces fermées, sans le démontrer. Il nous dit que les trois situations décrites plus haut couvrent toutes les possibilités :

**Théorème 3.4 :** *de classification des surfaces fermées*

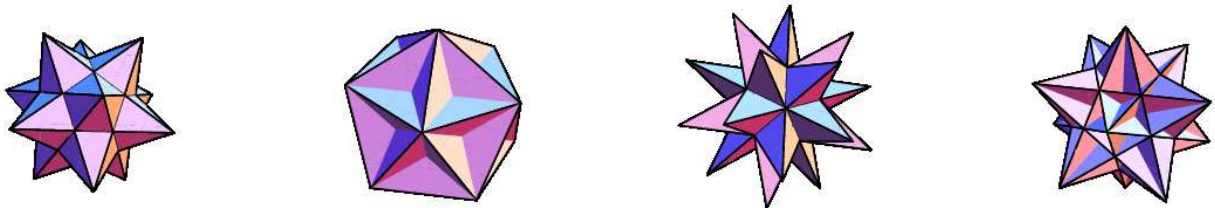
Toute surface fermée connexe est homéomorphe à une surface faisant partie de l'une des trois familles suivantes :

- ◇ la sphère  $S^2$ . Alors  $\chi = 0$ .
- ◇ la somme connexe de  $g$  tores  $\mathbb{T}^2$  ( $g \geq 1$ ) Alors  $\chi = 2 - 2g$ .
- ◇ la somme connexe de  $k$  plan projectifs réels  $P^2(\mathbb{R})$  tores  $\mathbb{T}^2$  ( $k \geq 1$ ). Alors  $\chi = 2 - k$ .

Les deux premières familles sont orientables, tandis que la troisième ne l'est pas. Ainsi, la classification des surfaces connexes fermées (à homéomorphisme près) est complètement déterminée par la valeur de sa caractéristique  $\chi$  et de son caractère orientable ou non :

- ◇ si la surface est orientable, connaissant sa caractéristique on en déduit qu'elle est la somme connexe de  $g = (2 - \chi)/2$  tores  $\mathbb{T}^2$ .
- ◇ si la surface est non orientable, on en déduit qu'elle est la somme connexe de  $k = 2 - \chi$  plan projectifs réels  $P^2(\mathbb{R})$ .

### 3.3 Les polyèdres réguliers non convexes



petit dodécaèdre étoilé    grand dodécaèdre    grand dodécaèdre étoilé    grand icosaèdre

FIGURE 3.16 – Les quatre polyèdres réguliers non convexes de Kepler-Poinsot.

Cauchy a démontré qu'il existe exactement quatre polyèdres réguliers non convexes, illustrés dans la fig. 3.16. Leur caractéristique d'Euler-Poincaré est égale à 2 à condition de prendre en compte tous les sommets, arêtes et faces, *i.e.* y compris internes.

A noter que le petit dodécaèdre étoilé est le dual du grand dodécaèdre, et le grand dodécaèdre étoilé celui du grand icosaèdre.

### 3.4 Caractéristique d'Euler-Poincaré

La caractéristique d'Euler-Poincaré est une généralisation de la caractéristique d'Euler-Descartes.

### 3.4.1 Polytopes

**Définition 3.5 :** *polytope convexe*

Un polytope convexe de l'espace euclidien  $\mathbb{R}^d$  (ou  $d$ -polytope) est le plus petit convexe contenant un nombre fini de points donnés  $x_i$ .

**Définition 3.6 :** *polytope convexe (bis)*

Une définition équivalente est donnée par

$$P = P(A, z) = \{x \in \mathbb{R}^d / Ax \leq z\} \quad (3.14)$$

où  $A \in \mathbb{R}^{m \times d}$  et  $z \in \mathbb{R}^m$ ,  $m$  étant un entier quelconque.

On peut également définir un polytope par l'intersection de  $m$  demi-espaces fermés, suivant la relation

$$P = \bigcap_1^m H_i \quad \text{où } H_i = \{x \in \mathbb{R}^d / a_i x \leq z_i\}, \quad (3.15)$$

$a_i$  correspondant à la  $i$ -ème ligne de la matrice  $A$ . Cette définition correspond à identifier la notion de polytope et la notion de région de Voronoï.

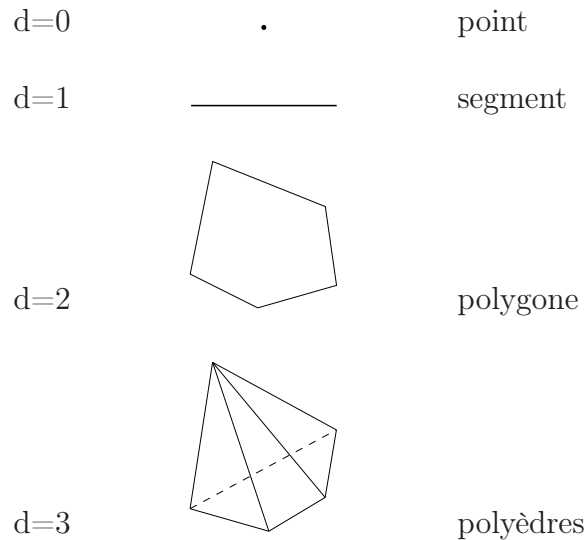


FIGURE 3.17 – Les polytopes en dimensions  $d$  pour  $0 \leq d \leq 3$ .

**Définition 3.7 :** *hyperface*

Les hyperfaces (ou facettes), de dimension  $n-1$ , d'un polytope de dimension  $n$ , sont définies par les relations

$$F_i = P \cap \{a_i x = z_i\}. \quad (3.16)$$

Les polytopes en dimension  $d$  constituent les hyperfaces des polytopes en dimension  $d+1$ . Les polytopes en dimension  $d$  pour  $0 \leq d \leq 3$  sont illustrés dans la fig. 3.17.

### 3.4.2 Simplexes

Un  $n$ -simplexe est une généralisation du triangle à  $n$  dimensions.

**Définition 3.8 :**  $n$ -simplexe

Un  $n$ -simplexe de l'espace affine de dimension  $n$  est l'enveloppe convexe de  $n + 1$  point formant un repère affine de cet espace.

De façon équivalente, c'est un  $n$ -polytope à  $n + 1$  points.

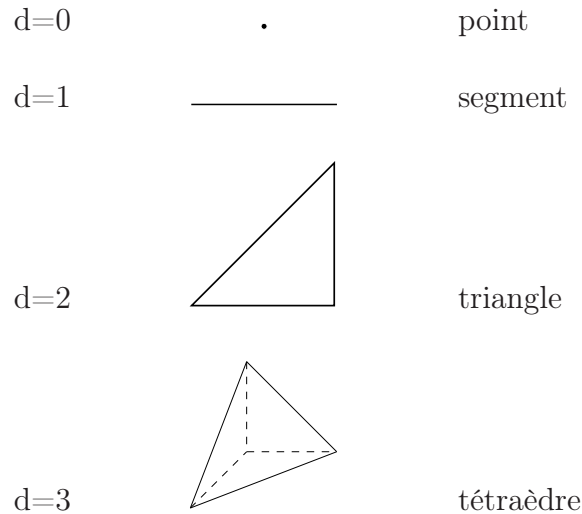


FIGURE 3.18 – Les  $n$ -simplexes pour  $0 \leq n \leq 3$ .

Les éléments d'un simplexe sont appelés ses  $n$ faces, où  $n$  est leur dimension. Ainsi

- ◇ les 0faces sont les sommets
- ◇ les 1faces sont les arêtes
- ◇ les 2faces sont les faces
- ◇ les 3faces sont les cellules.

L'ensemble des  $(n - 1)$ faces (au sens ci-dessus) d'un  $n$ -simplexe constituent son enveloppe. De façon générale, les  $i$ faces d'un  $n$ -simplexe ( $i \leq n$ ) sont elles-mêmes des simplexes de dimensions  $i \leq n$ .

#### Caractéristique d'Euler-Poincaré

Si l'on note  $N_i$  le nombre de  $i$ -face, alors de façon très remarquable, le tableau des  $N_i$  forme un triangle de Pascal, comme illustré par la table 3.3.

Pour un  $n$ -simplexe, ceci peut s'écrire sous la forme suivante :

$$\sum_{i=0}^{n-1} N_i (-1)^i = N_0 - N_1 + N_2 - N_3 + \cdots + N_{n-1} (-1)^{n-1}. \quad (3.17)$$

		nb de sommets	nb d'arêtes	nb de faces	nb de cellules	nb de 4faces	nb de 5faces	nb de 6faces
		$N_0$	$N_1$	$N_2$	$N_3$	$N_4$	$N_5$	$N_6$
0-simpl.	point	1						
1-simpl.	segment	2	1					
2-simpl.	triangle	3	3	1				
3-simpl.	tétraèdre	4	6	4	1			
4-simpl.	pentachore	5	10	10	5	1		
5-simpl.		6	15	20	15	6	1	
6-simpl.		7	21	35	35	21	7	1

TABLE 3.3 – Les différents simplexes et le nombre de leurs ifaces.

C'est la caractéristique d'Euler-Poincaré du  $n$ -simplexe. Elle vaut donc 2 pour  $n$  impair et 0 pour  $n$  pair.

*Exemples :*

◇ 0-simplexe (segment) :  $N_0 = 2$ .

◇ 1-simplexe (triangle) :  $N_0 - N_1 = 3 - 3 = 0$ .

◇ 2-simplexe (tétraèdre) :  $N_0 - N_1 + N_2 = 4 - 6 + 4 = 2$  : on retrouve ici un cas particulier de la caractéristique d'Euler-Descartes, ici pour le tétraèdre.

De façon plus générale, tout  $n$ -polytope peut se décomposer en  $n$ -simplexes.

Par exemple, un polyèdre se décompose en tétraèdres, qui sont des 3-simplexes. Le fait de coller deux tétraèdres  $a$  et  $b$  suivant un 2-simplexe (triangle) préserve la caractéristique d'Euler-Descartes :

$$\diamond N_0 = N_0(a) + N_0(b) - 3,$$

$$\diamond N_1 = N_1(a) + N_1(b) - 3,$$

$$\diamond N_2 = N_2(a) + N_2(b) - 2,$$

et donc  $N_0 - N_1 + N_2 = 2$  puisque  $N_0(a) - N_1(a) + N_2(a) = 2$  et  $N_0(b) - N_1(b) + N_2(b) = 2$ . Des considérations similaires pour des  $n$ -polytopes arbitraires permettent de montrer que la caractéristique d'Euler-Poincaré du  $n$ -simplexe est préservée (Henri Poincaré 1893).



# Chapitre 4

## Quelques éléments de théorie des groupes

### 4.1 Groupe

#### 4.1.1 Définitions

**Définition 4.1 :** *Groupe*

Un groupe est un couple  $(G, \cdot)$  constitué d'un ensemble  $G$  et d'une opération sur cet ensemble, qui à deux éléments  $a$  et  $b$  de  $G$  associe un élément  $a \cdot b$ .

Cette loi doit satisfaire à quatre axiomes :

**Loi de composition interne :**  $\forall a, b \in G, a \cdot b \in G$

**Associativité :**  $\forall a, b, c \in G, (a \cdot b) \cdot c = a \cdot (b \cdot c)$

**Existence d'un élément neutre :**  $\exists e \in G / \forall a \in G, e \cdot a = a \cdot e = a$ .

On dit que  $e$  est l'élément neutre du groupe.

**Existence d'un symétrique :**  $\forall a \in G, \exists \text{sym}(a)$  tel que  $a \cdot \text{sym}(a) = \text{sym}(a) \cdot a = e$ .

Selon le contexte, la loi  $\cdot$  sera notée différemment :

- ◇ lorsque la loi est notée additivement, on notera  $+$  cette loi, et le symétrique de  $a$  sera alors appelé opposé, et noté  $-$ . L'élément neutre sera alors noté  $0$ . Exemple : groupe  $(\mathbb{Z}, +)$ .
- ◇ lorsque la loi est notée multiplicativement, on notera  $\times$  cette loi, et le symétrique de  $a$  sera alors appelé inverse, et noté  $a^{-1}$ . L'élément neutre sera alors noté  $1$ . Exemple : pour  $n \in \mathbb{N}^*$ , le groupe des racines  $n$ -ièmes de l'unité, muni de la loi  $\times$ , est un groupe multiplicatif.
- ◇ on utilisera également les notations  $\circ$  et  $*$ .

**Définition 4.2 :** *Groupe abélien et non-abélien*

Si l'opération  $\cdot$  est commutative, on dira que le groupe est *abélien*. Dans le cas contraire, il est dit non-abélien.

**Définition 4.3 :** *Ordre d'un groupe*

L'*ordre* d'un groupe est le cardinal de  $G$ .

- ◇ Si ce cardinal est fini, on dira alors que  $G$  est un groupe fini (!) et on notera  $|G|$  son ordre.
- ◇ Dans le cas d'un groupe d'ordre infini, le groupe pourra être *discret* (au sens topologique, voir append. B), ou continu.

**4.1.2 Exemples de groupes**

- ◇ Le groupe trivial  $G = (\{0\}, +)$ , noté aussi 0, pour un groupe additif (ou  $G = (\{1\}, \times)$ , noté aussi 1, pour un groupe multiplicatif).
- ◇  $(\mathbb{Z}, +)$  est un groupe abélien discret.
- ◇  $U(1) = \{e^{i\theta} / \theta \in \mathbb{R}\}$ , le groupe des phases, est un groupe abélien.
- ◇ Si  $K$  est un corps (exemples :  $K = \mathbb{R}, \mathbb{C}$ ),  $(K, +)$  et  $(K^*, \times)$  sont des groupes abéliens.
- ◇ Pour  $n \in \mathbb{N}^*$ , l'ensemble des entiers modulo  $n$  muni de l'addition  $G = (\mathbb{Z}/n\mathbb{Z}, +)$  est un groupe abélien, d'ordre  $n$ .
- ◇  $G = (\mathcal{S}(E), \circ)$  :

Sur un ensemble  $E$ , l'ensemble  $\mathcal{S}(E)$  des bijections de  $E$  dans  $E$ , muni de la loi de composition  $\circ$  des applications est un groupe.

- ◇ *Groupe symétrique*  $\mathcal{S}_n$  :

Dans le cas particulier où  $E = \{1, \dots, n\}$ , on notera  $\mathcal{S}_n$  l'ensemble  $\mathcal{S}(E)$ , et on l'appellera le groupe symétrique à  $n$  éléments. Il est d'ordre  $n!$ , et non abélien pour  $n \geq 3$ , et il est constitué des permutations des  $n$  éléments  $\{1, \dots, n\}$ . Voir la section 4.8 pour son étude détaillée.

- ◇ *Groupe spécial linéaire*  $\text{GL}_n(K)$  :

L'ensemble des matrices inversibles  $n \times n$  à coefficients dans un corps  $K$ , muni du produit des matrices, est un groupe (non abélien pour  $n \geq 2$ ).

- ◇ *Groupe orthogonal*  $O(n)$  :

L'ensemble  $O(n)$  des matrices orthogonales  $n \times n$  est l'ensemble des matrices réelles vérifiant  $O \cdot O^t = O^t \cdot O = \mathbf{1}$ , qui laissent invariant la forme bilinéaire (le produit scalaire)

$$\vec{x} \cdot \vec{y} = \sum_{i=1}^n x_i y_i.$$

Muni du produit des matrices, c'est un groupe non-abélien pour  $n \geq 2$ . La relation de définition de  $O(n)$  impose que  $\det O = \pm 1$ .

◇ *Groupe spécial orthogonal*  $SO(n)$  :

Matrices de  $O(n)$  de déterminant  $+1$ .

◇ *Groupe unitaire*  $U(n)$  :

L'ensemble  $U(n)$  des matrices unitaires  $n \times n$  est constitué des matrices complexes vérifiant  $UU^\dagger = U^\dagger U = \mathbb{1}$ , qui laissent invariante la forme sesquilinéaire  $(x, y) = \sum_{i=1}^n x_i^* y_i$  (cette forme est sesquilinéaire à gauche : linéaire par rapport à  $y$ , antilinéaire par rapport à  $x$ , cf bra-ket en physique quantique).

Muni du produit des matrices, c'est un groupe non-abélien pour  $n \geq 2$ . La relation de définition de  $U(n)$  impose que  $\det U = \pm 1$ .

◇ *Groupe spécial unitaire*  $(SU(n), \cdot)$  :

Matrices de  $U(n)$  de déterminant  $+1$ .

◇ *Groupe symplectique*  $Sp(2n, K)$  :

Soit  $K$  un corps (typiquement  $\mathbb{R}$  ou  $\mathbb{C}$ ).

L'ensemble  $Sp(2n, K)$  des matrices symplectiques  $2n \times 2n$  est constitué des matrices dans  $K$  laissant invariante la forme antisymétrique  $x^t g y$  où  $g$  est la matrice antisymétrique

$$g = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}$$

*i.e.* satisfaisant la condition

$$S^t g S = g.$$

### 4.1.3 Groupes finis et table de Cayley

Pour tout ensemble fini  $E$  muni d'un loi  $\cdot$ , une table de Cayley est un tableau à double entrée, qui donne le résultat de  $a \cdot b$  pour tout  $a$  et  $b$  dans  $E$ , à la façon des tables de multiplication et d'addition. Ceci s'applique en particulier aux groupes finis, de sorte que les tables de Cayley permettent de faciliter leur étude.

#### Propriété 4.4 :

Soit  $G$  un groupe fini.

◇ Dans chaque ligne et chaque colonne de sa table de Cayley, chacun des éléments de  $G$  apparaît une et une seule fois. Autrement dit, chaque ligne, et chaque colonne d'un tableau de Cayley est une permutation des éléments du groupe.

◇ l'élément  $e$  apparaît de façon symétrique par rapport à la diagonale

◇ un groupe est abélien si et seulement si sa table de Cayley est symétrique par rapport à la diagonale.

Pour faciliter la classification (à isomorphisme près) des groupes finis à l'aide de leur table de Cayley, il est utile d'utiliser la convention suivante : on place d'abord dans chaque ligne et chaque colonne  $e$ , puis les éléments qui sont leur propre inverse, puis les éléments et leur inverse par paire. Classé de cette façon  $e$  apparaît soit sur la diagonale, soit juste en dessous ou juste en dessous.

### Exercice 4.1

Classifier les groupes finis d'ordres 1 à 4. On précisera leur caractère abélien ou non-abélien.

## 4.2 Morphismes de groupe, sous-groupe

### 4.2.1 Morphisme

**Définition 4.5 :** *Morphisme de groupe*

Soient  $G$  et  $G'$  deux groupes. Un morphisme (ou homomorphisme) entre ces deux groupes est une application  $f : G \rightarrow G'$  qui satisfait

$$\forall g_1, g_2 \in G, f(g_1 g_2) = f(g_1) f(g_2). \quad (4.1)$$

En d'autres termes, l'application  $f$  préserve les lois de groupe sur  $G$  et  $G'$ .

**Définition 4.6 :** *Isomorphisme de groupe*

Si en outre  $f$  est bijective, alors  $f^{-1}$  est aussi un morphisme de groupe. On dit alors que  $f$  est un isomorphisme. En d'autres termes, l'image du symétrique de  $g \in G$  est le symétrique dans  $G'$  de  $f(g)$ .

On dira alors que  $G$  et  $G'$  sont *isomorphes*, et l'on notera  $G \simeq G'$  ou encore  $G \cong G'$ .

**Définition 4.7 :** *Automorphisme de groupe*

Si  $G = G'$ , alors un isomorphisme de groupe  $f$  est appelé un automorphisme.

**Proposition 4.8 :**  $\text{Aut}(G)$

L'ensemble des automorphismes d'un groupe  $G$ , muni de la loi de composition  $\circ$  des applications, est un groupe noté  $\text{Aut}(G)$ .

**Définition 4.9 :** *Noyau d'un morphisme*

Le noyau d'un morphisme  $f$  est

$$\text{Ker}(f) = \{g \in G / f(g) = e\}. \quad (4.2)$$

**Définition 4.10 :** *Image d'un morphisme*

L'image d'un morphisme  $f$  est

$$\text{Im } f = \{f(g) / g \in G\}. \quad (4.3)$$

**Définition 4.11 :** *Morphisme injectif, morphisme surjectif*

- ◇ Un morphisme  $f$  est injectif si et seulement si  $\text{Ker } f = e$ .
- ◇ Un morphisme  $f$  est surjectif si et seulement si  $\text{Im } f = G'$ .
- ◇ Un morphisme est un isomorphisme si et seulement s'il est injectif et surjectif.

## 4.2.2 Sous-groupe

**Définition 4.12 :** *Sous-groupe*

Soit  $H$  un sous-ensemble de  $G$ . Un *sous-groupe*  $(H, \cdot)$  d'un groupe  $(G, \cdot)$  est un groupe dont la loi  $\cdot$  s'obtient par restriction de la loi  $\cdot$  sur  $H \times H$ . On note alors  $H \leq G$ .

On remarquera que la notation est en accord avec le fait que la notion de sous-groupe définit une relation d'ordre (le vérifier).

*Rappel :* Une relation d'ordre sur un ensemble  $E$  est une relation binaire  $\leq$  sur  $E$  qui est à la fois

**réflexive :** pour tout élément  $x$  de  $E$ , on a  $x \leq x$ .

**antisymétrique :**  $\forall x, y \in E, (x \leq y \text{ et } y \leq x) \Rightarrow x = y$

**transitive :**  $\forall x, y, z \in E, (x \leq y \text{ et } y \leq z) \Rightarrow x \leq z$ .

---

**Définition 4.13 :** Un sous-groupe sera dit *propre* s'il n'est pas identique à  $G$ . On note alors  $H < G$ .

On remarquera que la notation est en accord avec le fait que la notion de sous-groupe définit une relation d'ordre strict (le vérifier).

*Rappel :* A toute relation d'ordre  $\leq$  sur  $E$  on associe une relation d'ordre strict  $<$  sur  $E$  qui est la restriction de la relation d'ordre aux couples d'éléments distincts :

$$x < y \Leftrightarrow (x \leq y \text{ et } x \neq y).$$


---

**Proposition 4.14 :** Pour que  $H \subset G$  soit un sous-groupe, il faut et il suffit que  $\forall a, b \in H, a \cdot b^{-1} \in H$ .

*Exemples :*

- ◇ De façon évidente,  $G$  et  $\{e\}$  sont des sous-groupes de  $G$ .

- ◇ L'intersection d'une famille quelconque de sous-groupes d'un groupe  $G$  est un sous-groupe de  $G$ .
- ◇ La réunion de deux sous-groupes est un sous-groupe si et seulement si l'un des deux est inclus dans l'autre.
- ◇ Les sous-groupes de  $\mathbb{Z}$  sont les  $n\mathbb{Z}$  pour  $n \in \mathbb{N}$ .

*Remarque :*

On notera que le critère 4.14 permet de se dispenser de vérifier l'associativité, automatique par inclusion de  $H$  dans  $G$ .

### 4.2.3 Quelques sous-groupes particuliers

**Proposition 4.15 :** *Sous-groupe conjugué*

Si  $H$  est un sous-groupe de  $G$ , l'ensemble des éléments de  $G$  de la forme  $a^{-1} \cdot h \cdot a$  ( $a$  étant un élément fixé quelconque dans  $G$ ), que l'on notera  $a^{-1} \cdot H \cdot a$ , est un sous-groupe de  $G$ , appelé *sous-groupe conjugué* de  $H$ .

La preuve découle directement de la proposition 4.14.

**Définition 4.16 :** *Centre d'un groupe*

Le centre d'un groupe  $G$  est l'ensemble  $Z_G$  des éléments de  $G$  qui commutent avec tous les autres :

$$Z_G = \{z \in G / \forall g \in G, g \cdot z = z \cdot g\}. \quad (4.4)$$

**Propositions 4.17 :**

- ◇  $Z_G$  est un sous-groupe abélien de  $G$ .
- ◇ Le centre d'un groupe abélien est le groupe  $G$  tout entier.

**Définition 4.18 :** *Centralisateur d'un élément d'un groupe*

Soit  $G$  un groupe et  $a \in G$ . Le centralisateur  $Z_a$  de  $a$  dans  $G$  est l'ensemble des éléments de  $G$  qui commutent avec  $a$  :

$$Z_a = \{g \in G / g \cdot a = a \cdot g\}. \quad (4.5)$$

**Proposition :**  $Z_a$  est un sous-groupe de  $G$ .

**Définition 4.19 :** *Centralisateur d'une partie*

Plus généralement, le *centralisateur*  $Z_X$  d'une partie  $X$  d'un groupe  $G$  est l'ensemble des éléments de  $G$  qui commutent avec tous les éléments de  $X$  :

$$Z_X = \{g \in G / \forall x \in X, g \cdot x = x \cdot g\}. \quad (4.6)$$

**Propositions 4.20 :**

- ◇ Si  $X = \{x\}$ , alors  $Z_X = Z_x$ .
- ◇ Si  $X \neq \emptyset$ , alors  $Z_X = \bigcap_{x \in X} Z_x$ . Ainsi  $Z_X$  est l'intersection de sous-groupes de  $G$ , d'où l'on déduit que  $Z_X$  est lui-même un sous-groupe de  $G$ .
- ◇ Si  $X = G$ , alors de façon immédiate le centralisateur de la partie  $G$  de  $G$  est le centre de  $G$ .

Une condition de commutation globale plutôt qu'individuelle conduit à la notion de normalisateur.

**Définition 4.21 :** *Normalisateur d'une partie*

Le *normalisateur*  $N_X$  d'une partie  $X$  d'un groupe  $G$  est l'ensemble des éléments de  $G$  qui commutent globalement avec les éléments de  $X$  :

$$N_X = \{g \in G / g^{-1} \cdot X \cdot g = X\}, \quad (4.7)$$

ou de façon équivalente

$$N_X = \{g \in G / \forall x \in X, \exists y \in X, g^{-1} \cdot x \cdot g = y\}.$$

**Propositions 4.22 :**

- ◇ L'image et le noyau d'un morphisme  $f : G \rightarrow G'$  sont des sous-groupes de  $G'$  et  $G$  respectivement.

Plus généralement :

- ◇ L'image réciproque par  $f$  de tout sous-groupe de  $G'$  est un sous-groupe de  $G$ .
- ◇ L'image par  $f$  de tout sous-groupe de  $G$  est un sous-groupe de  $G'$ .

*Rappel :* Soit  $f$  une application de  $E$  dans  $F$ .

L'image (ou image directe) par  $f$  d'une partie  $A$  de  $E$  est le sous-ensemble  $B$  de  $F$  constitué des éléments qui ont au moins un antécédent par  $f$  dans  $A$  :

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

L'image réciproque par  $f$  d'une partie  $B$  de  $F$  est le sous-ensemble  $A$  de  $E$  des éléments dont l'image par  $f$  appartient à  $B$  :

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

Remarque : ne pas confondre cette application définie sur les parties de  $F$  avec la bijection réciproque de  $f$  (quand elle existe), notée également  $f^{-1}$ .

---

#### 4.2.4 Exemples de morphismes et sous-groupes associés

- ◇ Si  $a \in \mathbb{R}$ , alors l'application  $f_a : x \mapsto ax$  est un morphisme de  $(\mathbb{R}, +)$  sur lui-même (car  $f_a(x + y) = a(x + y) = ax + ay = f_a(x) + f_a(y)$ ). C'est un isomorphisme si et seulement si  $a \neq 0$ .

Ainsi  $a\mathbb{R}$  est un sous-groupe de  $(\mathbb{R}, +)$ . Les autres sous-groupes de  $(\mathbb{R}, +)$  sont denses dans  $\mathbb{R}$ .

En revanche, si  $G$  est un groupe et  $a \in G$ , alors l'application  $f_a : x \mapsto ax$  (appelée « translation à gauche ») est une bijection de  $G$  dans  $G$ . Mais en général, ce n'est pas un morphisme (car  $f_a(x \cdot y) = a \cdot (x \cdot y) = a \cdot x \cdot y$  diffère en général de  $f_a(x) \cdot f_a(y) = a \cdot x \cdot a \cdot y$ ).

- ◇ L'application

$$\begin{aligned} \text{Exp} : (\mathbb{C}, +) &\rightarrow (\mathbb{C}^*, \times) \\ z &\mapsto \exp z \end{aligned}$$

est un morphisme surjectif (mais non injectif puisque  $\exp(2in\pi) = 1$  pour  $n \in \mathbb{Z}$ ).

- ◇ Si  $E$  est un ensemble fini de cardinal  $n$ , alors  $\mathcal{S}(E) \cong \mathcal{S}_n$ .

Pour  $n \geq 2$ , le seul morphisme non trivial de  $\mathcal{S}_n$  vers  $\{\pm 1\}$  est la signature  $\epsilon$  de la permutation.

*Rappel* : Une transposition est un élément particulier de  $\mathcal{S}_n$ , qui laisse invariant  $\{1, \dots, n\}$  sauf deux entiers  $i$  et  $j$ , qui sont échangés. On note  $(ij)$  une telle permutation. Une permutation peut se décomposer (de façon non unique) en un produit de transpositions. Une permutation paire (resp. impaire) est une permutation qui peut être exprimée comme le produit d'un nombre pair (resp. impair) de transpositions. Par définition, la signature d'une permutation vaut 1 si elle est paire, et -1 si elle est impaire.

---

Le noyau de la signature est un sous-groupe de  $\mathcal{S}_n$ . C'est le *groupe alterné*  $\mathcal{A}_n$ .

- ◇ Soit  $\mathbb{K}$  un corps. Alors  $\det : \text{GL}_n(\mathbb{K}) \rightarrow \mathbb{K}^*$  est un morphisme.

Si  $E$  est  $\mathbb{K}$ -espace vectoriel de dimension  $n$ , alors  $\text{GL}_n(\mathbb{K}, \cdot) \cong (\text{GL}(E), \circ)$ , groupe des applications linéaires bijectives de  $E$  dans  $E$ .

Le noyau de  $\det$  est un sous-groupe de  $\text{GL}_n(\mathbb{K})$ , appelé le groupe spécial linéaire, et noté  $\text{SL}_n(\mathbb{K})$  (c'est donc l'ensemble des matrices de déterminant 1 à coefficients dans  $\mathbb{K}$ ).

### 4.3 Générateurs d'un groupe, groupe cyclique

#### Proposition 4.23 :

Soit  $A$  une partie d'un groupe  $G$ . Il existe un plus petit sous-groupe de  $G$  contenant  $A$ . On l'appelle sous-groupe engendré par  $A$  et on le note  $\langle A \rangle$ .



*Preuve :*

Il suffit de considérer l'intersection de l'ensemble des sous-groupes de  $G$  contenant  $A$ .

### Exercice 4.2

Soient  $H$  et  $F$  deux sous-groupes d'un groupe  $G$ .

1. Montrer que  $\langle H \cup F \rangle$  est l'ensemble des éléments de la forme  $h_1 f_1 \cdots h_r f_r$  ( $r \geq 1$ ) avec pour tout  $\forall i, h_i \in H$  et  $f_i \in F$ .
2. On note  $HF$  le sous-ensemble de  $G$  formé des éléments de la forme  $hf$  avec  $h \in H$  et  $f \in F$ .
  - a. Montrer que  $\langle HF \rangle = \langle H \cup F \rangle$ .
  - b. En général,  $HF$  est-il un sous-groupe de  $G$ ?

**Définition 4.24 :** *Partie génératrice*

Une partie  $A$  de  $G$  est une partie génératrice (on dit aussi qu'elle engendre  $G$ , ou que  $A$  est un ensemble de générateurs de  $G$ ), si  $\langle A \rangle = G$ .

**Définition 4.25 :** *Rang d'un groupe*

Le rang d'un groupe  $G$  est le plus petit cardinal d'une partie génératrice de  $G$ .

On dira que  $G$  est de type fini si  $G$  admet une partie génératrice finie, ou encore si son rang est fini. C'est bien sûr le cas des groupes finis.

**Définition 4.26 :** *Ordre d'un élément*

Soient  $G$  un groupe et  $g \in G$ . L'ordre de  $g$  est le plus petit entier  $n > 0$  tel que  $g^n = 1$ . Si  $g^n \neq 1$  pour tout  $n > 0$ , alors par définition  $g$  est d'ordre infini.

**Proposition 4.27 :**

Soient  $G$  un groupe et  $g \in G$ . Si  $\langle g \rangle$  est infini, il est isomorphe à  $\mathbb{Z}$ . S'il est de cardinal  $n$ , il est isomorphe à  $\mathbb{Z}/n\mathbb{Z}$ . Dans les deux cas, l'ordre de  $g$  est le cardinal de  $\langle g \rangle$  dans  $\mathbb{N}^* \cup \{\infty\}$ .

*Preuve :*

Supposons d'abord  $g$  d'ordre infini. Alors

$$\begin{aligned} \varphi : \mathbb{Z} &\rightarrow \langle g \rangle, \\ m &\mapsto g^m \end{aligned}$$

est un morphisme surjectif. Son noyau est trivial (car  $g$  est d'ordre infini, et  $g^m = 1$  est équivalent à  $g^{-m} = 1$  si  $m$  est un entier négatif) donc c'est un isomorphisme. Ainsi  $\mathbb{Z}$  est isomorphe à  $\langle g \rangle$ . Supposons maintenant  $g$  d'ordre  $n \in \mathbb{N}^*$ . Comme  $g^n = 1$ , l'application

$$\begin{aligned} \varphi : \mathbb{Z}/n\mathbb{Z} &\rightarrow \langle g \rangle, \\ \bar{m} &\mapsto g^m \end{aligned}$$

est bien définie, et par définition de  $g$  c'est un morphisme surjectif. Soit  $\bar{m} \in \text{Ker } \varphi$ . Effectuons la division euclidienne  $m = nq + r$  de  $m$  par  $n$  ( $0 \leq r < n$ ). On obtient  $g^r = 1$  d'où  $r = 0$  par définition de l'ordre. Ainsi  $\bar{m} = 0$  et  $\varphi$  est finalement un isomorphisme de  $\mathbb{Z}/n\mathbb{Z}$  sur  $\langle g \rangle$ .

**Définition 4.28 :** *Groupe monogène, groupe cyclique*

Un groupe est monogène s'il est engendré par un singleton, et cyclique s'il est de plus fini.

En notation additive, ceci signifie qu'il existe  $a \in G$  tel que tout élément de  $G$  peut s'exprimer sous forme d'un multiple de  $a$ , *i.e.*  $G = \{pa/p \in \{0, \dots, n-1\}\}$ . En notation multiplicative, ceci signifie qu'il existe  $a \in G$  tel que tout élément de  $G$  peut s'exprimer sous forme d'une puissance de  $a$ , *i.e.*  $G = \{a^p/p \in \{0, \dots, n-1\}\}$ .

De la proposition précédente découle alors le

**Théorème 4.29 :** *Théorème fondamental des groupes cycliques*

Tout groupe monogène infini est isomorphe à  $(\mathbb{Z}, +)$ .

Tout groupe cyclique d'ordre  $n$  est isomorphe à  $(\mathbb{Z}/n\mathbb{Z}, +)$ .

On notera de façon équivalente  $\mathbb{Z}/n\mathbb{Z}$  (par référence au groupe quotient de  $\mathbb{Z}$  par le sous-groupe des multiples de  $n$ , voir plus loin) ou encore  $\mathbb{Z}_n$  (par référence au groupe multiplicatif des racines  $n$ -ièmes de l'unité), ou  $C_n$ .

## 4.4 Classes, sous-groupes distingués, groupes quotient

### 4.4.1 Automorphismes intérieurs

Dans le groupe  $\text{Aut}(G)$  des automorphismes de  $G$ , l'un d'eux joue un rôle très important.

**Définition 4.30 :** *Automorphisme intérieur*

Soit  $g \in G$  fixé. L'application de *conjugaison* par  $g$ ,

$$\begin{aligned} \iota_g : G &\rightarrow G \\ x &\mapsto gxg^{-1}, \end{aligned} \tag{4.8}$$

est un automorphisme de  $G$ , appelé automorphisme intérieur de  $G$ .

On vérifie aisément que son inverse est  $\iota_{g^{-1}}$  défini par

$$\begin{aligned} \iota_{g^{-1}} : G &\rightarrow G \\ x &\mapsto g^{-1}xg. \end{aligned}$$

Si  $x$  et  $y$  sont deux éléments de  $G$  tel que  $x$  est l'image de  $y$  par un automorphisme intérieur, alors  $x$  et  $y$  sont dits conjugués. De même, deux sous-groupes de  $G$  qui sont isomorphes par automorphisme intérieur sont dits conjugués.

**Propositions 4.31 :**

◇ Le morphisme  $\iota$

$$\begin{aligned} \iota : G &\rightarrow \text{Aut}(G) \\ g &\mapsto \iota_g \end{aligned} \tag{4.9}$$

est un morphisme de groupes dont le noyau est le centre  $Z_G$  du groupe.

◊ L'image de  $G$  par  $\iota$  est le groupe des automorphismes intérieurs, noté  $\text{Int}(G)$ .

*Remarque :*

La notion d'automorphisme intérieur est très utile, car si  $G$  est muni de structures supplémentaires (exemple : groupe de Lie), les automorphismes intérieurs sont toujours des isomorphismes pour les structures considérées.

### 4.4.2 Classes de conjugaison d'un groupe

*Rappel :* Une relation d'équivalence sur un ensemble  $E$  est une relation binaire  $\sim$  sur  $E$  qui est à la fois

**réflexive :** pour tout élément  $x$  de  $E$ , on a  $x \sim x$ .

**symétrique :**  $\forall x, y \in E, x \sim y \Rightarrow y \sim x$

**transitive :**  $\forall x, y, z \in E, x \sim y$  et  $y \sim z \Rightarrow x \sim z$ .

---

Sur un groupe  $G$ , on définit la relation d'équivalence (le vérifier...) :

$a \sim b \Leftrightarrow \exists g \in G$  tel que  $a = gb g^{-1}$ . On dit alors que  $a$  et  $b$  sont conjugués, en accord avec la définition (4.30).

Cette relation d'équivalence réalise une partition de  $G$ , au sens où tout élément du groupe n'appartient qu'à une seule classe. En particulier, l'élément neutre constitue une classe à lui tout seul.

### 4.4.3 Classes par rapport à un sous-groupe

**Définition 4.32 :** *Classes à gauche*

Soit  $H$  un sous-groupe de  $G$ . Alors la relation  $x \sim y$  si et seulement si  $x^{-1}y \in H$  est une relation d'équivalence sur  $G$ . L'ensemble quotient s'appelle ensemble des classes à gauche selon  $H$ . Il est noté  $G/H$ . Ses éléments sont de la forme  $aH$  avec  $a \in G$ .

**Définition 4.33 :** *Classes à droite*

Soit  $H$  un sous-groupe de  $G$ . Alors la relation  $x \sim y$  si et seulement si  $xy^{-1} \in H$  est une relation d'équivalence sur  $G$ . L'ensemble quotient s'appelle ensemble des classes à droite selon  $H$ . Il est noté  $G \setminus H$ . Ses éléments sont de la forme  $Ha$  avec  $a \in G$ .

En particulier  $H$  est la classe de  $e$  (pour les deux types de classe).

Pour chacune des deux relations d'équivalence ci-dessus, on obtient ainsi une partition de  $G$  (ces deux partitions étant en général distinctes). Pour la preuve, voir ci-dessous.

**Théorème 4.34 :** *Théorème de Lagrange*

Si  $G$  est fini, l'ordre de tout sous-groupe  $H$  de  $G$  divise l'ordre de  $G$ .

*Preuve :*

Considérons par exemple les classes à gauche selon un sous-groupe  $H$  de  $G$ . Ces classes constituent une partition de  $G$ . En effet, soit  $x \in G$ . Alors de façon évidente,  $x \in xH$  puisque  $e \in H$ . D'autre part, si  $\exists y, z \in G$  tels que  $x \in yH \cap zH$ , alors  $\exists h, h' \in H$  tels que  $x = yh = zh'$ . Ainsi  $y = zh'h^{-1}$  et donc, puisque  $hh^{-1} \in H$ ,  $y$  et  $z$  sont dans la même classe, et  $yH = zH$ .

De plus, le cardinal de  $aH$  est le même que le cardinal de  $H$  puisque les translations à gauche sont des bijections de  $G$  sur  $G$ . Ceci prouve que  $G$  est partitionné en  $|G|/|H|$  classes, ayant chacune  $|H|$  éléments.

**Corollaire 4.35 :**

Dans un groupe fini  $G$ , l'ordre de tout élément est fini et divise l'ordre de  $G$ . En particulier, si  $m$  est l'ordre de  $G$ , alors  $\forall g \in G, g^m = 1$ .

*Preuve :*

La preuve est immédiate en utilisant le fait que  $\forall g \in G, \langle g \rangle$  est un sous-groupe de  $G$ , puis en appliquant théorème de Lagrange.

*Exemple :*

Pour  $\mathbb{Z}/6\mathbb{Z}$ , le tableau des ordres des éléments est donné par

$\bar{x}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$o(\bar{x})$	1	6	3	2	3	6

**Proposition 4.36 :** *Sous-groupes de  $\mathbb{Z}/n\mathbb{Z}$*

Soient  $n \in \mathbb{N}^*$  et  $d \in \mathbb{N}^*$  qui divise  $n$ . Alors  $G = \mathbb{Z}/n\mathbb{Z}$  possède un et un seul sous-groupe d'ordre  $d$ , cyclique d'ordre  $d$  et isomorphe à  $\mathbb{Z}/d\mathbb{Z}$ .

*Preuve :*

Posons  $m = n/d$  et considérons le sous-groupe engendré par  $\bar{m}$  :  $C_d = \langle \bar{m} \rangle$ . C'est un sous-groupe d'ordre  $d$  de  $G$ . En effet  $p\bar{m} = \bar{0}$  conduit à  $pm = kn$  ( $k \in \mathbb{N}^*$ ) et donc  $p = kd$ , ce qui montre que  $m$  est d'ordre  $d$  et donc que  $C_d = \{\bar{0}, \bar{m}, \dots, \overline{(d-1)m}\}$ .

D'autre part, si  $H$  est un sous-groupe d'ordre  $d$  de  $G$ , alors d'après le théorème de Lagrange,  $\forall x \in H, dx = \bar{0}$ , ce qui signifie que  $dx = kn$  et donc que  $x = km$ . On a donc  $x \in \langle \bar{m} \rangle$  ce qui prouve que  $H \subset \langle \bar{m} \rangle$ . Comme  $H$  et  $C_d$  sont de cardinal  $d$ , ceci montre que  $H = C_d$ .

**Proposition 4.37 :**

Le groupe  $(\mathbb{Z}/n\mathbb{Z}, +)$  est engendré par  $\bar{1}$ .

*Preuve :*

Il suffit de reprendre la preuve de la proposition précédente avec  $m = 1$ . Tout d'abord,  $\bar{1}$  est d'ordre  $n$  : en effet, l'ordre  $p$  ( $\leq n$ ) de  $\bar{1}$  vérifie  $p\bar{1} = \bar{0}$ . On a alors  $p\bar{1} = \overline{p} = \bar{0}$  soit  $n|p$  et donc  $p = n$ . D'après ce qui précède,  $\langle \bar{1} \rangle$  est donc un sous-groupe d'ordre  $n$  de  $\mathbb{Z}/n\mathbb{Z}$ , lui-même d'ordre  $n$ . Ces deux ensembles ont donc le même cardinal, d'où  $\langle \bar{1} \rangle = \mathbb{Z}/n\mathbb{Z}$ .

#### 4.4.4 Sous-groupe distingué

La distinction entre les classes à gauche et à droite disparaît dans le cas des classes par rapport à un sous-groupe distingué.

**Définition 4.38 :** *Sous-groupe distingué*

Un sous-groupe  $H$  est distingué dans  $G$  si et seulement s'il vérifie l'un des trois énoncés équivalents suivants :

- ◇  $H$  un sous-groupe tel que les deux relations d'équivalence définies ci-dessus soient identiques.
- ◇  $H$  est stable par tous les automorphismes intérieurs de  $G$  :  
 $\forall g \in G, \forall h \in H, g h g^{-1} \in H$ .
- ◇  $H$  est identique à tous ses sous-groupes conjugués.

On notera alors  $H \triangleleft G$ .

On parlera indifféremment de sous-groupe distingué ou sous-groupe normal ou sous-groupe invariant.

*Remarque :*

- ◇ Si  $G$  est abélien, tout sous-groupe de  $G$  est distingué.
- ◇  $\{e\}$  et  $G$  sont toujours des sous-groupes distingués de  $G$ .
- ◇ prendre garde au fait que  $\triangleleft$  n'est pas une relation transitive (ce n'est donc pas une relation d'ordre) : on peut avoir  $K \triangleleft H \triangleleft G$  sans que  $K \triangleleft G$  soit vraie.

#### Exercice 4.3

Reprendre l'exercice 4.2 dans le cas où  $H \triangleleft G$ . On montrera que  $\langle H \cup F \rangle = HF$ .

#### **Propositions 4.39 :**

Soit  $f : G \rightarrow G'$  un morphisme de groupes.

- ◇ Si  $H' \triangleleft G'$ , alors  $f^{-1}(H')$  est distingué dans  $G$ .
- ◇ Un cas particulier important est donné par  $\text{Ker}(f)$  qui est distingué dans  $G$ .
- ◇ Si  $H \triangleleft G$ , alors  $f(H)$  est distingué dans  $f(G)$ . En revanche, il n'y a pas de raison pour que  $f(H)$  soit distingué dans  $G'$  lui-même.
- ◇ L'intersection de deux sous-groupes distingués dans  $G$  est un sous-groupe distingué de  $G$ .

*Exemples :*

- ◇ Pour  $n \geq 2$ ,  $\mathcal{A}_n$ , le groupe alterné des permutations est distingué dans  $\mathcal{S}_n$ .
- ◇  $K$  étant un corps commutatif, alors  $\text{SL}_n(K)$  est distingué dans  $\text{GL}_n(K)$ .
- ◇ Le centre  $Z_G$  étant le noyau du morphisme  $\iota : G \rightarrow \text{Aut}(G)$ , il est donc distingué dans  $G$ .

### 4.4.5 Groupe quotient

La notion de sous-groupe distingué est centrale : elle permet en effet de munir  $G/H$  d'une structure de groupe.

#### Théorème 4.40 :

Soient  $G$  un groupe et  $H$  un sous-groupe distingué de  $G$ . Alors il existe une unique structure de groupe sur  $G/H$  telle que la surjection canonique

$$\begin{aligned} p : G &\rightarrow G/H \\ a &\mapsto \bar{a} = aH = Ha \end{aligned} \tag{4.10}$$

soit un morphisme de groupes.

On appelle le groupe  $G/H$  ainsi obtenu le groupe quotient de  $G$  par  $H$ .

*Preuve :*

Par morphisme, la loi de groupe sur  $G/H$  doit être définie par  $\bar{a}\bar{b} = \overline{ab}$ . Il nous faut donc vérifier que cette loi est bien définie, *i.e.* que  $\overline{ab}$  ne dépend pas du choix des représentants  $a$  et  $b$  dans les classes  $\bar{a}$  et  $\bar{b}$ . Si donc  $\bar{a}' = \bar{a}$  et  $\bar{b}' = \bar{b}$ , alors  $a' = h_1a$  et  $b' = bh_2$  pour  $h_1, h_2 \in H$ . On a donc  $a'b' = h_1(ab)h_2$ , soit encore  $a'b' \in H(abh_2) = (abh_2)H$ . Or  $h_2H = H$ , et donc  $a'b' \in (ab)H$ , ce qui permet de conclure que  $\bar{a}\bar{b} = \overline{a'b'}$ , *q.e.d.*

Il nous reste à montrer que l'on a bien une loi de groupe. Ceci résulte de la surjectivité de  $p$ , qui permet d'associer à deux éléments  $\bar{x}$  et  $\bar{y}$  de  $G/H$  deux antécédents  $x$  et  $y$  sur lesquels on peut faire agir la loi de groupe sur  $G$ . Il suffit ensuite d'utiliser le fait que  $p$  est un morphisme, *i.e.*  $p(xy) = p(x)p(y)$  pour conclure.

La multiplication sur les classes peut donc s'écrire

$$(aH) \cdot (bH) = (ab)H,$$

ou de façon équivalente  $\bar{a}\bar{b} = \overline{ab}$ . L'élément neutre de  $G/H$  est bien sûr  $\bar{e} = H$ .

*Remarques :*

On prendra garde au fait que  $G/H$  n'est pas un sous-groupe de  $G$ .

Dans la littérature,  $p$  est également appelé morphisme canonique ou projection canonique.

### 4.4.6 Exemples

- ◇ Le groupe  $\mathbb{Z}/n\mathbb{Z}$  est le quotient de  $\mathbb{Z}$  par le sous-groupe distingué  $n\mathbb{Z}$ . Ceci justifie *a posteriori* la notation utilisée depuis le début de ce chapitre pour le groupe additif  $(\mathbb{Z}/n\mathbb{Z}, +)$  des entiers modulo  $n$ .
- ◇ Le sous-groupe  $2\pi\mathbb{Z}$  est distingué dans le groupe  $(\mathbb{R}, +)$ . Le groupe quotient  $\mathbb{R}/2\pi\mathbb{Z}$  permet de définir les angles orientés.

#### Théorème 4.41 : de factorisation\*

Soit  $f : G \rightarrow G'$  un morphisme de groupe. Alors il existe un unique morphisme de groupes  $\tilde{f} : G/\text{Ker } f \rightarrow G'$  tel que  $f = \tilde{f} \circ p$  (où  $p$  est la surjection canonique de  $G$  sur  $G/\text{Ker } f$ ). De plus  $\tilde{f}$  est injectif d'image  $\text{Im } f$ , *i.e.*  $G/\text{Ker } f \cong \text{Im } f$ .

*Preuve :*

Notons tout d'abord que  $G/\text{Ker } f$  est bien un groupe puisque  $\text{Ker } f$  est distingué dans  $G$ .

On doit nécessairement définir  $\tilde{f}$  par  $\tilde{f}(\bar{a}) = f(a)$ , où l'on note comme plus haut  $\bar{a}$  la classe de  $a$  dans  $G/\text{Ker } f$ . Cette définition de  $\tilde{f}(\bar{a})$  a bien un sens car si  $\bar{a} = \bar{b}$  alors  $a = bn$  avec  $n \in \text{Ker } f$ , et donc  $f(a) = f(b)f(n) = f(b)$  par morphisme.

Si  $\bar{a}, \bar{b} \in G/\text{Ker } f$ , alors  $\tilde{f}(\bar{a}\bar{b}) = \tilde{f}(\overline{ab}) = f(ab) = f(a)f(b) = \tilde{f}(\bar{a})\tilde{f}(\bar{b})$  par morphisme de  $f$ , ce qui prouve que  $\tilde{f}$  est bien un morphisme.

Enfin, comme  $f = \tilde{f} \circ p$ , la surjectivité de  $p$  (*i.e.*  $\text{Im } p = G/\text{Ker } f$ ) implique que  $\text{Im } f = \text{Im } \tilde{f}$ . Reste l'injectivité de  $\tilde{f}$ , qui découle du fait que si  $\bar{a} \in \text{Ker } \tilde{f}$ , alors  $a \in \text{Ker } f$  et donc  $\bar{a} = e_{G/\text{Ker } f}$ .

*Remarque :*

Dans le cas où  $G$  est fini, une conséquence immédiate du théorème de factorisation est que  $|G| = |\text{Ker } f| |\text{Im } f|$ .

## 4.5 Suite exacte\*

### Définition :

Soient  $(G_i)_{i \in \mathbb{N}}$  des groupes et  $f_i : G_i \rightarrow G_{i+1}$  des morphismes de groupes. On dit que la suite

$$G_0 \xrightarrow{f_0} G_1 \xrightarrow{f_1} \dots \xrightarrow{f_{i-1}} G_i \xrightarrow{f_i} G_{i+1} \xrightarrow{f_{i+1}} \dots \quad (4.11)$$

est exacte si pour tout  $i$ ,  $\text{Im}(f_i) = \text{Ker}(f_{i+1})$ .

L'illustration ci-dessus correspond à une suite infinie à droite, donc indexée par  $\mathbb{N}$ . Cette indexation est arbitraire : on peut utiliser également  $-\mathbb{N}$  (suite exactes infinies à gauche), et  $\mathbb{Z}$  (suite infinie des deux côtés).

### 4.5.1 Exemples

Suivant le contexte, le groupe trivial (*i.e.* le singleton  $\{e\}$ ) est noté 0 ou 1.

◇ La suite  $0 \longrightarrow G_1 \xrightarrow{f} G_2$  est exacte si et seulement si  $f$  est injective.

En effet l'image de 0 est 0, donc par définition de la suite exacte, le noyau de  $f$  doit être égal à 0.

◇ La suite  $G_2 \xrightarrow{g} G_3 \longrightarrow 0$  est exacte si et seulement si  $g$  est surjective.

En effet, le noyau du morphisme qui envoie  $G_3$  sur 0 étant égal à  $G_3$ , par définition de la suite exacte l'image de  $G_2$  par  $g$  doit être égale à  $G_3$ , d'où la surjectivité.

◇ La suite  $0 \longrightarrow G_1 \xrightarrow{f} G_2 \longrightarrow 0$  est exacte si et seulement si  $f$  est un isomorphisme.

Ce dernier résultat est une conséquence immédiate des deux résultats précédents.

### 4.5.2 Suites exactes courtes

L'exemple le plus important de suite exacte est celui des suites exactes à trois termes appelées aussi suites exactes courtes, qui sont les suites de la forme

$$0 \longrightarrow A \xrightarrow{q} B \xrightarrow{r} C \longrightarrow 0. \quad (4.12)$$

Quand tous les groupes sont abéliens et notés additivement, on écrira 0 au lieu de 1 dans une suite exacte courte.

### 4.5.3 Extension de groupes

Une extension d'un groupe  $H$  par un groupe  $N$  est un groupe  $G$  qui s'insère dans une suite exacte courte,

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{p} H \longrightarrow 1, \quad (4.13)$$

comme illustré dans la figure (4.1). Donc si  $G$  est une extension de  $H$  par  $N$ , alors à isomor-

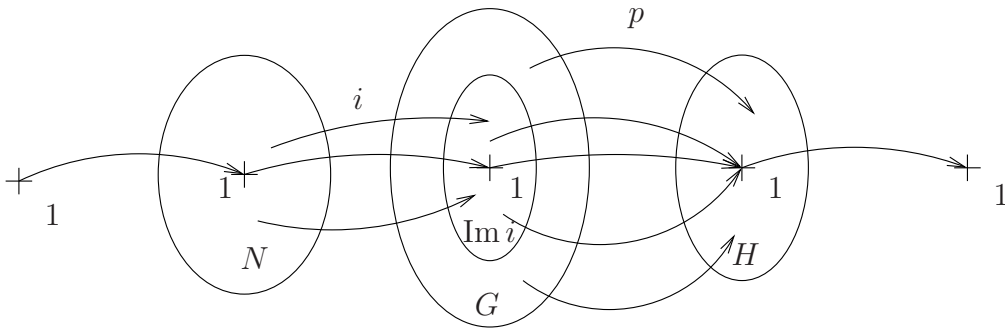


FIGURE 4.1 – Suite exacte courte.

phisme près,  $N$  est un sous-groupe normal de  $G$  et  $H$  est le groupe quotient  $G/N$ .

*Preuve :*



La suite est exacte si et seulement si  $i$  est injectif et  $p$  est surjectif. D'autre part,  $\text{Ker } i = 1$ ,  $\text{Im } i = \text{Ker } p$  et  $\text{Im } p = H$ . D'après le théorème de factorisation,  $p$  étant un morphisme de  $G$  sur  $H$ , il existe un morphisme injectif  $\tilde{p}$  de  $G/\text{Ker } p$  sur  $H$ , qui est surjectif puisque  $\text{Im } p = H$  (par surjection de  $p$ ). D'autre part, le morphisme injectif  $i$  est un isomorphisme de  $N$  sur son image  $\text{Im } i = \text{Ker } p$ . Il existe donc un isomorphisme de  $G/N$  sur  $H$ .

Les suites exactes courtes donnent donc un moyen compact d'écrire un groupe, un sous-groupe distingué et le quotient. Cette terminologie est beaucoup utilisée par exemple pour étudier le problème dit de l'extension : deux groupes  $H$  et  $N$  étant donnés, on cherche à déterminer tous les groupes  $G$  qui possèdent un sous-groupe distingué isomorphe à  $N$  avec un groupe quotient  $G/N$  isomorphe à  $H$ . Un tel groupe  $G$  est alors appelé une extension de  $H$  par  $N$ , ce qui explique la terminologie.

**Définition :** *Extension centrale*

L'extension est dite centrale si  $N$  est inclus dans le centre de  $G$ .

L'extension de  $H$  par  $N$  la plus simple est l'extension dite triviale qui consiste simplement à prendre le produit direct  $N \times H (= H \times N)$ . Ceci fournit au passage un moyen mnémotechnique pour le vocable «  $G$  est une extension de  $H$  par  $N$  », au sens  $G = H \times N$ .

**Définition :** *Section*

Une section de l'extension (4.13) est un morphisme  $s : H \rightarrow G$  tel que  $p \circ s = \text{id}_H$ . On dit alors que l'extension de  $H$  par  $N$  est scindée.

## 4.6 Groupe opérant sur un ensemble

**Définition 4.42 :** *Actions de groupe*

Un groupe  $G$  et un ensemble  $X$  étant donné, on dit que  $G$  agit sur  $X$  à gauche<sup>1</sup> s'il existe une application

$$\begin{aligned} G \times X &\rightarrow X \\ (g, x) &\mapsto g \cdot x \end{aligned}$$

avec

- ◇  $\forall x \in X, e \cdot x = x$  où  $e$  est l'élément neutre de  $G$ .
- ◇  $\forall x \in X, \forall g_1, g_2 \in G, g_1 \cdot (g_2 \cdot x) = (g_1 \cdot g_2) \cdot x$ .

*Remarque :*

---

1. On peut également introduire la notion d'action à droite :  $\forall g \in G, x \mapsto x \cdot g$ , qui satisfait  $X \cdot (g_1 g_2) = (x \cdot g_1) \cdot g_2$ .

La définition précédente implique immédiatement que  $\forall g \in G$ , l'application  $x \mapsto g \cdot x$  est une bijection de  $X$  sur  $X$ , de réciproque  $x \mapsto g^{-1} \cdot x$ . Ceci permet donc de donner la définition équivalente suivante :

**Définition 4.43 :**

Une action de groupe est un morphisme de groupes<sup>2</sup>

$$\begin{aligned} \Phi : G &\rightarrow \mathcal{S}(X) \\ g &\mapsto \Phi_g \end{aligned} \tag{4.14}$$

*Exemples :*

- ◇ Le groupe  $\mathcal{S}(X)$  agit sur  $X$ . Un exemple simple est fourni par  $\mathcal{S}_n$  qui agit sur l'ensemble  $\{1, \dots, n\}$
- ◇  $G$  opère sur lui-même par conjugaison :  $g \cdot x = g x g^{-1}$ . On notera que  $\Phi(G) \subset \text{Aut}(G)$ , condition plus forte que celle de la définition 4.43.
- ◇ Un groupe  $G$  opère sur lui-même par translations à gauche (appelée également action à gauche), suivant la relation  $g \cdot x = g x$ . Plus généralement, tout sous-groupe  $H$  de  $G$  opère par translations à gauche. De façon similaire, on définit l'action à droite.
- ◇ Si  $\mathbb{K}$  est un corps, alors le groupe  $\text{GL}_n(\mathbb{K})$  opère sur  $\mathbb{K}^n$ .
- ◇ Un cas très important d'action de groupe sur un ensemble (qui généralise l'exemple précédent) correspond à la situation où  $X$  est un  $\mathbb{K}$ -espace vectoriel, et  $\Phi$  un morphisme de  $G$  sur  $\text{GL}(X)$ . On dira alors que  $(X, \Phi)$  est une représentation linéaire de  $G$  dans  $X$ .

**Définition 4.44 :** *Orbite d'un élément*

L'orbite  $\mathcal{O}(x)$  d'un élément  $x$  de  $X$  est l'ensemble des éléments  $g \cdot x$  pour  $g$  parcourant  $G$ .

Si l'on introduit la relation d'équivalence  $x \sim y \Leftrightarrow \exists g \in G$  tel que  $y = g \cdot x$ , alors les orbites sont les classes d'équivalence pour cette relation.

On dira que  $X$  est un espace homogène pour l'action du groupe  $G$  s'il possède une seule orbite. On dit également que  $G$  agit transitivement sur  $X$ .

L'action à gauche est un exemple d'action transitive, de même que l'action à droite.

**Définition 4.45 :** *Groupe d'isotropie d'un élément* (également appelé petit groupe ou stabilisateur)

Pour tout  $x \in X$ , le groupe d'isotropie  $S(x)$  est l'ensemble des  $g$  de  $G$  qui laissent  $x$  invariant, *i.e.* vérifiant  $g \cdot x = x$ .

On notera que  $S(x)$  n'est pas distingué en général.

**Définition 4.46 :** *Action fidèle*

---

2. Dans le cas d'une action à droite, on introduira alors un anti-morphisme de  $G$  sur  $\mathcal{S}(X)$  au lieu d'un morphisme.

Une action est fidèle si le seul élément de  $G$  qui stabilise tous les éléments de  $X$  est l'élément neutre  $e$ .

**Définition 4.47 :** *Action libre*

Une action est libre si tous les stabilisateurs se réduisent à  $\{e\}$ .

*Exemple :* translation à gauche.

## 4.7 Produit semi-direct

Si  $G_1$  et  $G_2$  sont deux groupes, le *produit direct*  $G_1 \times G_2$  possède une structure de groupe en le munissant de la loi  $(g_1, g_2)(h_1, h_2) = (g_1g_2, h_2h_1)$ .

Le produit semi-direct est une généralisation de cette notion. Considérons  $N$  et  $H$  deux groupes, et soit  $\varphi : H \rightarrow \text{Aut } N$  un morphisme de groupes, qui permet en particulier de définir une action de  $H$  sur  $N$  par

$$h \cdot n = \varphi(h)(n), \quad (4.15)$$

avec une hypothèse plus forte que celle d'une action de groupe habituelle puisque l'on exige que  $\varphi(h) \in \text{Aut } N$ .

On définit une loi de groupes sur l'ensemble produit  $N \times H$  en posant

$$(n, h) \times (n', h') = (n(h \cdot n'), h h').$$

Le groupe ainsi obtenu est le produit semi-direct de  $N$  par  $H$ , pour l'action  $\varphi$ . On le notera  $N \rtimes_{\varphi} H$ , ou plus simplement  $N \rtimes H$  si l'action  $\varphi$  est implicite. On dit aussi que  $N \rtimes_{\varphi} H$  est le produit semi-direct de  $H$  agissant dans  $N$ .

*Preuve :*

Mettons tout d'abord en œuvre les notations :

$$(n, h) \times (n', h') = (n(h \cdot n'), h h') = (n(\varphi(h)(n')), h h').$$

- ◇ L'élément neutre est simplement  $(1, 1)$ . En effet  $(n, h) \times (1, 1) = (n(h \cdot 1), h) = (n \cdot 1, h) = (n, h)$  puisque  $h \cdot 1 = 1$ , étant donné que l'action est à valeur dans  $\text{Aut } N$ , et donc  $h \cdot 1 = \varphi(h)(1) = 1$ .
- ◇ L'inverse de  $(n, h)$  est  $(h^{-1} \cdot n^{-1}, h^{-1})$ . En effet

$$(n, h)(h^{-1} \cdot n^{-1}, h^{-1}) = (n(h \cdot (h^{-1} \cdot n^{-1})), 1).$$

Or

$$\begin{aligned} h \cdot (h^{-1} \cdot n^{-1}) &= \varphi(h)(\varphi(h^{-1})(n^{-1})) = (\varphi(h) \circ \varphi(h^{-1}))(n^{-1}) \\ &= \varphi(h h^{-1})(n^{-1}) = \varphi(1)(n^{-1}) = n^{-1}, \end{aligned}$$

les troisième et cinquième égalités résultant du fait que  $\varphi$  est un morphisme de groupe. On en déduit donc immédiatement que

$$(n, h)(h^{-1} \cdot n^{-1}, h^{-1}) = (n n^{-1}, 1) = (1, 1).$$

◇ L'associativité est un peu subtile. D'une part,

$$[(n_1, h_1)(n_2, h_2)](n_3, h_3) = (n_1(h_1 \cdot n_2), h_1 h_2)(n_3, h_3) = (n_1(h_1 \cdot n_2)[(h_1 h_2) \cdot n_3], h_1 h_2 h_3).$$

D'autre part

$$(n_1, h_1)[(n_2, h_2)(n_3, h_3)] = (n_1, h_1)(n_2(h_2 \cdot n_3), h_2 h_3) = (n_1[h_1 \cdot (n_2(h_2 \cdot n_3))], h_1 h_2 h_3)$$

Il nous reste à comparer  $(h_1 \cdot n_2)[(h_1 h_2) \cdot n_3]$  et  $h_1 \cdot (n_2(h_2 \cdot n_3))$ .

$$\begin{aligned} h_1 \cdot (n_2(h_2 \cdot n_3)) &= \varphi(h_1)(n_2 \varphi(h_2)(n_3)) = \varphi(h_1)(n_2) \varphi(h_1)(\varphi(h_2)(n_3)) \\ &= \varphi(h_1)(n_2) \varphi(h_1 h_2)(n_3). \end{aligned}$$

La deuxième égalité résulte du fait que  $\varphi(h_1)$  est un automorphisme de  $N$ , tandis que la troisième est due au fait que  $\varphi$  est un morphisme de groupe. De  $(h_1 \cdot n_2)[(h_1 h_2) \cdot n_3] = \varphi(h_1)(n_2) \varphi(h_1 h_2)(n_3)$  on tire finalement l'égalité recherchée.

---

*Remarque* : on notera que le produit direct est un cas particulier de la construction ci-dessus, correspondant à l'action triviale  $\varphi(h) = \text{id}$ .

### 4.7.1 Exemples

#### Groupe des isométries

Le groupe euclidien  $E(n)$  (noté également  $Is(n)$ ,  $ISO(n)$  ou  $Isom(n)$ ) à  $n$  dimensions est par définition le groupe des isométries dans l'espace affine euclidien  $\mathbb{R}^n$ , c'est-à-dire l'ensemble des transformations affines de  $\mathbb{R}^n$  qui conservent le produit scalaire, ou de façon équivalente qui préservent les longueurs.

C'est le produit semi-direct de  $O(n)$  agissant sur le groupe des translations  $\mathbb{T}(n)$ . Le sous-groupe  $E^+(n)$  des isométries directes préserve en outre l'orientation. C'est le produit semi-direct de  $SO(n)$  agissant sur le groupe des translations  $\mathbb{T}(n)$ . Ses éléments sont appelés déplacements.

La loi de groupe sur le produit semi-direct de  $O(n)$  (constitué des matrices orthogonales  $A$ ) agissant sur  $\mathbb{T}(n)$  (constitué des translations  $\tau_u$  de vecteur  $u$ ) s'écrit

$$(u, A) \cdot (v, B) = (u + Av, AB) \tag{4.16}$$

comme on le vérifie facilement en composant successivement la transformation  $x_i \mapsto v_i + B_{ij}x_j$  suivie de la transformation  $x_i \mapsto u_i + A_{ij}x_j$  (avec les conventions d'Einstein de sommation sur les indices). L'étude détaillée de ce groupe fait l'objet du chap.2.

On notera qu'à deux dimensions, alors que  $O(2)$  et  $\mathbb{T}(2)$  sont tous deux abéliens, leur produit semi-direct ne l'est pas.

#### Groupe de Galilée

Le groupe de Galilée est le groupe d'invariance d'espace-temps de la physique classique. Il possède comme sous-groupe les translations d'espace temps, les rotations, et les transformations

propres de Galilée, encore appelées transformations spéciales de Galilée. Rappelons que les transformations propres, caractérisées par un vecteur  $\vec{v} \in \mathbb{R}^3$ , sont données par

$$t' = t \quad (4.17)$$

$$\vec{r}' = \vec{r} - \vec{v}t. \quad (4.18)$$

Le sous-groupe de Galilée homogène s'obtient en enlevant les translations d'espace-temps. Si l'on se restreint au sous-groupe de Galilée propre (on conserve l'orientation de l'espace) et orthochrone (on conserve le sens du temps), alors ce sous-groupe de Galilée homogène propre et orthochrone est le produit semi-direct de  $SO(3)$  agissant sur le sous-groupe invariant (à vérifier...) des transformations propres, lui-même isomorphe à  $\mathbb{R}^3$ . Le groupe de Galilée propre et orthochrone lui-même (c'est la composante connexe de l'identité du groupe de Galilée, voir plus loin le chap. 12 pour ce concept) est le produit semi-direct du sous-groupe de Galilée homogène propre et orthochrone agissant sur le sous-groupe invariant (à vérifier de même) des translations d'espace-temps.

On peut donc écrire :

$$SGal(3) \cong \mathbb{R}^4 \rtimes (\mathbb{R}^3 \rtimes SO(3)).$$

Le groupe de Galilée complet s'obtient en ajoutant le renversement du temps  $T$  ( $t \mapsto -t$ ) et la parité  $P$  ( $\vec{x} \mapsto -\vec{x}$ ).

### Groupe de Poincaré

Le groupe de Poincaré  $\mathcal{P}$  est le groupe d'invariance d'espace-temps de la physique relativiste. La composante connexe de l'identité de  $\mathcal{P}$  est le produit semi-direct du groupe de Lorentz restreint engendré par les transformations de Lorentz pures et les rotations agissant sur le sous-groupe invariant des translations :

$$\mathcal{P}_+^\uparrow \cong \mathbb{R}^4 \rtimes SO(3, 1).$$

## 4.8 Le groupe symétrique $\mathcal{S}_n$

Ce groupe joue un rôle essentiel dans de nombreuses branches des mathématiques et de la physique. Historiquement, c'est le premier exemple de groupe qui a été introduit, par Evariste Galois dans son étude du groupe des permutations des racines d'un polynôme. Un théorème important, dû à Cayley que nous énoncerons plus loin assure que tout groupe fini d'ordre  $n$  est isomorphe à un sous-groupe de  $\mathcal{S}_n$ . Enfin, il joue un rôle central dans l'étude des représentations de nombreux groupes. Ceci justifie que nous lui consacrons une étude spécifique.

Rappelons tout d'abord que  $|\mathcal{S}_n| = n!$  comme il est facile de le montrer en dénombrant le nombre de permutations de  $n$  objets.

### 4.8.1 Notions élémentaires sur les permutations

#### Notation

Une permutation ou substitution est par définition une bijection  $j \mapsto \sigma(j) = i_j$ , notée conventionnellement

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \quad (4.19)$$

ou encore  $(i_1 i_2 \cdots i_n)$  avec à nouveau  $\sigma(j) = i_j$ .

**Définition 4.48 :** *Support d'une permutation*

Le support d'une permutation  $\sigma$  est l'ensemble des indices  $i$  tels que  $\sigma(i) \neq i$ .

#### Composition, structure de groupe

Les permutations se composent en écrivant  $\sigma'\sigma(j) = \sigma'(\sigma(j)) = \sigma'(i_j)$ .

On a ainsi par exemple

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad (4.20)$$

soit encore  $(2, 3, 1)(2, 1, 3) = (3, 2, 1)$  et

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad (4.21)$$

soit encore  $(2, 1, 3)(2, 3, 1) = (1, 3, 2)$ . Cet exemple montre la non commutativité de la loi de composition dans le cas général. L'associativité est immédiate.

La permutation identité est  $\text{Id} = (1, 2, \cdots, n)$ .

Enfin l'inverse de  $\sigma$  est défini par

$$\sigma^{-1} = \begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ 1 & 2 & \cdots & n \end{pmatrix}. \quad (4.22)$$

L'ensemble  $\mathcal{S}_n$  muni de cette loi de composition est un groupe multiplicatif.

#### Transposition

Une transposition est un élément particulier de  $\mathcal{S}_n$ , qui laisse invariant  $\{1, \cdots, n\}$  sauf deux entiers  $i$  et  $j$ , qui sont échangés. On note  $(ij)$  une telle permutation.

#### Permutation circulaire

Soit  $\sigma \in \mathcal{S}_n$ . C'est un  $k$ -cycle ou une permutation circulaire d'ordre  $k$  s'il existe des indices  $i_1, i_2, \cdots, i_k$  tels que  $\sigma$  soit de support  $\{i_1, i_2, \cdots, i_k\}$  et  $\sigma(i_1) = i_2, \sigma(i_2) = i_3, \cdots, \sigma(i_k) = i_1$ , autrement dit  $\sigma$  permute circulairement les indices  $i_1, i_2, \cdots, i_k$  et laisse les autres indices inchangés.

On notera  $(i_1, i_2, \dots, i_k)$  une telle permutation circulaire.

**Propriété 4.49 :**

Soit  $c_k$  un  $k$ -cycle. Alors

$$\forall p \text{ tel que } 1 \leq p \leq k-1, \quad c_k^p \neq \text{Id}, \quad c_k^k = \text{Id}, \quad c_k^{-1} = c_k^{k-1}.$$

On en déduit immédiatement que  $c_k$  engendre un sous-groupe de  $\mathcal{S}_n$  isomorphe au groupe cyclique multiplicatif  $\mathbb{Z}_k$ .

*Remarque :*

On prendra garde au fait que pour  $p$  fixé,  $c_k^p$  n'est pas une permutation circulaire en général. Ainsi, par exemple  $c = (1, 2, 3, 4)$  est tel que  $c^2 = (1, 3)(2, 4)$ . On vérifiera que  $c^p$  est une permutation circulaire si et seulement si  $p$  et  $k$  sont premiers entre eux.

## 4.8.2 Décomposition en produits de cycles

**Définition 4.50 :** *Orbite*

Soit une permutation  $\sigma \in \mathcal{S}_n$ . L'orbite d'un indice est l'ensemble des images successives de cet indice par action répétée de  $\sigma$ .

**Théorème 4.51 :**

Toute permutation  $\sigma \in \mathcal{S}_n$  se décompose en un produit commutatif de cycles.

*Preuve :*

Considérons en particulier l'orbite de 1. Il existe un entier  $k$  tel que  $\sigma^k(1) = 1$  (avec  $1 \leq k \leq n$ ). Alors  $(1, \sigma(1), \dots, \sigma^{k-1}(1))$  est un  $k$ -cycle.

Si  $k = n$ ,  $\sigma$  est un  $n$ -cycle et le résultat est démontré.

Si  $k < n$ , soit  $\alpha$  le plus petit entier n'appartenant pas à  $1, \sigma(1), \dots, \sigma^{k-1}(1)$ . On peut alors réitérer l'étape précédente, en considérant l'orbite de  $\alpha$ , ce qui permet d'identifier un nouveau cycle, qui commute avec le précédent puisque les ensembles d'indices sont disjoints.

L'algorithme s'arrête au bout d'au maximum  $n$  étapes (le cas où il faut  $n$  étapes correspondant à l'existence de  $n$  de longueur 1, *i.e.*  $\sigma = \text{Id}$ !).

Par cet algorithme, il est clair qu'on parvient ainsi à décomposer toute permutation en un produit commutatif de cycles.

---

L'algorithme qui précède conduit à une écriture particulière de la décomposition en produit de cycle, appelée la décomposition canonique. C'est par convention l'écriture d'une permutation en produit de cycle où chaque cycle est ordonné en plaçant en premier le plus petit indice, puis en ordonnant les différents cycles par rapport à leur premier élément.

**Propriété 4.52 :**

Supposons qu'une permutation  $\sigma$  a été décomposée en un produit de cycles disjoints de longueurs  $k_1, k_2, \dots, k_p$  avec  $\sum k_i = n$ . Alors pour toute permutation  $\tau$ , la permutation conjuguée  $\tau\sigma\tau^{-1}$  possède aussi une décomposition en cycles de longueur  $k_1, k_2, \dots, k_p$ .

Réciproquement, si deux permutations  $\sigma$  et  $\sigma'$  ont une décomposition en cycles de longueurs  $k_1, k_2, \dots, k_p$ , elles sont équivalentes, au sens où  $\exists \tau$  tel que  $\sigma' = \tau^{-1}\sigma\tau$ . Les classes d'équivalence de  $\mathcal{S}_n$  par la relation de conjugaison sont donc caractérisées par la décomposition en cycles de longueur donnée.

*Preuve :*

Soit  $\sigma$  une permutation décomposée en produit de cycles disjoints, i.e.  $\sigma = c_1 \cdots c_p$ , où  $c_i$  est de longueur  $k_i$ .

Soit  $\tau$  une permutation quelconque. Alors  $\tau\sigma\tau^{-1} = \tau c_1 \tau^{-1} \tau c_2 \tau^{-1} \cdots \tau c_p \tau^{-1}$ . Or chaque permutation  $\tau c_i \tau^{-1}$  conjuguée du cycle  $c_i = (a_1, \dots, a_{k_i})$  est égale au cycle  $(\tau(a_1), \dots, \tau(a_{k_i}))$  puisque

$$\tau = \begin{pmatrix} a_1 & \cdots & a_{k_i} & a_{k_i+1} & \cdots & a_n \\ \tau(a_1) & \cdots & \tau(a_{k_i}) & a_{k_i+1} & \cdots & a_n \end{pmatrix} \quad \text{et} \quad \tau^{-1} = \begin{pmatrix} \tau(a_1) & \cdots & \tau(a_{k_i}) & a_{k_i+1} & \cdots & a_n \\ a_1 & \cdots & a_{k_i} & a_{k_i+1} & \cdots & a_n \end{pmatrix}$$

d'où

$$\begin{aligned} \tau c_i \tau^{-1} &= \\ & \begin{pmatrix} a_1 & \cdots & a_{k_i} & a_{k_i+1} & \cdots & a_n \\ \tau(a_1) & \cdots & \tau(a_{k_i}) & a_{k_i+1} & \cdots & a_n \end{pmatrix} \begin{pmatrix} a_1 & \cdots & a_{k_i} & a_{k_i+1} & \cdots & a_n \\ a_2 & \cdots & a_1 & a_{k_i+1} & \cdots & a_n \end{pmatrix} \begin{pmatrix} \tau(a_1) & \cdots & \tau(a_{k_i}) & a_{k_i+1} & \cdots & a_n \\ a_1 & \cdots & a_{k_i} & a_{k_i+1} & \cdots & a_n \end{pmatrix} \\ &= \begin{pmatrix} \tau(a_1) & \tau(a_2) & \cdots & \tau(a_{k_i}) & a_{k_i+1} & \cdots & a_n \\ \tau(a_2) & \tau(a_3) & \cdots & \tau(a_1) & a_{k_i+1} & \cdots & a_n \end{pmatrix}, \end{aligned}$$

et donc  $\tau\sigma\tau^{-1}$  est le produit des cycles  $c'_i = \tau c_i \tau^{-1}$ , chacun de longueur  $k_i$ , et qui sont disjoints.

Réciproquement, soient  $\sigma = c_1 \cdots c_p$  et  $\sigma' = c'_1 \cdots c'_p$  deux permutations qui se décomposent en produit de cycles disjoints  $c_i$  et  $c'_i$ , chacun de même longueur. Les supports des cycles  $c_i$  étant disjoints, on peut construire au moins une permutation qui envoie le support de chaque  $c_i$  sur le support du cycle correspondant  $c'_i$  (cette permutation n'est bien sûr pas unique, puisque d'une part pour chaque  $c_i$  toute permutation circulaire est autorisée, et d'autre part l'ordre des  $c_i$  importe peu), de sorte que chaque  $c_i$  est conjugué de  $c'_i$ , cf l'argument utilisé ci-dessus pour la preuve dans le sens direct, d'où finalement le fait que  $\sigma$  et  $\sigma'$  sont conjugués.

On caractérisera donc une classe de  $\mathcal{S}_n$  par la longueur de ses cycles : en notant  $\nu_j$  le nombre de cycle de longueur  $j$  ( $j = 1$  correspond aux indices invariants,  $j = 2$  aux transpositions), on aura  $\sum j \nu_j = n$ . On notera  $\underline{\nu} = [1^{\nu_1} 2^{\nu_2} \cdots]$  cette classe.

On peut montrer que dans une classe  $\underline{\nu}$  donnée, il y a

$$n_{\underline{\nu}} = \frac{n!}{\prod \nu_j! j^{\nu_j}} \quad (4.23)$$

permutations distinctes.

#### Exercice 4.4

Démontrer la relation (4.23).



**Propriété 4.53 :**

En accord avec la définition donnée plus haut, l'ordre d'une permutation  $\sigma$  est le plus petit entier  $p$  tel quel  $\sigma^p = \text{Id}$ , qui est aussi l'ordre du sous-groupe cyclique engendré par  $\sigma$ . Un  $k$ -cycle étant d'ordre  $k$ , une permutation qui se décompose en cycles de longueurs  $k_1, \dots, k_q$  a donc pour ordre  $p = \text{ppcm}(k_1, \dots, k_q)$ . Les  $k_i$  étant des entiers tous distincts et compris entre 1 et  $n$ , leur produit divise  $n!$  et donc leur ppcm divise aussi  $n!$ , en accord avec le théorème de Lagrange qui assure que l'ordre d'un sous-groupe divise l'ordre du groupe.

**4.8.3 Décomposition en produits de transpositions****Lemme 4.54 :**

Tout  $k$ -cycle se décompose en produit de transpositions.

*Preuve :*

Il suffit de constater que tout  $k$ -cycle, par exemple  $(1, 2, \dots, k)$ , peut se mettre sous la forme  $(1, 2, \dots, k) = (1, k) \cdots (1, 3)(1, 2)$ .

**Théorème 4.55 :**

Toute permutation se décompose en produit de transposition.

*Preuve :*

En combinant le lemme précédent et le fait que toute permutation est un produit de cycles, le résultat est immédiat.

**4.8.4 Signature**

La décomposition précédente n'est pas unique. En revanche, la parité du nombre de transposition intervenant dans toute décomposition d'une permutation donnée est un invariant.

**Proposition 4.56 : Déterminant de Vandermonde**

Posons

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}. \quad (4.24)$$

On vérifie facilement que son déterminant, dit de Vandermonde, s'écrit

$$\Delta(x_1, x_2, \dots, x_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{i>j} (x_i - x_j). \quad (4.25)$$

*Preuve :*

Une façon aisée d'obtenir ce résultat est de constater que le déterminant est un polynôme dans les variables  $x_1, \dots, x_n$  qui s'annule lorsque deux de ces variables sont identiques. De

plus, lorsque l'on cherche à calculer le déterminant, on constate facilement qu'il est homogène, s'écrivant comme une somme de monômes qui sont tous de degré  $0+1+\dots+(n-1) = n(n-1)/2$ . Or le polynôme  $\prod_{i>j}(x_i - x_j)$  est un polynôme homogène de degré  $n(n-1)/2$  possédant les propriétés d'annulation mentionnées ci-dessus. Le déterminant de Vandermonde est donc proportionnel à ce polynôme. La constante de proportionnalité s'obtient en examinant parmi tous les monômes qui le constituent le terme obtenu par produit des termes diagonaux de  $M$ , qui s'écrit  $x_2x_3^2 \cdots x_{n-1}^{n-2}x_n^{n-1}$ . C'est le seul monôme de type. Or ce même monôme est présent à l'identique, avec le même coefficient (1) dans le développement de  $\prod_{i>j}(x_i - x_j)$ , ce qui achève de prouver le résultat.

Représentons l'action du groupe  $\mathcal{S}_n$  sous la forme de matrices  $n \times n$  agissant sur les vecteurs colonnes formés avec les  $x_i$ , en écrivant la transformation  $x_i \mapsto x_{\sigma^{-1}(i)}$  sous la forme

$$\begin{pmatrix} x_{\sigma^{-1}(1)} \\ x_{\sigma^{-1}(2)} \\ \vdots \\ x_{\sigma^{-1}(n)} \end{pmatrix} = (\Sigma_{i,j}(\sigma)) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (4.26)$$

avec  $\Sigma_{i,j}(\sigma) = \delta_{i,\sigma(j)}$  puisque  $\delta_{i,\sigma(j)} x_j = x_{\sigma^{-1}(i)}$ .

C'est bien une représentation au sens où  $\Sigma(\sigma\sigma') = \Sigma(\sigma)\Sigma(\sigma')$ .

Ceci se vérifie facilement sur les éléments de matrice : d'une part  $\Sigma_{i,j}(\sigma\sigma') = \delta_{i,(\sigma\sigma')(j)}$ , et d'autre part

$$[\Sigma(\sigma)\Sigma(\sigma')]_{i,j} = \Sigma_{ik}(\sigma)\Sigma_{kj}(\sigma') = \delta_{i,\sigma(k)}\delta_{k,\sigma'(j)} = \delta_{\sigma^{-1}(i),k}\delta_{k,\sigma'(j)} = \delta_{\sigma^{-1}(i),\sigma'(j)} = \delta_{i,(\sigma\sigma')(j)},$$

ce qui prouve le résultat.

Examinons maintenant l'action d'une permutation  $\sigma$  sur  $\Delta$  : En écrivant (4.26) sous forme transposée, on a, pour tout  $k$ ,

$$(x_{\sigma^{-1}(1)}^k x_{\sigma^{-1}(2)}^k \cdots x_{\sigma^{-1}(n)}^k) = (x_1^k x_2^k \cdots x_n^k) \Sigma^t(\sigma) \quad (4.27)$$

qui permet d'en déduire que la matrice  $M$  se transforme sous l'action de  $\mathcal{S}_n$  comme

$$M' = M \Sigma^t, \quad (4.28)$$

et donc que  $\det M' = \det M \det \Sigma^t$ . Or  $\det[\Sigma^t(\sigma)] = \det \Sigma(\sigma) = \epsilon_\sigma = \pm 1$ , la dernière égalité venant de la forme explicite de la matrice  $\Sigma(\sigma)$  dont la structure est telle que chaque ligne et chaque colonne que contient qu'un seul coefficient non nul, égal à 1.

Ainsi sous l'action d'une permutation  $\sigma$ ,  $\Delta$  change par  $\epsilon_\sigma$  appelée *signature* de  $\sigma$ .

### Propriété 4.57 :

La signature

$$\begin{aligned} \epsilon : \mathcal{S}_n &\rightarrow \mathbb{Z}_2 \\ \sigma &\mapsto \epsilon_\sigma \end{aligned}$$

est un morphisme.

*Preuve :*

La vérification est immédiate en écrivant

$$\epsilon_{\sigma'\sigma} = \det \Sigma(\sigma'\sigma) = \det[\Sigma(\sigma')\Sigma(\sigma)] = \det[\Sigma(\sigma')] \det[\Sigma(\sigma)] = \epsilon_{\sigma'} \epsilon_{\sigma} .$$

Une transposition ne change qu'un seul signe dans l'expression (4.26) du déterminant de Vandermonde, et sa signature est donc égale à -1. Comme toute permutation  $\sigma$  est le produit de  $p$  transpositions, par morphisme on peut écrire que

$$\epsilon_{\sigma} = (-1)^p . \quad (4.29)$$

Le fait que  $\epsilon$  soit un morphisme conduit d'autre part au fait que  $\epsilon_{\sigma}$ , et donc la parité de  $p$ , soit une propriété intrinsèque de  $\sigma$ , indépendante de sa réalisation explicite comme produit de transposition.

**Définition 4.58 :** *Groupe alterné*

Le noyau du morphisme  $\epsilon$  est le sous-groupe invariant  $\mathcal{A}_n$  des permutations paires. Il est d'ordre  $n!/2$ , comme il est facile de le montrer par récurrence.

**Propriété 4.59 :**

Pour toute permutation  $\sigma$  de  $\mathcal{S}_n$  décomposée en produit de cycle, on peut écrire

$$\epsilon_{\sigma} = (-1)^{n + \text{nombre de cycles de } \sigma} . \quad (4.30)$$

*Preuve :*

Tout  $k$ -cycle  $c_k$  est le produit de  $k - 1$  transposition, donc  $\epsilon_{c_k} = (-1)^{k-1}$ . On a donc pour la permutation  $\sigma$  :

$$\epsilon_{\sigma} = (-1)^{\sum(k_i-1)} = (-1)^{\sum k_i - \text{nombre de cycles de } \sigma} = (-1)^{n - \text{nombre de cycles de } \sigma} ,$$

puisque  $\sum_{\text{cycles}} k_i = n$ .

### 4.8.5 Théorème de Cayley

**Théorème 4.60 :**

Tout groupe  $G$  est isomorphe à un sous-groupe du groupe symétrique  $\mathcal{S}(G)$  des permutations de  $G$ . En particulier, si  $G$  est un groupe fini d'ordre  $n$ , il est isomorphe à un sous-groupe de  $\mathcal{S}_n$ .

*Preuve :*

Considérons la translation à gauche, étudiée dans la section 4.6, qui est l'action de groupe particulière définie par

$$\begin{aligned} T : G &\rightarrow \mathcal{S}(G) \\ g &\mapsto T(g) = f_g \end{aligned} \quad (4.31)$$

avec  $\forall x \in G, f_g(x) = gx$ . Nous avons bien défini une action de groupe de  $G$  sur lui-même : D'une part, la fonction  $f_g$  a pour inverse  $f_{g^{-1}}$ . Donc  $f_g \in \mathcal{S}(G)$ .

D'autre part, c'est un morphisme de groupes. En effet,

$$\forall g, h \in G, \text{ et } \forall x \in G, (f_g \circ f_h)(x) = f_g(f_h(x)) = f_g(hx) = g(hx) = (gh)x = f_{gh}(x).$$

Nous rencontrerons à nouveau le morphisme  $T$  dans le cadre de la théorie des représentations des groupes :  $T$  est en effet une représentation de  $G$  appelée représentation régulière.

$T$  est un morphisme injectif puisque  $T(g) = \text{Id}$  (l'identité sur  $\mathcal{S}(G)$ ) implique que  $\forall x \in G, gx = x$  et donc  $g = e$  en prenant  $x = e$ .

Nous avons donc construit un isomorphisme de  $G$  sur  $T(G)$ , qui est un sous-groupe de  $\mathcal{S}(G)$ .

*Remarques :*

- ◇ A l'élément  $e$  correspond la permutation identité.
- ◇ Si  $g \neq e$ , alors  $T(g)$  ne laisse aucun élément invariant (si  $\exists g, /gx = x$  alors  $x = e$  par multiplication à droite par  $x^{-1}$ ). Comme ceci s'applique bien sûr également à toute puissance d'un élément du groupe  $G$ , la puissance étant strictement inférieure à l'ordre de cet élément, on en déduit que la permutation  $T(g)$  est un produit de cycles de longueurs identiques, la longueur de ces cycles étant l'ordre de l'élément  $g$ .

*Exemples :*

- ◇  $\mathbb{Z}/2\mathbb{Z} = \{0, 1\}$  groupe additif des entiers modulo 2 :
  - à 0 correspond  $e$
  - à 1 correspond la permutation (12) puisque  $1 + 0 \equiv 1$  (i.e.  $1 \rightarrow 2$ ) et  $1 + 1 \equiv 0$  (i.e.  $2 \rightarrow 1$ )
- ◇  $\mathbb{Z}/3\mathbb{Z} = \{0, 1, 2\}$  groupe additif des entiers modulo 3 :
  - à 0 correspond  $e$
  - à 1 correspond la permutation circulaire (123)
  - à 2 correspond la permutation circulaire (132) puisque  $2 + 0 \equiv 2$  (i.e.  $1 \rightarrow 3$ ),  $2 + 1 \equiv 0$  (i.e.  $2 \rightarrow 1$ ) et  $2 + 2 \equiv 1$  (i.e.  $3 \rightarrow 2$ )

On vérifie le morphisme en écrivant par exemple la composition des deux éléments  $1+1 = 2$  qui correspond à  $(123)(123) = (132)$ .

- ◇  $\mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}$  groupe additif des entiers modulo 4 :
  - à 0 correspond  $e$
  - à 1 correspond la permutation circulaire (1234)
  - à 2 correspond la permutation (13)(24) : deux cycles de longueur 2
  - à 3 correspond la permutation circulaire (1432)
- ◇ On pourra traiter en exercice le cas du groupe diédral  $D_6$ .

## 4.9 Présentation d'un groupe

### 4.9.1 Groupe libre

**Définition 4.61 :** *Groupe libre*

Le groupe libre  $F_S$  sur un ensemble  $S$  (appelé ensemble des générateurs) est l'ensemble des mots formé de produit<sup>3</sup> d'éléments de  $S$  et d'éléments de  $S^{-1}$  (ensemble des inverses des éléments de  $S$ ). Après réduction (ce qui signifie que tous les produits de la forme  $xx^{-1}$  sont omis, l'élément neutre étant le mot vide, *i.e.* un mot sans aucun élément), les éléments ainsi obtenus sont par hypothèse tous distincts.

Réciproquement, un groupe est libre s'il existe un sous-ensemble  $S$  de  $G$  tel que  $G$  soit le groupe libre sur  $S$ .

**Propriété 4.62 :**

Le rang d'un groupe libre  $F_S$  est égal au cardinal de  $S$ .

### 4.9.2 Exemples

- ◇ Le groupe libre sur l'ensemble vide est le groupe trivial  $\{e\}$ , et le groupe libre sur un singleton est isomorphe à  $\mathbb{Z}$ .
- ◇ En topologie, voir chap.16, on définit la notion de lacet sur l'espace  $\mathcal{E}$  constitué du plan  $\mathbb{R}^2$  privé de  $n$  points  $M_i$ , chemins fermés de point de base par exemple  $O$ . Deux chemins sont dits *homotopes* s'ils peuvent être obtenus l'un de l'autre par déformation continue dans  $\mathcal{E}$ , ce qui définit une relation d'équivalence. Ils font alors partie de la même classe d'homotopie. Ces lacets peuvent être composés entre eux, et tout lacet possède un inverse (il suffit de le parcourir en sens inverse). Ces deux notions sont compatibles avec l'homotopie. Ceci permet finalement de définir le premier groupe d'homotopie, ou groupe fondamental  $\pi_1(\mathcal{E})$ .

On démontre que  $\pi_1(\mathcal{E})$  est le groupe libre engendré par les  $n$  lacets élémentaires  $\ell_i$  enserrant chacun uniquement  $M_i$ .

### 4.9.3 Présentation

#### Généralités, définitions

Considérons un groupe  $G$ , et  $A$  une famille génératrice de ce groupe. A priori, l'écriture d'un élément quelconque de  $G$  comme produit d'éléments de  $A$  n'est pas unique, et  $G$  n'est pas un groupe libre sur  $A$ .

Ceci est dû au fait qu'en général, les générateurs possèdent des relations entre eux, ou de façon équivalente, certains produits de générateurs sont égaux à l'élément neutre.

Deux exemples :

---

3. en notation multiplicative.

- ◇ Dans le cas d'un groupe cyclique d'ordre  $m$  (isomorphe à  $C_m$ ), un générateur  $a$  de ce groupe vérifie  $a^m = 1$ .
  - ◇ Considérons le groupe diédral  $D_{2m}$ , groupe des isométries d'un polygone régulier à  $m$  côtés, qui joue un rôle central en cristallographie. Il est engendré par deux symétries orthogonales  $s_1$  et  $s_2$ , la première par rapport à la médiatrice de l'un des côtés du polygone, la seconde par rapport à la droite joignant le centre du polygone à l'un de ses sommets. On vérifie facilement que  $s_1 s_2$  est une rotation d'angle  $2\pi/m$  (c'est donc un élément du groupe cyclique  $C_m$ ) par rapport au centre du polygone, qui est donc d'ordre  $m$  :  $(s_1 s_2)^m = 1$ . Une autre façon équivalente d'engendrer  $D_{2m}$  est donc de considérer une des deux symétries  $s_1$  ou  $s_2$ , ainsi que la rotation d'angle  $2\pi/m$ .
- On vérifiera que le groupe diédral  $D_{2m}$  est d'ordre  $2m$ .

En fait, il n'est pas nécessaire de spécifier toutes les relations entre ces générateurs, mais seulement un ensemble  $R$  de relations.

**Définition 4.63 :** *Présentation*

Soient  $S$  un ensemble,  $F_S$  le groupe libre sur cet ensemble, et  $R$  une partie de  $F_S$ . Soit  $N$  le plus petit sous-groupe distingué de  $F_S$  contenant  $R$  ( $N$  est appelé clôture normale de  $R$ ). On peut alors considérer le groupe quotient  $F_S/N$ , que l'on appellera le groupe défini par les générateurs  $S$  et les relations  $R$ , et que l'on notera  $\langle S|R \rangle$ .

Si  $G$  est un groupe quelconque, isomorphe à  $\langle S|R \rangle$ , on dit que  $G$  admet  $\langle S|R \rangle$  pour présentation.

*Remarque :*

- ◇ On notera que les éléments de  $N$  sont les conséquences algébriques des relations  $R$ .
- ◇ Tout groupe possède une présentation.  
Il suffit pour cela de considérer  $F_G$ , qui est le groupe libre sur  $G$ . Soit le morphisme de groupe  $\varphi : F_G \rightarrow G$  tel que  $\varphi|_G = \text{id}_G$ . Le noyau  $\text{Ker } \varphi$  est distingué dans  $F_G$ , et  $F_G/\text{Ker } \varphi \simeq \text{Im } \varphi = G$ .
- ◇ Les présentations d'un groupe  $G$  ne sont pas uniques en général.

**Définition 4.64 :** *Présentation finie*

Un groupe est de présentation finie s'il admet une présentation de la forme  $\langle S|R \rangle$ , avec  $S$  et  $R$  finis.

Tout groupe de présentation finie est de type fini, alors que la réciproque est fausse.

**Proposition 4.65 :**

Tout groupe fini a une présentation finie.

*Preuve :*

Il suffit de prendre les éléments du groupe comme générateurs, et sa table de Cayley comme ensemble de relations.

**Exemples**

- ◇ L'exemple le plus simple est le groupe libre sur  $S$  : c'est le groupe de présentation  $\langle S \mid \emptyset \rangle$ .
- ◇ Groupe cyclique  $C_n$  : il a pour présentation  $\langle a \mid a^n \rangle$ .
- ◇ Groupe diédral  $D_{2m}$  : il a pour présentation  $\langle s_1, s_2 \mid s_1^2, s_2^2, (s_1 s_2)^m \rangle$  comme expliqué ci-dessus.
- ◇ Groupe de Klein, isomorphe à  $D_4$  : c'est le plus petit groupe non trivial qui ne soit pas cyclique. Il est de cardinal 4. Il a pour présentation  $\langle s_1, s_2 \mid s_1^2, s_2^2, (s_1 s_2)^2 \rangle$  d'après ce qui précède.
- ◇ Groupe symétrique  $\mathcal{S}_n$  : il est engendré par les  $n - 1$  transpositions  $\sigma_i = (i, i + 1)$ . Les relations sont

$$(\sigma_i \sigma_{i+1})^3 = 1 \tag{4.32}$$

puisque  $\sigma_i \sigma_{i+1}$  s'identifie au cycle  $(i + 1, i + 2, i)$  qui est d'ordre 3,

$$\sigma_i \sigma_j = \sigma_j \sigma_i \quad |i - j| \geq 2, \tag{4.33}$$

puisque deux transpositions affectant des couples disjoints d'indices commutent entre elles, et

$$\sigma_i^2 = 1, \tag{4.34}$$

une transposition étant sa propre inverse.

On notera que la relation (4.33) peut encore s'écrire

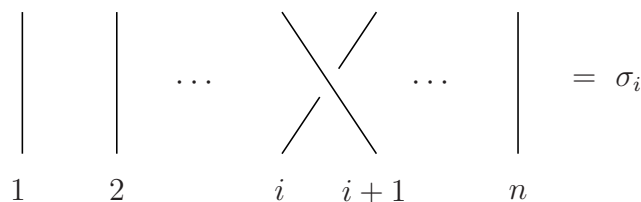
$$(\sigma_i \sigma_j)^2 = 1 \quad |i - j| \geq 2, \tag{4.35}$$

tandis que la relation (4.32) peut aussi s'écrire

$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \quad 1 \leq i < n - 1. \tag{4.36}$$

Par exemple,  $\mathcal{S}_4 \simeq \langle s_1 s_2 s_3 \mid s_1^2, s_2^2, s_3^2, (s_1 s_2)^3, (s_2 s_3)^3, (s_1 s_3)^2 \rangle$ .

- ◇ Partant des exemples précédents de  $D_{2m}$  et  $\mathcal{S}_n$ , on peut plus généralement considérer les groupes de Coxeter, définis par une présentation : les générateurs  $s_i$  satisfont à des relations de la forme  $s_i^2 = 1$  et  $(s_i s_j)^{m_{i,j}} = 1$  où les  $m_{i,j}$  sont des entiers naturels.
- ◇ Partant du groupe symétrique  $\mathcal{S}_n$ , on peut enlever la relation  $s_i^2 = 1$ . La présentation ainsi définie est une présentation du groupe des tresses  $B_n$ . Cette terminologie peut se comprendre aisément en décrivant cette présentation de façon géométrique. Imaginons que l'on dispose de  $n$  brins attachés au plafond et au sol. Codons l'action de  $\sigma_i$  par



qui correspond géométriquement à l'échange des brins  $i$  et  $i+1$ , et l'action de l'opération inverse  $\sigma_i^{-1}$  par

$$\begin{array}{ccccccc}
 | & | & \dots & \diagdown & \diagup & \dots & | \\
 & & & & & & \\
 | & | & & \diagup & \diagdown & & | \\
 1 & 2 & & i & i+1 & & n
 \end{array} = \sigma_i^{-1}.$$

On vérifie alors aisément que les relations algébriques de définition du groupe des tresses ont bien l'interprétation géométrique attendue. En particulier, la relation (4.36) correspond à l'équivalence topologique des deux successions d'échanges  $i \leftrightarrow i+1, i+1 \leftrightarrow i+2$  puis  $i \leftrightarrow i+1$  d'une part, et  $i+1 \leftrightarrow i+2, i \leftrightarrow i+1$  puis  $i+1 \leftrightarrow i+2$  d'autre part, comme on peut le constater sur la figure ci-dessous :

$$\begin{array}{ccccccc}
 | & | & \dots & \diagdown & \diagup & \dots & | \\
 & & & & & & \\
 | & | & & \diagup & \diagdown & & | \\
 1 & 2 & & i & i+1 & i+2 & n
 \end{array} = \sigma_i \sigma_{i+1} \sigma_i$$

est identique à

$$\begin{array}{ccccccc}
 | & | & \dots & \diagdown & \diagup & \dots & | \\
 & & & & & & \\
 | & | & & \diagup & \diagdown & & | \\
 1 & 2 & & i & i+1 & i+2 & n
 \end{array} = \sigma_{i+1} \sigma_i \sigma_{i+1}$$

Ce groupe des tresses joue un rôle important dans de nombreux domaines de la physique, en particulier dans la théorie des modèles intégrables.



## Deuxième partie

# Symétries discrètes et cristallographie



# Chapitre 5

## Introduction à la cristallographie

### 5.1 L'état cristallin

#### 5.1.1 Les cristaux parfaits

### 5.1.2 Les défauts dans les cristaux

## 5.2 Qu'est-ce qu'un cristal du point de vue mathématique ?

### 5.2.1 Réseau

### 5.2.2 Maille primitive

**5.2.3 Quelques choix possibles de maille primitive**





# Chapitre 6

## Groupe ponctuel à deux dimensions

### 6.1 Définitions et propriétés principales

**Rappel :** groupe du papier-peint = sous-groupe  $G$  de  $Isom(\mathbb{R})$  qui contient un réseau 2d de translation  $T$  comme sous-groupe normal

★ on étudie le groupe quotient  $G/T$

★ on choisit une base de  $T$  et on considère  $G/T$  comme un sous-groupe de  $O_2(\mathbb{R})$ . On étudie comment les éléments de  $G/T$  agissent sur cette base **i.e.** on étudie les représentations de  $G/T$  sur  $GL_2(\mathbb{R})$ .

#### Définition et propriétés principales

Soit  $G$  un groupe du papier-peint de réseau de translation  $T$ .

**Notation :** Si  $\varphi$  est une isométrie,  $\exists A \in O_2(\mathbb{R})$  et  $b \in \mathbb{R}^2$  tel que  $\varphi(n) = A(n) + b$ . On note  $\varphi = (A, b)$ .

La loi de composition dans  $Isom(\mathbb{R}^2)$  s'écrit alors  $(A, b)(c, d) = (AC, A(d) + b)$ .

Inverse :  $(A, b)^{-1} = (A^{-1}, -A^{-1}(b))$  [en effet :  $(A, b)(A^{-1}, -A^{-1}(b)) = (I, -b + b) = (I, 0)$ ]

**donc :**

★ translation de vecteur  $b = (I, b)$

★ élément de  $O_2(\mathbb{R})$  : de la forme  $(A, 0)$

#### Définition

Soit  $G$  un groupe du papier-peint.

Le groupe ponctuel  $G_0$  de  $G$  est l'ensemble  $\{A \in O_2(\mathbb{R}) / \exists b \in \mathbb{R}^2 \text{ tel que } (A, b) \in G\}$ .

On en déduit que  $G_0$  est un sous-groupe de  $O_2(\mathbb{R})$ .

**Remarque :**  $b$  n'est pas forcément un élément de  $T$ . Exemple : réflexions, pour lesquelles,  $b$  est un demi-vecteur du réseau.

#### Proposition

Si  $G$  est un groupe du papier-peint de réseau de translation  $T$  et de groupe ponctuel  $G_0$ , alors  $G_0 \cong G/T$ .

**Preuve :**

Si  $\varphi$  est l'homomorphisme de groupe

$$\begin{aligned} \text{Isom}(\mathbb{R}^2) &\rightarrow O_2(\mathbb{R}) \\ (A, b) &\mapsto A \end{aligned}$$

alors  $\varphi(G) = G_0$ .

Le noyau de  $\varphi|_G$  est  $T$ , puisque  $T = G \cap \Pi$ , et que  $\Pi$  est le noyau de  $\varphi$ .

---

Le groupe  $O_2(\mathbb{R})$  agit sur  $\Pi$  par conjugaison : pour  $t \in \mathbb{R}^2$  et  $A \in O_2(\mathbb{R})$ ,

$$(A, 0)(I, t)(A, 0)^{-1} = (I, A(t))$$

(car  $(A, 0)(I, t) = (A, A(t))$  et  $(A, A(t))(A, 0)^{-1}(A, A(t))(A^{-1}, 0) = (I, A(t))$ )

or  $\Pi \cong \mathbb{R}^2$ , donc cette action par conjugaison est en fait l'action naturelle de  $O_2(\mathbb{R})$  sur  $\mathbb{R}^2$  (au sens "naturel"  $t \mapsto A(t)$ ).

**Cette action peut se restreindre à l'action de  $G_0$  sur  $T$  :**

si  $A \in G_0$  et  $t \in T$ , alors  $\exists b \in \mathbb{R}^2$  tel que  $(A, b) \in G$ , donc  $(A, b)(I, t)(A, b)^{-1} \in G$   
 or  $(A, b)(I, t) = (A, A(t) + b)$  et  $(A, A(t) + b)(A, b)^{-1} = (A, A(t) + b)(A^{-1}, -A^{-1}(b))$   
 $= (I, -b + b + A(t)) = (I, A(t))$

Ainsi  $(I, A(t)) \in G$ , donc  $\boxed{A(t) \in T}$ .

L'existence de cette action, combinée à la structure de groupe de  $T$ , va nous permettre de déterminer les candidats  $G_0$  comme groupes ponctuels d'un groupe de papier-peint  $G$ .

$C_n$  : sous-groupe de  $O_2(\mathbb{R})$  cyclique, engendré par une rotation de  $\frac{2\pi}{n}$

$D_n$  : groupe diédral d'ordre  $2n$ . Engendré par deux éléments  $r$  et  $f$  vérifiant

$$r^n = f^2 = id \quad \text{et} \quad frf = r^{-1}$$

On peut voir  $D_n$  comme un sous-groupe de  $O_2(\mathbb{R})$  en considérant

$$\begin{cases} r = \text{rotation d'angle } \frac{2\pi}{n} \\ f = \text{une réflexion} \end{cases}$$

(tout autre choix pour  $f$  conduit à d'autres sous-groupes de  $O_2(\mathbb{R})$  isomorphes à  $D_n$ ).

**Lemme**

*Le groupe ponctuel  $G_0$  d'un groupe de papier-peint  $G$  est fini.*

En effet, il suffit de considérer  $\{t_1, t_2\}$  une base de  $T$ , et  $C$  un cercle de centre  $O$  qui contient  $t_1$  et  $t_2$  dans son intérieur. Le nombre d'éléments de  $T$  dans  $C$  est fini.

Comme  $G_0$  est un sous-groupe de  $O_2(\mathbb{R})$ , l'action **d'un de ses éléments** sur un vecteur intérieur à  $C$  donne un vecteur intérieur à  $C$ , élément de  $T$ , en nombre fini. **Il** agit donc comme une permutation, donc  $G_0$  est fini.

## 6.2 Les 10 groupes ponctuels bidimensionnels

### Théorème

Soit  $G_0$  le groupe ponctuel d'un groupe de papier-peint  $G$ .

Alors  $G_0$  est isomorphe à l'un des 10 groupes suivants :  $\left\{ \begin{array}{l} C_1, C_2, C_3, C_4, C_6 \\ D_1, D_2, D_3, D_4, D_6 \end{array} \right\}$ .

**Preuve** : on sait (cf Lemme) que  $G_0$  est un groupe fini.

On sait donc également que  $G_0 \cong C_n$  ou  $D_n$  pour un certain  $n$  (cf proposition 2.8). Déterminons les volumes possibles pour  $n$ .

Nous avons vu dans la preuve de 2.8 que  $N = G_0 \cap SO_2(\mathbb{R})$  est un groupe cyclique engendré par une rotation d'angle possible normale. De plus,  $|N| = n$ , donc  $r$  est d'ordre  $n$ .

Représentons  $r$  de deux façons différentes :

★ dans la base standard, si  $r$  est une rotation d'angle  $\theta$ , alors la matrice représentant  $r$  est

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

★ dans une base entière  $\{t_1, t_2\}$  de  $T$  : comme  $r(T) = T$  et  $T = \mathbb{Z} t_1 \oplus \mathbb{Z} t_2$ , la matrice de  $r$  dans cette base est de la forme

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

avec  $a, b, c, d$  entiers.

Ces deux matrices représentent la même transformation linéaire, mais dans des bases différentes : elles sont conjuguées et ont donc **même trace**. Donc  $2 \cos \theta = a + d \in \mathbb{Z} \Leftrightarrow |\theta| = \frac{2\pi}{n}$  avec  $n = 1, 2, 3, 4, 6$ . Ceci prouve que  $N = \langle r \rangle$  est d'ordre  $n \in \{1, 2, 3, 4, 6\}$ . Comme  $G_0 \cong C_n$  ou  $D_n$  avec  $n = |N|$ , ceci termine la preuve.

**Remarque** : si un réseau **tri**-dimensionnel est invariant par une rotation d'angle  $\theta$ , le résultat reste vrai en complétant une base orthonormée du plan de rotation par un vecteur unitaire  $\vec{u}$  le long de l'axe de rotation, et en écrivant donc

$$\text{rot } r = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

### Lemme 3.5

Soit  $G$  un groupe de symétrie de réseau de translation  $T$  et  $G_n = \{g \in G / xg^n = g^n x \ \forall g \in G\}$ .

Alors  $T = G_n$  pour tout  $n$  multiple de  $[G : T]$ .

De plus, si  $G$  et  $G'$  sont des groupes de papier-peint de réseau de translation  $T$  et  $T'$  respectivement et si  $\varphi : G \rightarrow G'$  est un isomorphisme, alors  $\varphi(T) = T'$ .

**Preuve :**

1. On sait (cf. 3.3) que  $G/T$  est fini. Soit  $n$  multiple de  $[G : T]$ .

◇ Si  $g \in G$ , alors  $g^n \in T$  (par Lagrange : l'ordre d'un élément divise l'ordre du groupe). Comme  $T$  est abélien, si  $t \in T$  alors  $\forall g \in G, tg^n = g^n t$ , donc  $t \in G_n$ . Ainsi  $\boxed{T \in G_n}$ .

◇ réciproquement, supposons que  $x \in G_n$ .

On peut donc écrire  $x = (A, b)$  avec  $b \in \mathbb{R}^2$ . Soit  $g = (I, t)$  avec  $t \in T$ . Alors  $g^n = (I, nt) \in T$ . Comme  $x \in G_n$ , on a  $xg^n x^{-1} = g^n$ . Or l'action de  $O_2(\mathbb{R})$  sur  $\Pi$  s'écrit  $xg^n x^{-1} = (I, A(nt))$ . Donc  $A(nt) = nt$ , pour tout  $t \in T$ . Soit  $\{t_1, t_2\}$  une base entière de  $T$ . Alors  $\{nt_1, nt_2\}$  est une base de  $\mathbb{R}^2$ . Comme  $A \in O_2(\mathbb{R})$  est complètement déterminée par son action sur une base, en particulier sur celle-là, on en déduit que  $A = I$ . D'où  $x = (I, b) \in T$  et donc  $\boxed{G_n \subset T}$ .

Conclusion  $\boxed{G_n = T}$ .

2. Soient  $G \cong G'$  deux groupes du papier peint isomorphes. Soit  $\varphi : G \rightarrow G'$  un isomorphe. Alors  $\forall n, \varphi(G_n) = G'_n$  (par morphisme). Posons  $m = [G : T]$  et  $m' = [G' : T']$ . Avec  $n = m$ .

**Corollaire 3.6**

*Si  $g \cong g'$  sont deux groupes de papier peint isomorphes, alors leurs groupes ponctuels sont isomorphes.*

**Preuve :**

Soit  $\varphi : G \rightarrow G'$  un isomorphisme.

Si  $T$  et  $T'$  sont les réseaux de translation de  $G$  et  $G'$  respectivement, alors  $\varphi(T) = T'$  (lemme 3.5). Donc  $\varphi$  induit un isomorphisme entre  $G/T$  et  $G'/T'$ . Or  $G/T \cong G_0$  et  $G'/T' \cong G'_0$ , donc  $G_0 \cong G'_0$  **c.q.f.d.**

Résultat plus fort : on va maintenant établir un critère nécessaire pour que deux groupes de papier-peint soient isomorphes. On utilisera ce critère pour montrer que certains groupes du papier-peint ne sont pas isomorphes.

Soit  $G$  un groupe de papier peint de réseau de translation  $T$  et de groupe ponctuel  $G_0$ . L'action de  $G_0$  sur  $T$  produit un homomorphisme de groupe :

$$G_0 \rightarrow \text{Aut}(T) \simeq \text{Aut}(\mathbb{Z}^2) .$$

Prenons une base  $\{t_1, t_2\}$  de  $T$ . Les éléments de  $\text{Aut}(\mathbb{Z}^2)$  peuvent être représentés par des matrices  $2 \times 2$  à coefficients entiers. Donc

$$\text{Aut}(\mathbb{Z}^2) \cong \underbrace{GL_2(\mathbb{Z})}_{\text{matrices inversibles } 2 \times 2 \text{ à coefficients entiers}} .$$

Or pour qu'une matrice soit irréversible sur  $\mathbf{Z}$ , son déterminant doit valoir  $\pm 1$ .  $GL_2(\mathbf{Z})$  est donc constitué des matrices à coefficient entier, de déterminant  $\pm 1$ .

**Conclusion :** l'action de  $G_0$  sur  $T$  permet de représenter  $G_0$  comme un sous-groupe de  $GL_2(\mathbf{Z})$ .

**Proposition 3.7** *Soit  $\varphi : G \rightarrow G'$  un isomorphisme des groupes de papier peints  $G$  et  $G'$ . Soient  $T$  et  $T'$  les réseaux de translation de  $G$  et  $G'$  respectivement. Soient  $G_0$  et  $G'_0$  les groupes ponctuels de  $G$  et  $G'$  respectivement. En choisissant des bases entières pour  $T$  et  $T'$ , l'application  $\varphi|_T$  est un isomorphisme linéaire, donné par une base  $U \in GL_2(\mathbf{Z})$ , et l'isomorphisme  $\bar{\varphi} : G_0 \rightarrow G'_0$  est la conjugaison par  $U$ .*

**Preuve :** Supposons que  $\varphi : G \rightarrow G'$  est un isomorphisme. Par 3.5,  $\varphi|_T$  est un isomorphisme de  $T$  sur  $T'$ . Soit  $\{t_1, t_2\}$  une base de  $T$ , et  $\{s_1, s_2\}$  une base de  $T'$ .

$$\text{Alors } \begin{cases} \varphi(t_1) = \alpha s_1 + \beta s_2 \\ \varphi(t_2) = \gamma s_1 + \delta s_2 \end{cases} \quad \text{avec } \alpha, \beta, \gamma, \delta \text{ entiers.}$$

Comme  $\varphi|_T$  est un isomorphisme de

$$\underbrace{\mathbf{Z} - \text{module}}_{\text{i.e. espace vectoriel avec le corps remplacé par l'anneau } \mathbf{Z}}$$

il est déterminé par son action sur la base choisie de  $T$ , i.e. représenté par une matrice

$$U = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}.$$

De plus  $\varphi^{-1}$  est un isomorphisme  $T' \rightarrow T$ . Donc  $U^{-1}$  a des coefficients entiers  $\Rightarrow U \in GL_2(\mathbf{Z})$ .

Soit  $(A, b) \in G$ . Notons  $(C, d) = \varphi(A, b)$ .

Pour  $t \in T$ ,  $\varphi(I, t) = (I, Ut)$ .

$$\text{Donc } (C, d)(I, Ut)(C, d)^{-1} = \varphi \left( \underline{(A, b)(I, t)(A, b)^{-1}} \right)$$

(par morphisme)

$$\text{Or } (C, d)(I, Ut) = (C, d + CUt)$$

$$\begin{aligned} \text{et donc } (C, d)(I, Ut) \underbrace{(C, d)^{-1}}_{(C^{-1}, -C^{-1}d)} &= (C, d + CUt)(C^{-1}, -C^{-1}d) \\ (C^{-1}, -C^{-1}d) &= (I, d + CUt - d) = (I, CUt) \end{aligned}$$

[ **Rappel :**  $(A, u)(B, v) = (AB, u + Av)$  ]. D'autre part  $\underline{(A, b)(I, t)(A, b)^{-1}} = (I, At)$ .

Nous avons donc montré que  $\varphi(I, At) = (I, CUt)$ . Or  $\varphi(I, At) = (I, UAt)$ .

Finalement :

$$(I, UAt) = (I, CUt) \quad \text{soit } CU = UA \quad \text{i.e. } \underline{C = UAU^{-1}}_{\text{conjugaison par } U}$$

**Corollaire 3.8**

Soient  $G$  et  $G'$  deux groupes du papier-peint de groupes ponctuels  $G_0$  et  $G'_0$  respectivement. En identifiant  $G_0$  et  $G'_0$  à des sous-groupes de  $GL_2(\mathbf{Z})$ , en choisissant pour bases celles des réseaux de translation de  $G_0$  et  $G'_0$ , il existe  $U \in GL_2(\mathbf{Z})$  tel que  $G'_0 = UG_0U^{-1}$ .

**La réciproque est vraie :** si les réseaux de translation de deux groupes du papier peint sont isomorphes par  $U$ , et que la conjugaison par  $U$  est un isomorphisme entre leurs groupes ponctuels, alors on obtient un isomorphisme de ces deux groupes de papier-peint par

$$(g, t) \mapsto (U_g U^{-1}, U(t)) .$$

La condition est donc plus forte que celle consistant à demander que les groupes ponctuels soient isomorphes.

**exemple :**  $C_2 \cong D_1$  mais ils ne sont pas conjugués sous  $GL_2(\mathbf{Z})$  pas de surprise, réseau avec une symétrie de rotation  $180^\circ$  et sans symétrie de réflexion ( $C_2$ )  $\neq$  réseau avec une symétrie de réflexion et sans symétrie de rotation ( $D_1$ ).

**6.3 Les 5 types de réseau**

**idée :** partant des 10 groupes ponctuels on fait agir  $G_0$  sur  $T$

**pour cela :** on fixe une base  $\{t_1, t_2\}$  de  $T \rightarrow T \cong \mathbf{Z}^2$ .

Dans cette base l'action de  $G_0$  sur  $T$  induit un homomorphisme

$$G_0 \rightarrow \text{Aut}(\mathbf{Z}^2) \cong GL_2(\mathbf{Z})$$

**i.e. :** un choix de base combiné avec l'action de  $G_0$  sur  $T$  donne une représentation de  $G$  sur un sous-groupe spécifique de  $GL_2(\mathbf{Z})$ .

**6.3.1 Groupes  $C_1, C_2$  : les réseaux parallélogrammes**

★ Si  $\underline{G_0 = C_1}$ , alors

$$C_1 = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle$$

★ Si  $\underline{G_0 = C_2}$ , alors la rotation d'angle  $\Pi$  est la multiplication par  $-1$  sur  $T$  donc

$$C_2 = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle$$

FIGURE 6.1 – réseau parallélogramme

**Lemma 3.9**

Supposons que  $G_0$  contient une rotation  $r$  d'angle  $\frac{2\pi}{n}$  ( $n \geq 3$ ). Si  $t \neq 0 \in T$ , de longueur minimale, alors  $\{t, r(t)\}$  est une base de  $T$ .

**Preuve :** Soit  $\{t_1, t_2\}$  une base de  $T$ . Alors

$$\begin{cases} t = at + bt_2 \\ c(t) = ct - 1 + dt_2 \end{cases} \quad \text{avec } a, b, c, d \in \mathbf{Z}$$

$\{t, r(t)\}$  est linéairement indépendante car  $n > 2$ . On peut donc résoudre en  $t_1$  :

$$t_1 = \alpha t + \beta r(t) \quad \text{avec } \alpha, \beta \in \mathbb{R}$$

On écrit  $\begin{cases} \alpha = \alpha_0 + \varepsilon \\ \beta = \beta_0 + \varepsilon' \end{cases} \quad \alpha_0, \beta_0 \in \mathbf{Z} \quad \text{et } |\varepsilon|, |\varepsilon'| \leq \frac{1}{2}$

$$S = \alpha_0 t + \beta_0 r(t) \in T \quad \text{et } t_1 \in T$$

$$\swarrow \in \mathbf{Z} \nearrow$$

donc  $t_1 - S = \varepsilon t + \varepsilon' r(t) \in T$ . Comme  $t$  et  $r(t)$  ne sont pas parallèles ( $n > 2$ ),

$$\|t_1 - S\| = \|\varepsilon t + \varepsilon' r(t)\| < \|\varepsilon t\| + \|\varepsilon' r(t)\| \leq \frac{1}{2} \left( \|t\| + \underbrace{\|r(t)\|}_{= \|t\|} \right) = \|t\|$$

contradictoire avec  $\|t\|$  minimal, sauf si  $s = t_1$ , donc  $t_1$  est une combinaison  $\mathbf{Z}$ -linéaire de  $t$  et de  $r(t)$ , et idem :  $t_2$  est une combinaison  $\mathbf{Z}$ -linéaire de  $t$  et de  $r(t)$ .

Comme  $\{t_1, t_2\}$  est une base de  $T$ ,  $\{t, r(t)\}$  est aussi une base de  $T$ .

### 6.3.2 Groupes $C_4, D_4$ : les réseaux carrés

Soit  $r$  une rotation de  $\frac{\pi}{2}$ . Si  $t_1$  est un vecteur de  $T$  de norme minimale,  $\{t_1, r(t_1)\}$  est une base de  $T \rightarrow$  **réseau carré**

★ si  $\mathbf{G}_0 = \mathbf{C}_4 = \langle r \rangle$ , une représentation de  $G_0$  dans cette base est

$$C_4 = \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle$$

★ si  $\mathbf{G}_0 = \mathbf{D}_4$ , alors  $G_0$  contient une réflexion  $f$ . Les quatre éléments  $f, rf, r^2f, r^3f$  sont des réflexions de  $G_0$ . Elles doivent préserver l'ensemble des vecteurs de  $T$  de norme minimale. Quatre de ces vecteurs sont  $\pm t_1, \pm t_2$ .

Considérons un point quelconque sur le cercle de centre  $O$ , de rayon  $\|t_1\|$ . Tout autre point de ce cercle (i.e. autre que  $\pm t_1, \pm t_2$ ) est à une distance de l'un de ces 4 points **inférieurs** à  $\|t_1\| \rightarrow$  **contradiction**.

On a donc  $\pm t_1, \pm t_2$  qui **constituent** l'ensemble des vecteurs de norme minimale de  $T$ .  $D_4$  est engendré par  $r$  et par une réflexion. On peut choisir la réflexion par rapport à l'axe porté par  $t_1$ .

$$D_4 = \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle$$

FIGURE 6.2 – Les quatre axes de réflexion dans le cas  $D_4$ .

### 6.3.3 Groupes $C_3, D_3, C_6, D_6$ : réseaux hexagonaux

Soit  $r$  une rotation de  $\frac{2\pi}{2}$  (elle fait partie de  $C_3, D_3, C_6, D_6$ ). Si  $t_1$  est un vecteur de  $T$  de norme minimale, en posant  $t_2 = r(t_1)$ ,  $\{t_1, t_2\}$  est une base de  $T \rightarrow$  **réseau hexagonal**

FIGURE 6.3 – axes des réflexions

★ si  $\mathbf{G}_0 = \mathbf{C}_3$  : engendré par  $r$

$$C_3 = \left\langle \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \right\rangle$$

★ si  $\mathbf{G}_0 = \mathbf{C}_6$  : engendré par une rotation de  $\frac{\pi}{3}$

$$C_6 = \left\langle \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle.$$

On constate sur la figure que  $T$  comporte 6 vecteurs de longueur minimale. Tout autre point sur le cercle centré en  $O$  et de rayon  $\|t_1\|$  serait à une distance  $\|t_1\|$  de l'un de ces 6 points. Il y a donc exactement 6 vecteurs de longueur minimale.

★ si  $\mathbf{G}_0 = \mathbf{D}_3$  ou  $\mathbf{G}_0 = \mathbf{D}_6$ , alors  $G_0$  comporte 3 ou 6 réflexions, respectivement. Ces réflexions permutent les 6 vecteurs discutés ci-dessus.

$\mathbf{G}_0 = \mathbf{D}_6$  : 6 axes de réflexions (voir figure). On utilise la réflexion par rapport à  $t_1$  comme générateur

$$D_6 = \left\langle \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} \right\rangle$$

$\mathbf{G}_0 = \mathbf{D}_3$ , alors  $G_0$  comporte 3 réflexions, séparées par des angles  $\frac{\pi}{3}$  : si  $f$  est une réflexion de  $D_3$ , alors  $rf$  est une réflexion d'axe faisant un angle  $\frac{D}{3}$  par rapport à celui de  $f$ .

Les axes des réflexions de  $D_3$  sont des axes de réflexions de  $D_6$  car  $D_3$  est un sous-groupe de  $D_6$ .

→ **deux possibilités :**

★ angles  $\frac{\pi}{6}, \frac{\pi}{2}, \frac{5\pi}{6}/t_1 \rightarrow D_{3,\ell}$

★ angles  $0, \frac{\pi}{3}, \frac{2\pi}{2}/t_1 \rightarrow D_{3,s}$

**$D_{3,\ell}$  et  $D_{3s}$  ne sont pas conjugués sous  $GL(2, \mathbb{Z})$**

$$C_f M_A + n : \boxed{D_3 \cdot nb}$$

$$A_s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

réflexion suivant  
la petite diagonale

$$A_\ell = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

réflexion suivant  
la grande diagonale



**Figure**

$$\mathbf{rot} \left| \begin{array}{l} t_1 \rightarrow t_2 \\ t_2 \rightarrow -t_1 - t_2 \end{array} \right.$$

$$\mathbf{rot} = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$$

$$\mathbf{ref} \ell = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

**Figure**  $\mathbf{ref} \ell \mathbf{cot} \mathbf{rot} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = A_\ell$

$\mathbf{ref} s = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}$  **Figure**  $\mathbf{rot} \cdot \mathbf{ref} s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A_s$

$D_{3\ell} = \langle \mathbf{rot}, \mathbf{ref}_\ell \rangle$   $D_{3s} = \langle \mathbf{rot}, \mathbf{ref}_s \rangle$  on cherche à conjuguer  $A_\ell$  et  $A_s$

*i.e.*  $\underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_A \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{A_s} = \underbrace{\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}}_{A_\ell} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  avec  $a, b, c, d \in \mathbb{Z}$

$$A \cdot A_s = \begin{pmatrix} -b & a \\ -d & -c \end{pmatrix} \quad A_\ell \cdot A = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

$$\Rightarrow \left| \begin{array}{l} c = -b \\ d = -a \end{array} \right. \quad \text{donc} \quad ad - bc = b^2 - a^2 = \overbrace{(b-a)}^{\swarrow \text{entiers} \searrow} \overbrace{(b+a)}^{\swarrow \text{entiers} \searrow} = \pm 1$$

↑ det = ±1

**det = +1**

$$\Rightarrow \left| \begin{array}{l} b - a = \pm 1 \\ b + a = \pm 1 \end{array} \right. \Rightarrow 2b = \pm 2 \quad \text{donc} \quad b = \pm 1 \text{ et } a = 0 \Rightarrow \underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{\text{conj 1}} \underbrace{\text{et}}_{\text{opposées}} \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix}$$

**→ 2 cas**

**det = -1**

$$\Rightarrow \left| \begin{array}{l} b - a = \pm 1 \\ b + a = \mp 1 \end{array} \right. \Rightarrow 2b = 0 \quad \text{donc} \quad b = 0 \text{ et } a = \mp 1 \Rightarrow \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\text{conj 2}} \underbrace{\text{et}}_{\text{opposées}} \begin{pmatrix} -1 & 0 \\ 0 & +1 \end{pmatrix}$$

**→ 2 cas**

Or en fait la conjugaison par  $-I_2$  est en fait identique à la conjugaison par  $I_2$ . On ne considère donc que 2 cas : conj 1 et conj 2

$$\text{conj 2} \cdot \text{ref } S \cdot \text{conj } 2^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \notin D_{3\ell} \quad \text{conj 1} \cdot \text{ref } S \cdot \text{conj } 1^{-1} = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \notin D_{3\ell}$$

$$D_{3\ell} = \left\{ \begin{array}{l} P \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & -0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \end{array} \right\}$$

$$= \left\{ \begin{array}{l} \text{rot} \\ \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & -0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \end{array} \right\}$$

$$D_{3S} = \left\{ \begin{array}{l} \text{ref } S \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & -0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array} \right\}$$

### 6.3.4 Groupes $D_1, D_2$ : réseaux rectangulaires ou orthorhombiques

★ Si  $G_0 = D_1$  ou  $D_2$ , alors  $G_0$  ne contient pas de rotation d'ordre  $\geq 3$ . On ne peut donc appliquer le lemme pour trouver une base. Soit  $f$  une réflexion non triviale de  $G_0$ . Soit  $t \in T$  un vecteur non nul non parallèle à l'axe de la réflexion  $f$ . Comme  $f$  est une application de  $T$  sur  $T$ ,  $\begin{cases} t + f(t) \\ t - f(t) \end{cases}$  sont des éléments de  $T$ . Ils sont parallèles et perpendiculaires à l'axe de la réflexion  $f$ .

FIGURE 6.4 –

Soient  $S_1$  et  $S_2$  deux vecteurs non nuls de  $T$ , de norme minimale avec  $S_1$  parallèle à l'axe,  $S_2$  perpendiculaire à l'axe,  $T$  discret  $\Rightarrow$

★  $S_1$  et  $S_2$  existent

★ tout vecteur parallèle à l'axe est de la forme  $\alpha S_1$   $\alpha \in \mathbf{Z}$

★ tout vecteur perpendiculaire à l'axe est de la forme  $\beta S_2$   $\beta \in \mathbf{Z}$

donc  $\forall t \in T$ ,  $\begin{cases} t + f(t) = m_t S_1 & m_t \in \mathbf{Z} \\ t - f(t) = n_t S_2 & n_t \in \mathbf{Z} \end{cases}$ . On résout en  $t$  :  $t = \frac{m_t}{2} S_1 + \frac{n_t}{2} S_2$ .

★ Si,  $\forall t \in T$ ,  $m_t$  et  $n_t$  sont pairs, alors  $\{S_1, S_2\}$  est une base de  $T$ .

★ Si  $\exists t \in T$  tel que  $m_t$  ou  $n_t$  sont impairs, alors  $m_t$  et  $n_t$  sont impairs :

$\left[ \begin{array}{l} \text{sinon, par exemple avec } m_t = 2p + 1 \text{ et } n_t = 2q (p, q \in \mathbf{Z}) \text{ on aurait} \\ t = pS_1 + \frac{S_1}{2} + qS_2 \in T \text{ avec } pS_1 \text{ et } qS_2 \in T \text{ donc } \frac{S_1}{2} \in T \text{ absurde.} \end{array} \right.$

En écrivant  $m_t = 2p + 1$  et  $n_t = 2q + 1$  ( $p, q \in \mathbf{Z}$ ) on déduit que

$$t_1 = \frac{1}{2}(S_1 + S_2) \in T \quad \text{et} \quad t_2 = \underbrace{\frac{1}{2}(S_1 - S_2)}_{= f(t_1)} \in T$$

$$\begin{aligned} \text{D'autre part} \quad t = \frac{m_t}{2} S_1 + \frac{n_t}{2} S_2 &= \left( \frac{m_t + n_t}{2} \right) \frac{S_1 + S_2}{2} + \left( \frac{m_t - n_t}{2} \right) \frac{S_1 - S_2}{2} \\ &= m'_t t_1 + n'_t t_2 \quad \text{avec} \quad m'_t, n'_t \in \mathbf{Z}. \end{aligned}$$

Ceci prouve que  $\{t_1, t_2\}$  est une base de  $T$ .

**Résumé :**

★ Soit  $\{t_1, t_2\}$  est une base de deux vecteurs orthogonaux

FIGURE 6.5 – Réseau rectangulaire “ $p$ ”

★ soit  $\{t_1, t_2\}$  est une base de deux vecteurs de même norme qui s'échangent par réflexion.

FIGURE 6.6 – Réseau orthorhombique “ $C$ ” (on dessinera aussi  $\begin{matrix} \nearrow t_2 \\ \searrow t_1 \end{matrix}$ )

Représentations de  $D_1$  et  $D_2$ 

$$D_{1,p} = \left\langle \begin{pmatrix} 1 & -0 \\ 0 & -1 \end{pmatrix} \right\rangle$$

$$D_{1,c} = \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle$$

$$D_{2,p} = \left\langle \begin{pmatrix} -1 & -0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle$$

$$D_{2,c} = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\rangle$$

\*  $D_{1,p}$  et  $D_{1,c}$  ne sont pas conjugués sous  $GL(2, \mathbb{Z})$

On cherche

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

tel que  $A \cdot M_{f,p} = M_{f,a} \cdot A$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{avec } ad - bc = \pm 1$$

$M_{f,p} \qquad M_{f,e}$

donc  $\begin{cases} d = -b \\ c = -a \end{cases}$  et donc  $ad - bc = -2ab \neq \pm 1$  car  $a$  et  $b$  sont entiers.

Ainsi  $D_{1,p}$  et  $D_{1,c}$  ne sont pas conjugués sous  $GL_2(\mathbb{Z})$ .

\*  $D_{2,p}$  et  $D_{2,c}$  ne sont pas conjugués sous  $GL(2, \mathbb{Z})$

idem

# Chapitre 7

## Classification des 17 groupes du papier peint

### 7.1 Les 17 groupes du papier peint et leur nomenclature

## 7.2 Des groupes ponctuels aux groupes du papier peint : le problème de l'extension

### 7.3 La théorie des groupes en marche : construction des isométries du réseau à partir du groupe ponctuel

## 7.4 Mise en oeuvre : des 10 groupes ponctuels aux 17 groupes du papier peint

### 7.4.1 $G = 0 = D_{1,p}$



7.4. MISE EN OEUVRE : DES 10 GROUPES PONCTUELS AUX 17 GROUPES DU PAPIER PEINT101

**7.4.2**  $G = 0 = D_{1,c}$

**7.4.3**  $G = 0 = D_{3,\ell}$  ou  $D_{3,s}$

7.4. MISE EN OEUVRE : DES 10 GROUPES PONCTUELS AUX 17 GROUPES DU PAPIER PEINT103

7.4.4  $G = 0 = D_6$

**7.4.5**  $G = 0 = D_4$

**7.4.6**  $G = 0 = D_2$

$$G = 0 = D_{2,c}$$

$$G = 0 = D_{2,p}$$

## 7.5 Quelques compléments sur les centres et axes de symétries des 17 pavages





# Chapitre 8

## Quelques éléments sur la cristallographie à 3 dimensions

### 8.1 Les 32 groupes ponctuels

#### 8.1.1 Notation de Schönflies

### 8.1.2 Notation d'Hermann- Mauguin

## 8.2 Des 32 groupes ponctuels aux 230 groupes d'espace

### 8.2.1 Les 7 systèmes cristallins

triclinique, monoclinique, orthorhombique, tétragonal (quadratique), trigonal, hexagonal, cubique

### **8.2.2 Les 7 systèmes réticulaires**

triclinique, monoclinique, orthorhombique, tétragonal (quadratique), rhomboédrique, hexagonal, cubique

### 8.2.3 Les 6 familles cristallines

Triclinique, Monoclinique, Orthorhombique, Tétragonale (quadratique), Hexagonale, cubique

#### **8.2.4 Les 14 réseaux de Bravais**

Troisième partie

Symétries dynamiques en physique  
classique





# Chapitre 9

## Mécanique analytique

Dans ce chapitre, nous allons rappeler quelques concepts essentiels de mécanique analytique, en décrivant les approches hamiltonienne et lagrangienne, pour les systèmes comportant un nombre fini de degrés de liberté. On trouvera un exposé détaillé de ces notions dans le Landau de mécanique [7] dont ce chapitre s'inspire fortement.

### 9.1 Principe de moindre action

#### 9.1.1 Equations d'Euler-Lagrange

Dans la mécanique lagrangienne (Lagrange, 1787), tout système mécanique est caractérisé par sa fonction de Lagrange (ou lagrangien)  $L(q_1, q_2, \dots, q_s; \dot{q}_1, \dot{q}_2, \dots, \dot{q}_s; t)$ , notée par la suite  $L(q, \dot{q}, t)$ .  $q_1, q_2, \dots, q_s$  représentent les  $s$  coordonnées généralisées du système. Ces coordonnées, qui ne sont pas forcément cartésiennes, caractérisent complètement la position d'un système à  $s$  degrés de liberté. Leurs dérivées par rapport aux temps  $\dot{q}_1, \dot{q}_2, \dots, \dot{q}_s$  sont les vitesses généralisées du système. Le fait que la connaissance à un instant donné des positions et des vitesses d'un système suffise à caractériser complètement son évolution ultérieure vient du fait que de façon générale, les équations du mouvement d'un système sont des équations différentielles du second ordre.

Supposons que la position du système soit bien déterminée aux instants  $t_1$  et  $t_2$ . On note  $q^{(1)}, q^{(2)}$  ces positions.

Le principe d'Hamilton, ou principe de moindre action, postule que le système se déplace de telle sorte que l'action

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t) dt \quad (9.1)$$

soit minimale (plus généralement extrémale pour l'ensemble de la trajectoire, la condition de minimalité n'étant valable que pour des portions suffisamment petites de la trajectoire). La trajectoire du système peut alors être déterminée en résolvant les équations d'Euler-Lagrange, qui s'établissent par une méthode variationnelle.

Soit  $q(t)$  la fonction (supposée unique) qui rend  $S$  extrémale. Considérons une variation  $\delta q(t)$  de la trajectoire, qui vérifie  $\delta q(t_1) = \delta q(t_2) = 0$ . Cette variation est supposée petite, dans un sens que l'on peut préciser mathématiquement à condition de définir correctement l'espace

de fonction sur lequel agit la fonctionnelle action, et de le munir d'une norme [?]. Nous laissons de côté ici tout soucis de rigueur mathématique. La variation de l'action s'écrit alors

$$\begin{aligned}\delta S &= \int_{t_1}^{t_2} L(q + \delta q, \dot{q} + \delta \dot{q}, t) dt - \int_{t_1}^{t_2} L(q, \dot{q}, t) dt \\ &= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) dt = 0.\end{aligned}\quad (9.2)$$

En utilisant le fait que  $\delta \dot{q} = \frac{d}{dt} \delta q$  et en intégrant par partie, ceci s'écrit encore, puisque  $\delta q(t_1) = \delta q(t_2) = 0$ ,

$$\delta S = \left[ \frac{\partial L}{\partial \dot{q}} \delta q \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q dt = \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q dt = 0, \quad (9.3)$$

pour toute variation  $\delta q$ . On obtient donc les équations d'Euler Lagrange (pour un système à  $s$  degrés de liberté)

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad (i = 1, \dots, s). \quad (9.4)$$

Ce sont les équations du mouvement du système : si  $L$  est connu, alors les équations (9.4) permettent de relier accélérations, vitesses et coordonnées. Elles forment un système de  $s$  équations différentielles du second ordre à  $s$  fonctions inconnues  $q_i(t)$ . Leur solution générale contient  $2s$  constantes arbitraires qui doivent être précisées par les conditions initiales.

### Exercice 9.1

Montrer que les équations du mouvement ne sont pas modifiées si l'on ajoute au lagrangien la dérivée totale d'une fonction quelconque des coordonnées et du temps.

Considérons à titre d'exemple le cas d'un point matériel libre. Plaçons nous dans un référentiel galiléen, dans lequel par définition l'espace est homogène et isotrope, et le temps uniforme. Il est alors clair que le lagrangien du point matériel ne peut dépendre que du carré de sa vitesse.

### Exercice 9.2

En déduire le principe d'inertie, à savoir que le mouvement d'un point matériel libre dans un référentiel galiléen est rectiligne et uniforme.

## 9.1.2 Principe de relativité galiléen :

Il existe une infinité de référentiels galiléens, qui sont animés les uns par rapport aux autres d'un mouvement rectiligne et uniforme, dans lesquels les propriétés du temps et de l'espace sont les mêmes, ainsi que les lois de la dynamique. Les lois de transformation sont données par un groupe à 10 paramètres :

$$\begin{cases} t' = t + t_0 \\ x'_i = R_{ij} x_j + V_i t + a_i \end{cases} \quad (9.5)$$

où  $R_{ij}$  est une matrice orthogonale et  $V_i$  représente les 3 composantes de la vitesse.

### Exercice 9.3

En utilisant le principe de relativité galiléen, montrer que le lagrangien d'un point matériel libre est de la forme

$$L = av^2 . \quad (9.6)$$

On pourra considérer deux référentiels galiléens qui se déplacent l'un par rapport à l'autre avec une vitesse infiniment petite, et utiliser le résultat de l'exercice 9.1.

Par convention, on fixe  $a = m/2$ , ce qui définit la masse  $m$  du point matériel. Cette quantité doit être positive d'après le principe de moindre action

### 9.1.3 Fonction de lagrange pour un système de points matériels

Pour un système fermé de points matériels en interaction, le lagrangien s'écrit

$$L = T - U \quad (9.7)$$

où  $T = \frac{1}{2} \sum_a m_a \dot{x}_a^2$  est l'énergie cinétique, la somme portant sur les points du système matériel, et  $U(\vec{x}_1, \vec{x}_2, \dots)$  est l'énergie potentielle. Si au lieu de coordonnées cartésiennes on utilise des coordonnées généralisées pour décrire le mouvement du système, il faut effectuer la transformation

$$x_a = f_a(q_1, q_2, \dots, q_s), \quad (9.8)$$

et donc

$$\dot{x}_a = f_a(\dot{q}_1, \dot{q}_2, \dots, \dot{q}_s). \quad (9.9)$$

La fonction de Lagrange aura alors la forme

$$L = \frac{1}{2} \sum_{i,j} a_{i,j}(q) \dot{q}_i \dot{q}_j - U(q), \quad (9.10)$$

où les fonctions  $a_{i,j}$  ne dépendent que des coordonnées. Ainsi dans un système de coordonnées généralisées, l'énergie cinétique est toujours fonction quadratique des vitesses, mais elle peut dépendre aussi des coordonnées.

### Exercice 9.4

En utilisant les équations d'Euler-Lagrange, retrouver la relation fondamentale de la dynamique entre force et accélération.

## 9.2 Intégrales du mouvement

Nous avons vu plus haut (voir 9.1.1) que la solution des équations du mouvement d'un système fermé à  $s$  degrés de libertés fait intervenir  $2s$  constantes. En fait, à cause de l'uniformité du

temps, une des constantes du mouvement peut être éliminée, de sorte qu'il reste  $2s - 1$  équations du mouvement. Nous allons considérer ici successivement trois types d'intégrales premières tout à fait fondamentales : l'énergie, l'impulsion et le moment cinétique, qui sont directement reliées respectivement aux propriétés d'uniformité du temps, de l'espace, et d'isotropie spatiale, dans les référentiels galiléens.

Nous verrons plus loin grâce au théorème de Noether comment de façon générale toute symétrie continue de l'action entraîne l'existence d'une quantité conservée.

### Exercice 9.5

En utilisant l'uniformité du temps pour un système fermé et en se servant des équations d'Euler-Lagrange, montrer que l'énergie définie par

$$E = \dot{q} \frac{\partial L}{\partial \dot{q}} - L \quad (9.11)$$

est conservée.

### **Exercice 9.6**

En utilisant l'uniformité de l'espace pour un système fermé et en se servant des équations d'Euler-Lagrange, montrer que l'impulsion définie par

$$\vec{p} = \sum_a \frac{\partial L}{\partial \dot{q}_a} \quad (9.12)$$

est conservée (la somme sur  $a$  représente la somme sur tous les points du matériels système). On pourra considérer une translation globale infinitésimale du système et calculer la variation correspondante du lagrangien. Quelle est l'expression de l'impulsion dans le cas d'une particule libre ? Ecrire les équations d'Euler-Lagrange en terme des impulsions et des forces.

Dans le cas où l'on utilise des coordonnées généralisées  $q_i$  ( $i=1, \dots, s$ ), les impulsions généralisées sont définies par

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (9.13)$$

et les forces généralisées par

$$F_i = \frac{\partial L}{\partial q_i} \quad (9.14)$$

Dans le cas de coordonnées cartésiennes en mécanique, les coordonnées  $p_i$  sont simplement les composantes du vecteur  $\vec{p}_a$ . Mais dans le cas général, l'impulsion est une fonction linéaire homogène des vitesses généralisée qui ne se ramène pas au produit d'une masse par une vitesse.

### Exercice 9.7

En utilisant l'isotropie de l'espace pour un système fermé et en utilisant les équations d'Euler-Lagrange, montrer que le moment cinétique défini par

$$\vec{L} = \sum_a \vec{r}_a \wedge \vec{p}_a \quad (9.15)$$

est conservé. On pourra considérer une rotation globale infinitésimale du système et calculer la variation correspondante du lagrangien.

## 9.3 Equations canoniques

### 9.3.1 Equations de Hamilton

Dans l'approche lagrangienne étudiée plus haut, les lois de la mécanique sont formulées en supposant que l'état du système est décrit par les coordonnées et les vitesses généralisées de celui-ci. Il est possible de passer d'un choix de variables indépendantes à un autre en effectuant une transformation de Legendre. Elle consiste à remplacer une variable dynamique par sa variable conjuguée. Considérons pour cela la différentielle totale du lagrangien :

$$dL = \sum_i \frac{\partial L}{\partial q_i} dq_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i. \quad (9.16)$$

En utilisant la définition des impulsions généralisées et les équations d'Euler-Lagrange, on peut remplacer cette expression par

$$dL = \sum_i \dot{p}_i dq_i + \sum_i p_i d\dot{q}_i. \quad (9.17)$$

On définit la fonction de Hamilton (ou hamiltonien) du système par la transformée de Legendre du lagrangien par rapport au couple  $(p_i, \dot{q}_i)$

$$H(p, q, t) = \sum_i p_i \dot{q}_i - L. \quad (9.18)$$

C'est l'énergie du système. On constate que la différentielle de  $H$  s'écrit

$$dH = - \sum_i \dot{p}_i dq_i + \sum_i \dot{q}_i dp_i, \quad (9.19)$$

ce qui est exactement le but recherché : on a remplacé la description du système en terme de coordonnées et de vitesses  $(q_i, \dot{q}_i)$  par une description en terme de coordonnées et d'impulsion  $(q_i, p_i)$ .

Cette transformation est très utilisée en thermodynamique par exemple, lorsque l'on remplace l'étude de l'énergie interne  $U(\Omega, S, N)$  par celle de l'énergie libre  $F(\Omega, T, N) = U - TS$  (où  $S$  est l'entropie,  $\Omega$  est le volume,  $N$  le nombre de particules, et  $T$  la température), remplaçant ainsi  $dU = -Pd\Omega + TdS + \mu dN$  par  $dF = -Pd\Omega - SdT + \mu dN$ .

Les équations du mouvement dans ce nouveau couple de variables sont maintenant

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = - \frac{\partial H}{\partial q_i} \quad (9.20)$$

comme on peut le voir immédiatement d'après l'expression de la différentielle de  $H$  (9.19). Ces équations dites d'Hamilton (ou canoniques à cause de leur simplicité et de leur symétrie) forment un système de  $2s$  équations différentielles du premier ordre à  $2s$  fonctions inconnues  $q(t)$  et  $p(t)$ , qui remplacent les  $s$  équations du second ordre d'Euler-Lagrange.

### Exercice 9.8

Montrer que

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} . \quad (9.21)$$

Retrouver ainsi la loi de conservation de l'énergie.

Dans le cas où le système dépend également d'un paramètre  $\lambda$  qui n'est pas une variable dynamique (exemple : champ extérieur dans lequel le point matériel est plongé, voir chapitre ??), montrer que

$$\left. \frac{\partial H}{\partial \lambda} \right|_{p,q} = - \left. \frac{\partial L}{\partial \lambda} \right|_{q,\dot{q}} \quad (9.22)$$

En déduire que dans le cas où le lagrangien dépend explicitement du temps,

$$\left. \frac{\partial H}{\partial t} \right|_{p,q} = - \left. \frac{\partial L}{\partial t} \right|_{q,\dot{q}} \quad (9.23)$$

## 9.3.2 Fonction de Routh

Il est parfois utile de n'effectuer la transformation de Legendre précédente que pour certains couples de variables  $(\dot{q}, p)$ . Supposons par exemple que la dynamique du système est décrite par deux variables  $q$  et  $\xi$ , et effectuons une transformée de Legendre par rapport au couple  $(\dot{q}, p)$  uniquement. Il est alors naturel d'introduire la fonction de Routh

$$R(q, p, \xi, \dot{\xi}) = p\dot{q} - L . \quad (9.24)$$

### Exercice 9.9

Montrer que la fonction de Routh est une fonction de Hamilton par rapport à la coordonnée  $q$  et une fonction de Lagrange par rapport à la coordonnée  $\xi$  :

$$\dot{q} = \frac{\partial R}{\partial p} , \quad \dot{p} = - \frac{\partial R}{\partial q} \quad (9.25)$$

$$\frac{d}{dt} \frac{\partial R}{\partial \dot{\xi}} = \frac{\partial R}{\partial \xi} . \quad (9.26)$$

Montrer que l'énergie s'écrit

$$E = R - \dot{\xi} \frac{\partial R}{\partial \dot{\xi}} . \quad (9.27)$$

La généralisation au cas de plus de 2 variables est immédiate. L'intérêt de cette fonction de Routh apparaît dans les systèmes possédant une coordonnée cyclique. On dit que  $q$  est une coordonnée cyclique si elle n'entre pas explicitement dans le lagrangien. Il est alors immédiat

que l'impulsion  $p$  est constante, comme on peut le voir d'après les équations canoniques (9.25) ou d'après l'équation d'Euler-Lagrange pour la coordonnée  $q$ . Les équations d'Euler-Lagrange pour les coordonnées  $\xi_i$  (9.26) se réduisent alors simplement à des équations contenant les coordonnées  $\xi_i$  (notées symboliquement  $\xi$ )

$$\frac{d}{dt} \frac{\partial R(p, \xi, \dot{\xi})}{\partial \dot{\xi}} = \frac{\partial R(p, \xi, \dot{\xi})}{\partial \xi} . \quad (9.28)$$

Les coordonnées cycliques disparaissent donc complètement des équations du mouvement,  $p$  jouant simplement le rôle d'un paramètre. La méthode consiste donc à trouver les fonctions  $\xi(t)$  solutions des équations (9.28) puis à obtenir la coordonnée cyclique  $q(t)$  par intégration de l'équation canonique

$$\dot{q} = \frac{\partial R(p, \xi, \dot{\xi})}{\partial p} . \quad (9.29)$$

### 9.3.3 Crochets de Poisson

Il est intéressant de reformuler les équations canoniques à l'aide des crochets de Poisson. Ce formalisme s'avère en particulier très utile pour comprendre le passage de la dynamique classique à la dynamique quantique, que le système possède un nombre fini de degrés de liberté, comme c'est le cas dans ce chapitre, ou qu'il possède un nombre infini de degrés de liberté, situation que nous examinerons dans le chapitre 10.

On définit le crochet de Poisson de deux fonctions  $f$  et  $g$  agissant sur l'espace  $R^{2s}$  par

$$\{f, g\} = \sum_{i=1}^s \left( \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i} \right) . \quad (9.30)$$

#### Exercice 9.10

Montrer que

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{f, H\} . \quad (9.31)$$

Une condition pour que  $f$  soit intégrale première du mouvement est donc que

$$\frac{\partial f}{\partial t} + \{f, H\} = 0 , \quad (9.32)$$

soit, dans le cas où cette intégrale première ne dépend pas explicitement du temps,

$$\{f, H\} = 0 . \quad (9.33)$$

#### Exercice 9.11

1. Montrer que l'ensemble des fonctions  $f(p, q)$  différentiable sur l'espace  $R^{2s}$  muni de l'opérateur crochet de Poisson définit une algèbre de Lie de dimension  $\infty$  (voir chap. 12).

2. Montrer que

$$\{f, \text{constante}\} = 0 \quad (9.34)$$

$$\{fg, h\} = f\{g, h\} + g\{f, h\} \quad (9.35)$$

$$\frac{\partial}{\partial t}\{f, g\} = \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} \quad (9.36)$$

$$\{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij} \quad (9.37)$$

$$\{q, f\} = \frac{\partial f}{\partial p}, \quad \{p, f\} = -\frac{\partial f}{\partial q} \quad (9.38)$$

$$\dot{q} \equiv \frac{dq}{dt} = \{q, H\}, \quad \dot{p} \equiv \frac{dp}{dt} = \{p, H\}. \quad (9.39)$$

### Théorème de Poisson

Si  $f$  et  $g$  sont des intégrales premières, alors  $\{f, g\}$  est aussi une intégrale première.

#### Exercice 9.12

Démontrer ce théorème. On pourra utiliser la propriété (9.35) et l'identité de Jacobi satisfaite par le crochet de Poisson (voir exercice 9.11).

#### Exercice 9.13

- 1- Calculer les crochets de Poisson formés à partir des composantes du moment cinétique.
- 2- Calculer les crochets de Poisson formés à partir des composantes du moment cinétique et de l'impulsion.
- 3- Montrer que le crochet de Poisson d'une fonction scalaire des coordonnées et des impulsions avec les composantes du moment cinétique est nul.
- 4- Soit  $\vec{n}$  un vecteur unitaire. Montrer que

$$\{f, \vec{L} \cdot \vec{n}\} = \vec{n} \wedge \vec{f}, \quad (9.40)$$

où  $\vec{f}$  est une fonction vectorielle des coordonnées et des impulsions.

## 9.4 L'action en fonction des coordonnées

### 9.4.1 Variation de l'action pour deux trajectoires réelles voisines

Lorsque nous avons formulé le principe de moindre action dans la partie 9.1, nous avons envisagé des variations de l'action entre deux valeurs fixées  $q(t_1)$  et  $q(t_2)$  des coordonnées du



ystème considéré à deux instants fixés  $t_1$  et  $t_2$ . La trajectoire réelle correspond alors à celle qui minimise l'action  $S$ , les contraintes précédentes étant satisfaites.

Nous allons maintenant étudier les variations de l'action lorsque l'on considère différentes trajectoires réelles. On peut tout d'abord considérer la variation de l'action lorsque l'on impose aux trajectoires de passer par la même position  $q(t_1) = q^{(1)}$  au même instant  $t_1$  mais en des positions différentes à l'instant  $t_2$ . On peut également considérer l'action comme une fonction explicite du temps, c'est à dire étudier ses variations lorsque l'on impose aux trajectoires réelles de passer par la même position  $q(t_1) = q^{(1)}$  au même instant  $t_1$ , et à la même position  $q^{(2)}$  mais en des instants différents  $t_2$ .

Nous allons ici directement envisager une transformation générale infinitésimale de la trajectoire, pour laquelle les coordonnées et le temps varient tous deux à la fois au début et à la fin du mouvement, comme illustré dans la fig. 9.1. Soit donc  $q(t)$  la trajectoire définie de  $t_1$  à  $t_2$ , so-

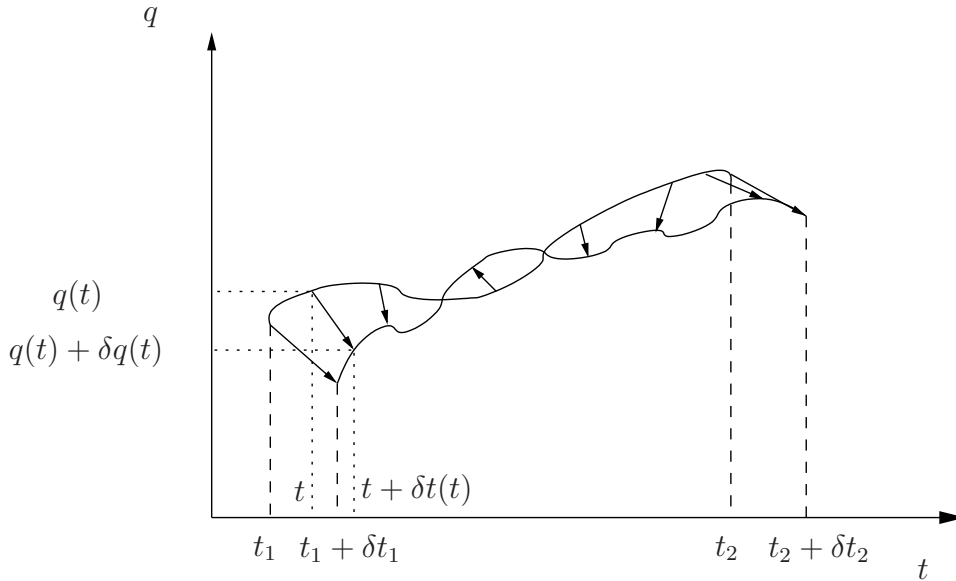


FIGURE 9.1 – Variation de la trajectoire.

lution des équations d'Euler-Lagrange. Considérons une transformation générale infinitésimale de cette trajectoire de la forme

$$\begin{aligned} t' &= t + \delta t(t) \\ q'(t') &= q(t) + \delta q(t) . \end{aligned} \tag{9.41}$$

Notons  $\bar{\delta}q(t)$  la différence entre les coordonnées de la trajectoire transformée et la trajectoire initiale mesurées au même instant  $t$ .

$$\begin{aligned} \bar{\delta}q(t) &= q'(t) - q(t) \\ &= q(t - \delta t) + \delta q(t - \delta t) - q(t) \\ &= \delta q(t) - \dot{q}(t) \delta t + \dots (\text{termes d'ordre supérieur en } \delta t) \end{aligned} \tag{9.42}$$

L'action de la trajectoire initiale s'écrit, en notant  $\delta t_{1,2} = \delta t(t_{1,2})$ ,

$$S = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt. \quad (9.43)$$

Après la transformation (9.41), elle s'écrit, le long de la trajectoire ainsi obtenue,

$$\begin{aligned} S' &= \int_{t_1+\delta t_1}^{t_2+\delta t_2} L(q'(t), \dot{q}'(t), t) dt \\ &= \int_{t_1+\delta t_1}^{t_1} L(q'(t), \dot{q}'(t), t) dt + \int_{t_2}^{t_2+\delta t_2} L(q'(t), \dot{q}'(t), t) dt + \int_{t_1}^{t_2} L(q'(t), \dot{q}'(t), t) dt, \end{aligned} \quad (9.44)$$

soit encore, en négligeant les termes d'ordre supérieur à  $\delta t$ ,

$$\begin{aligned} S' &= [\delta t L(q'(t), \dot{q}'(t), t)]_{t_1}^{t_2} + \int_{t_1}^{t_2} L(q'(t), \dot{q}'(t), t) dt \\ &= [\delta t L(q(t), \dot{q}(t), t)]_{t_1}^{t_2} + \int_{t_1}^{t_2} L(q(t) + \bar{\delta}q(t), \dot{q}(t) + \bar{\delta}\dot{q}(t), t) dt. \end{aligned} \quad (9.45)$$

Examinons le dernier terme. Il s'écrit encore, en développant au premier ordre en  $\bar{\delta}q$  et  $\bar{\delta}\dot{q}$ ,

$$\begin{aligned} &\int_{t_1}^{t_2} \left\{ L(q(t), \dot{q}(t), t) + \frac{\partial L}{\partial q} \bar{\delta}q + \frac{\partial L}{\partial \dot{q}} \bar{\delta}\dot{q} \right\} dt \\ &= \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t) dt + \int_{t_1}^{t_2} \left\{ \frac{\partial L}{\partial q} \bar{\delta}q - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) \bar{\delta}q \right\} dt + \int_{t_1}^{t_2} dt \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \bar{\delta}q \right) \\ &= \frac{\partial L}{\partial \dot{q}} \bar{\delta}q \Big|_{t_1}^{t_2} + S + \int_{t_1}^{t_2} \left\{ \frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right\} \bar{\delta}q dt. \end{aligned} \quad (9.46)$$

Le dernier terme de l'équation précédente s'annule puisque  $q(t)$  est solution des équations du mouvement. Ainsi

$$dS = \left[ \frac{\partial L}{\partial \dot{q}} \bar{\delta}q + L \delta t \right]_{t_1}^{t_2}. \quad (9.47)$$

Il est intéressant de remarquer que dans le calcul menant à l'expression précédente, on a uniquement utilisé le fait que la trajectoire initiale est solution des équations du mouvement. Il semble au premier abord que la trajectoire transformée puisse être quelconque. En fait, puisque  $S$  est extrémale dans l'espace des trajectoires pour la trajectoire réelle initiale qui va de  $(q^{(1)}, t_1)$  à  $(q^{(2)}, t_2)$ , une trajectoire voisine quelconque (i.e a priori non solution des équations du mouvement) ne différant que d'une "petite" quantité  $\bar{\delta}q$  (au sens de la norme dans l'espace sur lequel agit la fonctionnelle action) rendra l'action extrémale (à des termes en  $\bar{\delta}q^2$  près), et sera donc solution des équations du mouvement à l'ordre  $\bar{\delta}q$  considéré.

Ceci se traduit par le fait que la contribution correspondante au calcul précédent est d'ordre  $\bar{\delta}q^2$ . La seule contribution à la variation de l'action ne peut provenir que des termes de bord, comme le montre effectivement l'équation (9.47). Celle-ci peut encore s'écrire, en restaurant la

dépendance explicite dans les  $2s$  degrés de liberté, et en utilisant la définition du hamiltonien et de l'impulsion,

$$dS = \sum_i \frac{\partial L}{\partial \dot{q}_i} \delta q_i - \left( \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \right) \delta t \Big|_{t_1}^{t_2} = \sum_i p_i \delta q_i - H \delta t \Big|_{t_1}^{t_2}. \quad (9.48)$$

Considérons une transformation qui ne met en jeu que les coordonnées  $q = q^{(2)}$  et l'instant  $t = t_2$  finals. De l'expression correspondante

$$dS = \sum_i p_i dq_i - H dt \quad (9.49)$$

on voit alors immédiatement que les dérivées partielles de l'action par rapport aux coordonnées sont égales aux impulsions correspondantes :

$$\frac{\partial S}{\partial q_i} = p_i, \quad (9.50)$$

et que la dérivée partielle de l'action par rapport au temps est égale à l'opposé du hamiltonien :

$$\frac{\partial S}{\partial t} = -H, \quad (9.51)$$

Le fait que l'expression (9.49) doive être une différentielle impose des contraintes sur les mouvements possibles. En particulier, les lois de l'optique géométrique peuvent s'établir de cette façon. On trouvera un exposé détaillé de ces techniques dans le Landau de théorie des champs, §53-55 [7].

### Exercice 9.14

En utilisant l'expression (9.49) pour calculer l'action et en considérant les coordonnées et les impulsions comme des variables indépendantes, montrer que le principe de moindre action conduit aux équations de Hamilton (9.20).

## 9.4.2 Théorème de Noether

Ce théorème constitue une conséquence importante de la formule (9.48) donnant la variation de l'action sous une transformation continue quelconque : à toute transformation générale continue qui laisse invariante l'action stationnaire correspond une constante du mouvement.

En effet si l'action est invariante sous la transformation infinitésimale (9.41), on en déduit que

$$p \delta q - H \delta t \quad (9.52)$$

est une constante du mouvement, et la généralisation pour toute transformation *continue* laissant l'action invariante est immédiate. Notons que le lagrangien lui-même n'est pas forcément

invariant sous la transformation. On exige uniquement l'invariance de l'action.

### Exercice 9.15

- 1- En appliquant le théorème précédent, retrouver la loi de conservation de l'énergie pour un hamiltonien indépendant du temps.
- 2- En utilisant l'uniformité de l'espace pour un système fermé dans un référentiel galiléen, retrouver la conservation de l'impulsion.
- 3- En utilisant l'isotropie de l'espace pour un système fermé dans un référentiel galiléen, retrouver la conservation du moment cinétique.

Nous verrons sa généralisation aux systèmes comportant un nombre infini de degrés de liberté dans le chapitre 10. Ce théorème joue un rôle central en théorie quantique des champs car il permet, connaissant les symétries d'un système, de deviner la forme du lagrangien possédant une telle symétrie au niveau classique. Cette symétrie peut éventuellement disparaître au niveau quantique, lorsque l'on quantifie le champ, laissant alors apparaître des anomalies.

### Exercice 9.16

Montrer que la quantité conservée  $Q = \sum_i p_i \delta q_i - H \delta t$  est le générateur des transformations infinitésimales :

$$\{Q, q_j\} = -\bar{\delta} q_j \quad (9.53)$$

et

$$\{Q, p_j\} = -\bar{\delta} p_j . \quad (9.54)$$

### 9.4.3 Principe de Maupertuis

Nous avons vu que le principe de moindre action détermine complètement le mouvement d'un système mécanique. En résolvant les équations du mouvement, on obtient ainsi la relation entre les coordonnées et le temps. En particulier on peut obtenir la forme des trajectoires. Si l'on s'intéresse seulement à la détermination des trajectoires, il est possible de simplifier le principe de moindre action. Nous allons supposer dans ce qui suit que le lagrangien, et donc l'hamiltonien, *ne dépendent pas explicitement du temps*. Alors l'énergie du système se conserve

$$H(p, q) = E = cste . \quad (9.55)$$

Si l'on considère la variation de l'action pour des valeurs initiales et finale des coordonnées données et pour une valeur initiale donnée  $t_1$  du temps, avec un temps final  $t$  variable, on obtient, d'après l'équation (9.48),

$$\delta S = -H \delta t , \quad (9.56)$$

qui s'écrit encore, si l'on se restreint aux trajectoires satisfaisant à la conservation de l'énergie,

$$\delta S + E \delta t = 0 . \quad (9.57)$$

De même, la restriction aux trajectoires qui conservent l'énergie donne pour l'action l'expression

$$S = \int \sum_i p_i dq_i - E(t - t_1) \quad (9.58)$$

d'après l'expression générale (9.48) de la différentielle totale de l'action comme fonction des coordonnées et du temps à la limite supérieure de l'intégrale. En substituant l'expression précédente dans l'équation (9.57) il vient

$$\delta S_0 = 0, \quad (9.59)$$

où l'action réduite est définie par

$$S_0 = \int \sum_i p_i dq_i. \quad (9.60)$$

Ainsi nous pouvons énoncer le Principe de Maupertuis : l'action réduite possède un minimum sur l'ensemble des trajectoires satisfaisant à la conservation de l'énergie et passant par le point final donné à un instant arbitraire. Cet énoncé répond donc bien à la question initiale puisqu'il nous dit comment déterminer la trajectoire, indépendamment du mouvement effectif du système. Le problème se ramène finalement à trouver le minimum ne dépendant que des variables  $q$  et de leurs différentielles  $dq$ , avec l'énergie comme paramètre.

D'après la définition de  $p$ ,

$$p_i = \frac{\partial}{\partial \dot{q}_i} L \left( q, \frac{dq}{dt} \right). \quad (9.61)$$

La conservation de l'énergie s'écrit

$$E \left( q, \frac{dq}{dt} \right) = E, \quad (9.62)$$

d'où l'on tire l'expression de la différentielle  $dt$  en fonction de  $q$  et  $dq$ ,

$$dt = f(q, dq). \quad (9.63)$$

De l'équation (9.61) on tire alors  $p = g(q, dq, E)$ , où  $E$  joue le rôle d'un paramètre. Examinons en particulier comment cette suite de transformations s'opère dans le cas du lagrangien d'un système de points matériels (9.10). Partant de

$$L = \frac{1}{2} \sum_{i,j} a_{i,j}(q) \dot{q}_i \dot{q}_j - U(q), \quad (9.64)$$

on tire

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \sum_j a_{ij}(q) \dot{q}_j \quad (9.65)$$

et

$$E = \frac{1}{2} \sum_{ij} a_{ij}(q) \dot{q}_i \dot{q}_j + U(q), \quad (9.66)$$

d'où

$$dt = \sqrt{\frac{\sum_{ij} a_{ij}(q) dq_i dq_j}{2(E - U)}} \quad (9.67)$$

et

$$\sum_i p_i dq_i = \sum_{ij} a_{ij}(q) \frac{dq_j}{dt} dq_i. \quad (9.68)$$

Ceci permet donc d'écrire l'action réduite sous la forme

$$S_0 = \int \sum_i p_i dq_i = \int \sqrt{2(E - U) \sum_{ij} a_{ij} dq_i dq_j}. \quad (9.69)$$

Prenons l'exemple d'un point matériel, dont la partie cinétique du lagrangien est donnée par l'expression (9.6)

$$T = \frac{m}{2} \left( \frac{dl}{dt} \right)^2, \quad (9.70)$$

où  $dl$  est l'élément de longueur sur la trajectoire. D'après l'équation (9.59),

$$\delta \int \sqrt{2m(E - U)} dl = 0, \quad (9.71)$$

l'intégrale étant prise entre deux points donnés de l'espace. Dans le cas du mouvement libre,  $U = 0$  donc l'équation (9.71) devient

$$\delta \int dl = 0, \quad (9.72)$$

c'est-à-dire que la particule suit le plus court chemin, qui est un segment de droite.

Si l'on revient à l'expression

$$S = \int \sum_i p_i dq_i - E(t - t_1) = S_0 - E(t - t_1), \quad (9.73)$$

et que l'on fait une variation du paramètre  $E$ , on obtient

$$\delta S = \frac{\partial S_0}{\partial E} \delta E - (t - t_1) \delta E - E \delta t. \quad (9.74)$$

Or  $\delta S + E \delta t = 0$  donc

$$\frac{\partial S_0}{\partial E} = t - t_1. \quad (9.75)$$

En utilisant l'expression (9.69) on obtient donc

$$t - t_1 = \int \sqrt{\frac{\sum_{ij} a_{ij}(q) dq_i dq_j}{2(E - U)}}, \quad (9.76)$$

qui n'est autre que l'intégrale de l'expression (9.67) obtenue pour  $dt$ .

Ainsi nous avons obtenu non seulement la trajectoire du système, mais également la loi de son mouvement au cours du temps, puisque nous avons pu exprimer  $t$  comme fonction des positions du système au même instant.

## 9.5 Transformations canoniques

Les transformations

$$(q_1, q_2, \dots, q_s) \rightarrow (Q_1, Q_2, \dots, Q_s) \quad (9.77)$$

ne modifient pas les équations d'Euler-Lagrange. Plus généralement, les transformations, dites ponctuelles, du type  $Q_i = Q_i(q, t)$  laissent les équations du mouvement invariées.

Elles laissent aussi les équations d'hamilton invariées. En fait ces équations d'Hamilton ont une classe d'invariance beaucoup plus grande. Considérons donc une transformation du type

$$\begin{cases} Q_i = Q_i(p, q, t) \\ P_i = P_i(p, q, t) \end{cases} \quad (9.78)$$

A quelle condition cette transformation laisse-t-elle invariée la forme canonique des équations du mouvement ?

### 9.5.1 Fonction génératrice d'une transformation canonique

La condition s'exprime par le fait que  $Q_i$  et  $P_i$  doivent vérifier

$$\begin{cases} \dot{Q}_i = \frac{\partial H'}{\partial P_i} \\ \dot{P}_i = -\frac{\partial H'}{\partial Q_i} \end{cases} \quad (9.79)$$

Or les équations d'Hamilton découlent du principe variationnel

$$\delta \int (\sum_i p_i dq_i - H dt) = 0, \quad (9.80)$$

d'après l'exercice (9.14). On doit donc également avoir

$$\delta \int (\sum_i P_i dQ_i - H' dt) = 0, \quad (9.81)$$

c'est-à-dire que les expressions sous le signe somme ne doivent différer que par une différentielle totale  $dF$  :

$$\sum_i p_i dq_i - H dt = \sum_i P_i dQ_i - H' dt + dF. \quad (9.82)$$

$F$  est la fonction génératrice de la transformation, dite *canonique*. De l'expression

$$dF = \sum_i p_i dq_i - \sum_i P_i dQ_i + (H' - H)dt \quad (9.83)$$

on tire

$$\begin{cases} p_i = \frac{\partial F}{\partial q_i} \\ P_i = -\frac{\partial F}{\partial Q_i} \\ H' = H + \frac{\partial F}{\partial t} \end{cases} \quad (9.84)$$

pour  $F = F(q, Q, t)$ . Si la transformation est écrite en terme de  $q, P, t$ , il suffit d'effectuer une transformation de Legendre :

$$d(F + \sum_i P_i Q_i) = \sum_i p_i dq_i + \sum_i Q_i dP_i + (H' - H) dt. \quad (9.85)$$

En posant  $\Phi(q, P, t) = F + \sum_i P_i Q_i$ , on obtient donc

$$\begin{cases} p_i = \frac{\partial \Phi}{\partial q_i} \\ Q_i = \frac{\partial \Phi}{\partial P_i} \\ H' = H + \frac{\partial \Phi}{\partial t} \end{cases}. \quad (9.86)$$

Ces résultats appellent plusieurs remarques :

i) Si l'on prend  $\Phi = \sum_i f_i(q, t) P_i$ , on retrouve la classe des transformations ponctuelles  $Q_i = f_i(q, t)$ .

ii) Si la fonction génératrice ne dépend pas du temps, alors  $H' = H$ , donc il suffit de remplacer dans  $H$  les quantités  $p, q$  exprimées en fonction de  $P$  et  $Q$ .

iii) Considérons  $F = \sum_i q_i Q_i$ . Alors  $Q_i = p_i$  et  $P_i = -q_i$ . Cet exemple montre donc que la distinction coordonnées / impulsions est arbitraire. Mieux vaut donc parler pour  $(p_i, q_i)$  de grandeurs canoniquement conjuguées.

### 9.5.2 Condition en terme de crochets de Poisson

Montrons que condition pour que la transformation soit canonique peut s'écrire

$$\{f, g\}_{p,q} = \{f, g\}_{P,Q}, \quad (9.87)$$

pour toutes fonctions  $f$  et  $g$ .

Avant de démontrer ce résultat, notons que le temps  $t$  joue le rôle d'un paramètre dans les transformations canoniques, donc il suffit pour démontrer le résultat de considérer  $f$  et  $g$  ne dépendant pas explicitement de  $t$ .

Considérons  $g$  comme une fonction de hamilton. Alors

$$\frac{df}{dt} = \{f, g\}_{p,q}. \quad (9.88)$$

$\frac{df}{dt}$  ne peut dépendre que des propriétés du mouvement, et non du choix de variables. Donc

$$\frac{df}{dt} = \{f, g\}_{P,Q}. \quad (9.89)$$

Ainsi  $\{f, g\}_{p,q} = \{f, g\}_{P,Q}$ , ce qui prouve le résultat. Conséquence immédiate :

$$\begin{cases} \{Q_i, Q_j\}_{p,q} = 0 \\ \{P_i, P_j\}_{p,q} = 0 \\ \{P_i, Q_j\}_{p,q} = \delta_{ij} \end{cases} \quad (9.90)$$



On peut montrer que la variation de  $p$  et  $q$  lors du mouvement est une transformation canonique. En effet, notons  $q_t, p_t$  les coordonnées et les impulsions du système à l'instant  $t$ , et  $q_{t+\tau}, p_{t+\tau}$  leur valeur à l'instant  $t + \tau$ . Alors la transformation

$$\begin{cases} q_{t+\tau} = q(q_t, p_t, \tau) \\ p_{t+\tau} = p(q_t, p_t, \tau) \end{cases} \quad (9.91)$$

est canonique. En effet, pour  $t$  et  $t + \tau$  donnés, la différentielle de l'action s'écrit

$$dS = \sum p_{t+\tau} dq_{t+\tau} - p_t dq_t. \quad (9.92)$$

En comparant avec

$$dF = \sum_i p_i dq_i - \sum_i P_i dQ_i + (H' - H) dt, \quad (9.93)$$

on constate que  $F = -S$  (et  $H = H'$ ) : on a bien effectué une transformation canonique, la fonction génératrice de la transformation étant l'action elle-même.

### 9.5.3 Théorème de Liouville

L'élément de volume de l'espace des phases  $d\Gamma = dq_1 \dots dq_s dp_1 \dots dp_s$  est invariant sous les transformations canoniques.

Preuve : on veut montrer que  $\int \dots \int dq_1 \dots dq_s dp_1 \dots dp_s = \int \dots \int dQ_1 \dots dQ_s dP_1 \dots dP_s$ . Or cette dernière intégrale peut encore s'écrire, par changement de variable,

$$\int \dots \int dQ_1 \dots dQ_s dP_1 \dots dP_s = \int \dots \int \frac{\partial(Q_1, \dots, Q_s, P_1, \dots, P_s)}{\partial(q_1, \dots, q_s, p_1, \dots, p_s)} dq_1 \dots dq_s dp_1 \dots dp_s. \quad (9.94)$$

En notant  $J$  le jacobien de la transformation précédente, il nous faut donc montrer que  $J = 1$ . Or

$$J = \frac{\partial(Q_1, \dots, Q_s, P_1, \dots, P_s)}{\partial(q_1, \dots, q_s, P_1, \dots, P_s)} / \frac{\partial(q_1, \dots, q_s, p_1, \dots, p_s)}{\partial(q_1, \dots, q_s, P_1, \dots, P_s)} // = \frac{\partial(Q_1, \dots, Q_s)}{\partial(q_1, \dots, q_s)} \Big|_{P=cste} / \frac{\partial(p_1, \dots, p_s)}{\partial(P_1, \dots, P_s)} \Big|_{q=cste} \quad (9.95)$$

En écrivant la transformation canonique à l'aide de la fonction génératrice  $\Phi(q, P)$  on peut écrire

$$\frac{\partial Q_i}{\partial q_k} = \frac{\partial^2 \Phi}{\partial q_k \partial P_i} \quad \text{et} \quad \frac{\partial p_i}{\partial P_k} = \frac{\partial^2 \Phi}{\partial q_i \partial P_k}. \quad (9.96)$$

Le déterminant de ces deux matrices est donc identique. Donc  $J = 1$ , ce qui prouve le résultat annoncé.

Remarque : comme nous avons vu que la variation de  $p$  et  $q$  au cours du temps est une transformation canonique, on déduit du théorème de Liouville que le volume d'une portion d'espace des phases que l'on suit dans son mouvement est constant.

Ce résultat se généralise à n'importe quelle  $p$ -forme définie sur des variétés  $p$ -dimensionnelles de l'espace des phases. Exemple :  $\int \int \sum_i dq_i dp_i$ .



# Chapitre 10

## Théorie classique des champs

### 10.1 Introduction

Dans ce chapitre nous allons examiner comment généraliser le formalisme de la mécanique analytique aux systèmes comportant un nombre infini de degrés de liberté. Un champ est un système possédant un nombre infini de degrés de liberté. C'est le cas par exemple de la mécanique des fluides ou encore de l'électromagnétisme. Dans la suite, nous allons nous intéresser aux théories des champs relativistes, pour lesquels le groupe de symétrie d'espace-temps sera le groupe de Poincaré. Notons cependant que la forme des équations du mouvement pour le champ, de même que l'écriture de la seconde variation de l'action, ne font pas explicitement intervenir le groupe de symétrie d'espace temps envisagé (groupe de Galilée pour la mécanique des fluides non relativistes par exemple). Le passage de la mécanique analytique à la théorie des champs consiste à remplacer les coordonnées  $q_i(t)$ , en nombre fini (indexées par  $i$ ) par une collection infinie de coordonnées, indexées par un paramètre qui pourra être discret (exemple : théorie des champs sur réseau) ou continu. Dans la suite, on supposera que ces coordonnées sont indexées par les coordonnées spatiales supposées continues, et pour simplifier les notations, on discutera le cas d'un champ scalaire, qui sera noté génériquement  $\phi(\vec{x}, t)$ . Dans le cas relativiste, on notera  $\phi(x)$  ce champ,  $x$  étant alors la coordonnée quadridimensionnelle. Bien entendu, le champ pourra être vectoriel, spinoriel, tensoriel... Les lois de transformation de tels champs sous le groupe de symétrie d'espace-temps seront discutées en détail dans les chapitres suivants.

### 10.2 Formulation lagrangienne

On se limitera dans tout ce qui suit aux théories des champs *locales*, pour lesquelles l'action pourra s'écrire sous la forme

$$S = \int d^4x \mathcal{L}(x) \tag{10.1}$$

où  $\mathcal{L}(x)$  est la densité lagrangienne qui sera supposée ne dépendre que d'un nombre *fini* de dérivées des champs (en pratique, en général  $\mathcal{L}(x)$  ne dépendra que des champs et de leur dérivée première). Ceci interdit par définition la non localité : pour que  $\mathcal{L}(x)$  dépendent de la valeur des champs au point  $y$  distinct de  $x$  il faudrait que l'on puisse reconstruire, par développement en série, le champ  $\phi(y)$  à partir de  $\phi(x)$  et de ses dérivées en  $x$ , ce qui bien sûr

nécessite en général un nombre infini de dérivées.  $\mathcal{L}(x)$  sera donc de la forme

$$\mathcal{L}(x) = \mathcal{L}(\phi_i(x), \partial_\mu \phi_i(x), x). \quad (10.2)$$

L'indice  $i$  permet d'étiqueter les champs de différente nature ou, si les champs ne sont pas scalaire, leurs composantes (exemple :  $A_\mu(x)$  dans le cas de l'électromagnétisme). En général, on supposera que les champs s'annulent assez vite à l'infini, ce qui permet de se débarrasser des termes de bord lors des manipulation habituelles basées sur des intégrations par partie (exception notable : théories topologiques).

### 10.3 Equations d'Euler-Lagrange

On note  $\Omega$  le domaine d'espace temps<sup>1</sup> sur lequel on intègre la densité lagrangienne  $\mathcal{L}$  pour obtenir l'action  $S$ . La méthode pour obtenir les équations du mouvement est exactement la même qu'en mécanique : on varie cette fois l'action sur une trajectoire, en effectuant la transformation

$$\phi(x) \rightarrow \phi(x) + \delta\phi(x), \quad (10.3)$$

qui conduit à une variation de l'action de la forme

$$\delta S = \int_{\Omega} d^4x \left[ \frac{\delta\mathcal{L}(x)}{\delta\phi_i(x)} \delta\phi_i(x) + \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} \delta(\partial_\mu\phi_i(x)) \right]. \quad (10.4)$$

En intégrant par partie, on obtient donc

$$\delta S = \int_{\Omega} d^4x \left[ \frac{\delta\mathcal{L}(x)}{\delta\phi_i(x)} - \partial_\mu \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} \right] \delta\phi_i(x) + \int_{\partial\Omega} \delta\phi_i(x) \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} d^3\sigma_\mu \quad (10.5)$$

où  $\partial\Omega$  est le bord de  $\Omega$  et  $d^3\sigma_\mu$  est l'élément d'intégration sur ce bord. Le dernier terme de l'équation précédente est nul par hypothèse de décroissance des champs à l'infini. L'action devant être stationnaire au voisinage des solutions des équations du mouvement (donc pour toute variation  $\delta\phi_i(x)$ ), on déduit de (10.5) que

$$\frac{\delta\mathcal{L}(x)}{\delta\phi_i(x)} - \partial_\mu \frac{\delta\mathcal{L}(x)}{\delta\partial_\mu\phi_i(x)} = 0 \quad (10.6)$$

qui sont les équations d'Euler-Lagrange pour les champs  $\phi_i(x)$ . Notons que l'ajout d'une divergence totale à la densité lagrangienne ne modifie pas l'action.

## 10.4 Les symétries globales

### 10.4.1 Seconde variation de l'action

Le calcul de la seconde variation de l'action est à mettre en parallèle avec celui effectué en mécanique, voir 9.4.1. De la même façon, ce calcul permet d'obtenir la variation de l'action

---

1. En général,  $\Omega$  sera l'espace temps tout entier (et donc chaque composante de  $x$  variera sur  $]-\infty, +\infty[$ , mais il est utile de garder  $\Omega$  en toute généralité, en particulier en vue du théorème de Noether).

sous une transformation quelconque des coordonnées d'espace-temps et des champs, que ces transformations soient ou non des symétries de l'action.

Considérons donc une transformation quelconque

$$\begin{aligned} x'^{\mu} &= x^{\mu} + \delta x^{\mu} \\ \phi'(x') &= \phi(x) + \delta\phi(x) \end{aligned} \quad (10.7)$$

qui est l'analogie de la transformation (9.41). On obtient alors la variation locale du champ

$$\begin{aligned} \bar{\delta}\phi(x) &= \phi'(x) - \phi(x) \\ &= \phi(x - \delta x) + \delta\phi(x - \delta x) - \phi(x) \\ &= \delta\phi(x) - \delta x^{\mu} \partial_{\mu} \phi(x) + \dots (\text{termes d'ordre supérieur}) \end{aligned} \quad (10.8)$$

L'action pour les champs solutions des équations du mouvement qui s'écrit, avant transformation,

$$S = \int_{\Omega} d^4 \mathcal{L}(\phi(x), \partial_{\mu} \phi(x), x) \quad (10.9)$$

devient donc

$$S' = \int_{\Omega'} d^4 \mathcal{L}(\phi'(x), \partial_{\mu} \phi'(x), x) \quad (10.10)$$

où  $\Omega'$  est le transformé de  $\Omega$  par (10.7). Or  $S'$  s'écrit encore

$$S' = \int_{\Omega} d^4 x \mathcal{L}(\phi', \partial_{\mu} \phi', x) + \int_{\delta\Omega} d^3 \sigma_{\mu} \delta x^{\mu} \mathcal{L}(\phi', \partial_{\mu} \phi', x), \quad (10.11)$$

ce qui donne, par développement de Taylor à l'ordre 1, et en négligeant les termes d'ordre 2,

$$S' = S + \int_{\Omega} d^4 x \left( \frac{\delta \mathcal{L}}{\delta \phi} \bar{\delta} \phi + \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \partial_{\mu} \bar{\delta} \phi \right) + \int_{\delta\Omega} d^3 \sigma_{\mu} \delta x^{\mu} \mathcal{L}(\phi, \partial_{\mu} \phi, x) + \dots \quad (10.12)$$

En intégrant par partie, on obtient alors

$$\begin{aligned} S' &= S + \int_{\Omega} d^4 x \left( \frac{\delta \mathcal{L}}{\delta \phi} - \partial_{\mu} \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \right) \bar{\delta} \phi + \int_{\Omega} d^4 x \partial_{\mu} \left( \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \bar{\delta} \phi \right) \\ &\quad + \int_{\delta\Omega} d^3 \sigma_{\mu} \delta x^{\mu} \mathcal{L}(\phi, \partial_{\mu} \phi, x) + \dots \end{aligned} \quad (10.13)$$

Le second terme s'annule puisque les champs satisfont les équations du mouvement. En intégrant le troisième terme, on obtient ainsi

$$\begin{aligned} S' &= S + \int_{\delta\Omega} d^3 \sigma_{\mu} \left( \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \bar{\delta} \phi + \delta x^{\mu} \mathcal{L} \right) + \dots \\ &= S + \int_{\delta\Omega} d^3 \sigma_{\mu} \left( \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \delta \phi - \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \partial_{\nu} \phi \delta x^{\nu} + \delta x^{\mu} \mathcal{L} \right) + \dots \end{aligned} \quad (10.14)$$

où on a utilisé l'eq. (10.8). On obtient finalement

$$\delta S = \int_{\delta\Omega} d^3 \sigma_{\mu} \left( \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \delta \phi - \left( \frac{\delta \mathcal{L}}{\delta(\partial_{\mu} \phi)} \partial_{\nu} \phi - g^{\mu}_{\nu} \mathcal{L} \right) \delta x^{\nu} \right) + \dots, \quad (10.15)$$

en utilisant le fait que  $g^{\mu}_{\nu} \equiv \delta^{\mu}_{\nu}$ , voir chap. ??.

### 10.4.2 Théorème de Noether

Dans le cas particulier où les transformations considérées laissent invariante l'action, on obtient alors le théorème de Noether, qui énonce qu'à *tout groupe continu de symétrie de l'action est associé une quantité conservée*. Le courant conservé correspondant à cette symétrie s'écrit

$$j^\mu = \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} \delta \phi - \left( \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} \partial_\nu \phi - g^\mu{}_\nu \mathcal{L} \right) \delta x^\nu \quad (10.16)$$

puisque d'après la relation (10.15),

$$\delta S = 0 = \int_{\delta \Omega} d^3 \sigma_\mu j^\mu = \int_\Omega d^4 x \partial_\mu j^\mu$$

quelque soit  $\Omega$ , ce qui montre que le courant  $j^\mu$  est conservé :

$$\partial_\mu j^\mu = 0.$$

La charge conservée correspondante s'écrit

$$Q = \int d^3 x j^0(x). \quad (10.17)$$

En effet,

$$\frac{\partial Q}{\partial t} = \int d^3 x \partial_0 j^0(x) = - \int d^3 x \partial_i j^i(x) = 0$$

pour des champs rapidement décroissants à l'infini. On peut plus généralement définir la charge conservée par la relation

$$Q = \int_{N_3} d^3 \sigma_\mu j^\mu(x), \quad (10.18)$$

où  $N_3$  est une 3-surface de genre espace.

Les transformations envisagées précédemment sont a priori globales (appelées encore *rigides*). Aucune hypothèse sur le caractère *local* de ces transformations n'est ici nécessaire. Le fait d'imposer une invariance de l'action sous les transformation locales est une exigence supplémentaire que l'on impose dans le cas des théories de jauge, ce qui fournit un principe dynamique pour construire le lagrangien (exemple : électromagnétisme, théories de Yang-Mills).

### 10.4.3 Applications

#### Invariance par translation

Si la densité lagrangienne  $\mathcal{L}$  ne dépend pas explicitement de  $x^\mu$  (invariance par translation d'espace-temps), alors on peut considérer les transformations particulières

$$\begin{aligned} \delta x^\mu(x) &= \text{constante} = \delta x^\mu \\ \delta \phi &= 0. \end{aligned} \quad (10.19)$$

On en déduit que le courant

$$\left( \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} \partial^\nu \phi - g^{\mu\nu} \mathcal{L} \right) \delta x_\nu \quad (10.20)$$

est conservé, pour tout  $\delta x^\nu$  constant arbitraire, c'est-à-dire que le tenseur

$$T^{\mu\nu} = \left( \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)} \partial^\nu\phi - g^{\mu\nu}\mathcal{L} \right), \quad (10.21)$$

appelé *tenseur canonique d'énergie-impulsion du champ*, est conservé.

Par analogie immédiate avec la définition de l'impulsion en mécanique analytique, l'*impulsion du champ* est définie par

$$\Pi = \frac{\delta\mathcal{L}}{\delta(\partial_0\phi)} \quad (10.22)$$

d'où l'on déduit que

$$T^{0\nu} = \Pi \partial^\nu\phi - g^{0\nu}\mathcal{L} \quad (10.23)$$

et

$$T^{00} = \Pi \partial^0\phi - \mathcal{L} \equiv \mathcal{H}. \quad (10.24)$$

ce qui permet d'interpréter la charge  $\int T^{00} d^3x$  comme l'énergie totale du champ. Ainsi

$$P^\nu = \int T^{0\nu} d^3x, \quad (10.25)$$

qui se transforme comme un quadrivecteur, peut être identifié au quadrivecteur énergie-impulsion totale du champ.

On dispose ainsi à présent d'une formule générale pour le courant conservé par une transformation laissant invariante l'action : en combinant (10.16) et (10.21) on obtient en effet

$$j^\mu = \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)} \delta\phi - T^{\mu\nu} \delta x_\nu. \quad (10.26)$$

### Cas des lagrangiens indépendants des champs

Le cas des lagrangiens indépendants des champs présente un intérêt particulier. C'est le fait par exemple des champs de jauge, qui sont de masse nulle (on verra plus loin qu'un terme de masse dans le lagrangien briserait l'invariance de jauge ; les bosons massifs ne peuvent acquérir leur masse que par un mécanisme de brisure spontanée de symétrie d'un champ auxiliaire scalaire : c'est le mécanisme de Higgs).

Considérons par exemple le cas de QED en l'absence de matière. Le lagrangien correspondant s'écrit

$$\mathcal{L}_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (10.27)$$

avec

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (10.28)$$

Sous la transformation globale

$$\begin{aligned} \delta x^\mu &= 0 \\ \delta A^\mu(x) &= \text{constante} = \delta A^\mu, \end{aligned} \quad (10.29)$$

l'action du champ est laissée invariante, d'où l'on déduit la conservation du courant

$$j^\mu = \frac{\delta \mathcal{L}}{\delta(\partial_\mu A_\nu)} \delta A_\nu$$

pour tout  $\delta A^\mu$  constant, ce qui conduit donc à la conservation de

$$\frac{\delta \mathcal{L}}{\delta(\partial_\mu A_\nu)}.$$

Le calcul de ce courant pour le lagrangien (22.4) est élémentaire. L'antisymétrie de  $\partial_\mu A_\nu - \partial_\nu A_\mu$  permet de réécrire (22.4) sous la forme

$$\mathcal{L}_{QED} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) = -\frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu)\partial^\mu A^\nu, \quad (10.30)$$

qui mène immédiatement à

$$\frac{\delta \mathcal{L}}{\delta(\partial_\mu A_\nu)} = -F^{\mu\nu}.$$

La conservation de ce courant s'écrit

$$\partial^\mu F^{\mu\nu} = 0, \quad (10.31)$$

ce qui constitue la forme covariante du premier couple d'équations de Maxwell dans le vide. Notons que la même équation s'obtient en écrivant l'équation du mouvement satisfaite par le champ  $A^\mu$ .

### Invariance de Lorentz

Si le lagrangien est invariant sous les transformations de Lorentz, alors on peut envisager (dans le cas des champs scalaires) la transformation

$$\begin{aligned} \delta x^\nu &= \omega^{\nu\mu} x_\mu \\ \delta \phi &= 0, \end{aligned} \quad (10.32)$$

qui laisse invariante l'action, pour tout tenseur infinitésimal  $\omega^{\nu\mu}$  antisymétrique. D'après le théorème de Noether, on peut donc immédiatement en déduire que le courant

$$\left( \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} \partial_\nu \phi - g^{\mu\nu} \mathcal{L} \right) \omega_{\nu\rho} x^\rho \quad (10.33)$$

est conservé, ou encore, puisque  $\omega^{\nu\rho}$  est antisymétrique et en utilisant la définition du tenseur d'énergie-impulsion (10.21), que

$$(T^{\mu\nu} x^\rho - T^{\mu\rho} x^\nu) \omega^{\nu\rho} \quad (10.34)$$

est conservé, pour tout  $\omega^{\nu\rho}$ . On en tire finalement la conservation du tenseur de moment cinétique

$$J^{\mu,\nu\rho} = x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu}. \quad (10.35)$$

La charge correspondante s'écrit

$$J^{\nu\rho} = \int d^3x J^{0,\nu\rho} = \int d^3x (x^\nu T^{0\rho} - x^\rho T^{0\nu}). \quad (10.36)$$



C'est le moment angulaire total du champ, qui s'obtient bien, comme on pouvait s'y attendre, en intégrant la contribution en chaque point  $x$  du moment orbital. L'écriture explicite de la conservation de  $J^{\mu,\nu\rho}$  est instructive. En utilisant la conservation du tenseur  $T^{\mu\nu}$ , on obtient

$$\partial_\mu J^{\mu,\nu\rho} = \partial_\mu(x^\nu T^{\mu\rho}) - \partial_\mu(x^\rho T^{\mu\nu}) = g_\mu^\nu T^{\mu\rho} - g_\mu^\rho T^{\mu\nu} \quad (10.37)$$

soit encore

$$T^{\nu\rho} - T^{\rho\nu} = 0. \quad (10.38)$$

ce qui prouve que la conservation de  $J^{\mu,\nu\rho}$  a pour conséquence la symétrie de  $T^{\nu\rho}$ . Dans le cas général d'un champ qui n'est pas scalaire, la loi de transformation du champ sous le groupe de Lorentz  $\phi'(x') = \phi(x)$  doit être généralisée en  $\phi'_a(x') = S(\Lambda)_{ab}\phi_b(x)$  où  $S(\Lambda)$  est la matrice de la représentation considérée (spinorielle, vectorielle, ...), qui se réduit à l'identité dans le cas des champs scalaires. La structure du tenseur moment cinétique obtenu est alors de la forme

$$J^{\mu,\nu\rho} = x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu} + \Delta^{\mu\nu\rho} \quad (10.39)$$

où  $\Delta^{\mu\nu\rho}$  est antisymétrique en  $\nu$  et  $\rho$ , et correspond à la contribution du moment cinétique intrinsèque, i.e du spin. Le tenseur canonique d'énergie-impulsion n'est alors plus symétrique dans le cas général d'un champ de spin arbitraire. Or le tenseur d'énergie-impulsion n'est pas défini de manière univoque : de façon plus générale, le courant  $j^\mu$  associé à une symétrie continue de l'action peut être modifié sans changer la charge associée.

### Exercice 10.1

Vérifier que le courant  $j'^\mu = j^\mu + \partial_\rho X^{\rho\mu}$  (où  $X^{\rho\mu}$  est un tenseur antisymétrique fonction des champs et des coordonnées d'espace-temps, supposé s'annuler rapidement à grande distance) est également conservé, et que les charges associées à ces deux courants sont identiques.

Il est possible de préserver l'écriture (10.35) de la densité de moment cinétique du champ en fonction de la densité d'impulsion (*a priori* uniquement valable pour un champ scalaire), pour un champ de spin quelconque, en redéfinissant le tenseur énergie-impulsion. Le tenseur moment cinétique est alors également modifié, mais la relation entre ce tenseur modifié et la densité d'impulsion modifiée est la même que dans le cas scalaire. Il faut pour cela rendre le tenseur d'énergie-impulsion symétrique (on a vu en effet plus haut que c'est une condition nécessaire à la conservation du tenseur moment cinétique correspondant). Le tenseur d'énergie-impulsion obtenu porte le nom de tenseur de Belinfante.



# Chapitre 11

## Symétries en Mécanique Quantique

### 11.1 Introduction

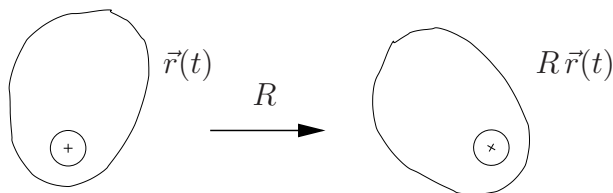
La première description du concept de symétrie est due à Galilée. Dans « Dialogue sur les deux plus grands systèmes du monde, celui de Ptolémée et celui de Copernic » publié à Florence en 1632, il introduit cette notion en prenant comme exemple le mouvement d'un objet à l'intérieur d'un bateau. Il montre que le mouvement ne dépend pas du fait que le bateau est immobile ou bien en mouvement uniforme.

De façon moderne, il montre l'invariance des lois de la mécanique sous les transformations de Galilée dite de translation uniforme  $\vec{x}' = \vec{x} - \vec{V}t$ . C'est le premier exemple d'une symétrie d'espace-temps.

Le lien entre les propriétés d'invariance et les lois de conservation a été formalisé beaucoup plus tard par E. Noether (1882-1935). Elle a montré comment exploiter les propriétés de symétrie dans le cadre Lagrangien : à une transformation de symétrie d'espace temps ou interne on peut associer des quantités conservées  $\vec{P}$ ,  $\vec{J}$ ,  $Q$ , comme nous l'avons vu aux chap. 9 et 10.

Ce théorème trouve son application aussi bien en physique classique qu'en physique quantique. Il existe cependant des différences considérables entre la mécanique classique et la mécanique quantique pour ce qui est du rôle joué par les symétries.

1) En mécanique classique les symétries des équations du mouvement permettent de construire de nouvelles solutions. Exemple : si les équations sont invariantes par rotation (cas du mouvement sous l'action d'une force centrale) et si  $\vec{r}(t)$  est une solution alors  $R\vec{r}(t)$  est encore solution



2) En mécanique quantique cette propriété reste vraie : si  $|\psi\rangle$  est solution,  $R|\psi\rangle$  est encore solution. La nouveauté par rapport au cas classique est le principe de superposition qui n'a pas

d'équivalent en mécanique classique. On peut ainsi considérer l'état  $|\psi\rangle + R|\psi\rangle$ . En particulier, l'état  $|\phi\rangle = \sum_R R|\psi\rangle$  est invariant par rotation :

$$R'|\phi\rangle = \sum_R R'R|\psi\rangle = \sum_{R''} R''|\psi\rangle = |\phi\rangle .$$

On obtient ainsi une représentation irréductible particulière du groupe des rotations (c'est la représentation scalaire). Wigner a montré qu'on pouvait considérer n'importe quel état comme une superposition d'états élémentaires se transformant selon une représentation irréductible du groupe de symétrie.

3) Les symétries discrètes  $P$ ,  $T$ , qui ne peuvent être obtenues continûment à partir de l'identité jouent un rôle essentiel (règles de sélection).

## 11.2 Description des états

Pour décrire un système physique en mécanique quantique on introduit un espace de Hilbert  $\mathcal{H}$ . L'état du système est décrit par un vecteur  $\psi \in \mathcal{H}$  de norme unité, i.e vérifiant  $\langle\psi|\psi\rangle = 1$ .

A toute grandeur physique mesurable on associe un opérateur autoadjoint  $A$  (i.e. tel que  $A = A^+$ ).

$$\left[ \begin{array}{l} \text{rappel : opérateur linéaire :} \\ \text{adjoint d'un opérateur linéaire :} \end{array} \quad \begin{array}{l} A(\lambda|\psi\rangle + \mu|\phi\rangle) = \lambda A|\psi\rangle + \mu A|\phi\rangle \\ \langle\phi|A^+\psi\rangle = \langle A\phi|\psi\rangle = \langle\psi|A\phi\rangle^* \end{array} \right.$$

L'élément de matrice  $\langle\psi|A\psi\rangle = \langle A\psi|\psi\rangle = \langle\psi|A|\psi\rangle$  représente donc la valeur moyenne de cette grandeur physique dans l'état  $|\psi\rangle$ . La probabilité de transition  $W_{12}$  d'un état  $|\psi_1\rangle$  à un état  $|\psi_2\rangle$  est donnée par  $W_{12} = |\langle\psi_1|\psi_2\rangle|^2$ .

## 11.3 Lois de symétrie

**Définition 11.1 :** *Opération de symétrie*

Une opération de symétrie est une correspondance entre *vecteurs d'états définis à une phase près* qui conserve les probabilités de transition : la transformation  $S$

$$\begin{aligned} \omega_1|\psi_1\rangle &\longrightarrow {}^S\omega_1|{}^S\psi_1\rangle \\ \omega_2|\psi_2\rangle &\longrightarrow {}^S\omega_2|{}^S\psi_2\rangle, \end{aligned} \tag{11.1}$$

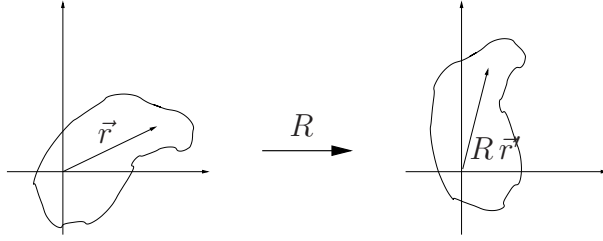
où  $|\omega_1| = |{}^S\omega_1| = |\omega_2| = |{}^S\omega_2| = 1$ , est telle que  $|\langle\psi_1|\psi_2\rangle|^2 = |\langle{}^S\psi_1|{}^S\psi_2\rangle|^2$ .

L'arbitraire de phase sur les vecteurs d'états nous conduit à définir un *rayon*  $|\underline{\psi}\rangle$  comme l'ensemble des vecteurs d'états qui ne diffèrent entre eux que par une phase  $e^{i\alpha}$ .

Une symétrie  $S$  est donc une correspondance entre rayons satisfaisant la contrainte  $|\langle \underline{\psi}_1 | \underline{\psi}_2 \rangle|^2 = |\langle {}^S \underline{\psi}_1 | {}^S \underline{\psi}_2 \rangle|^2$ .

*Exemple :*

On considère une particule non relativiste de spin 0. Elle est décrite par sa fonction d'onde  $\psi(\vec{r})$ . Considérons une rotation *active*<sup>1</sup> du système



Dans une telle rotation la fonction d'onde se transforme selon

$${}^R \psi(\vec{r}') = \psi(\vec{r}) \quad \text{où} \quad \vec{r}' = R\vec{r}, \quad (11.2)$$

et l'on a donc

$$\psi(\vec{r}) \xrightarrow{R} {}^R \psi(\vec{r}) = \psi(R^{-1}\vec{r}). \quad (11.3)$$

Une telle transformation satisfait manifestement à la conservation des probabilités de transition :

$$\langle \psi_1 | \psi_2 \rangle = \int \psi_1^*(\vec{r}) \psi_2(\vec{r}) d^3\vec{r} \quad (11.4)$$

devient

$$\begin{aligned} \langle {}^R \psi_1 | {}^R \psi_2 \rangle &= \int \psi_1^*(R^{-1}\vec{r}) \psi_2(R^{-1}\vec{r}) d^3\vec{r} \\ &= \int \psi_1^*(\vec{r}) \psi_2(\vec{r}) d^3\vec{r} = \langle \psi_1 | \psi_2 \rangle. \end{aligned} \quad (11.5)$$

Dans ce cas particulier, on a en fait un résultat plus fort, à savoir la conservation des amplitudes de transition elles-mêmes

$$\langle {}^R \psi_1 | {}^R \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle.$$

Introduisons l'opérateur  $U(R) \in H$  (ensemble des opérateurs linéaires sur  $\mathcal{H}$ ) tel que  $U(R)|\psi\rangle = |{}^R \psi\rangle$ . Alors

$$\langle {}^R \psi | = \langle \psi | U^+(R).$$

---

1. Au sens où c'est le système qui tourne, et non le repère (on parlerait alors de rotation *passive*); les deux points de vue sont bien sûr équivalents.

et l'on a donc

$$\langle \psi_1 | U^+(R) U(R) | \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle.$$

$U(R)$  est donc un opérateur unitaire.

Toutefois, exiger de façon générale qu'à toute loi de symétrie soit associée un opérateur unitaire est une exigence beaucoup trop forte. On peut aussi se contenter de  $\langle \psi_1 | \psi_2 \rangle = \langle {}^S \psi_1 | {}^S \psi_2 \rangle^*$ .

## 11.4 Théorème de Wigner

**Théorème 11.2 :** *de Wigner*

Soit  $S$  une transformation des rayons de  $\mathcal{H}$  qui conserve les probabilités de transition. Alors il existe un opérateur  $U(S)$  agissant sur  $\mathcal{H}$  qui

◇ induit la transformation  $S$ .

Ceci signifie que si  $|\psi\rangle$  est un représentant du rayon  $\underline{|\psi\rangle}$ , alors  $U(S)|\psi\rangle$  est un représentant du rayon  $\underline{{}^S \psi}$ .

◇ cet opérateur est :

– soit unitaire et linéaire :

$$U(S)(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle) = \lambda_1 U(S)|\psi_1\rangle + \lambda_2 U(S)|\psi_2\rangle \quad (11.6)$$

$$\langle \psi_1 | \psi_2 \rangle = \langle U(S)\psi_1 | U(S)\psi_2 \rangle \quad (11.7)$$

– soit antiunitaire et antilinéaire

$$U(S)(\lambda|\psi_1\rangle + \lambda_2|\psi_2\rangle) = \lambda_1^* U(S)|\psi_1\rangle + \lambda_2^* U(S)|\psi_2\rangle \quad (11.8)$$

$$\langle \psi_1 | \psi_2 \rangle^* = \langle U(S)\psi_1 | U(S)\psi_2 \rangle \quad (11.9)$$

*Remarque :*

L'adjoint d'un opérateur linéaire  $A$  est défini par

$$\langle \psi_1 | A^+ \psi_2 \rangle = \langle A \psi_1 | \psi_2 \rangle.$$

Cette condition ne peut être vérifiée par un opérateur antilinéaire car le membre de droite serait linéaire en  $|\psi_1\rangle$  alors que le membre de gauche est antilinéaire en  $|\psi_1\rangle$ , d'où la définition  $\langle \psi_1 | A^+ \psi_2 \rangle = \langle A \psi_1 | \psi_2 \rangle^* = \langle \psi_2 | A \psi_1 \rangle$ .

Pour toute transformation de symétrie  $S$ , plusieurs remarques peuvent être faites sur l'opérateur  $U(S)$ , qu'il soit unitaire ou antiunitaire :

◇  $U(S)$  est défini à une phase globale près.

◇ Dans les deux cas,  $U(S)U^+(S) = U^+(S)U(S) = \mathbf{1}$ .

- ◇ La symétrie triviale  $|\underline{\psi}\rangle \rightarrow |\underline{\psi}\rangle$  est représentée par l'opérateur identité  $U = \mathbb{1}$ , qui est linéaire et unitaire. Par continuité toute symétrie qui peut être rendue triviale par changement continu de paramètres (ex : angle pour une rotation, distance pour une translation, vitesse pour une transformation de Lorentz pure) doit être représentée par un opérateur linéaire et unitaire.

Les symétries représentées par un opérateur antilinéaire et antiunitaire font intervenir le renversement du temps.

On trouvera une démonstration du théorème de Wigner dans l'appendice A du chap. 1 de la ref. [8].

## 11.5 Conséquence du théorème de Wigner

Considérons un système physique qui admet un groupe de symétrie  $G$ . Soit  $g_1, g_2, \dots$  les éléments de  $G$ . D'après le théorème de Wigner à tout élément  $g \in G$  correspond un opérateur  $U(g)$  agissant sur  $\mathcal{H}$  qui est soit unitaire soit antiunitaire. Nous disposons donc de deux représentants du rayon  $|^{g_1 g_2} \underline{\psi}\rangle$  qui sont  $U(g_1)U(g_2)|\psi\rangle$  et  $U(g_1 g_2)|\psi\rangle$ . Ces deux représentants ne peuvent différer que par un facteur de phase, donc  $U(g_1)U(g_2)|\psi\rangle = e^{i\varphi_\psi(g_1, g_2)}U(g_1, g_2)|\psi\rangle$ . En fait la phase  $\varphi_\psi(g_1, g_2)$  ne dépend pas (sauf exception due à une règle de supersélection) de l'état  $|\psi\rangle$ .

*Preuve :*

Soient  $|\psi_a\rangle, |\psi_b\rangle$  deux états qui ne sont pas proportionnels entre eux. Considérons l'état  $|\psi_{ab}\rangle = |\psi_a\rangle + |\psi_b\rangle$ . Par linéarité (ou antilinéarité) de  $U$ ,

$$e^{i\varphi_{ab}}U(g_1 g_2)(|\psi_a\rangle + |\psi_b\rangle) = U(g_1)U(g_2)(|\psi_a\rangle + |\psi_b\rangle) \quad (11.10)$$

$$\begin{aligned} &= U(g_1)U(g_2)|\psi_a\rangle + U(g_1)U(g_2)|\psi_b\rangle \\ &= e^{i\varphi_a}U(g_1 g_2)|\psi_a\rangle + e^{i\varphi_b}U(g_1 g_2)|\psi_b\rangle. \end{aligned} \quad (11.11)$$

Comme tout opérateur unitaire (resp. antiunitaire) possède un inverse (son adjoint), lui-même unitaire (resp. antiunitaire), en multipliant l'équation précédente par  $U^{-1}(g_1 g_2)$  on obtient

$$e^{\pm i\varphi_{ab}}(|\psi_a\rangle + |\psi_b\rangle) = e^{\pm i\varphi_a}|\psi_a\rangle + e^{\pm i\varphi_b}|\psi_b\rangle$$

où le signe + (resp. -) correspond au cas linéaire (resp. antilinéaire). Comme  $|\psi_a\rangle$  et  $|\psi_b\rangle$  sont linéairement indépendants, on en déduit que  $e^{i\varphi_{ab}} = e^{i\varphi_a} = e^{i\varphi_b}$ .

Ainsi on peut écrire

$$U(g_1)U(g_2) = e^{i\varphi(g_1, g_2)}U(g_1 g_2) \quad (11.12)$$

*Quelques remarques :*

- ◇ si  $\varphi = 0$ ,  $U(g)$  fournit par définition une représentation du groupe  $G$

- ◇ dans le cas général on parle de représentation à une phase près, ou représentation projective.

*Exemple :*

Une particule chargée plongée dans un champ magnétique bidimensionnel donne lieu à une représentation projective du groupe des translations (niveaux de Landau).

- ◇ la démonstration précédente est prise en défaut s'il n'est pas possible de préparer le système dans l'état  $|\psi_a\rangle + |\psi_b\rangle$ . Par exemple on ne peut préparer un système à partir de la somme de deux états de moment angulaire total entier et demi-entier respectivement.

Il existe alors une règle de supersélection qui divise les états en différentes sous-classes. Les phases  $\varphi(g_1, g_2)$  peuvent alors dépendre de la classe sur laquelle  $U(g_1)U(g_2)$  et  $U(g_1g_2)$  agissent.

En fait, tout groupe de symétrie qui possède intrinsèquement des représentations projectives peut toujours être étendu (groupe de recouvrement) de façon à ce que ses représentations puissent être définies comme non projectives, comme on le verra au chap. 16.

*Exemple :* voir plus loin le cas de  $SO(3)$  et  $SU(2)$ .



Quatrième partie  
Symétries continues



# Chapitre 12

## Groupes et algèbres de Lie

Nous avons vu dans le chapitre précédent le rôle particulier joué par les symétries en mécanique quantique. Les groupes discrets jouent un rôle important dans de nombreux domaines de la physique (cristallographie par exemple), y compris en physique des hautes énergies (groupe de permutation et principe de Pauli). Les groupes continus ont cependant joué un rôle particulier dans l'histoire de la physique subatomique, et nous allons essentiellement nous limiter à leur étude dans ce chapitre. L'exposé sera loin d'être systématique, tant sur le plan de l'exhaustivité que de la rigueur mathématique. On pourra se référer aux ouvrages cités en référence pour plus de détail, en particulier à l'ouvrage de Barut et Raczka [9] qui est très exhaustif et rigoureux tout en étant accessible à un physicien.

### 12.1 Généralités

#### 12.1.1 Groupe continu, groupe de Lie, algèbre de Lie

**Définition 12.1 :** *Groupe continu*

$G$  est un groupe continu (ou groupe topologique) de dimension  $n$  si et seulement si

- ◇ il existe une correspondance bi-univoque entre  $G$  et un sous-ensemble de  $\mathbb{R}^n$  :  $g \in G \leftrightarrow \vec{\alpha} \in \mathbb{R}^n$ . *ex.* : rotation de  $\mathbb{R}^3$  caractérisée par un angle et un vecteur unitaire
- ◇  $\forall \alpha, \beta \in \mathbb{R}^n$ , si  $g(\vec{\alpha}), g(\vec{\beta}) \in G$ ,  $g(\vec{\alpha})g(\vec{\beta}) \in G$ . Il existe donc  $\vec{\gamma} \in \mathbb{R}^n$  t.q.  $g(\vec{\alpha})g(\vec{\beta}) = g(\vec{\gamma})$   $\vec{\gamma} = \phi(\vec{\alpha}, \vec{\beta})$ . Par hypothèse,  $\phi$  est une fonction *continue* de  $\vec{\alpha}, \vec{\beta}$
- ◇  $g^{-1}(\vec{\alpha})$  peut s'écrire  $g(\vec{\alpha}')$  avec  $\vec{\alpha}' = f(\vec{\alpha})$  où par hypothèse  $f$  est fonction *continue* de  $\vec{\alpha}$

**Définition 12.2 :** *Groupe de Lie*

$G$  est un groupe de Lie si et seulement si  $G$  est un groupe continu pour lequel  $\phi$  et  $f$  sont analytiques. Pour simplifier les notations, on notera  $\alpha$  le vecteur  $\vec{\alpha}$ .

Notons que les définitions qui viennent d'être données ne sont pas les plus générales. On s'est en fait limité ici aux groupes linéaires. Pour un mathématicien, un groupe se définit sur

une *variété*. Pour donner une image géométrique simple de cette notion, imaginons que le groupe considéré soit de dimension 2, et que sa variété soit la sphère  $\mathbb{S}^2$ . Cela signifie qu'un élément donné du groupe et son voisinage peuvent être repérés de façon continue et bijective (on dira qu'il y a homéomorphisme) par une sous partie de  $\mathbb{R}^2$ . Ce voisinage et la bijection correspondante jouent le rôle d'une *carte* permettant de se repérer au voisinage d'un point de la sphère. Bien entendu, au voisinage d'un autre point de la sphère, il faudra utiliser une autre carte. L'ensemble des cartes constitue un atlas (c'est le même vocabulaire qu'en géographie!). Pour un groupe de Lie, les changements de cartes associées à la variété constituée par le groupe sont supposés analytiques. Dans de nombreux domaines de la physique, une telle définition est trop générale, et se limiter au cas où la variété est  $\mathbb{R}^n$  lui-même suffit (l'atlas est alors constitué d'une seule carte sur  $\mathbb{R}^n$  en entier, et on parle alors de groupe linéaire). Le lecteur attentif pourra cependant constater qu'une grande partie des résultats énoncés dans ce chapitre sont des résultats basés sur des propriétés *locales* des groupes de Lie, pour lesquelles le passage des groupes de Lie linéaires au cas général est uniquement un jeu d'écriture.

*Remarque :*

un choix usuel pour l'application de  $\mathbb{R}^n$  sur  $G$  est de choisir que  $g(0) = e$  où  $e$  est l'élément neutre de  $G$ , que l'on notera également 1 en utilisant une notation multiplicative pour la loi de groupe.

### 12.1.2 Générateurs

D'après la définition donnée ci-dessus d'un groupe de Lie, ayant fixé un élément de  $\mathbb{R}^n$  et l'élément correspondant du groupe, on peut effectuer un développement en série localement au voisinage de ce point de  $\mathbb{R}^n$  et donc de l'élément correspondant du groupe. A cause de la structure de groupe, il suffira de mettre en pratique cette idée au voisinage de l'identité. On pose donc

$$g(\alpha) = 1 + i\alpha^a T_a + O(\alpha^2). \quad (12.1)$$

Les  $T_a$  (au nombre de  $n$  puisque l'on peut différentier dans  $n$  directions possibles sur  $\mathbb{R}^n$ ) sont appelés *générateurs* du groupe. Avec cette définition incluant un facteur  $i$ , ils sont hermitiens si le groupe est unitaire, ce qui sera souvent le cas en physique.

### 12.1.3 Application exponentielle

Pour un groupe de matrices  $G$ , l'application exponentielle est aisée à définir.

L'application exponentielle de l'algèbre de Lie  $M(n, \mathbb{R})$  du groupe de Lie  $GL(n, \mathbb{R})$  vers son groupe de Lie  $GL(n, \mathbb{R})$  peut se définir par le développement en série

$$\begin{aligned} \exp : M(n, \mathbb{R}) &\rightarrow GL(n, \mathbb{R}) \\ A &\mapsto \exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \mathbb{1} + A + \frac{1}{2}A^2 + \dots \end{aligned} \quad (12.2)$$

Plus généralement, si l'on considère un sous-groupe  $G$  de  $GL(n, \mathbb{R})$ , alors  $\exp$  est une application de l'algèbre de Lie de  $G$  vers  $G$ .

La question est maintenant d'étendre cette définition au cas où le groupe de Lie  $G$  n'est pas un groupe de matrices. D'autre part, une définition intrinsèque permet de ne pas faire dépendre

cette application de la représentation du groupe de Lie par un groupe de matrice. Une telle construction fait appel au concept de variété, que nous ne discuterons pas ici. Voir l'appendice C pour une introduction au sujet.

### 12.1.4 Algèbre de Lie

Nous allons nous limiter ici aux groupes de matrices, ce qui permet d'utiliser la notion d'exponentielle de matrice, définie de façon naturelle par le développement en série comme discuté précédemment.

Notons  $\vec{e}^a = (0, \dots, 0, 1, 0 \dots 0)$  les vecteurs de base de  $\mathbb{R}^n$ , où le coefficient non nul est en  $a$ -ième position. On peut alors considérer l'exponentielle d'un générateur le long d'une direction (en gardant un paramètre libre de façon à pouvoir différentier par rapport à ce paramètre), et donc définir, pour tout vecteur  $\alpha \in \mathbb{R} \setminus 0$  (de composantes  $\alpha^a$ ),

$$g(\sqrt{t} \vec{\alpha}) = \exp(i\sqrt{t} \alpha^a T_a). \quad (12.3)$$

et donc en particulier

$$g(\sqrt{t} \vec{e}^a) = \exp(i\sqrt{t} T_a). \quad (12.4)$$

Dans la limite où  $|t| \ll 1$ , considérons le produit

$$\begin{aligned} & g(\sqrt{t} \vec{e}^b) g(\sqrt{t} \vec{e}^a) g^{-1}(\sqrt{t} \vec{e}^b) g^{-1}(\sqrt{t} \vec{e}^a) = \\ & \left[ 1 + i\sqrt{t} T_b - \frac{t}{2} T_b^2 + o(t) \right] \left[ 1 + i\sqrt{t} T_a - \frac{t}{2} T_a^2 + o(t) \right] \left[ 1 - i\sqrt{t} T_b - \frac{t}{2} T_b^2 + o(t) \right] \\ & \times \left[ 1 - i\sqrt{t} T_a - \frac{t}{2} T_a^2 + o(t) \right]. \end{aligned} \quad (12.5)$$

A l'ordre  $t$ , un calcul immédiat montre que le produit (12.5) est égal à  $1 + t(T_a T_b - T_b T_a)$ . Comme ce produit appartient lui-même au groupe  $G$ , il existe  $\vec{\gamma} \in \mathbb{R}^n$  t.q.

$$g(\sqrt{t} \vec{e}^b) g(\sqrt{t} \vec{e}^a) g^{-1}(\sqrt{t} \vec{e}^b) g^{-1}(\sqrt{t} \vec{e}^a) = g(t\vec{\gamma}). \quad (12.6)$$

En composant membre à membre, on obtient donc

$$1 + t(T_a T_b - T_b T_a) + o(t) = 1 + it \gamma^c T_c + o(t)$$

i.e.  $T_a T_b - T_b T_a = \gamma^c T_c$ . Le coefficient  $\gamma^c$  dépend de  $a$  et  $b$  : on le notera donc  $f_{ab}^c$ , d'où

$$\boxed{[T_a, T_b] = i f_{ab}^c T_c} \quad (12.7)$$

Les coefficients  $f_{ab}^c$  sont appelés *constantes de structure du groupe*.

**Définition 12.3 :** *Algèbre de Lie*

L'ensemble des générateurs  $T_a$ , muni du commutateur  $[\ , \ ]$  qui joue le rôle d'un produit, forme une  $\mathbb{R}$ -algèbre, appelée algèbre de Lie du groupe, et notée  $\mathfrak{g}$ .

### Identité de Jacobi

On déduit immédiatement de la définition (12.7), en utilisant les propriétés d'antisymétrie du commutateur, que

$$[T_a, [T_b, T_c]] + [T_b, [T_c, T_a]] + [T_c, [T_a, T_b]] = 0 \quad (12.8)$$

i.e., en terme des constantes de structure du groupe :

$$f_{bc}^e f_{ae}^d + f_{ca}^e f_{be}^d + f_{ab}^e f_{ce}^d = 0. \quad (12.9)$$

*Remarques :*

- ◇ D'après la définition (12.7),  $f_{ab}^c = -f_{ba}^c$ .
- ◇ Contrairement à ce que le nom « constante de structure » pourrait sous-entendre, les  $f_{ab}^c$  ne sont pas des constantes : elles dépendent de la base  $T_1, \dots, T_n$  choisie pour l'algèbre  $g$ . En effet, soient  $X, Y \in g$  avec  $X = \alpha^a T_a$  et  $Y = \beta^b T_b$ . Les  $\alpha^a$  (resp.  $\beta^b$ ) doivent être compris comme les composantes de  $X$  dans la base  $T_a$  (resp.  $T_b$ ).

Posons  $Z = [X, Y]$ . Alors  $Z = [\alpha^a T_a, \beta^b T_b] = \alpha^a \beta^b i f_{ab}^c T_c = \gamma^c T_c$ , d'où  $\gamma^c = i f_{ab}^c \alpha^a \beta^b$ .

$\alpha^a : \text{coordonnée} = \text{vecteur contravariant}$ $T_a : \text{vecteur de base} = \text{vecteur covariant}$	}	se transforment par changement de base (voir plus loin l'exemple du groupe de Lorentz pour la notion de calcul covariant)
--	---	--

### 12.1.5 Quelques définitions utiles

Nous présentons ici quelques structures qui seront utiles dans la suite, en particulier pour comprendre l'importance de la compacité discutée plus loin.

**Définition 12.4 :** *Algèbre de Lie réelle*

Une algèbre de Lie réelle  $\mathfrak{g}$  est un  $\mathbb{R}$ -espace vectoriel de dimension  $n$  ( $n \geq 1$ ) muni d'un produit bilinéaire  $[\cdot, \cdot]$  antisymétrique et vérifiant l'identité de Jacobi.  $[\cdot, \cdot]$  doit donc vérifier :

1.  $[a, b] \in \mathfrak{g} \quad \forall a, b \in \mathfrak{g}$
2.  $\forall (a, b, c) \in \mathfrak{g}^3, \quad \forall (\alpha, \beta) \in \mathbb{R}^2, \quad [\alpha a + \beta b, c] = \alpha [a, c] + \beta [b, c]$
3.  $[a, b] = -[b, a] \quad \forall (a, b) \in \mathfrak{g}^2$
4. identité de Jacobi :  $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$

On parle alors d'algèbre de Lie abstraite, puisqu'il n'est pas nécessaire de la construire à partir d'un groupe de Lie.

**Théorème 12.5 :** *Théorème d'Ado*

Toute algèbre de Lie abstraite sur un corps commutatif  $\mathbb{K}$  de caractéristique nulle est isomorphe à une algèbre de Lie de matrices carrée dont le crochet de Lie est défini par le crochet de Lie usuel  $[a, b] = ab - ba$ .

La théorie des algèbres de Lie abstraite ne fait donc pas apparaître de nouvelles structures.

**Définition 12.6 :** *Algèbre de Lie complexe*

Une algèbre de Lie complexe  $\mathfrak{g}$  est un  $\mathbb{C}$ -espace vectoriel de dimension  $n$  ( $n \geq 1$ ) muni d'un produit bilinéaire  $[\ , \ ]$  antisymétrique et vérifiant l'identité de Jacobi.

Cela signifie que dans le point 2. de la définition d'une algèbre de Lie réelle, il faut maintenant considérer  $(\alpha, \beta) \in \mathbb{C}^2$ .

*Conséquence :* les constantes de structure peuvent maintenant être complexes.

Exemple de  $\mathbb{R}$ -algèbre de Lie :  $\text{SO}(3)$  et  $\text{SU}(2)$  ont pour algèbre de Lie  $\mathfrak{su}(2)$  :

$$\left[ \frac{\sigma_i}{2}, \frac{\sigma_j}{2} \right] = i \varepsilon_{ijk} \frac{\sigma_k}{2}$$

*Remarques :*

- ◇ autres notations, utilisées en particulier en mathématiques :

$$X_a = \left. \frac{\partial}{\partial t_a} X(0, \dots, t_a, \dots, 0) \right|_{t_a=0} = iT_a.$$

D'une part

$$[T_a, T_b] = i f_{ab}^c T_c$$

et d'autre part

$$[X_a, X_b] = [iT_a, iT_b] = C_{ab}^c X_c = C_{ab}^c iT_c \quad \text{donc } f_{ab}^c = -C_{ab}^c$$

donc

$$[X_a, X_b] = C_{ab}^c X_c \xleftrightarrow{X_a = iT_a} [T_a, T_b] = i f_{ab}^c T_c, \quad \text{avec } f_{ab}^c = -C_{ab}^c.$$

Les  $T_a$  sont des générateurs hermitiens pour un groupe unitaire.

De manière équivalente, on peut définir une algèbre de Lie comme l'espace vectoriel (sur  $\mathbb{R}$  ou  $\mathbb{C}$ ) engendré par les  $X_a$ .

- ◇ lexicque : lorsque l'on ne précise pas le corps de l'algèbre de Lie considérée, il est sous-entendu que l'on considère la  $\mathbb{R}$ -algèbre de Lie.

**Définition 12.7 :** *Extension complexe  $\mathfrak{g}^c$  d'une algèbre de Lie  $\mathfrak{g}$*

Elle est définie de la façon suivante :

- ◇  $\mathfrak{g}^c$  est l'extension complexe de l'espace vectoriel  $\mathfrak{g}$ , i.e. formée des éléments  $z = x + iy$  pour  $x, y \in \mathfrak{g}$
- ◇ Le crochet de lie dans  $\mathfrak{g}^c$  est

$$\begin{aligned} Z &= [Z_1, Z_2] = [X_1 + iY_1, X_2 + iY_2] \\ &= [X_1, X_2] - [Y_1, Y_2] + i[X_1, Y_2] + i[Y_1, X_2] \\ &\equiv X + iY. \end{aligned}$$

L'algèbre de Lie *complexe*  $\mathfrak{g}$  de dimension  $n$ , de base  $\{X_1, \dots, X_n\}$ , peut être considérée comme une algèbre de Lie *réelle* de dimension  $2n$ , de base  $\{X_1, iX_1, \dots, X_n, iX_n\}$ . On note cette algèbre de Lie  $\mathfrak{g}^{\mathbb{R}}$ .

**Définition 12.8 :** *Forme réelle  $\mathfrak{g}^r$  d'une algèbre de Lie  $\mathfrak{g}^c$*

C'est une algèbre de lie dont l'extension complexe est  $\mathfrak{g}^c$  ( $\mathfrak{g}^r$  et  $\mathfrak{g}^c$  ont même dimension, l'une par rapport à  $\mathbb{R}$ , l'autre par rapport à  $\mathbb{C}$ ). Une algèbre de lie sur  $\mathbb{C}$  possède de nombreuses formes réelles !

On se servira un peu plus loin de cette liberté, en particulier pour la notion de compacité.

*Remarque :*

Dans le cas où l'on considère l'algèbre de Lie sur  $\mathbb{R}$ , le changement  $X_a \leftrightarrow iT_a$  consiste simplement à prendre deux formes réelles de la même algèbre de Lie sur  $\mathbb{C}$ . Ceci n'a aucun effet sur la structure (en particulier cela ne change pas la compacité : voir plus loin cette notion).

## 12.2 Quelques résultats sur les groupes et algèbres de Lie

### 12.2.1 Lien entre groupe de Lie et algèbre de Lie correspondante

Nous avons construit la notion d'algèbre de Lie comme un espace tangent au groupe. réciproquement, on peut se demander si l'on peut reconstruire tout ou partie du groupe connaissant son algèbre de Lie. Nous allons successivement examiner cette question localement puis globalement.

#### Sous-groupe à un paramètre d'un groupe de Lie

**Définition 12.9 :** *Sous-groupe à un paramètre d'un groupe de Lie*

C'est un sous-groupe de Lie du groupe de Lie  $G$  constitué des éléments  $T(t)$  avec  $t \in \mathbb{R}$  variant entre  $-\infty$  et  $+\infty$ , t.q.  $\forall s, t, T(s)T(t) = T(s+t)$ .

**Proposition 12.10 :**

Tout sous-groupe à un paramètre d'un groupe de Lie est abélien.

*Preuve :*

Ceci découle immédiatement de la relation  $T(t)T(s) = T(s+t) = T(s)T(t)$ .

*Exemple :*  $SO(3)$

Considérons les matrices

$$A(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix}.$$

Posons  $a_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ . Alors  $a_1^{2k+1} = (-1)^k a_1$  et  $a_1^{2k} = (-1)^k \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  donc  $A(t) = \exp t a_1$ .



Ce résultat est général, comme le montre le théorème suivant.

**Théorème 12.11 :**

Tout sous-groupe à un paramètre d'un groupe de Lie  $G$  formé de matrices  $n \times n$  s'obtient par exponentiation de matrices  $n \times n$  :  $A(t) = \exp t\dot{A}(0)$ .

*Preuve :*

Soit  $B(t) = A(t) \exp[-t\dot{A}(0)]$

Alors  $\dot{B}(t) = \left( \dot{A}(t) - A(t)\dot{A}(0) \right) \exp[-t\dot{A}(0)]$ . Or

$$\dot{A}(t) = \lim_{s \rightarrow 0} \frac{A(t+s) - A(t)}{s} = \lim_{s \rightarrow 0} A(t) \frac{A(s) - A(0)}{s} = A(t)\dot{A}(0)$$

Ainsi  $\forall t, \dot{B}(t) = 0$ , et donc  $B(t) = B(0) = 1$ , d'où  $A(t) = \exp t\dot{A}(0)$ .

**Relation entre l'algèbre de Lie réelle  $\mathfrak{g}$  d'un groupe de Lie  $G$  et ses sous-groupes à un paramètre**

**Théorème 12.12 :**

$\forall a \in \mathfrak{g}, A(t) = \exp t a$  pour  $t \in ]-\infty, +\infty[$  forme un sous-groupe à un paramètre de  $G$ .

**Théorème 12.13 :**

Tout élément d'un groupe de Lie  $G$  dans un petit voisinage de l'identité appartient à un sous-groupe à un paramètre de  $G$ . Cela signifie que tout élément de  $G$  dans un voisinage de l'identité peut s'écrire comme l'exponentielle d'un élément de son algèbre de Lie. Cette correspondance est bi-univoque.

Les théorèmes suivants permettent d'étendre le résultat à l'ensemble du sous-groupe connexe de  $G$  (problème *global*).

**Composante connexe d'un groupe continu**

**Définition 12.14 :** *Composante connexe d'un groupe continu  $G$*

C'est l'ensemble des éléments du groupe qui peuvent être obtenus l'un de l'autre par variation continue de ses paramètres (i.e. par homéomorphisme).

**Théorème 12.15 :**

La composante connexe  $C$  de l'identité  $e$  d'un groupe  $G$  est un sous-groupe de  $G$ .

*Preuve :*

Il suffit de considérer, pour un  $c$  quelconque dans  $C$ , l'homéomorphisme  $x \mapsto c^{-1}x$ . L'image de  $C$  par cet homéomorphisme est  $C$  puisqu'elle contient  $e$  et  $C$ . On a donc  $C^{-1}C = C$ , *q.e.d.*

**Théorème 12.16 :**

La composante connexe  $C$  de l'identité  $e$  de  $G$  est un sous-groupe invariant de  $G$ , i.e.  $\forall A \in C, \forall X \in G, XAX^{-1} \in C$ . On l'appelle souvent le sous-groupe connexe de  $G$ .

*Preuve :*

Pour un  $x$  quelconque dans  $G$ ,  $x C^{-1} x$  est l'image par un homéomorphisme de  $C$  et contient  $e$ , donc est inclus dans  $C$ .

**Problème global****Théorème 12.17 :**

Si  $G$  est un groupe de Lie compact, tout élément  $X$  de son sous-groupe connexe s'écrit sous la forme  $X = \exp a$  où  $a \in \mathfrak{g}$ . En particulier si  $G$  est compact et connexe, tout élément de  $G$  est l'exponentielle d'un élément de  $\mathfrak{g}$ .

*Remarque :*

La correspondance n'est pas forcément bi-univoque : on peut avoir  $e^a = e^b$  pour  $a \neq b$ .

*Exemple :*

Pour  $\text{SO}(3)$ ,  $\exp(\theta a_1) = \exp[(\theta + 2\pi n)a_1]$  pour  $n = \pm 1, \pm 2, \dots$

**Théorème 12.18 :**

Tout élément du sous-groupe connexe d'un groupe de Lie  $G$  peut s'écrire comme le produit fini d'exponentielles de son algèbre de Lie réelle  $\mathfrak{g}$ .

*Exemple :*

Groupe de Lorentz restreint  $L_+^\uparrow$ .

## 12.2.2 Structure des algèbres de Lie et groupes de Lie ; caractérisation de Cartan

Commençons par donner quelques définitions et propriétés concernant les structures des groupes et algèbres de Lie.

**Définition 12.19 :** *Algèbre de Lie abélienne*

Une algèbre de Lie  $\mathfrak{g}$  est abélienne si et seulement si  $\forall a, b \in \mathfrak{g}, [a, b] = 0$ .

On en déduit facilement la proposition suivante :

**Proposition 12.20 :**

L'algèbre de Lie  $\mathfrak{g}$  d'un groupe abélien est abélienne.

**Définition 12.21 :** *Sous-algèbre de Lie*

Une sous-algèbre de Lie  $\mathfrak{g}'$  d'une algèbre de Lie  $\mathfrak{g}$  est un sous-ensemble de  $\mathfrak{g}$ , de base  $e_1, \dots, e_k$ , muni du même crochet de Lie et formant lui-même une algèbre :  $C_{ij}^s = 0$ ,  $i, j \leq k$ ,  $s > k$ . La sous-algèbre  $\mathfrak{g}'$  est dite propre si au moins un élément de  $\mathfrak{g} \notin \mathfrak{g}'$ , i.e.  $\dim \mathfrak{g}' < \dim \mathfrak{g}$ .

Ceci implique alors facilement la proposition suivante :

**Proposition 12.22 :**

Si  $G$  et  $G'$  sont deux groupes de Lie et  $\mathfrak{g}$  et  $\mathfrak{g}'$  leurs algèbres de Lie correspondantes, et si  $G'$  est un sous-groupe de  $G$ , alors  $\mathfrak{g}'$  est une sous-algèbre de  $\mathfrak{g}$ .

**Définition 12.23 :** *Sous-algèbre invariante d'une algèbre de Lie (ou idéal)*

Une sous-algèbre invariante  $\mathfrak{g}'$  d'une algèbre de Lie  $\mathfrak{g}$  est une sous-algèbre tel que  $\forall a \in \mathfrak{g}'$  et  $\forall b \in \mathfrak{g}$ ,  $[a, b] \in \mathfrak{g}'$ , soit encore  $C_{ij}^s = 0$ ,  $i \leq k$ ,  $s > k$ ,  $j$  quelconque.

**Théorème 12.24 :**

Si  $G'$  est un sous-groupe invariant de  $G$ , alors  $\mathfrak{g}'$  est une sous-algèbre invariante de  $\mathfrak{g}$ .

*Preuve :*

Soient

- ◇  $b$  un vecteur tangent à la courbe analytique  $B(t)$  de  $G$
- ◇  $a'$  un vecteur tangent à la courbe analytique  $A(t)$  de  $G'$ .

Alors  $[a', b]$  est tangent à la courbe  $C(t) = A(\sqrt{t})B(\sqrt{t})A(\sqrt{t})^{-1}B(\sqrt{t})^{-1}$ . Comme  $G'$  est un sous-groupe invariant,  $B(\sqrt{t})A(\sqrt{t})^{-1}B(\sqrt{t})^{-1} \in G'$ , et donc  $C(t)$  est une courbe analytique de  $G'$ . Ceci prouve que  $[a', b]$  est tangent à une courbe analytique de  $G'$ , et donc que  $[a', b] \in G'$ .

**Théorème 12.25 :**

Soit  $\mathfrak{g}$  l'algèbre de Lie réelle correspondant au groupe de Lie linéaire  $G$ . Alors chaque sous-algèbre de  $\mathfrak{g}$  est l'algèbre de Lie d'exactly un sous-groupe de Lie connexe de  $G$ .

**Définition 12.26 :** *Algèbre de Lie simple*

Une algèbre de Lie  $\mathfrak{g}$  est dite simple si elle n'est pas abélienne et ne possède pas de sous-algèbre de Lie propre invariante.

**Définition 12.27 :** *Algèbre de Lie semi-simple*

Une algèbre de Lie  $\mathfrak{g}$  est dite semi-simple si elle ne possède pas de sous-algèbre de Lie abélienne invariante.

**Définition 12.28 :** *Sous-groupe discret d'un groupe de Lie*

Un sous-groupe  $H$  d'un groupe de Lie  $G$  est dit discret si

a)  $H$  est un groupe fini

ou

b)  $H$  possède un nombre infini dénombrable d'éléments, tel qu'il existe un voisinage de l'identité de  $G$  qui ne contienne aucun élément de  $H$  (excepté  $e$  elle-même). De façon équivalente,  $H$  possède un nombre infini dénombrable d'éléments entourés chacun d'un voisinage sans éléments dans  $H$ .

**Définition 12.29 :** *Groupe de Lie simple*

Un groupe de Lie  $G$  est simple s'il ne possède pas de sous-groupe connexe de Lie propre invariant. Le groupe  $G$  peut donc posséder des sous-groupes discrets invariants.

*Exemple :*

Le groupe  $SU(N)$  est simple pour  $N > 1$ .

L'élément  $g = e \exp\left(\frac{2\pi i}{N}\right)$  appartient à  $SU(N)$ . Cet élément  $g$  engendre un sous-groupe discret cyclique invariant d'ordre  $N$ ,  $Z_N$ .

**Définition 12.30 :** *Groupe de Lie semi-simple*

Un groupe de Lie  $G$  est semi-simple s'il ne possède pas de sous-groupe de Lie *abélien* propre connexe et invariant.

Bien entendu, un groupe simple est semi-simple !

*Exemples :*

- ◇ Le groupe de Poincaré n'est pas semi-simple, puisque le sous-groupe des translations est invariant

$$\begin{aligned} [P_\mu, P_\rho] &= 0 \\ [P_\rho, J_{\mu\nu}] &= i(g_{\mu\rho} P_\nu - g_{\nu\rho} P_\mu) \end{aligned}$$

- ◇  $SO(3)$  est simple

- ◇  $SO(4) = SO(3) \times SO(3)$  est semi-simple.

Nous allons maintenant présenter une caractérisation, due à Cartan, des algèbres de Lie semi-simples.

### Tenseur de Cartan-Killing

**Définition 12.31 :** *Tenseur de Cartan-Killing*

Le tenseur de Cartan-Killing  $g$  peut être défini de deux façons équivalentes :

◇ Par ses éléments de matrice, une base de l'algèbre de Lie étant fixée, suivant

$$g_{\rho\sigma} = C_{\rho\alpha}^{\beta} C_{\sigma\beta}^{\alpha} (= f_{\rho\alpha}^{\beta} f_{\sigma\beta}^{\alpha}). \quad (12.10)$$

◇ En introduisant l'opérateur *adjoint* (noté  $\text{ad}$ ), défini par<sup>1</sup>

$$\left| \begin{array}{l} X \quad \mapsto \quad \text{ad}X \\ \text{ad}X(Y) \quad = \quad [X, Y] \end{array} \right., \quad (12.11)$$

ce tenseur de Cartan-Killing peut être défini par

$$g_{\rho\sigma} = \text{Tr}(\text{ad}X_{\rho} \text{ad}X_{\sigma}). \quad (12.12)$$

*Preuve :*

L'élément de matrice  $(\alpha, \beta)$  de  $\text{ad}X_{\rho}$  est donné par  $[X_{\rho}, X_{\beta}]_{\alpha}$  (c'est le coefficient  $\alpha$  de l'image du vecteur  $X_{\beta}$  par l'opérateur  $\text{ad}X_{\rho}$ ), qui vaut  $C_{\rho\beta}^{\alpha}$  et donc la trace de  $\text{ad}X_{\rho} \text{ad}X_{\sigma}$  est simplement

$$C_{\rho\beta}^{\alpha} C_{\sigma\beta'}^{\alpha'} \delta_{\alpha}^{\beta'} \delta_{\alpha'}^{\beta} = C_{\rho\beta}^{\alpha} C_{\sigma\alpha}^{\beta} \quad (12.13)$$

qui est identique à  $C_{\rho\alpha}^{\beta} C_{\sigma\beta}^{\alpha}$ .

On vérifie immédiatement d'après les définitions équivalentes (12.10) et (12.12) que  $g$  est symétrique.

**Théorème 12.32 :** (*Cartan*)

Une algèbre de Lie est semi-simple si et seulement si  $\det g \neq 0$ .

*Preuve :*

On trouvera une preuve de ce théorème p. 15-16 de la ref. [9]. Nous allons seulement montrer que  $\det g \neq 0 \Rightarrow$  algèbre de Lie semi-simple. Montrons donc que si  $\mathfrak{H}$  est une sous-algèbre abélienne invariante de  $\mathfrak{g}$  alors  $\det g = 0$ . On indexe par un indice latin  $i, j, \dots \leq p$  les éléments de  $\mathfrak{H}$ . Considérons les éléments  $g_{\rho i}$ ,  $\forall i \leq p$ . Comme  $\mathfrak{H}$  est une sous-algèbre invariante,  $C_{i\beta}^{\alpha} = 0$  si  $\alpha > p$ , et donc

$$g_{\rho i} = C_{\rho\alpha}^{\beta} C_{i\beta}^{\alpha} = C_{\rho j}^{\beta} C_{i\beta}^j, \quad (12.14)$$

ce qui conduit encore, en utilisant le même argument pour  $C_{\rho j}^{\beta}$ , à

$$g_{\rho i} = C_{\rho j}^k C_{ik}^j. \quad (12.15)$$

Comme  $\mathfrak{H}$  est abélienne,  $C_{ik}^j = 0$ , ce qui montre que  $\forall \alpha, \forall i \leq p, g_{\alpha i} = 0$ . D'où  $\det g = 0$ .

---

1. On notera également  $\text{ad}X Y = \text{ad}X(Y)$ .

Le tenseur  $g$  permet de définir un calcul covariant. Ainsi  $C_{\rho\sigma\tau} = g_{\alpha\tau} C_{\rho\sigma}{}^{\alpha}$ .

### Exercice 12.1

On considère les transformations  $E_2$  du plan euclidien dans lui-même

$$\begin{cases} x' = x \cos \theta - y \sin \theta + a \\ y' = x \sin \theta + y \cos \theta + b . \end{cases}$$

Introduisant le vecteur  $X = (x, y, 1)$ , on a  $X' = MX$  où  $M = \begin{pmatrix} \cos \theta & -\sin \theta & a \\ \sin \theta & \cos \theta & b \\ 0 & 0 & 1 \end{pmatrix}$ . Montrer que les générateurs s'écrivent

$$X_{\theta} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad X_a = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad X_b = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

et qu'il vérifient

$$\begin{cases} [X_{\theta}, X_a] = X_b \\ [X_{\theta}, X_b] = -X_a \\ [X_a, X_b] = 0 \end{cases}$$

Montrer que  $g_{\mu\nu} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  et que  $E_2$  n'est pas semi-simple.

### 12.2.3 Algèbre de Lie compacte

Afin de définir la notion d'algèbre de Lie compacte, quelques résultats préliminaires vont nous être utiles.

#### Proposition 12.33 :

L'opérateur  $\text{ad}$  vérifie, pour tout  $X$  et  $Y$  de  $\mathfrak{g}$ ,

$$[\text{ad}X, \text{ad}Y] = \text{ad}[X, Y]. \quad (12.16)$$

*Preuve :*

Il suffit pour cela de montrer que  $\forall Z \in \mathfrak{g}$ ,

$$[\text{ad}X, \text{ad}Y]Z = \text{ad}[X, Y]Z. \quad (12.17)$$

D'une part,

$$\text{ad}[X, Y]Z = [[X, Y], Z], \quad (12.18)$$

et d'autre part

$$[\text{ad } X, \text{ad } Y] Z = \text{ad } X \text{ad } Y Z - \text{ad } Y \text{ad } X Z = [X, [Y, Z]] - [Y, [X, Z]]. \quad (12.19)$$

En utilisant l'identité de Jacobi

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \quad (12.20)$$

on a donc

$$[\text{ad } X, \text{ad } Y] Z = [[X, Y], Z], \quad (12.21)$$

*q. e. d.*

Le tenseur de Cartan-Killing (12.10,12.12) permet de définir une forme bilinéaire symétrique, dite forme de Killing.

**Définition 12.34 :** *Forme de Killing*

La forme de Killing est définie par son action sur les vecteurs de l'algèbre de Lie  $\mathfrak{g}$  par

$$(X, Y) = \text{Tr}(\text{ad } X \text{ad } Y). \quad (12.22)$$

En décomposant  $X = x^\rho X_\rho$  et  $Y = y^\sigma X_\sigma$  dans une base  $X_\alpha$ , on peut bien sûr écrire

$$(X, Y) = x^\rho y^\sigma g_{\rho\sigma} = x^\rho y^\sigma C_{\rho\alpha}{}^\beta C_{\sigma\beta}{}^\alpha. \quad (12.23)$$

**Proposition 12.35 :** *Invariance de la forme de Killing*

La forme de Killing (12.22) est invariante sous l'action de  $\text{ad } Z$ , *i.e.*

$$\forall X, Y, Z \in \mathfrak{g}, (\text{ad } Z X, Y) + (X, \text{ad } Z Y) = 0, \quad (12.24)$$

qui s'écrit encore

$$\forall X, Y, Z \in \mathfrak{g}, ([Z, X], Y) + (X, [Z, Y]) = 0. \quad (12.25)$$

*Preuve :*

On applique la relation  $[\text{ad } X, \text{ad } Y] Z = \text{ad } [X, Y] Z$  à chacun des deux termes de l'éq. (12.25) : le premier s'écrit

$$\begin{aligned} ([Z, X], Y) &= \text{Tr}(\text{ad } [Z, X] \text{ad } Y) = \text{Tr}([\text{ad } Z, \text{ad } X] \text{ad } Y) \\ &= \text{Tr}(\text{ad } Z \text{ad } X \text{ad } Y) - \text{Tr}(\text{ad } X \text{ad } Z \text{ad } Y), \end{aligned}$$

et le second conduit à

$$(X, [Z, Y]) = \text{Tr}(\text{ad } X [\text{ad } Z, \text{ad } Y]) = \text{Tr}(\text{ad } X \text{ad } Z \text{ad } Y) - \text{Tr}(\text{ad } X \text{ad } Y \text{ad } Z).$$

En combinant ces deux résultats on obtient

$$(\operatorname{ad} Z X, Y) + (X, \operatorname{ad} Z Y) = \operatorname{Tr}(\operatorname{ad} Z \operatorname{ad} X \operatorname{ad} Y) - \operatorname{Tr}(\operatorname{ad} X \operatorname{ad} Y \operatorname{ad} Z) = 0 \quad (12.26)$$

par cyclicité de la trace, *q.e.d.*

---

On notera que cette invariance est somme toute naturelle, puisqu'elle correspond à l'invariance du produit scalaire défini à l'aide de la forme de Killing sous l'action de  $\operatorname{ad} Z$  qui est un générateur infinitésimal de la représentation de l'algèbre de Lie sur elle-même (voir plus loin la notion de représentation adjointe).

**Définition 12.36 :** *Algèbre de Lie compacte*

Une algèbre de Lie  $\mathfrak{g}$  est dite compacte s'il existe sur  $\mathfrak{g}$  une forme quadratique  $(\cdot, \cdot)$  définie négative qui satisfait l'invariance (12.25).

La forme de Cartan-Killing satisfait la condition (12.25), comme nous venons de le montrer. On démontre que dans une algèbre de Lie simple, les formes bilinéaires invariantes au sens (12.25) sont toutes multiples de la forme de Killing. Il nous suffit donc d'étudier la forme de Killing pour connaître la nature compacte ou non de l'algèbre de Lie. Ceci nous amène au second théorème de Cartan.

**Théorème 12.37 :** *(Cartan)*

Une algèbre de Lie semi-simple  $\mathfrak{g}$  est compacte si et seulement si sa forme de Killing est définie négative.

**Théorème 12.38 :**

L'algèbre de Lie  $\mathfrak{g}$  d'un groupe de Lie  $G$  compact (au sens topologique) est compacte.

On trouvera une preuve de ce théorème (qui utilise la mesure de Haar sur un groupe compact, voir plus loin) p. 109 de la ref. [9].

Dans le cas des groupes de matrices auquel nous nous limitons ici, il suffit que la partie de  $\mathbb{R}^n$  qui est en bijection avec  $G$  soit un compact pour que  $G$  soit compact.

*Rappel :* sur  $\mathbb{R}^n$  un compact est un fermé borné.

---

De façon générale, une forme quadratique définie sur une algèbre de Lie complexe est indéfinie. Donc (puisque c'est le cas en particulier pour la forme associée à  $g$ ), toute algèbre de Lie complexe est non compacte.

Une algèbre de Lie compacte est une forme réelle particulière  $L^r$  d'une algèbre de Lie complexe.

**Théorème 12.39 :**



Pour une algèbre de Lie compacte semi-simple,  $C_{rs}^\ell$  peut être représenté par un tenseur d'ordre 3 covariant, complètement antisymétrique.

*Preuve :*

$$\begin{aligned} C_{rs\ell} &= C_{rs}^t g_{t\ell} \\ &= C_{rs}^t C_{tm}^n C_{\ell n}^m = -C_{sm}^t C_{tr}^n C_{\ell n}^m - C_{mr}^t C_{ts}^n C_{\ell n}^m \\ &= C_{sm}^t C_{rt}^n C_{\ell n}^m + C_{mr}^t C_{ts}^n C_{n\ell}^m. \end{aligned}$$

Cette dernière expression est invariante sous les permutations circulaires, et antisymétrique en  $r, s$ . Donc  $C_{rs\ell}$  est complètement antisymétrique.

D'autre part, puisque  $L$  est compacte, il est toujours possible par changement de base d'écrire  $g_{t\ell} = -\delta_{t\ell}$  (en effet  $g$  est symétrique, réelle et définie négative), et donc  $C_{rs\ell} = -C_{rs}^\ell$ , (ou encore  $C_{rs\ell} = C_{rs}^\ell$  en changeant le signe des générateurs)).

### Exercice 12.2

Montrer que  $C = g^{\rho\sigma} X_\rho X_\sigma$  est un opérateur de Casimir, c'est-à-dire qu'il commute avec tous les éléments de l'algèbre de Lie

$$\begin{aligned} [C, X_\tau] &= g^{\rho\sigma} [X_\rho X_\sigma, X_\tau] = g^{\rho\sigma} X_\rho [X_\sigma, X_\tau] + g^{\rho\sigma} [X_\rho, X_\tau] X_\sigma \\ &= g^{\rho\sigma} X_\rho C_{\sigma\tau}^\lambda X_\lambda + g^{\rho\sigma} C_{\rho\tau}^\lambda X_\lambda X_\sigma = g^{\rho\sigma} C_{\sigma\tau}^\lambda X_\rho X_\lambda + g^{\sigma\rho} C_{\sigma\tau}^\lambda X_\lambda X_\rho \\ &= g^{\rho\sigma} C_{\sigma\tau}^\lambda (X_\rho X_\lambda + X_\lambda X_\rho) = g^{\rho\sigma} \underbrace{g^{\lambda\nu} C_{\nu\sigma\tau}}_{\text{antisymétrique en } \rho, \lambda} \underbrace{(X_\rho X_\lambda + X_\lambda X_\rho)}_{\text{symétrique en } \rho, \lambda} = 0 \end{aligned}$$

### Exercice 12.3

Dans le cas particulier de la  $\mathbb{R}$ -algèbre de Lie  $\mathfrak{su}(2)$ , vérifier qu'avec les normalisations (14.36),  $g_{\rho\sigma} = -2\delta_{\rho\sigma}$ .

Examinons le cas plus général d'une algèbre de Lie semi-simple réelle quelconque.

Rappelons que par une transformation réelle orthogonale on peut toujours diagonaliser la matrice réelle symétrique  $A$  (ici  $g$ ) qui définit la forme quadratique (ici la forme de Cartan-Killing), celle-ci ayant des valeurs propres non nulles par hypothèse de semi-simplicité.

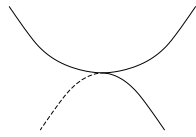
En normalisant les vecteurs propres (ici les générateurs  $X_i$  qui forme la nouvelle base dans laquelle  $g$  est diagonale), on se ramène ensuite à une forme quadratique dont la matrice est uniquement constituée de coefficients diagonaux, valant 0 (en nombre  $n_0$ ),  $-1$  (en nombre  $n_-$ ) ou  $+1$  (en nombre  $n_+$ ). La signature de la forme quadratique est le triplet  $(n_0, n_-, n_+)$ , qui est un invariant (par changement de base) d'après le théorème d'inertie de Sylvester.

Dans le cas présent d'une algèbre de Lie semi-simple, la forme est non dégénérée, *i.e.*  $n_0 = 0$ . Si  $g$  possède des valeurs propres positives, le théorème de Sylvester interdit de pouvoir changer

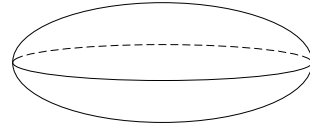
leur signe, et  $\mathfrak{g}$  est non-compacte. Cela sera possible à condition de multiplier chacun des générateurs de valeur propre  $+1$  par  $i$ , de façon à obtenir  $g$  sous la forme  $g_{t\ell} = -\delta_{t\ell}$ .

Une telle manipulation correspond à prendre une forme réelle particulière de l'extension complexe de la  $\mathbb{R}$ -algèbre de Lie initiale. Cela change la compacité de la  $\mathbb{R}$ -algèbre ainsi construite par rapport à la  $\mathbb{R}$ -algèbre initiale.

*Exemple :* Considérons les deux groupes  $O(3, 1)$  et  $O(4)$ . La topologie de ces deux variétés diffère notablement, la première étant non-compacte alors que la seconde est compacte, comme illustré dans la fig. 12.1.



La variété de  $O(3, 1)$  à 4 dimensions



La variété de  $O(4)$  à 4 dimensions

FIGURE 12.1 – Visualisation symbolique des variétés correspondant aux groupes  $O(3, 1)$  et  $O(4)$ . A gauche, hyperboloïde (correspondant à  $O(3, 1)$ ). À droite, ellipsoïde (correspondant à  $O(4)$ ).

Le chapitre 17 sera consacré à l'étude détaillée du *groupe de Lorentz*. En particulier, nous construirons son algèbre de Lie. Elle est constituée de 6 générateurs :

- ◇  $J_1, J_2, J_3$       générateurs des rotations
- ◇  $K_1, K_2, K_3$     générateurs des boosts

qui satisfont les relations de commutation :

$$\begin{aligned} [J_i, J_j] &= i \varepsilon_{ijk} J_k \\ [K_i, K_j] &= -i \varepsilon_{ijk} J_k \\ [J_i, K_j] &= i \varepsilon_{ijk} K_k. \end{aligned}$$

Cette  $\mathbb{R}$ -algèbre  $\mathfrak{so}(3, 1; \mathbb{R})$  est non compacte. Ceci est directement relié au fait que le groupe de Lorentz est non compact : les rapidités  $\phi$  qui codent les boosts  $e^{i\phi\vec{n}\cdot\vec{K}}$  suivant l'axe  $\vec{n}$  varient sur tout  $\mathbb{R}$ , qui n'est pas compact, contrairement aux rotations pour lesquelles l'angle de rotation peut être ramené à varier sur un domaine compact de  $\mathbb{R}$ .

On pose

$$\begin{cases} \vec{M} &= \frac{1}{2} (\vec{J} + i\vec{K}) \\ \vec{N} &= \frac{1}{2} (\vec{J} - i\vec{K}) . \end{cases}$$

Avec  $\vec{M}$  et  $\vec{N}$  on construit une nouvelle algèbre de Lie sur  $\mathbb{R}$ . C'est la *forme réelle compacte* de la  $\mathbb{C}$ -algèbre de Lie  $\mathfrak{so}(4)$ , dont  $\mathfrak{so}(3, 1; \mathbb{R})$  est une autre forme réelle.

L'algèbre de Lie ainsi obtenue satisfait les relations de commutation suivantes :

$$\begin{cases} [M_i, M_j] = i \varepsilon_{ijk} M_k \\ [N_i, N_j] = i \varepsilon_{ijk} N_k \\ [M_i, N_j] = 0. \end{cases}$$

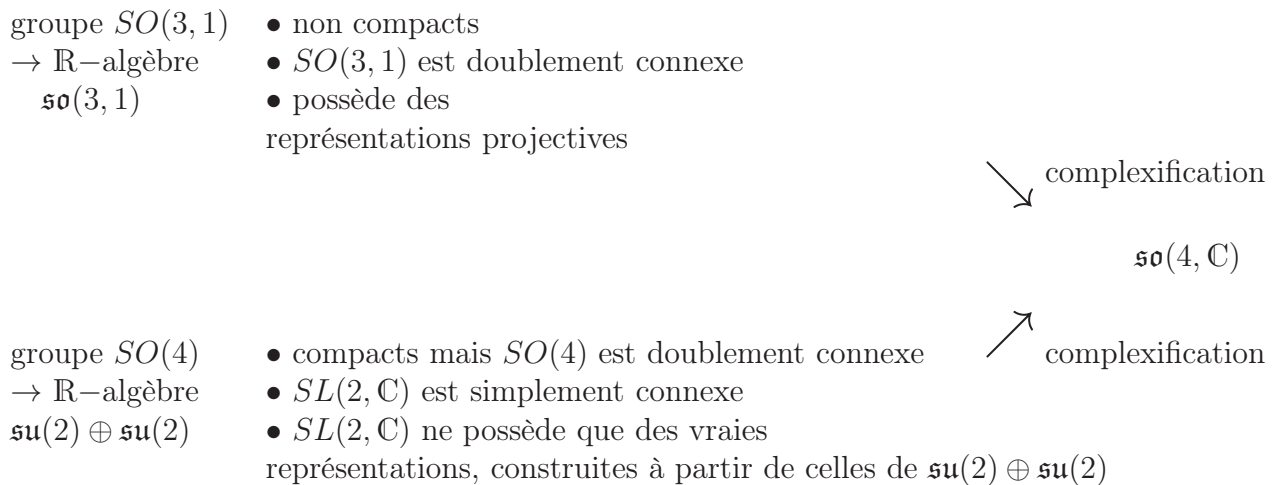
L'algèbre de Lie obtenue est donc tout simplement celle de  $SU(2) \times SU(2)$ , c'est-à-dire  $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ . On notera que la  $\mathbb{C}$ -algèbre de Lie  $\mathfrak{sl}(2)$  du groupe  $SL(2, \mathbb{C})$  peut être vue comme une  $\mathbb{R}$ -algèbre de Lie de dimension 6, qui est précisément  $\mathfrak{su}(2) \oplus i\mathfrak{su}(2) \simeq \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ .

Cette dernière  $\mathbb{R}$ -algèbre de Lie  $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$  de dimension 6 est la forme réelle compacte de l'algèbre de Lie du groupe non-compact  $SO(3, 1)$ .

Le groupe  $SL(2, \mathbb{C})$  est le groupe de recouvrement de  $SO(3, 1)$  et de  $SO(4)$ , de la même façon que  $SU(2)$  est le groupe de recouvrement de  $SO(3)$ .

La construction des représentations correspondantes est alors standard (voir chapitres 14, 15 et 18).

**Résumé :**



Formellement, le point clef qui assure l'absence de représentation projectives dans les groupes  $SU(2)$  et  $SL(2, \mathbb{C})$  est le fait que ces groupes sont semi-simples, et simplement connexes, *i.e.* que leur groupe d'homotopie est trivial. Ceci signifie que tout chemin sur le groupe est réductible à un chemin, ce qui n'est pas le cas de  $SO(3)$ , de  $SO(3, 1)$  ou de  $SO(4)$  par exemple. Ceci justifie le passage de  $SO(3)$  à  $SU(2)$  ou de  $SO(3, 1)$  à  $SL(2, \mathbb{C})$  puisque ces groupes n'ont que des vraies représentations, ce qui est beaucoup plus simple à manipuler mathématiquement, au prix de l'introduction des spins demi-entiers qui caractérisent ces représentations.

D'autre part, le passage à des groupes compacts (« trick unitaire » inventé par Hermann Weyl) facilite grandement l'étude des représentations.

Nous espérons avoir convaincu le lecteur sur cet exemple que l'introduction de notions aussi complexes (dans tous les sens du terme) est pertinente physiquement !

La compacité est donc un cadre puissant :

- ◇ Elle permet d'obtenir *toute* la composante connexe de l'identité d'un groupe à partir de l'exponentielle de son algèbre de Lie (dans le cas général on a seulement accès à un voisinage de l'identité).
- ◇ L'étude des représentations des groupes compacts est beaucoup plus simple (voir plus loin).

En outre, tous les résultats classiques valables pour les groupes finis (application à la cristallographie par exemple) s'étendent sans difficulté aux groupes compacts. En effet, la somme sur les éléments d'un groupe fini peut être remplacée par une intégrale : pour tout groupe  $G$  compact, il existe une *mesure* de Haar unique t.q.

- ◇ La mesure est invariante à gauche et à droite, *i.e.*  $\forall f$  continue sur  $G, \forall h \in G, \int_G f(g)dg = \int_G f(hg)dg = \int_G f(gh)dg$
- ◇ Le volume du groupe est fini. On peut normaliser la mesure de Haar de sorte que  $\int_G dg = 1$ .

On notera que les théories de jauge non abéliennes sont basées sur les groupes compacts.

### 12.2.4 Classification

Le théorème suivant permet de ramener l'étude de la classification des algèbres de Lie à celle des algèbres simples.

#### **Théorème 12.40 :**

Toute algèbre de Lie  $\mathfrak{g}$  peut s'écrire sous la forme  $\mathfrak{g} = \oplus$  générateurs abéliens  $\oplus$  sous-algèbres simples.

Les algèbres de Lie simples réelles et complexes sont complètement classifiées (Killing, Cartan). Cette classification dépasse le cadre de ce cours. Nous allons simplement donner la classification des  $\mathbb{C}$ -algèbres de Lie, ce qui nous permettra de caractériser les groupes de Lie classiques et leurs algèbres de Lie.

#### **Transformations unitaires : $U(N)$ et $SU(N)$**

Soient  $\xi$  et  $\eta \in \mathbb{C}^N$  qui se transforment sous l'action de matrices complexes  $N \times N$  suivant :

$$\forall U \in GL(N, \mathbb{C}), \quad \eta^a \rightarrow \eta'^a = U^a_b \eta^b \quad \text{et} \quad \xi^a \rightarrow \xi'^a = U^a_b \xi^b. \quad (12.27)$$

La loi de transformation est ici écrite de façon covariante : les composantes de  $\xi$  et  $\eta$  sont par définition des quantités contravariantes, notées  $\xi^a$  et  $\eta^a$ , qui se transforment sous l'action de la matrice  $U$  dont les éléments de matrice  $(U)_{ab}$  permettent de définir

$$U^a_b \equiv (U)_{ab} \quad (12.28)$$

de sorte que les vecteurs images  $\xi'$  et  $\eta'$  sont bien eux aussi contravariants.

La transposée  $U^t$  de la matrice  $U$  a pour éléments de matrice

$$(U^t)_{ab} \equiv (U)_{ba} = U^b_a. \quad (12.29)$$

Par définition,  $U(N)$  est l'ensemble des matrices qui laissent invariant  $\bar{\eta}^t \xi = \bar{\eta}^a \xi^a$ .

Sous l'action d'une matrice de  $U(N)$ ,  $\bar{\eta}^a \xi^a$  devient  $\bar{\eta}^a \xi'^a = \bar{U}^a_b \bar{\eta}^b U^a_c \xi^c = \bar{\eta}^a \xi^a$  donc  $\bar{U}^a_b U^a_c = \delta^b_c$  soit encore  $(\bar{U}^t)^b_a U^a_c = \delta^b_c$ , ce qui signifie matriciellement que  $\bar{U}^t U = \mathbb{1}$ , i.e.  $U^+ U = \mathbb{1}$ , où

$$\mathbb{1} = \begin{pmatrix} 1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & 1 \end{pmatrix}_{N \times N}.$$

On montre de la même façon que  $U U^+ = \mathbb{1}$ .

*Remarque* : il n'est bien sûr pas indispensable d'utiliser les notations covariantes pour ce calcul élémentaire. Le seul intérêt est ici de se familiariser avec ces manipulations, utiles pour l'étude générale des représentations irréductibles de  $SU(N)$  à l'aide des tableaux de Young. Nous verrons en détail dans le chap. 15 le cas particulier des spineurs de  $SU(2)$ , qui présentent un grand intérêt en physique quantique.

L'algèbre de Lie de  $U(N)$  s'obtient aisément en écrivant la condition d'unitarité sous la forme

$$(1 + X_a)(1 + X_a^+) \approx 1 + (X_a + X_a^+) \quad (12.30)$$

qui conduit donc à  $X_a = -X_a^+$  : les générateurs de  $U(N)$  sont donc formés des *matrices antihermitiennes*, au nombre de  $N^2$  (voir ci-dessous).

Les transformations particulières  $\xi^a \rightarrow e^{i\alpha} \xi^a$  appartiennent à  $U(N)$ , et correspondent au sous-groupe  $U(1)$ , de générateur  $i\mathbb{1}$ .

On se ramène à un groupe *simple*, en imposant que les générateurs soient orthogonaux à  $\mathbb{1}_{N \times N}$ , i.e.  $\delta_{bc}(X_a)_{bc} = 0$  soit  $(X_a)_{bb} = 0$  ou encore  $\text{Tr } X_a = 0$ .

Ceci définit l'algèbre de Lie de  $SU(N)$ , qui est formée de  $N^2 - 1$  générateurs (voir ci-dessous).

*Remarques* :

- ◇ Pour  $U(N)$ ,  $U^+ U = 1$  conduit à  $|\det U|^2 = 1$  donc  $\det U = e^{i\alpha}$ .
- ◇ à la condition  $\text{Tr } X_a = 0$  au niveau de l'algèbre correspond la condition  $\det U = 1$  pour le groupe.
- ◇ Pour toute matrice  $n \times n$  (supposée diagonalisable),

$$\boxed{\det(1 + \varepsilon X) = 1 + \varepsilon \text{Tr } X + O(\varepsilon^2)}. \quad (12.31)$$

*Preuves* :

Pour la seconde remarque, on écrit, dans un voisinage de l'identité,  $U = e^X$  et l'on diagonalise  $X$ , sous la forme  $X = P\Lambda P^{-1}$ , où  $\Lambda = \text{Diag}(\lambda_1, \dots, \lambda_n)$ . Alors  $\exp X = P \exp \Lambda P^{-1}$  et

$$\det U = \det \exp \Lambda = \exp \text{Tr} \Lambda = \exp \text{Tr} X, \text{ q.e.d.}$$

Pour la troisième remarque, il suffit de diagonaliser la matrice  $X$ . On écrit donc  $X = P \Lambda P^{-1}$ , où  $\Lambda = \text{Diag}(\lambda_1, \dots, \lambda_n)$ , ce qui conduit à  $\det(\mathbb{1} + \varepsilon X) = \det(P(\mathbb{1} + \varepsilon \Lambda)P^{-1}) = \det(\mathbb{1} + \varepsilon \Lambda) \sim 1 + \varepsilon \text{Tr} \Lambda + O(\varepsilon^2)$ , ce qui mène au résultat puisque  $\text{Tr} D\Lambda = \text{Tr} X$ .

Nombre de générateurs :

$$\left. \begin{array}{l} N \text{ coefficients imaginaires purs diagonaux} \\ \frac{N(N-1)}{2} \text{ coefficients non diagonaux complexes} \end{array} \right\} \Rightarrow N + 2 \frac{N(N-1)}{2} = N^2 \text{ paramètres réels,}$$

ce qui donne le nombre de générateurs de  $U(N)$ ,

Pour  $SU(N)$ , la condition supplémentaire  $\det U = 1$ , qui est équivalente à  $\text{Tr} X = 0$ , correspond donc à un paramètre en moins, ce qui laisse finalement  $N^2 - 1$  paramètres réels.

**Résumé**

spécial  
(det = 1)

↘

SU(N)

↑

unitaire

matrices  $N \times N$

↙

$N^2 - 1$  générateurs = matrices anti-hermitiennes  $N \times N$ ,  
de trace nulle

La  $\mathbb{R}$ -algèbre de Lie  $\mathfrak{su}(N, \mathbb{R})$  est la seule forme réelle compacte de la  $\mathbb{C}$ -algèbre de Lie  $\mathfrak{sl}(N, \mathbb{C})$ , constituée des matrices complexes  $N \times N$  de trace nulle. C'est une sous-algèbre de Lie de l'algèbre de Lie de l'algèbre  $\mathfrak{gl}(N, \mathbb{C})$  constituée des matrices complexes  $N \times N$ . Dans la classification de Cartan, c'est l'algèbre de Lie  $A_{N-1}$ .

**Transformations orthogonales :  $O(N)$  et  $SO(N)$**

Soient  $\eta, \xi \in \mathbb{R}^n$ . Par définition,  $O(N)$  est l'ensemble des matrices réelles  $A$  qui laissent invariant  $\eta^a \xi^a = \eta^t \xi$ . Ceci s'écrit  $A^t A = A A^t = 1$ , c'est-à-dire, au niveau de l'algèbre de Lie,  $T^a + (T^a)^t = 0$ , qui correspondent aux matrices antisymétriques  $N \times N$ .

*Remarque :*  $O(N)$  contient des transformations discrètes, qui n'ont pas d'analogue dans l'algèbre de Lie. On se débarrasse de ces transformations en considérant  $SO(N)$  qui impose la contrainte  $\det A = 1$ .

Nombre de générateurs :

Chaque élément de  $O(N)$  possède  $N^2$  éléments de matrice.

La contrainte  $A^T A = A A^T = 1$  impose donc  $\frac{N(N-1)}{2}$  conditions d'orthogonalité  $\vec{e}_i \cdot \vec{e}_j = 0$  ( $i \neq j$ ).

Les  $N$  conditions d'orthogonalité  $\vec{e}_i \cdot \vec{e}_i = 1$  réduisent le nombre de générateurs : au total,  $\mathfrak{o}(N)$  possède

$$N^2 - \left[ N + \frac{N(N-1)}{2} \right] = \frac{N(N-1)}{2} \text{ générateurs.}$$

L'algèbre  $\mathfrak{o}(N, \mathbb{R})$  est la seule forme réelle compacte de la  $\mathbb{C}$ -algèbre de Lie  $\mathfrak{o}(N, \mathbb{C})$ , qui est constituée des matrices complexes  $N \times N$  antisymétriques. Dans la classification de Cartan,  $\mathfrak{o}(2N+1, \mathbb{C})$  est notée  $B_N$  et  $\mathfrak{o}(2N, \mathbb{C})$  est notée  $D_N$ .

*Remarques :*

- ◇  $O(N)$  n'est pas connexe (à cause des symétries miroirs, discrètes).
- ◇  $SO(N)$  est connexe
- ◇  $SO(N)$  n'est pas simplement connexe pour  $N \geq 3$

On a vu dans le chapitre 11 qu'en mécanique quantique, les états physiques sont décrits par des rayons (qui sont par définition des vecteurs de l'espace de Hilbert à une phase près). Les opérations de symétrie, qui préservent par définition les probabilités de transition (mais pas forcément les amplitudes de probabilité), sont codées sur l'espace de Hilbert par des opérateurs unitaires ou anti-unitaires (c'est le théorème de Wigner). L'espace de Hilbert fournit alors un espace de représentation à *une phase près* du groupe de symétrie  $G$ , ce qui est très pénible mathématiquement. On étudie plutôt, comme on le verra dans le chapitre 16, le groupe de recouvrement  $\tilde{G}$  du groupe de symétrie qui est (en un sens à préciser mathématiquement) le groupe le plus simple qui ait la même algèbre de Lie que  $G$ , et qui est *simplement connexe* (tous les chemins sur le groupe peuvent se contracter en un point). On montre alors que les *représentations de  $\tilde{G}$*  induisent automatiquement les *représentations projectives de  $G$* . Le groupe  $\tilde{G}$  peut en principe posséder lui-même des représentations projectives, s'il possède ce que l'on appelle des charges centrales (des opérateurs qui commutent avec tous les éléments du groupe de symétrie). Une telle situation n'arrive pas dans le cas des algèbres de Lie semi-simples.

*Exemples :* (qui seront détaillés respectivement aux chapitres 14,15 et 17,18)

non simplement connexe	simplement connexe
$SO(3)$	$SU(2)$
spin entier = représentation	spin entier = représentation
spin demi-entier $\neq$ représentation	spin demi-entier = représentation

$SO(3, 1)$	$SL(2, \mathbb{C}) \leftrightarrow SU(2) \times SU(2)$ (forme réelle compacte)
(groupe de Lorentz)	spineurs    spineurs $(\frac{1}{2}, 0)$ $(0, \frac{1}{2})$
	$\underbrace{\hspace{10em}}$ $\downarrow$

bispineurs (indispensables si l'on rajoute la symétrie  $P$ )  
 → équations d'onde relativistes (ex. : éq. de Dirac)

### Groupes symplectiques :

Soient  $\eta, \xi \in \mathbb{R}^{2N}$ . Par définition,  $S_p(2N)$  est le groupe des matrices qui préserve le produit scalaire antisymétrique  $\eta^a E_{ab} \xi^b$  avec

$$E_{ab} = \left( \begin{array}{c|c} 0 & \mathbf{1} \\ \hline - & - \\ -\mathbf{1} & 0 \end{array} \right) \quad (12.32)$$

↙ ↘  
blocs  $N \times N$  (12.33)

L'algèbre de Lie correspondante possède  $N(2N + 1)$  générateurs. Elle est non compacte. La  $\mathbb{C}$ -algèbre de Lie est notée  $\mathfrak{sp}(2N, \mathbb{C})$ , ou encore  $C_N$  dans la classification de Cartan.

Un exemple de structure symplectique est fourni par le crochets de Poisson que nous avons étudiés au Chap. 9.

### Groupes exceptionnels : $G_2, F_4, E_6, E_7, E_8$

Ces groupes sont en particulier utilisés dans certains modèles d'unification des interactions fondamentales, par exemple en théorie des cordes.

Bien entendu, nous n'avons pas épuisé, comme le lecteur attentif l'aura compris, toute les  $\mathbb{R}$ -algèbres de Lie simples dans ce qui vient d'être présenté. En effet, à chaque  $\mathbb{C}$ -algèbre de Lie simple que nous venons de décrire correspondent plusieurs formes réelles. On trouvera le détail de cette classification dans [9] (Chap. 1, 5).



# Chapitre 13

## Représentations

Dans le chapitre 4, nous avons introduit le concept d'action d'un groupe  $G$  sur un ensemble  $X$ . Le cas particulier où l'ensemble  $X$  est un espace vectoriel joue un rôle central en physique. On parle alors de représentation linéaire d'un groupe, ou plus simplement d'une représentation d'un groupe. Leur étude générale fait l'objet de ce chapitre.

### 13.1 Généralités

#### 13.1.1 Définition, premiers concepts

**Définition 13.1 :** *Représentation d'un groupe*

Soit  $E$  un espace vectoriel sur  $\mathbb{R}$  ou  $\mathbb{C}$ , de dimension  $n$ .

Une représentation linéaire d'un groupe  $G$  dans  $E$  est un homomorphisme  $D$  de  $G$  dans un sous-groupe  $\mathcal{G}$  du groupe  $GL(n)$  des opérateurs linéaires et inversibles de  $E$ . On a donc

$$\boxed{\forall g_1, g_2 \in G, D(g_1 g_2) = D(g_1) D(g_2)}. \quad (13.1)$$

En particulier  $D(e) = \mathbb{1}$  et  $D(g^{-1}) = D(g)^{-1}$ , où  $e$  et  $\mathbb{1}$  sont les éléments neutres de  $G$  et de  $GL(n)$ .

$E$  est appelé espace de représentation, et  $n = \dim V$  est la dimension de la représentation.  $D(g)$  s'écrit donc comme une matrice  $n \times n$ .

Il ne faut surtout pas confondre cette dimension avec celle du groupe. De façon générale, un groupe donné possède plusieurs représentations, de dimensions différentes.

**Définition 13.2 :** *Représentation triviale*

La représentation de  $G$  qui à tout  $g \in G$  associe  $\mathbb{1}$  est appelée représentation triviale, ou représentation identité.

**Définition 13.3 :** *Représentation fidèle*

Une représentation de  $G$  est dite fidèle si  $G$  est isomorphe à  $\mathcal{G}$ , soit de façon équivalente que  $\text{Ker } D = \{e\}$ , ou  $D(g) = D(g') \Leftrightarrow g = g'$ .

Dans le cas contraire, puisque  $\ker D$  est un sous-groupe invariant de  $G$ , on se ramène à une représentation fidèle en considérant la représentation de  $G/\ker D$  dans  $E$ .

**Définition 13.4 :** *Matrice représentative*

Dans le cas où  $E$  est de dimension finie, soit  $e_i, i \in \{1, \dots, n\}$  une base de  $E$ .

A tout  $g \in G$  on associe la matrice  $\mathcal{D}(g)$  représentative de  $D(g)$  définie, de façon usuelle en algèbre linéaire, par (en utilisant la convention d'Einstein de sommation sur les indices répétés)

$$D(g)e_j = e_i \mathcal{D}_{ij}(g), \quad (13.2)$$

ce qui signifie que l'on lit sur la matrice, dans la colonne  $j$ , les coefficients dans la base  $e_i$  de l'image du vecteur  $e_j$ .

On vérifie facilement le fait que  $D$  est une représentation si et seulement si la relation matricielle suivante est satisfaite :

$$\mathcal{D}_{ij}(g \cdot g') = \mathcal{D}_{ik}(g) \mathcal{D}_{kj}(g'). \quad (13.3)$$

Exercice 13.1 Exemples de représentations

1. *Groupe  $SO(2)$  des rotations du plan*

Vérifier que les matrices

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (13.4)$$

forment une représentation de dimension 2 du groupe des rotations du plan. Quelle est l'interprétation géométrique de ces matrices ?

2. *Groupe des permutations  $\mathcal{S}_n$*

a. Vérifier qu'une représentation de dimension 2 de  $\mathcal{S}_2$  est fournie par

$$\begin{aligned} e &\mapsto \mathcal{D}(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ p &\mapsto \mathcal{D}(p) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Plus généralement, une matrice de permutation de taille  $n$  est une matrice carrée  $n \times n$  dont les éléments de matrice sont constitués de 0 et de 1, avec la contrainte que dans chaque ligne et chaque colonne, 1 apparaît une et une seule fois.

A toute permutation  $\sigma$  de  $\mathcal{S}_n$  on associe alors la matrice de permutation  $P_\sigma$  de taille  $n$  dont les éléments de matrice sont définis par

$$[P_\sigma]_{ij} = \delta_{i, \sigma(j)}. \quad (13.5)$$

b. Montrer que l'on définit ainsi une représentation de  $\mathcal{S}_n$ .

c. En déduire que l'ensemble des matrices de permutation de taille  $n$  est un sous-groupe de  $GL(n, \mathbb{R})$ .

d. Cette représentation est-elle fidèle ?

e. Quel est le cardinal du groupe des matrices de permutation de taille  $n$  ?

3. *groupe des tresses*  $B_3$

Nous avons rencontré ce groupe page 75. Vérifier qu'une représentation de  $B_3$  est fournie par

$$\sigma_1 = \begin{pmatrix} -t & 1 \\ 0 & 1 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 1 & 0 \\ t & -t \end{pmatrix}$$

**Définition 13.5 :** *Représentations équivalentes*

Deux représentations  $D$  et  $D'$  dans les espaces  $E$  et  $E'$  sont dites équivalentes si et seulement si il existe un opérateur  $S$  inversible t.q.

$$\forall g \in G, D'(g) = S^{-1}D(g)S. \quad (13.6)$$

$S$  est appelé opérateur d'entrelacement.

Dans le cas où  $E$  et  $E'$  sont de dimension finie, on peut identifier  $E$  et  $E'$ , et la relation (13.6) traduit simplement un changement de base au niveau des matrices représentatives. Cela justifie par avance le fait de ne pas distinguer des représentations équivalentes.

**Définition 13.6 :** *Caractère d'une représentation*

On appelle caractère d'une représentation de dimension finie la trace de l'opérateur  $D(g)$  :

$$\chi(g) = \text{Tr } D(g). \quad (13.7)$$

**Propriétés 13.7 :**

- 1) Deux représentations équivalentes ont même caractère.
- 2) Un caractère est une fonction de classe, ou fonction centrale : il est constant sur chaque classe de conjugaison.
- 3) La dimension de la représentation est donnée par  $\chi(e)$ .

*Preuve :*

1) est immédiat puisque par cyclicité de la trace, un changement de base n'affecte pas la valeur du caractère.

2) s'obtient par cyclicité de la trace et en utilisant le fait que  $D$  est un morphisme :

$$\chi(aga^{-1}) = \text{Tr}D(aga^{-1}) = \text{Tr}[D(a)D(g)D(a)^{-1}] = \text{Tr}D(g) = \chi(g).$$

3) est due au fait que  $\text{Tr}D(e) = \text{Tr } \mathbf{1}_n = n = \dim D$ .

---

### 13.1.2 Représentations réductibles et irréductibles

**Définition 13.8 :** *Somme directe de deux représentations*

Soient  $D_1$  et  $D_2$  deux représentations d'un même groupe  $G$ , respectivement dans les espaces  $E_1$  et  $E_2$ . On peut alors construire une représentation qui agit sur l'espace somme directe  $E = E_1 \oplus E_2$ . On note cette représentation  $D_1 \oplus D_2$ .

Par construction, chacun des deux sous-espace  $E_1$  et  $E_2$  de  $E$  sont laissés invariants par cette représentation  $D = D_1 \oplus D_2$ .

Dans le cas où  $E_1$  et  $E_2$  sont de dimensions finies  $p_1$  et  $p_2$ , on peut alors considérer les matrices représentatives  $\mathcal{D}_1(g)$  et  $\mathcal{D}_2(g)$  de  $D_1(g)$  et  $D_2(g)$ , dans des bases données de  $E_1$  et  $E_2$ . La matrice représentative de la somme directe de ces représentations, dans la base correspondante de  $E_1 \oplus E_2$ , est donnée par

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 \\ 0 & \mathcal{D}_2(g) \end{pmatrix} \quad (13.8)$$

de dimension  $(p_1 + p_2) \times (p_1 + p_2)$ .

Nous allons à présent, partant d'un représentation donnée, examiner la situation inverse.

**Définition 13.9 :** *Réductibilité, irréductibilité et complète réductibilité*

*Cas général :*

- ◇ Une représentation d'un groupe  $G$  est dite réductible si elle laisse invariant un sous-espace  $E_1$  de  $E$ .
- ◇ Dans le cas contraire elle est dite irréductible, ce qui est équivalent au fait qu'il n'existe pas de sous-espace invariant par  $D(g)$ ,  $\forall g \in G$ , autre que  $\emptyset$  et  $E$ .
- ◇ Une représentation sur  $E$  est complètement réductible si elle peut s'écrire comme la somme directe de représentations irréductibles :  $E$  peut s'écrire comme somme directe de sous-espace laissés invariants par la représentation et sur lesquels la représentation est irréductible.

*Cas d'une représentation de dimension finie :*

Si  $E$  est de dimension finie :

- ◇ Une représentation est réductible si elle est équivalente (par un changement de base d'après ce qui a été dit plus haut) à une représentation  $D$  de  $G$  dont la matrice représentative peut s'écrire,  $\forall g \in G$ , sous la forme

$$\mathcal{D}(g) = \left( \begin{array}{c|c} A & B \\ \hline 0 & C \end{array} \right)$$

où  $A$  est une matrice  $n_A \times n_A$ , cette écriture de  $\mathcal{D}$  étant équivalente au fait qu'il existe un sous-espace de  $E$  de dimension  $n_A$  invariant sous  $\mathcal{D}(g)$ ,  $\forall g \in G$ .

- ◇ Une représentation est dite complètement réductible si elle est équivalente à une représentation de la forme

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_{11}(g) & 0 & & 0 \\ 0 & \mathcal{D}_{22}(g) & & \\ \vdots & & \ddots & \\ 0 & & & \mathcal{D}_{nn}(g) \end{pmatrix}, \forall g \in G,$$

où  $\mathcal{D}_{11}, \mathcal{D}_{22}, \dots, \mathcal{D}_{nn}$  sont irréductibles.

*Remarques :*

- ◇ On prendra bien garde au fait que les structures en blocs ci-dessus doivent être valables dans une base donnée une fois pour toute,  $g$  étant arbitraire (ce qui est bien sûr beaucoup plus restrictif que demander, pour chaque choix de  $g$ , l'existence d'une base telle que  $D(g)$  possède une telle structure en blocs...).
- ◇ Le corps du  $\mathbb{K}$ -espace vectoriel  $E$  joue un rôle essentiel dans la discussion de l'irréductibilité ci-dessus. En particulier, une représentation peut être réductible sur  $\mathbb{C}$  sans l'être sur  $\mathbb{R}$ . Ainsi la représentation (13.4) est irréductible sur  $\mathbb{R}$  mais est réductible sur  $\mathbb{C}$  : par changement de base on peut la mettre sous la forme

$$\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}. \quad (13.9)$$

Plus généralement, pour  $M \in SO(N)$ , on peut écrire  $M = \exp A$  où  $A$  est une matrice antisymétrique (la composante connexe de l'identité d'un groupe compact s'obtient exponentiation de son algèbre de Lie). Or  $iA$  est hermitienne, donc est diagonalisable à valeur réelles (notons  $\Lambda$  cette matrice diagonale réelle), et donc  $A$  est diagonalisable sur  $\mathbb{C}$  à valeur imaginaire pure. On a donc  $M = \exp[Pi\Lambda P^{-1}] = P \exp(i\Lambda) P^{-1}$ .

**Définition 13.10 :** *Représentation conjuguée*

Soit  $D$  une représentation d'un groupe  $G$  et  $\mathcal{D}$  sa matrice représentative dans une certaine base. Les matrices  $\mathcal{D}^*$  complexes conjuguées de  $\mathcal{D}$  forment une autre représentation  $D^*$  appelée conjuguée.

En effet,  $\mathcal{D}_{ij}(gg') = \mathcal{D}_{ik}(g)\mathcal{D}_{kj}(g')$  conduit immédiatement à

$$\mathcal{D}_{ij}^*(gg') = \mathcal{D}_{ik}^*(g)\mathcal{D}_{kj}^*(g'). \quad (13.10)$$

**Définition 13.11 :** *Représentation réelle*

Une représentation  $D$  est dite réelle s'il existe une base dans laquelle  $\mathcal{D} = \mathcal{D}^*$ . Cela a pour conséquence que le caractère de la représentation est réel.

**Propriété 13.12 :** *Caractère réel*

Réciproquement, dans le cas des représentations irréductibles des groupes finis ou des groupes compacts, si un caractère est réel, alors  $D$  et  $D^*$  sont équivalentes. Nous montrerons ce résultat dans le cas des groupes finis un peu plus loin.

**Définition 13.13 :** *Représentation pseudo-réelle*

Une représentation  $D$  est appelée pseudo-réelle si  $D$  et  $D^*$  sont équivalentes sans qu'il existe une base dans laquelle  $\mathcal{D} = \mathcal{D}^*$ .

Un exemple de représentation pseudo-réelle est fourni par la représentation de spin 1/2 de  $SU(2)$  : on verra aux chapitres 14 et 15 que la représentation fondamentale de spin 1/2 et sa complexe conjuguée sont équivalentes (en utilisant la matrice  $i\sigma_2$  comme matrice d'entrelacement). En revanche il n'est pas possible de rendre la matrice représentative réelle, car on aurait alors construit une représentation de  $SO(2)$  (hermitique conjugué et transposition se confondant alors...), ce qui est bien sûr impossible puisque  $SU(2)$  est non-abélien alors que  $SO(2)$  est abélien !

**Définition 13.14 :** *Représentation contragradiante*

La représentation contragradiante d'une représentation  $D$  est définie par

$$\bar{D}(g) = {}^t D^{-1}(g). \quad (13.11)$$

soit encore, au niveau des matrices représentatives,

$$\bar{\mathcal{D}}_{ij}(g) = \mathcal{D}_{ji}(g^{-1}). \quad (13.12)$$

Il est immédiat que  $\bar{D}$  est bien une représentation :  $D(g_1 g_2) = D(g_1) D(g_2)$  conduit à

$$\begin{aligned} \bar{D}(g_1 g_2) &= {}^t D^{-1}(g_1 g_2) = {}^t [D(g_1) D(g_2)]^{-1} = {}^t [D(g_2^{-1}) D(g_1^{-1})] = {}^t D(g_1^{-1}) {}^t D(g_2^{-1}) \\ &= \bar{D}(g_1) \bar{D}(g_2). \end{aligned} \quad (13.13)$$

D'autre part,  $D$ ,  $D^*$ ,  $\bar{D}$  sont de façon simultanée soit réductibles, soient irréductibles.

**Définition 13.15 :** *Représentations unitaires*

Si l'espace  $E$  est muni d'un produit scalaire  $\phi(x, y) = \langle x|y \rangle$ , forme bilinéaire symétrique sur  $\mathbb{R}$  ou sesquilinéaire sur  $\mathbb{C}$ , définissant une norme définie positive (*i.e.*  $x \neq 0 \Rightarrow \langle x|x \rangle > 0$ ), on peut toujours choisir une base orthonormale, ce qui signifie que la matrice de  $\phi$  se réduit à  $\mathbf{1}$ . Les opérateurs sont alors dits unitaires s'ils vérifient  $UU^\dagger = \mathbf{1}$ .

Considérons à présent une représentation  $D$ . Elle est dite unitaire si  $\forall g \in G$ , les matrices  $\mathcal{D}(g)$  sont unitaires. Ceci se traduit par

$$\forall g \in G, \forall x, y \in E, \quad \langle x|y \rangle = \langle D(g)x|D(g)y \rangle \quad (13.14)$$

ou de façon équivalente

$$\mathcal{D}(g)^\dagger \mathcal{D}(g) = \mathbf{1}. \quad (13.15)$$

On aura alors

$$\mathcal{D}(g^{-1}) = \mathcal{D}^{-1}(g) = \mathcal{D}^\dagger(g). \quad (13.16)$$

et donc  $\bar{D} = D^*$  d'après (13.11).

**Théorème 13.16 :**

Toute représentation unitaire réductible est complètement réductible.

*Preuve :*

Soit  $E_1$  un sous-espace de  $E$  invariant par  $D$ , et  $E_2$  un sous-espace supplémentaire de  $E_1$ . Le sous-espace  $E_2$  est lui-même invariant car  $\forall g \in G, \forall x \in E_1, \forall y \in E_2$ ,

$$\langle x|D(g)y \rangle = \langle D(g^{-1})x|y \rangle = 0, \quad (13.17)$$

et donc  $D(g)y \in E_2$ .

**Théorème 13.17 :**

Toute représentation d'un groupe fini sur un espace muni d'un produit scalaire est équivalente à une représentation unitaire.

*Preuve :*

Considérons

$$Q = \sum_{g' \in G} D^\dagger(g')D(g'). \quad (13.18)$$

On a alors

$$D^\dagger(g)QD(g) = \sum_{g' \in G} D^\dagger(g'g)D(g'g) = Q, \quad (13.19)$$

où l'on a utilisé le lemme de réarrangement (*i.e* la somme sur  $g'$  ou sur  $g'g$  parcourt une et une seule fois l'ensemble des éléments du groupe fini  $G$ , lemme qu'il est facile d'établir).

D'autre part,  $Q$  est hermitienne, et ses éléments de matrice diagonaux sont de la forme

$$\langle x|Q|x \rangle = \sum_{g' \in G} \|D(g')x\|^2 \quad (13.20)$$

qui est une somme de carrés, positifs ou nul. Ainsi, d'une part  $\langle x|Q|x \rangle \geq 0$  et d'autre part  $\langle x|Q|x \rangle = 0 \Leftrightarrow x = 0$ , ce qui justifie le fait que  $Q$  est définie positive.

La matrice  $Q$  peut alors s'écrire sous la forme

$$Q = V^\dagger V \quad (13.21)$$

où  $V$  est inversible (si l'on demande en plus à  $V$  d'être triangulaire inférieure, cette décomposition, appelée décomposition de Cholesky<sup>1</sup>, est même unique).  $V$  joue alors le rôle d'entrelaceur permettant de définir une représentation  $D'$ ,

$$D'(g) = VD(g)V^{-1}, \quad (13.22)$$

équivalente à  $D$  et unitaire :

$$D'^{\dagger}(g)D'(g) = V^{\dagger-1}D^{\dagger}(g)V^{\dagger}VD(g)V^{-1} = V^{\dagger-1}D^{\dagger}(g)QD(g)V^{-1} = V^{\dagger-1}QV^{-1} = \mathbb{1}. \quad (13.23)$$

En combinant les deux propriétés ci-dessus, on en déduit le

**Corollaire 13.18 :**

Toute représentation réductible d'un groupe fini est complètement réductible.

Ce résultat se généralise à plusieurs situations d'intérêt pratique :

**Théorème 13.19 :**

Si  $G$  est un groupe de Lie *compact*, alors toute représentation réductible est complètement réductible. La propriété est également vraie pour toute représentation réductible d'un groupe de Lie connexe, non compact et semi-simple, de même que pour toute représentation réductible *unitaire* d'un groupe quelconque.

### 13.1.3 Lemme de Schur

**Lemme 13.20 :**

Soient deux représentations irréductibles  $D$  et  $D'$ . S'il existe un opérateur  $V$  tel que  $\forall g \in G$ ,  $VD(g) = D'(g)V$ , alors soit  $V$  est un isomorphisme (et alors  $V^{-1}D'V = D$  i.e.  $D$  et  $D'$  sont équivalentes), soit  $V = 0$ .

**Corollaire 13.21 :**

Soit une représentation irréductible sur  $\mathbb{C}$ , de *dimension finie*. Si  $M$  commute avec toutes les matrices de cette représentation, alors  $M = \lambda\mathbb{1}$  ( $M$  est une homothétie).

*Preuve :*

◇ Commençons par démontrer le corollaire. Par hypothèse,  $\forall g \in G$ ,  $D(g)M = MD(g)$ .

Soit  $E_{\lambda}$  le sous-espace propre de  $M$  pour la valeur propre  $\lambda$  (qui existe car le corps est  $\mathbb{C}$ , qui est algébriquement clos, ce qui signifie que tout polynôme de degré supérieur ou égal à un possède au moins une racine).  $\forall v \in E_{\lambda}$ ,  $MD(g)v = D(g)Mv = \lambda D(g)v$ , donc  $E_{\lambda}$  est stable par  $D(g)$ ,  $\forall g$ .

---

1. Du nom d'André-Louis Cholesky, officier et ingénieur français.



Donc si  $E_\lambda \neq E$ ,  $D(g)$  est de la forme

$$D(g) = \left( \begin{array}{c|c} \text{---} & \text{---} \\ \hline 0 & \end{array} \right), \quad (13.24)$$

avec un bloc nul. La représentation est alors réductible, ce qui est absurde.

Donc  $E_\lambda = E$ , i.e.  $M = \lambda E$ .

◇ Démontrons maintenant le lemme lui-même. Il est clair que  $V = 0$  est solution. Supposons maintenant que  $V \neq 0$ .

– Noyau :  $VD(g) = D'(g)V \Rightarrow \forall x \in \text{Ker } V, VD(g)(x) = D'(g)V(x) = 0$ . Donc  $D(g)\text{Ker } V \subset \text{Ker } V$ . Ceci montre que  $\text{Ker } V$  est un sous-espace invariant de  $E$ . Comme  $D(g)$  est irréductible,  $\text{Ker } V = 0$  (sinon  $\text{Ker } V = E$ , et alors  $V = 0$ , ce qui est absurde).

– Image :  $\forall y \in \text{Im } V, \exists x$  t.q.  $y = V(x)$  donc  $D'(g)(y) = D'(g)V(x) = VD(g)(x)$ .

Ainsi  $D'(g)(y) = V(y')$  avec  $y' = D(g)(x)$ , i.e.  $D'(g)\text{Im } V \subset \text{Im } V$ . Ceci montre que  $\text{Im } V$  est un sous-espace invariant par  $D'(g)$  (pour tout  $g$ ). Comme  $V \neq 0, \text{Im } V \neq 0$ . L'irréductibilité de  $D'$  implique que  $\text{Im } V = E$ .

Donc  $V$  est bijective.

### Réciproque 13.22 :

Soit une représentation  $D$  de dimension finie, unitaire (ou plus généralement satisfaisant aux conditions du théorème de complète irréductibilité).

Si  $\forall M, \forall g \in G, [D(g), M] = 0 \Rightarrow M = \lambda I$ , alors  $D$  est irréductible.

On a ainsi un critère puissant pour prouver l'irréductibilité d'une représentation.

*Preuve :*

Si la représentation est réductible, elle l'est complètement par hypothèse, et  $\exists H \subset E$  invariant par  $D(g), \forall g \in G$  tel que

$$\forall g, D(g) = \left( \begin{array}{c|c} \text{---} & 0 \\ \hline 0 & \text{---} \end{array} \right).$$

Les matrices qui commutent avec les  $D(g)$  sont de la forme  $\left( \begin{array}{c|c} \lambda_1 I & 0 \\ \hline \text{---} & \lambda_2 I \end{array} \right)$  qui diffère de  $\lambda I$  en général. Donc la représentation ne peut pas être réductible.

*Autre preuve du corollaire :* sur  $\mathbb{C}$ ,  $M$  a au moins une valeur propre  $\lambda$  ( $\lambda \neq 0$  car  $M$  est inversible d'après le lemme de Schur).

$$MD = DM \text{ donc } (M - \lambda \mathbf{1})D = D(M - \lambda \mathbf{1}).$$

Mais comme  $M - \lambda \mathbf{1}$  est singulier,  $M = \lambda \mathbf{1}$  d'après le lemme de Schur.

**Corollaire 13.23 :**

Une représentation irréductible sur  $\mathbb{C}$  d'un groupe abélien est nécessairement de dimension 1.

*Preuve :*

Soit  $g' \in G$ . Alors  $[D(g'), D(g)] = 0, \forall g \in G$ . Donc  $D(g') = \lambda(g')\mathbf{1}$ . La représentation se décompose donc en  $\dim D$  copies de la représentation de dimension 1 :  $g \mapsto \lambda(g)$ .

*Remarque :* Il est essentiel que le corps soit  $\mathbb{C}$  (algébriquement clos) et non  $\mathbb{R}$ .

Contre-exemple : la représentation sur  $\mathbb{R}$  de  $SO(2)$  par  $D(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$  est irréductible (pourtant  $\forall \alpha, [D(\alpha), D(\theta)] = 0$ ).

## 13.2 Produit tensoriel de représentations

Plusieurs méthodes existent pour fabriquer des représentations irréductibles de dimensions élevées à partir de représentations irréductibles déjà connues de basses dimensions, qui jouent en quelque sorte le rôle de briques de base. La plus immédiate consiste à faire le produit tensoriel de telles représentations, et à décomposer le résultat en représentations irréductibles.

Par ailleurs, d'un point de vue physique, cette situation est particulièrement importante puisqu'on la rencontre par exemple en mécanique quantique, où l'on est notamment amené à examiner comment le moment cinétique total d'un système possédant un moment cinétique orbital et un moment cinétique de spin est quantifié.

### 13.2.1 Produit tensoriel de représentations

**Proposition 13.24 :** *Produit tensoriel*

Soient  $E$  et  $F$  deux espaces vectoriels, de dimensions arbitraires (finies ou infinies). L'espace vectoriel produit tensoriel  $E \otimes F$  est par définition l'espace engendré par les vecteurs de base constitués des produits tensoriels deux à deux des vecteurs de base de  $E$  et  $F$ . Si  $(e_i)_{i \in I}$  et  $(f_j)_{j \in J}$  sont des bases respectives de  $E$  et  $F$ , alors  $(e_i \otimes f_j)_{(i,j) \in I \times J}$  est une base de  $E \otimes F$ . En particulier, si  $E$  et  $F$  sont de dimensions finies,

$$\dim(E \otimes F) = \dim(E) \times \dim(F). \quad (13.25)$$

**Définition 13.25 :** *Produit tensoriel de deux représentations*

Considérons deux représentations  $D_1$  et  $D_2$  d'un groupe<sup>2</sup>  $G$ , agissant respectivement sur les espaces vectoriels  $E_1$  et  $E_2$ . L'espace  $E = E_1 \otimes E_2$  porte alors une représentation produit tensoriel de  $D_1$  et  $D_2$ , notée  $D = D_1 \otimes D_2$  qui agit de la façon suivante sur  $E_1 \otimes E_2$  :

$$\begin{aligned} \forall g \in G, \forall v_1 \in E_1, \forall v_2 \in E_2, \\ D(g)(v_1 \otimes v_2) = (D_1 \otimes D_2)(g)(v_1 \otimes v_2) = D_1(g)v_1 \otimes D_2(g)v_2. \end{aligned} \quad (13.26)$$

Dans le cas où  $E_1$  et  $E_2$  sont de dimensions finies, et où  $D_1$  et  $D_2$  ont pour matrices représentatives  $\mathcal{D}_1$  et  $\mathcal{D}_2$ ,  $D$  a pour matrice représentative  $\mathcal{D}$ , caractérisées par les éléments de matrice

$$\mathcal{D}_{n_1 n_2; n'_1 n'_2} = \langle n_1 n_2 | D(g) | n'_1 n'_2 \rangle = \langle n_1 | D_1(g) | n'_1 \rangle \langle n_2 | D_2(g) | n'_2 \rangle = \mathcal{D}_{n_1 n'_1} \mathcal{D}_{n_2 n'_2}. \quad (13.27)$$

**Propriété 13.26 :** *Caractère d'un produit tensoriel*

Le caractère du produit tensoriel de deux représentations est égal au produit de leur caractère.

*Preuve :*

Ceci est immédiat d'après la relation (13.27). En particulier, en faisant  $g = e$  on retrouve le fait que

$$\dim E = \dim(E_1 \otimes E_2) = \dim E_1 \times \dim E_2. \quad (13.28)$$

### 13.2.2 Décomposition et coefficients de Clebsch-Gordan

En général, le produit direct de deux représentations  $D_1$  et  $D_2$  n'est pas irréductible. Dans le cas où elle est complètement réductible, on écrira ce produit sous la forme d'une somme de représentation irréductible, appelée décomposition de Clebsch-Gordan, sous la forme

$$D_1 \otimes D_2 = \oplus_j D_j \quad (13.29)$$

où les  $D_j$  sont des représentations irréductibles, décomposition dans laquelle des représentations équivalentes peuvent apparaître plusieurs fois. Si les différentes représentations irréductibles inéquivalentes sont indexées par un indice  $\rho$ , on peut réécrire cette décomposition en faisant apparaître la multiplicité de chacune de ces représentations inéquivalentes  $D^{(\rho)}$ , sous la forme

$$D_1 \otimes D_2 = \oplus_\rho m_\rho D^{(\rho)} \quad (13.30)$$

On tire immédiatement des relations (13.29) et (13.30) les

**Propriétés 13.27 :**

$$\chi_{D_1} \chi_{D_2} = \sum_j \chi_j = \sum_\rho m_\rho \chi^{(\rho)} \quad (13.31)$$

---

2. On pourrait plus généralement considérer le produit tensoriel de deux représentations  $D_1$  et  $D_2$  respectivement de deux groupes  $G_1$  et  $G_2$  agissant sur  $E_1$  et  $E_2$ .

et

$$\dim D_1 \dim D_2 = \sum_j \dim D_j = \sum_\rho m_\rho \dim D^{(\rho)}. \quad (13.32)$$

**Exemples :**

1)  $SO(3)$  :

Représentons ce groupe par son action sur l'espace euclidien à 3 dimensions, et considérons le produit tensoriel de deux copies de cet espace euclidien à 3 dimensions. Partant de deux vecteurs  $\vec{x}$  et  $\vec{y}$ , on peut former le tenseur à deux indices  $x_i x_j$ , qui peut se décomposer en la somme de 3 structures irréductibles :

◇ le produit scalaire  $\vec{x} \cdot \vec{y}$ , invariant sous  $SO(3)$  (c'est la représentation triviale)

◇ un tenseur antisymétrique à deux indices :

$$A_{ij} = x_i y_j - x_j y_i \quad (13.33)$$

correspondant à une représentation irréductible de dimension 3, de spin 1 (voir le chapitre 14)

◇ un tenseur symétrique de trace nulle à deux indices :

$$S_{ij} = x_i y_j + x_j y_i - \frac{2}{3} \vec{x} \cdot \vec{y} \delta_{ij} \quad (13.34)$$

correspondant à une représentation irréductible de dimension 5, de spin 2 (voir le chapitre 14).

Nous verrons au chapitre 14 comment obtenir de façon systématique une telle décomposition.

On vérifie explicitement que la dimension de la représentation produit tensoriel vérifie bien  $3 \times 3 = 9 = 1 + 3 + 5$ . En utilisant la dimension des représentations comme étiquette, on a la décomposition

$$D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus D^{(3)} \oplus D^{(5)}. \quad (13.35)$$

2)  $SO(n)$  :

La décomposition ci-dessus reste valable pour  $SO(n)$  : partant d'une matrice arbitraire  $M_{ij}$  agissant sur l'espace  $n \otimes n$ , on peut toujours écrire la décomposition

$$M_{ij} = \frac{1}{2} \left( M_{ij} + M_{ji} - \frac{2}{n} \delta_{ij} \delta^{pq} M_{pq} \right) + \frac{1}{2} (M_{ij} - M_{ji}) + \frac{1}{n} \delta_{ij} \delta^{pq} M_{pq} \quad (13.36)$$

qui montre que

$$D^{(n)} \otimes D^{(n)} = D^{(1)} \oplus D^{(n(n-1)/2)} \oplus D^{(n(n+1)/2-1)}, \quad (13.37)$$

puisque l'espace des tenseurs de rang 2 antisymétriques est de dimension  $n(n-1)/2$ , l'espace des tenseurs de rang 2 symétriques de trace nulle est de dimension  $n(n+1)/2-1$ , et la trace correspond à la représentation triviale. Ces trois tenseurs correspondent bien à des représentations

irréductibles de  $SO(n)$  puisque les seules opérations algébriques possibles pouvant permettre de réduire le rang sont la symétrisation ou l'antisymétrisation, et la contraction sur deux indices (qui donne la trace du tenseur), opérations qui sont justement celles qui ont été mises en œuvre ici.

Nous allons à présent donner le cadre général permettant de formaliser la décomposition d'un produit tensoriel de vecteurs de base de deux représentations irréductibles en la somme de vecteurs de base de représentations irréductibles.

**Définition 13.28 :** *Coefficients de Clebsch-Gordan*

Soit  $\Psi_{\alpha, \alpha \in \{1, \dots, \dim D^{(\rho)}\}}^{(\rho)}$  une base de vecteurs de la représentation  $D^{(\rho)}$ . On peut alors écrire le produit de deux vecteurs sous la forme générique

$$\Psi_{\alpha}^{(\rho)} \Psi_{\beta}^{(\sigma)} = \sum_{\tau, \gamma, i} C_{\rho, \alpha; \sigma, \beta | \tau_i, \gamma} \Psi_{\gamma}^{(\tau_i)}. \quad (13.38)$$

Les coefficients  $C_{\rho, \alpha; \sigma, \beta | \tau_i, \gamma}$  sont appelés coefficients de Clebsch-Gordan.

Précisons la signification des différents indices :

- ◇  $\rho$  et  $\sigma$  sont des étiquettes pour les deux représentations initiales,  $\tau$  est une étiquette pour la représentation finale.
- ◇ comme chaque représentation  $\tau$  peut apparaître  $m_{\tau}$  fois, il est nécessaire d'introduire un indice supplémentaire  $i$  variant de 1 à  $m_{\tau}$ .
- ◇  $\gamma$  est une étiquette pour le vecteur de la représentation finale  $\tau_i$

Il est d'usage d'utiliser des notations de mécanique quantique, et d'écrire cette décomposition sous la forme

$$|\rho, \alpha; \sigma, \beta\rangle = |\rho\alpha\rangle |\sigma\beta\rangle = \sum_{\tau, \gamma, i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle |\tau_i \gamma\rangle. \quad (13.39)$$

Alors que les coefficients  $m_{\rho}$  sont entiers, les coefficients de Clebsch-Gordan sont a priori des nombres complexes.

### Cas où les représentations sont unitaires

Dans le cas de représentations unitaires, en supposant que les bases sont orthonormées, les coefficients de Clebsch-Gordan sont simplement des coefficients de changement de base qui satisfont à deux relations d'orthonormalité et de complétude :

- l'orthonormalité, qui s'écrit ici

$$\langle \rho, \alpha'; \sigma, \beta' | \rho, \alpha; \sigma, \beta \rangle = \delta_{\alpha, \alpha'} \delta_{\beta, \beta'} \quad (13.40)$$

et donc

$$\sum_{\tau, \gamma, i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau_i \gamma | \rho, \alpha'; \sigma, \beta' \rangle^* = \delta_{\alpha, \alpha'} \delta_{\beta, \beta'} \quad (13.41)$$

- la complétude, qui s'écrit

$$\sum_{\alpha, \beta} |\rho, \alpha; \sigma, \beta\rangle \langle \rho, \alpha; \sigma, \beta| = \mathbf{1} \quad (13.42)$$

soit encore

$$\sum_{\alpha, \beta} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta\rangle \langle \tau'_j \gamma' | \rho, \alpha'; \sigma, \beta'\rangle^* = \delta_{\tau, \tau'} \delta_{\gamma, \gamma'} \delta_{i, j}. \quad (13.43)$$

La relation (13.39) s'inverse alors en

$$|\tau_i \gamma\rangle = \sum_{\alpha, \beta} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta\rangle^* |\rho, \alpha; \sigma, \beta\rangle = \sum_{\alpha, \beta} \langle \rho, \alpha; \sigma, \beta | \tau_i \gamma\rangle |\rho, \alpha; \sigma, \beta\rangle \quad (13.44)$$

*Produit de matrices représentatives :*

En insérant  $D^{(\rho)} \otimes D^{(\sigma)}$  dans la relation (13.39) et sa conjuguée, on obtient finalement

$$\mathcal{D}_{\alpha\alpha'}^{(\rho)} \mathcal{D}_{\beta\beta'}^{(\sigma)} = \sum_{\tau, \gamma, \gamma', i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta\rangle^* \langle \tau'_i \gamma' | \rho, \alpha'; \sigma, \beta'\rangle \mathcal{D}_{\gamma\gamma'}^{(\tau_i)}. \quad (13.45)$$

## 13.3 Représentation régulière

### 13.3.1 Généralités

Une représentation joue un rôle très important, aussi bien pour les groupes finis que pour les groupes de Lie (nous la rencontrerons au chapitre 17 dans l'étude du groupe de Lorentz).

**Définition 13.29 :**

Soit  $E$  le  $\mathbb{K}$ -espace vectoriel des fonctions de  $G$  dans  $\mathbb{K}$ . L'action de  $G$  sur cet espace définit la représentation régulière (à gauche<sup>3</sup>) de la façon suivante :

$$\forall g \in G, D^{\text{reg}}(g) : E \rightarrow E \quad (13.46)$$

$$f \mapsto D^{\text{reg}}(g)f = gf \quad (13.47)$$

$$\text{où } gf \text{ est définie par } (gf)(x) = f(g^{-1}x). \quad (13.48)$$

*Preuve :*

Le passage à l'inverse permet de s'assurer de la propriété de morphisme :

$$(g'(gf))(x) = (gf)(g'^{-1}x) = f(g^{-1}g'^{-1}x) = f((g'g)^{-1}x) = ((g'g)f)(x) \quad (13.49)$$

et donc  $D^{\text{reg}}(g')D^{\text{reg}}(g) = D^{\text{reg}}(g'g)$ , *q.e.d.*

**Propriété 13.30 :**

La représentation régulière est fidèle.

---

3. On définit de la même façon l'action à droite par  $(D_d^{\text{reg}}(g)f)(x) = f(xg)$ , qui est bien une représentation puisque  $(D_d^{\text{reg}}(g')(D_d^{\text{reg}}(g)f)(x) = (D_d^{\text{reg}}(g)f)(xg') = f(xg'g) = (D_d^{\text{reg}}(g'g)f)(x)$ , ce qui montre que  $D_d^{\text{reg}}(g'g) = D_d^{\text{reg}}(g')D_d^{\text{reg}}(g)$ .

### 13.3.2 Représentation régulière d'un groupe fini

Dans le cas d'un groupe fini, l'espace  $E$  des fonctions de  $G$  sur  $\mathbb{K}$  est de dimension  $n$ , l'ordre du groupe.

Cela se vérifie en considérant la base canonique de  $E$  constituée des  $n$  fonctions  $\delta_g$  définies par

$$\delta_g(x) = \delta_{g,x}. \quad (13.50)$$

Cette famille est libre. En effet, soit  $\{\lambda_k\}$  telle que  $\sum_g \lambda_g \delta_g = 0$ . Alors  $\forall x \in G$ ,  $\sum_g \lambda_g \delta_g(x) = 0$  implique que  $\sum_g \lambda_g \delta_{g,x} = 0$  soit  $\lambda_g = 0, \forall g \in G$ .

Dans cette base, la représentation régulière agit comme une permutation (de la même façon que la translation à gauche agit sur  $G$  comme une permutation, voir la preuve du théorème de Cayley 4.60) :

$$g' \delta_g(x) = \delta_g(g'^{-1}x) = \delta_{g,g'^{-1}x} = \delta_{g'g}(x), \quad (13.51)$$

et donc  $g' \delta_g = \delta_{g'g}$  comme annoncé.

Le caractère de la représentation régulière a une forme très simple. En effet, remarquons tout d'abord que les éléments diagonaux de la matrice  $\mathcal{D}^{(\text{reg})}(g')$  sont tous nuls dans le cas où  $g' \neq e$ , et tous égaux à 1 si  $g' = e$ .

En prenant la trace, on en déduit donc que

$$\chi^{(\text{reg})}(g) = n \delta_{g,e}. \quad (13.52)$$

## 13.4 Représentations des groupes finis

On s'intéresse maintenant aux représentations des groupes finis sur le corps  $\mathbb{C}$ . Le point clef sera de pouvoir sommer sur les éléments du groupe. Les résultats de cette section pourront donc s'étendre aux groupes infinis s'il est possible de remplacer cette somme par une intégrale sur les éléments du groupe. Ce sera le cas s'il existe une mesure de Haar sur le groupe. Ceci est assuré pour les groupes localement compacts, ce qui est le cas des groupes de Lie, mais cette mesure ne sera normalisable sur l'ensemble du groupe que pour les groupes compacts.

### 13.4.1 Orthogonalité et complétude

Considérons un groupe fini  $G$  d'ordre  $n$ . On note  $D^{(\rho)}$  les représentations irréductible inéquivalentes, et  $n_\rho$  leur dimension. D'après la propriété 13.17, on peut supposer les matrices  $\mathcal{D}_{\alpha\beta}^{(\rho)}$  unitaires.

**Théorème 13.31 :** *Orthogonalité des matrices représentatives*

Les matrices représentatives  $\mathcal{D}_{\alpha\beta}^{(\rho)}$  satisfont les propriétés d'orthogonalité

$$\boxed{\frac{1}{n} \sum_g \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha'\beta'}^{(\rho)*}(g) = \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'}} \quad (13.53)$$

d'où l'on déduit le

**Théorème 13.32 :** *Orthogonalité des caractères*

$$\boxed{\frac{1}{n} \sum_g \chi^{(\rho)}(g) \chi^{(\rho')^*}(g) = \delta_{\rho\rho'}} \quad (13.54)$$

*Preuves :*

La preuve la plus directe repose sur le lemme de Schur. Partant d'une matrice arbitraire  $M$ , de dimension  $n_\rho \times n_{\rho'}$ , formons la matrice  $V$  définie par

$$V = \frac{1}{n} \sum_{g'} \mathcal{D}^{(\rho)}(g') M \mathcal{D}^{(\rho')\dagger}(g'). \quad (13.55)$$

Elle vérifie la relation

$$\forall g \in G, V \mathcal{D}^{(\rho')}(g) = \mathcal{D}^{(\rho)}(g) V. \quad (13.56)$$

En effet,

$$\begin{aligned} \mathcal{D}^{(\rho)}(g) V &= \frac{1}{n} \sum_{g'} \mathcal{D}^{(\rho)}(g) \mathcal{D}^{(\rho)}(g') M \mathcal{D}^{(\rho')\dagger}(g') = \frac{1}{n} \sum_{g'} \mathcal{D}^{(\rho)}(gg') M [\mathcal{D}^{(\rho')}(g^{-1}) \mathcal{D}^{(\rho')}(gg')]^\dagger \\ &= \frac{1}{n} \sum_{g'} \mathcal{D}^{(\rho)}(gg') M \mathcal{D}^{(\rho')\dagger}(gg') \mathcal{D}^{(\rho')}(g) = \frac{1}{n} \sum_{g'} \mathcal{D}^{(\rho)}(g') M \mathcal{D}^{(\rho')\dagger}(g') \mathcal{D}^{(\rho')}(g) \\ &= V \mathcal{D}^{(\rho')}(g) \end{aligned} \quad (13.57)$$

où l'on a utilisé pour la troisième égalité le fait que les représentations sont unitaires, et donc que  $\mathcal{D}^\dagger(g) = \mathcal{D}(g^{-1})$ , et pour la quatrième égalité le lemme de réarrangement.

D'après le lemme de Schur,  $V$  est donc nulle si  $\rho \neq \rho'$ , et multiple de l'identité si  $\rho = \rho'$ .

Choisissons  $M$  dont les éléments de matrices sont définis par  $M_{\alpha_1\alpha_2} = \delta_{\alpha_1\beta} \delta_{\alpha_2\beta'}$ .

- ◇ Dans le cas où  $\rho \neq \rho'$ , il suffit de déterminer l'élément de matrice  $V_{\alpha\alpha'}$  pour obtenir la relation (13.53).
- ◇ Dans le cas où  $\rho = \rho'$ ,  $V = c\mathbf{1}$ . En prenant alors la trace de  $V$ , et en utilisant la relation (13.55) on obtient par cyclicité de la trace (et en tenant compte de la somme sur les  $n$  éléments  $g'$  qui donne un facteur  $n$ ) l'égalité

$$cn_\rho = \text{Tr} M = \delta_{\beta\beta'} \quad (13.58)$$

soit  $c = \frac{1}{n_\rho} \delta_{\beta\beta'}$ , ce qui conduit à la relation (13.53).

La relation satisfaite par les caractères est immédiate en prenant simultanément les traces de  $\mathcal{D}^{(\rho)}(g)$  et  $\mathcal{D}^{(\rho')^*}(g)$  dans la relation (13.53).



**Théorème 13.33 :**

Les dimensions  $n_\rho$  vérifient l'égalité

$$\boxed{\sum_{\rho} n_{\rho}^2 = n.} \quad (13.59)$$

*Preuve :*

Considérons la représentation régulière, qui *a priori* n'est pas irréductible, et décomposons son caractère sur les caractères irréductibles :

$$\chi^{(\text{reg})}(g) = \sum_{\rho} m_{\rho} \chi^{(\rho)}(g). \quad (13.60)$$

En utilisant la relation (13.52), on a

$$\frac{1}{n} \sum_g \chi^{(\text{reg})}(g) \chi^{(\rho)*}(g) = \chi^{(\rho)*}(e) = n_{\rho} \quad (13.61)$$

et d'autre part, en utilisant l'orthogonalité des caractères (13.54), la même somme s'écrit aussi

$$\frac{1}{n} \sum_g \chi^{(\text{reg})}(g) \chi^{(\rho)*}(g) = \sum_{\rho'} m_{\rho'} \frac{1}{n} \sum_g \chi^{(\rho')}(g) \chi^{(\rho)*}(g) = \sum_{\rho'} m_{\rho'} \delta_{\rho\rho'} = m_{\rho}, \quad (13.62)$$

d'où  $m_{\rho} = n_{\rho}$  : ceci montre d'une part que chaque représentation irréductible, caractérisée par sa dimension  $n_{\rho}$ , apparaît dans la représentation régulière, et d'autre part que cette dimension est égale à sa multiplicité dans la décomposition de cette représentation régulière. On a donc

$$\chi^{(\text{reg})}(g) = \sum_{\rho} n_{\rho} \chi^{\rho}(g). \quad (13.63)$$

Il suffit alors de faire  $g = e$  dans cette expression, et d'utiliser l'expression (13.52) du caractère de la représentation régulière pour en déduire le résultat (13.59).

**Corollaire 13.34 :**

Le nombre de représentations irréductibles d'un groupe abélien est égal à l'ordre du groupe.

*Preuve :*

Il suffit d'appliquer la relation (13.59) au cas d'un groupe abélien, pour lequel toutes les représentations irréductibles sont de dimension 1.

**Théorème 13.35 :** *Complétude des matrices représentatives*

$$\boxed{\sum_{\rho, \alpha, \beta} \frac{n_{\rho}}{n} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') = \delta_{gg'}.} \quad (13.64)$$

*Preuve :*

Ce résultat s'établit en considérant, pour une représentation  $\rho$  donnée, et pour chaque couple  $(\alpha, \beta)$  (au nombre de  $n_\rho \times n_\rho$ ), la collection des  $n$  éléments  $\sqrt{\frac{n_\rho}{n}} \mathcal{D}_{\alpha\beta}^{(\rho)}(g)$  (lorsque  $g$  parcourt tout le groupe  $G$ ) comme les composantes d'un vecteur de dimension  $n$ . La relation d'orthogonalité (13.53) montre alors que ces vecteurs sont tous orthogonaux entre eux, et de norme 1. Etant au nombre de  $n$  d'après la relation (13.59), ils forment donc une base, d'où la relation de complétude (13.64).

**Théorème 13.36 :** *Complétude des caractères*

Les caractères étant des fonctions de classe, on note  $C_i$  les classes,  $n_i$  le nombre d'éléments de chacune des classes, et  $\chi_i^{(\rho)}$  la valeur du caractère de la représentation  $\rho$  pour la classe  $C_i$ . On a alors la relation de complétude

$$\boxed{\frac{n_i}{n} \sum_{\rho} \chi_i^{(\rho)} \chi_j^{(\rho)*} = \delta_{ij}.} \quad (13.65)$$

*Preuve :*

La preuve est basée sur le lemme de Schur. Considérons l'opérateur

$$X_i^{(\rho)} = \sum_{g' \in C_i} D^{(\rho)}(g'). \quad (13.66)$$

Il vérifie la relation

$$D^{(\rho)}(g) X_i^{(\rho)} = X_i^{(\rho)} D^{(\rho)}(g). \quad (13.67)$$

En effet,

$$\begin{aligned} D^{(\rho)}(g) X_i^{(\rho)} &= \sum_{g' \in C_i} D^{(\rho)}(gg') = \sum_{g' \in C_i} D^{(\rho)}(gg'g^{-1}g) \\ &= \sum_{g_1 \in C_i} D^{(\rho)}(g_1g) = \sum_{g_1 \in C_i} D^{(\rho)}(g_1) D^{(\rho)}(g) = X_i^{(\rho)} D^{(\rho)}(g), \end{aligned} \quad (13.68)$$

où dans la deuxième égalité on a posé  $g_1 = gg'g^{-1}$  et utilisé le lemme de réarrangement. Le lemme de Schur implique donc que  $X_i^{(\rho)} = c \mathbf{1}^{(\rho)}$ . Le passage à la trace conduit à  $\text{Tr} X_i^{(\rho)} = n_i \chi_i^{(\rho)} = c \text{Tr} \mathbf{1}^{(\rho)} = c n_\rho$  d'où  $c = \frac{n_i}{n_\rho} \chi_i^{(\rho)}$ . On a ainsi

$$X_i^{(\rho)} = \frac{n_i}{n_\rho} \chi_i^{(\rho)} \mathbf{1}^{(\rho)}. \quad (13.69)$$

En partant de la relation de complétude (13.64) et en sommant sur les éléments des classe  $C_i$  et  $C_j$  on obtient alors,

$$\sum_{g \in C_i} \sum_{g' \in C_j} \sum_{\rho} \frac{n_\rho}{n} \text{Tr} [\mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(g')] = \sum_{\rho} \frac{n_\rho}{n} \text{Tr} [X_i^{(\rho)} X_j^{(\rho)\dagger}] = \sum_{\rho} \frac{n_i n_j}{n} \chi_i^{(\rho)} \chi_j^{(\rho)*} = n_i \delta_{ij}, \quad (13.70)$$

où l'on a utilisé l'expression (13.69) pour obtenir la deuxième égalité, et le second membre de la relation de complétude (13.64) pour la dernière égalité.

On notera que l'orthogonalité des caractères (13.54) s'écrit encore, classe par classe,

$$\boxed{\frac{1}{n} \sum_i n_i \chi_i^{(\rho)} \chi_i^{(\rho')*} = \delta_{\rho\rho'}} \quad (13.71)$$

### Théorème 13.37 :

Le nombre de représentations irréductibles d'un groupe fini est égal au nombre de classes.

*Preuve :*

Ceci résulte simplement du fait que la dimension de l'espace des  $\hat{\chi}$ , de composantes  $\hat{\chi}_i^{(\rho)} = \sqrt{\frac{n_i}{n}} \chi_i^{(\rho)}$ , est égale au nombre de ces vecteurs (le nombre d'indices  $\rho$ ), lui-même égal au nombre de composantes (le nombre d'indices  $i$ ).

### Propriété 13.38 : Produit hermitien

L'ensemble des résultats précédents sur les caractères peut se reformuler en introduisant un produit hermitien<sup>4</sup> sur les fonctions de  $G$  dans  $\mathbb{C}$

$$\langle f_1 | f_2 \rangle = \frac{1}{n} \sum_{g \in G} f_1(g)^* f_2(g). \quad (13.72)$$

Les caractères irréductibles forment, pour ce produit hermitien, une base orthonormée de l'espace des fonctions centrales :

$$\langle \chi^{(\rho')} | \chi^{(\rho)} \rangle = \delta_{\rho\rho'}, \quad (13.73)$$

$$\sum_{\rho} |\chi^{(\rho)} \rangle \langle \chi^{(\rho)}| = \mathbf{1}. \quad (13.74)$$

## 13.4.2 Applications

### Calcul des multiplicités

Le caractère d'une décomposition arbitraire se décompose sur la base des caractères irréductibles, suivant

$$\chi(g) = \sum_{\rho} m_{\rho} \chi^{(\rho)}(g), \quad (13.75)$$

les multiplicités  $m_{\rho}$  se calculant grâce à la relation

$$m_{\rho} = \frac{1}{n} \sum_i n_i \chi_i \chi_i^{(\rho)*} = \frac{1}{n} \sum_g \chi(g) \chi^{(\rho)*}(g) = \langle \chi^{(\rho)} | \chi \rangle. \quad (13.76)$$

---

4. On utilise ici la convention bra et ket de la physique, *i.e.* linéarité sur  $f_2$  et antilinéarité en  $f_1$  ; noter que dans les ouvrages de mathématiques, la convention est généralement inverse.

On pourra aussi écrire la décomposition (13.75) de façon équivalente :

$$|\chi\rangle = \sum_{\rho} |\chi^{(\rho)}\rangle \langle \chi^{(\rho)}|\chi\rangle . \quad (13.77)$$

### Indicatrice et fonctions de classe

Plus généralement, toute fonction de classe se décompose sur la base des caractères.

#### Définition 13.39 : Indicatrice

L'indicatrice  $P_i$  de la classe  $C_i$ , qui prend la valeur 1 pour  $g \in C_i$  et 0, s'écrit sous la forme

$$P_i(g) = \frac{n_i}{n} \sum_{\rho} \chi^{(\rho)}(g) \chi_i^{(\rho)*} \quad (13.78)$$

d'après la relation (13.65).

Toute fonction de classe  $f$  se décompose alors suivant l'expression

$$f(g) = \sum_i f_i P_i(g) = \frac{1}{n} \sum_{\rho} \sum_i f_i n_i \chi_i^{(\rho)*} \chi^{(\rho)}(g) . \quad (13.79)$$

### Décomposition d'un produit direct de deux représentations

Partons de la décomposition

$$\mathcal{D}^{(\rho)} \otimes \mathcal{D}^{(\sigma)} = \bigoplus_{\tau} m_{\tau} \mathcal{D}^{(\tau)} \quad (13.80)$$

ce qui se traduit au niveau des caractères par

$$\chi^{(\rho)} \chi^{(\sigma)} = \sum_{\tau} m_{\tau} \chi^{(\tau)} . \quad (13.81)$$

La multiplicité  $m_{\tau}$  vaut donc

$$m_{\tau} = \frac{1}{n} \sum_i n_i \chi_i^{(\rho)} \chi_i^{(\sigma)} \chi_i^{(\tau)*} , \quad (13.82)$$

relation identique à celle permettant d'obtenir la multiplicité de la représentation  $\sigma^*$  dans le produit  $\rho \otimes \tau^*$ . Ainsi la multiplicité de  $\tau$  dans  $\rho \otimes \sigma$  est la même que celle de  $\sigma^*$  dans  $\rho \otimes \tau^*$ .

*Remarque :*

La relation (13.81) permet de munir l'ensemble des caractères d'une structure d'algèbre.

#### Propriété 13.40 : Norme des caractères

La norme d'un caractère arbitraire s'écrit

$$\|\chi\|^2 = \langle \chi|\chi\rangle = \frac{1}{n} \sum_{g \in G} |\chi(g)|^2 = \sum_{\rho} m_{\rho}^2 . \quad (13.83)$$

*Preuve :*

Il suffit d'utiliser la relation (13.75), ce qui donne

$$\frac{1}{n} \sum_{g \in G} |\chi(g)|^2 = \frac{1}{n} \sum_{g \in G} \left[ \sum_{\rho} m_{\rho} \chi^{(\rho)} \right] \left[ \sum_{\rho'} m_{\rho'} \chi^{(\rho')} \right] = \sum_{\rho} m_{\rho}^2, \quad (13.84)$$

où l'on a utilisé la propriété d'orthogonalité des caractères (13.54).

La norme d'un caractère est donc un entier supérieur ou égal à un. Il est égal à un si et seulement si la représentation est irréductible.

## 13.5 Représentation des algèbres et groupes de Lie

### 13.5.1 Représentation unitaires d'un groupe de Lie

Examinons sous quelle condition le résultat important que nous avons démontré pour les groupes finis, à savoir que toutes les représentations sont équivalentes à des représentations unitaires, reste valable pour les groupes de Lie. Le point clef est de pouvoir remplacer la somme sur les éléments du groupe par une intégration sur les éléments, à l'aide d'une mesure qui doit vérifier une propriété d'invariance.

#### Définition 13.41 :

Une mesure sur un groupe  $G$  invariante à gauche est une mesure qui vérifie

$$\forall g_1 \in G, \forall g \in G, \mu(g_1 g) = \mu(g). \quad (13.85)$$

De même une mesure sur un groupe  $G$  invariante à droite est une mesure qui vérifie

$$\forall g_1 \in G, \forall g \in G, \mu(g g_1) = \mu(g). \quad (13.86)$$

#### Théorème 13.42 : *Mesure de Haar*

Sur un groupe topologique  $G$ , il existe une unique (à une constante multiplicative près) mesure invariante à gauche, qui est finie sur toute partie compacte du groupe. Cette mesure est appelée mesure de Haar. Il existe également une unique (à une constante multiplicative près) mesure invariante à droite.

En général, ces deux mesures ne sont pas identiques. Ce sera le cas dans de nombreuses situations d'intérêt pratique, comme l'énonce le théorème suivant.

#### Théorème 13.43 : *Mesure de Haar invariante à gauche et à droite*

Dans le cas où le groupe est abélien, ou compact, ou discret (exemple : groupes finis), ou un groupe de Lie semi-simple, la mesure de Haar est à la fois invariante à gauche et à droite.

#### Théorème 13.44 : *Mesure de Haar sur un groupe de Lie compact*

Sur un groupe de Lie compact, la mesure de Haar du groupe est finie.

Ceci est une conséquence immédiate de la définition de la mesure de Haar. C'est le pendant du fait que le nombre d'éléments d'un groupe fini est fini(!).

**Théorème 13.45 :** *Représentation d'un groupe de Lie compact*

Si  $G$  est un groupe de Lie compact, alors toute représentation de  $G$  est équivalente à une représentation unitaire. Elle est complètement réductible en une somme de représentations irréductibles. De plus toute représentation irréductible est de dimension finie.

*Preuve :*

L'existence d'une mesure de Haar finie sur le groupe permet de mener l'ensemble des raisonnements menés au théorème 13.17, et donc d'arriver au résultat, en remplaçant les sommes par des intégrales. La complète réductibilité a été établie au théorème 13.16. On admettra le fait que ces représentations irréductibles soient de dimension finie. On verra au chapitre 14 une illustration de ces résultats pour le groupe  $SU(2)$ .

Dans le cas où le groupe n'est pas compact, on a en revanche le

**Théorème 13.46 :** *Représentation d'un groupe de Lie non compact*

Si  $G$  est un groupe de Lie simple non compact, alors  $G$  ne possède pas de représentation unitaire de dimension finie en dehors de la représentation triviale  $D(A) = 1, \forall A \in G$ .

### 13.5.2 Lien entre représentation des algèbres et groupes de Lie

**Définition 13.47 :** *Représentation d'une algèbre de Lie*

Soit  $V$  un espace vectoriel sur  $\mathbb{R}$  ou  $\mathbb{C}$ . On dit que  $D$  est une représentation linéaire de  $\mathfrak{g}$  dans l'espace vectoriel  $V$  si

$$\begin{aligned} [D(A), D(B)] &= D([A, B]) \\ D(\alpha A + \beta B) &= \alpha D(A) + \beta D(B). \end{aligned} \quad (13.87)$$

**Théorème 13.48 :**

Une représentation  $D_G$  d'un groupe  $G$  induit par différentiation une représentation  $D_A$  de son algèbre de Lie.

*Preuve :*

Il suffit de définir  $D_A(T_a)$  par  $D_G(1 + tT_a + o(t)) = 1 + tD_A(T_a) + o(t)$ , ce qui est équivalent à poser que

$$D_A(T_a) = \left. \frac{d}{dt} D_G(e^{tT_a}) \right|_{t=0}.$$

Pour montrer que l'on a ainsi construit une représentation de  $\mathfrak{g}$ , considérons la courbe

$$C(t) = A(\sqrt{t})B(\sqrt{t})A(\sqrt{t})^{-1}B(\sqrt{t})^{-1} = 1 + t[T_a, T_b] + o(t),$$

où  $A(t)$  et  $B(t)$  sont des courbes analytiques de  $G$  de vecteurs tangents respectifs  $T_a$  et  $T_b$ . On a donc  $D_G(C(t)) = D_G(1 + t[T_a, T_b] + o(t))$ , soit encore

$$D_A([T_a, T_b]) = \left. \frac{d}{dt} D_G(C(t)) \right|_{t=0}.$$

Or par morphisme de  $D_G$ ,

$$D_G(C(t)) = D_G(A(\sqrt{t})) D_G(B(\sqrt{t})) D_G^{-1}(A(\sqrt{t})) D_G^{-1}(B(\sqrt{t}))$$

qui s'écrit encore, après développement en série,

$$D_G(C(t)) = 1 + t[D_A(T_a), D_A(T_b)] + o(t),$$

d'où

$$\left. \frac{d}{dt} D_G(C(t)) \right|_{t=0} = [D_A(T_a), D_A(T_b)],$$

ce qui prouve que  $D_A([T_a, T_b]) = [D_A(T_a), D_A(T_b)]$ .

*Remarque :*

On notera le fait très important suivant, qui justifie la définition (13.87) d'une représentation d'une algèbre de Lie : les relations de commutation  $(T_a, T_b) = if_{ab}^c T_c$  des générateurs de l'algèbre de Lie sont préservées au niveau de sa représentation :

$$\boxed{[D(T_a), D(T_b)] = if_{ab}^c D(T_c)}. \quad (13.88)$$

### 13.5.3 Exemples importants de représentations

#### Représentation fondamentale

Pour un groupe de matrices sous-groupe de  $GL(N, \mathbb{R})$  (ou de  $GL(N, \mathbb{C})$ ), elle est définie par  $D(A) = A$ . L'espace vectoriel de la représentation est donc  $\mathbb{R}^N$  (ou  $\mathbb{C}^N$ ), et sa dimension est  $N$ , qu'il ne faut pas confondre avec la dimension du groupe. Cette relation définit aussi bien la représentation fondamentale du groupe que de son algèbre.

#### Représentation adjointe

◇ Sur le groupe :

Sur le groupe  $G$ , l'application  $g \mapsto ad\ g$  définie par  $ad\ g(h) = g h g^{-1}$  est une représentation de  $G$  sur lui-même.

*Preuve :*

$$\begin{aligned} \forall g_1, g_2 \in G, \forall h \in G, ad(g_1 g_2)(h) &= (g_1 g_2) h (g_1 g_2)^{-1} = g_1 (g_2 h g_2^{-1}) g_1^{-1} \\ &= ad_{g_1}(ad_2(h)) \end{aligned}$$

ce qui prouve que l'on a bien construit un morphisme.





Remarque :

- ◇ Pour une algèbre de Lie semi-simple compacte, comme on l'a vu plus haut, on pourra écrire plus simplement  $D(T_a)_{bc} = -if_{abc}$  que l'on note aussi  $(T_a^A)_{bc} = -if_{abc}$  (avec  $A$  comme « adjoint »).
- ◇ La dimension de la représentation adjointe est égale à la dimension du groupe.

En revanche les générateurs  $X_a$  peuvent être dans une représentation complètement arbitraire. La seule chose qui compte finalement est que leur nombre soit la dimension du groupe (qu'il ne faut surtout pas confondre avec la dimension  $d$  de la représentation des générateurs  $X_a$  (i.e. les  $X_a$  sont des matrices  $d \times d$ ). On obtient ainsi différentes formes isomorphes de la représentation adjointe.

### Exemple : $SU(2)$

Le cas de  $SU(2)$ , qui sera étudié en détail dans le Chap. 14), permet d'illustrer les deux représentations précédentes.

- ◇ Représentation fondamentale : matrices  $\frac{\vec{\sigma}}{2}$  (3 matrices  $2 \times 2$ )
- ◇ Représentation adjointe : c'est la représentation vectorielle que l'on peut aussi fabriquer à partir des rotations infinitésimales  $(J_k)_{ij} = -i\varepsilon_{kij} = -i\varepsilon_{ijk}$  (3 matrices  $3 \times 3$ ).

En tant que matrice,  $T_a^A$  agit comme une matrice  $n \times n$  ( $n =$  dimension du groupe). En revanche si l'on décide de décrire  $T_a^A$  comme un opérateur construit à partir d'une représentation donnée (en général la représentation fondamentale) de dimension  $d$ , alors

$$T_a^A(T_b) = [T_a, T_b] \quad T_a : \text{pour } a \text{ fixé, c'est une matrice } d \times d$$

↙  $n$  indices possibles

Notations usuelles dans la littérature :  $\text{ad} \Leftrightarrow T^A \Leftrightarrow T$  pour l'adjointe

$T^f \Leftrightarrow t$  pour la fondamentale.

Il est assez usuel d'utiliser les indices  $i, j, k$  pour les coefficients des générateurs de la (des) représentation(s) fondamentale(s), et  $a, b, c$  pour les coefficients des générateurs de la représentation adjointe.

Dans les théories de Yang-Mills, nous verrons que les champs de jauge apparaissent naturellement comme se transformant sous le groupe de jauge suivant la représentation adjointe du groupe.

### 13.5.4 Normalisation

$$g_{ab} = \text{Tr} (T_a^A T_b^A) \quad \text{par définition de } g$$

Dans le cas d'une algèbre de Lie semi-simple compacte, on a vu que l'on peut toujours choisir les générateurs pour que  $g_{ab}$  soit multiple de l'identité :

$$\text{Tr} (T_a^A T_b^A) = C(A) \delta_{ab}$$

On peut montrer que pour toute représentation  $r$ , le fait d'avoir fait ce choix pour les générateurs de la représentation adjointe conduit à ce que

$$\text{Tr} (T_a^r T_b^r) = C(r) \delta_{ab}.$$

Ceci repose sur le fait que  $\text{Tr} T_a^r T_b^r$  est une forme invariante. Or un théorème montre que la seule forme invariante (à un facteur multiplicatif près) est la forme de Killing.

*Remarque* :  $[T_a^r, T_b^r] = i f_{abc} T_c^r$  donc  $f_{abc} = -\frac{i}{C(r)} \text{Tr} \{ [T_a^r, T_b^r] T_c^r \}$

# Chapitre 14

## Les Groupes $SO(3)$ et $SU(2)$

### 14.1 Le groupe des rotations à 3 dimensions

#### 14.1.1 Définition

Par définition, le groupe des rotations  $SO(3)$  est le groupe des transformations de  $\mathbb{R}^3$  qui préservent le produit scalaire et l'orientation :

$$\diamond \det R = 1 \quad \leftarrow S$$

$$\diamond \forall x, y \quad R_{ik} x_k R_{il} y_l = x_k y_l \delta_{kl} \Leftrightarrow R_{ik} R_{il} = \delta_{kl} \Leftrightarrow R^T R = \mathbb{1} \quad \leftarrow O$$

*Remarque :*

Le groupe  $O(N)$  est défini en exigeant que  $\forall x \in \mathbb{R}^N$ , la norme de  $x$  soit invariante, ou de façon équivalente en exigeant que  $\forall x, y \in \mathbb{R}^N$ , le produit scalaire  $x \cdot y$  soit invariant.

#### 14.1.2 Forme explicite

Considérons une rotation autour de  $\vec{n}$  d'angle  $\omega$ , comme illustré dans la fig. 14.1. Dans le cas où  $\vec{r}$  est perpendiculaire à l'axe de rotation,

$$\vec{r}'_{\perp} = \vec{r}_{\perp} \cos \omega + \sin \omega \vec{n} \wedge \vec{r}_{\perp}. \quad (14.1)$$

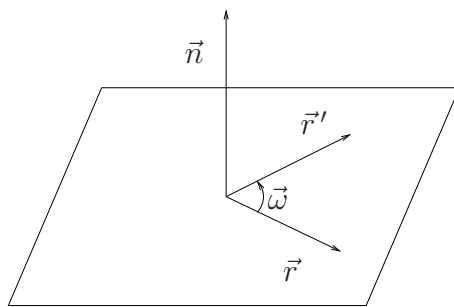


FIGURE 14.1 – Rotation d'angle  $\omega$  autour d'un axe orienté suivant le vecteur  $\vec{n}$ .

Dans le cas général, en utilisant la décomposition

$$\vec{r} = \vec{r}_{\parallel} + \vec{r}_{\perp} \quad \text{avec} \quad \vec{r}_{\parallel} = \vec{n} \cdot (\vec{r} \cdot \vec{n}) \quad \text{et} \quad \vec{r}_{\perp} = \vec{r} - \vec{n}(\vec{r} \cdot \vec{n}), \quad (14.2)$$

la rotation d'axe  $\vec{n}$  laisse invariante  $\vec{r}_{\parallel}$  et l'on a donc

$$\vec{r}' - \vec{n}(\vec{r}' \cdot \vec{n}) = [\vec{r} - \vec{n}(\vec{r} \cdot \vec{n})] \cos \omega + \sin \omega \vec{n} \wedge [\vec{r} - \vec{n}(\vec{r} \cdot \vec{n})]$$

soit

$$\boxed{\vec{r}' = \vec{n}(\vec{r}' \cdot \vec{n}) + \cos \omega [\vec{r} - \vec{n}(\vec{r} \cdot \vec{n})] + \sin \omega \vec{n} \wedge \vec{r}.} \quad (14.3)$$

Une rotation est complètement déterminée par la donnée d'un angle  $\omega$  tel que  $0 \leq \omega \leq \pi$  et d'un vecteur unitaire  $\vec{n}$ .

### 14.1.3 Topologie

Comme  $R_{\pi}(\vec{n}) = R_{\pi}(-\vec{n})$ , on en déduit que la variété du groupe est une boule de diamètre  $\pi$  dont on identifie les points diamétralement opposés. En utilisant les concepts vus au chap. 3, on montre que cette boule est homéomorphe à l'espace projectif  $P^3(\mathbb{R})$ , qui est par définition l'ensemble des droites vectorielles de  $\mathbb{R}^4$ . Ainsi

$$\boxed{SO(3) \cong P^3(\mathbb{R}).} \quad (14.4)$$

### 14.1.4 Algèbre de Lie de $SO(3)$

Le groupe  $SO(3)$  est déterminé par les conditions  $R^t R = \mathbb{1}$  avec  $\det R = 1$ . D'après le chapitre 12, l'algèbre de Lie de  $SO(N)$  est constituée des matrices antisymétriques  $N \times N$ . On prendra comme base

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad L_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (14.5)$$

soit  $(L_k)_{ij} = -\varepsilon_{ijk}$ . On vérifie que  $[L_i, L_j] = \varepsilon_{ijk} L_k$ . En terme des générateurs hermitiens,  $J_i = i L_i$ , et

$$\boxed{[J_i, J_j] = i \varepsilon_{ijk} J_k.} \quad (14.6)$$

## 14.2 Groupe $SU(2)$ : définition et paramétrisation

### 14.2.1 Définition

Le groupe de Lie  $SU(N)$  est par définition le groupe des matrices unimodulaires  $N \times N$ , i.e.

$$\begin{cases} U^\dagger U = 1 \\ \det U = 1. \end{cases}$$

Nous allons ici nous intéresser au cas particulier du groupe  $SU(2)$ .

### 14.2.2 Caractérisation

Nous allons montrer qu'une paramétrisation de  $SU(2)$  est donnée par

$$\boxed{U = \cos \frac{\theta}{2} - i \vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} = \exp \left( -i \theta \frac{\vec{n} \cdot \vec{\sigma}}{2} \right)} \quad (14.7)$$

où  $\theta$  est un angle réel,  $\vec{n}$  est vecteur tri-dimensionnel unitaire ( $\vec{n}^2 = 1$ ) et  $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$  sont les matrices de Pauli :

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (14.8)$$

On notera dans toute la suite

$$C = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (14.9)$$

#### Propriétés des matrices de Pauli

◇ Hermiticité :

$$\sigma_i^\dagger = \sigma_i. \quad (14.10)$$

◇ Produit :

$$\sigma_i \sigma_j = \delta_{ij} + i \varepsilon_{ijk} \sigma_k \quad (14.11)$$

d'où l'on déduit la relation très utile

$$\boxed{(\vec{\sigma} \cdot \vec{a})(\vec{\sigma} \cdot \vec{b}) = \vec{a} \cdot \vec{b} + i \vec{\sigma} \cdot (\vec{a} \wedge \vec{b})} \quad (14.12)$$

puisque

$$(\vec{\sigma} \cdot \vec{a})(\vec{\sigma} \cdot \vec{b}) = (\sigma_i a_i)(\sigma_j b_j) = a_i b_j (\delta_{ij} + i \varepsilon_{ijk} \sigma_k) = \vec{a} \cdot \vec{b} + i \vec{\sigma} \cdot (\vec{a} \wedge \vec{b}).$$

◇ Propriétés de  $C$  :

$$C = C^* = -C^{-1} = -C^\dagger = -C^T. \quad (14.13)$$

◇ Transposition et conjugaison complexe :

$$C^{-1} \sigma_i^T C = -\sigma_i \quad \text{et donc} \quad C^{-1} \sigma_i^* C = -\sigma_i. \quad (14.14)$$

◇ Pour toute matrice  $A$  ( $2 \times 2$ ), un calcul explicite permet de vérifier que

$$C A C^{-1} A^T = (\det A) \mathbf{1}. \quad (14.15)$$

Donc pour un élément de  $SL(2, \mathbb{C})$ , la condition  $\det U = 1$  donne alors  $CUC^{-1}U^T = \mathbf{1}$ , soit encore, en multipliant à gauche par  $C^{-1}$  et à droite par  $C$ ,  $UC^{-1}U^T C = \mathbf{1}$ , d'où

$$U^{-1} = C^{-1}U^T C. \quad (14.16)$$

Dans le cas particulier où  $U \in SU(2)$ , on aura donc

$$U^\dagger = U^{-1} = C^{-1}U^T C. \quad (14.17)$$

Venons-en maintenant à la preuve de la paramétrisation (14.7).

*Première caractérisation : exponentiation de l'algèbre de Lie*

D'après le chap. 12, la forme exponentielle s'obtient aisément si l'on admet que le groupe est compact et connexe et qu'il peut donc s'obtenir par exponentiation de son algèbre de Lie. En anticipant sur ce qui va suivre, où l'on étudiera l'algèbre de Lie de  $SU(2)$ , considérons donc la courbe de  $SU(2)$  définie par

$$t \mapsto e^{tA} \quad \text{avec} \quad e^{tA^\dagger} e^{tA} = 1.$$

On a, en différentiant en 0,  $A^\dagger + A = 1$  soit, en posant  $A = i\beta$ ,  $\beta = \beta^\dagger$ . Une base des matrices hermitiennes est fournie par les matrices  $\{\sigma_1, \sigma_2, \sigma_3\}$ . On en déduit donc que

$$U = \exp\left(-i\theta \frac{\vec{n} \cdot \vec{\sigma}}{2}\right). \quad (14.18)$$

A noter que la façon de paramétrer explicitement  $U$  à l'aide de  $\theta$  et  $\vec{n}$  est guidée par le morphisme de  $SU(2)$  sur  $SO(3)$  que nous allons étudier plus loin.

*Deuxième caractérisation : calcul direct*

Toute matrice  $2 \times 2$  à coefficients dans  $\mathbb{C}$  peut se mettre sous la forme  $U = a_0 - i\vec{a} \cdot \vec{\sigma}$  (par indépendance linéaire de  $1, \sigma_1, \sigma_2, \sigma_3$  sur le corps  $\mathbb{C}$ ). En utilisant la relation (14.17),  $U^\dagger$  s'écrit, puisque  $\sigma^\dagger = \sigma$ ,

$$U^\dagger = C^{-1}(a_0 - i\vec{a} \cdot \vec{\sigma}^T)C = a_0 + i\vec{a} \cdot \vec{\sigma} = a_0^* + i\vec{a}^* \cdot \vec{\sigma},$$

où l'on a utilisé la relation (14.14) pour la deuxième égalité, et le calcul direct de  $U^\dagger$  pour la dernière. On en déduit donc que  $a_0$  et  $\vec{a}$  sont réels. En écrivant explicitement

$$U = a_0 \mathbb{1} - i\vec{a} \cdot \vec{\sigma} = \begin{pmatrix} a_0 - ia_3 & -a_2 - ia_1 \\ a_2 - ia_1 & a_0 + ia_3 \end{pmatrix}, \quad (14.19)$$

on vérifie immédiatement que  $\det U = a_0^2 + a_3^2 + a_1^2 + a_2^2$  soit encore

$$a_0^2 + \vec{a}^2 = 1, \quad \text{i.e.} \quad (a_0, \vec{a}) \in \mathbb{S}^3. \quad (14.20)$$

On peut donc poser

$$\begin{cases} a_0 = \cos \frac{\theta}{2} \\ \vec{a} = \vec{n} \sin \frac{\theta}{2} \end{cases} \quad 0 \leq \theta \leq 2\pi. \quad (14.21)$$

En combinant cette paramétrisation avec l'expression (14.19) on a donc

$$U(\theta, \vec{n}) = \cos \frac{\theta}{2} - i\vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} = \begin{pmatrix} \cos \frac{\theta}{2} - in_3 \sin \frac{\theta}{2} & -n_2 \sin \frac{\theta}{2} - in_1 \sin \frac{\theta}{2} \\ n_2 \sin \frac{\theta}{2} - in_1 \sin \frac{\theta}{2} & \cos \frac{\theta}{2} + in_3 \sin \frac{\theta}{2} \end{pmatrix}. \quad (14.22)$$

Le groupe  $SU(2)$  est donc un groupe à 3 paramètres réels : l'angle  $\theta$  et les deux angles sphériques caractérisant le vecteur  $\vec{n}$ .

On notera en passant, comme conséquence immédiate du développement (14.19), la relation très utile

$$\boxed{\vec{x}^2 = -\det[\vec{x} \cdot \vec{\sigma}]} \quad (14.23)$$

*Lien entre les deux caractérisations*

On peut relier les deux formes obtenues dans (14.7), ce qui permettra de justifier que l'angle  $\theta$  est le même. Pour évaluer le développement en série

$$\exp\left(-i\theta\frac{\vec{n} \cdot \vec{\sigma}}{2}\right) = \sum_{k=0}^{\infty} \left(\frac{-i\theta}{2}\right)^k \frac{(\vec{n} \cdot \vec{\sigma})^k}{k!}, \quad (14.24)$$

il suffit de calculer les puissances successives de  $(\vec{n} \cdot \vec{\sigma})$ . En utilisant la relation (14.12), on obtient immédiatement que  $(\vec{n} \cdot \vec{\sigma})(\vec{n} \cdot \vec{\sigma}) = \vec{n}^2 = 1$ . Ainsi

$$\begin{aligned} \exp\left(-i\theta\frac{\vec{n} \cdot \vec{\sigma}}{2}\right) &= \sum_{k=0}^{\infty} \left(\frac{-i\theta}{2}\right)^{2k} \frac{1}{(2k)!} + \sum_{k=0}^{\infty} \left(\frac{-i\theta}{2}\right)^{2k+1} \frac{1}{(2k+1)!} \vec{n} \cdot \vec{\sigma} \\ &= \sum_{k=0}^{\infty} (-1)^k \left(\frac{\theta}{2}\right)^{2k} \frac{1}{(2k)!} - i \sum_{k=0}^{\infty} (-1)^k \left(\frac{\theta}{2}\right)^{2k+1} \frac{1}{(2k+1)!} \vec{n} \cdot \vec{\sigma} \\ &= \cos\frac{\theta}{2} - i \vec{n} \cdot \vec{\sigma} \sin\frac{\theta}{2}, \end{aligned}$$

ce qui démontre le résultat (14.7).

### 14.2.3 Topologie

Nous verrons dans la section 14.3 qu'à un élément de  $SU(2)$  correspond une rotation, caractérisée par l'angle  $\theta$  et le vecteur  $\vec{n}$ . Notons que les transformations particulières de  $SU(2)$  correspondant à  $\theta = 2\pi$  et  $\vec{n}$  quelconque correspondent toutes à l'élément  $U = -\mathbf{1}$  qui n'est pas égal à l'identité, alors que la rotation correspondante, lorsqu'elle agit sur des vecteurs de  $\mathbb{R}^3$ , est clairement égale à l'identité! Nous reviendrons plus loin sur ce paradoxe apparent.

D'après la paramétrisation (14.21), et en tenant compte de la remarque précédente, la variété du groupe  $SU(2)$  est homéomorphe à une boule de rayon  $2\pi$  de  $\mathbb{R}^4$  dont on identifie tous les points du bord (puisque ceux-ci sont caractérisés par  $\theta = 2\pi$  et  $\vec{n}$  quelconque). Elle est également homéomorphe à la sphère  $\mathbb{S}^3$ , d'après l'eq. (14.20).

On notera que l'homéomorphisme entre  $\mathbb{S}^3$  et la boule de rayon  $2\pi$  dont on identifie tous les points du bord peut se voir directement. En effet, l'application

$$\begin{aligned} [0, 2\pi] \times \mathbb{S}^2 &\rightarrow \mathbb{S}^3 \\ (\theta, \vec{n}) &\mapsto (a_0, \vec{a}) \end{aligned} \quad (14.25)$$

où  $a_0$  et  $\vec{a}$  sont donnés par l'eq. (14.21), définit un homéomorphisme de la boule ouverte  $[0, 2\pi[ \times \mathbb{S}^2 \subset \mathbb{R}^4$  sur la sphère privée d'un point  $\mathbb{S}^3 \setminus (-1, 0, 0, 0)$ . Le point  $(-1, 0, 0, 0)$  est en bijection avec le bord de la boule fermée  $[0, 2\pi] \times \mathbb{S}^2$ , ce qui établit finalement l'homéomorphisme de  $\mathbb{S}^3$ , avec la boule de rayon  $2\pi$  dont on identifie tous les points du bord. On rapprochera cette construction de la compactification d'Alexandroff, voir append. B.4.4), le point à l'infini étant ici  $(-1, 0, 0, 0)$ .

### Compacité

Il est clair que  $SU(2)$  est compacte d'après la construction précédente, puisque  $\mathbb{S}^3$  est compacte.

### Connexité

D'après ce qui précède,  $SU(2)$  est connexe par connexité de  $\mathbb{S}^3$  (alors que  $U(2)$  n'est pas connexe : elle possède deux composantes connexes, celle de  $\mathbb{1}$  qui n'est autre que  $SU(2)$ , et celle de  $\sigma_3$ , matrice unitaire de déterminant  $-1$ ).

### Simple connexité

Le groupe  $SU(2)$  est simplement connexe, puisque  $\mathbb{S}^3$  l'est : tout lacet est contractile en un point.

Ceci distingue  $SU(2)$  de  $SO(3)$ , qui a un groupe d'homotopie, ou groupe fondamental  $\pi_1(SO(3))$  (*i.e.* le groupe des classes de chemins déformables continûment (dits "homotopes", voir chap. 16 pour plus de détails) isomorphe à  $\mathbb{Z}_2$ ). Ceci peut se voir de la façon suivante. Rappelons tout d'abord que  $SO(3)$  est homéomorphe à la boule unité  $E^3$  dont on identifie les points diamétralement opposés.

Soient  $P$  et  $P'$  et  $Q$  et  $Q'$  deux couples de points diamétralement opposés de la boule unité  $E^3$ . Nous allons maintenant considérer deux lacets, illustrés dans la fig. 14.2 :

1) Le premier est constitué du segment  $c_1$  partant de  $P$  vers  $P'$ . Il correspond donc à une succession de rotation d'axe  $(PP')$ , et d'angle variant de  $-\pi$  à  $\pi$ . C'est bien un lacet dans  $SO(3)$ , par identification des points opposés, mais il n'est pas contractile à un point.

2) Le second chemin  $c_2$  est constitué d'un chemin arbitraire joignant  $P$  à  $Q$  à la surface de  $E^3$ , suivi d'un chemin joignant  $Q'$  au point  $P'$ . Par identification des points diamétralement opposés, c'est bien un lacet. Il est contractile en un lacet constant  $c'_2$  puisque les chemins  $(PQ)$  et  $(Q'P')$  sont contractiles respectivement aux chemins constants  $P$  et  $P'$ , qui sont bien diamétralement opposés, donc identifiés à un seul point, ce qui n'est autre qu'un lacet constant (l'identité dans le groupe fondamental).

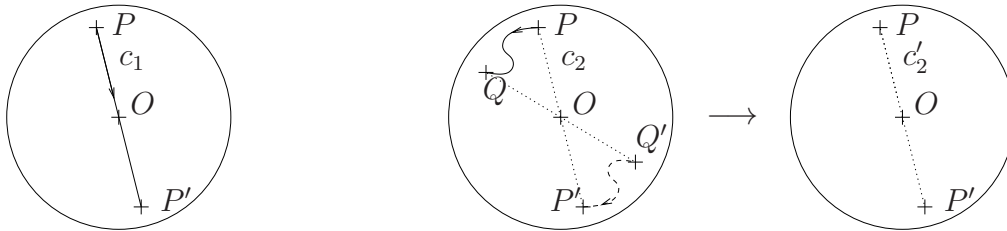


FIGURE 14.2 – A gauche : lacet  $c_1$  joignant en ligne droite les deux points diamétralement opposés  $P$  et  $P'$ . A droite : lacet  $c_2$  joignant en ligne droite les deux points diamétralement opposés  $P$  et  $P'$ , et le chemin constant  $c'_2$  qui lui est homotope.

On montre qu'il n'y a que deux classes de lacets. Nous allons maintenant vérifier que la structure de  $\mathbb{Z}_2$  est bien présente.



Pour cela, remarquons tout d'abord que le chemin  $(PQ)$  peut être continûment déformé en un demi grand cercle (*i.e.* de centre 0), en particulier en le chemin  $PP'$  (cela signifie simplement que l'on a choisi le cas particulier  $Q = P'$ ), voir fig. 14.3. Dans cette même déformation, le chemin  $(Q'P')$  se déforme continûment en le chemin  $(PP')$ , demi grand cercle complémentaire du premier.

D'une part, ce nouveau chemin  $c_3$  est homotope à un chemin constant, puisque le fait d'avoir choisi  $Q = P'$  ne change en effet rien à ce que qui a été dit en 2). D'autre part, le demi grand cercle  $PP'$  ci-dessus est homotope au lacet  $c_1$  joignant  $P$  à  $P'$ , cf 1). De même pour le demi grand cercle complémentaire joignant  $P$  à  $P'$ , homotope au même lacet  $c_1$  joignant  $P$  à  $P'$ .

En conclusion,  $c_2$ , homotope à l'identité, est homotope à la succession de deux lacets  $c_1$  (leur produit au sens du groupe fondamental). Le produit de  $c_1$  avec lui-même est donc homotope à l'identité : on retrouve la structure du groupe  $\mathbb{Z}_2$ .

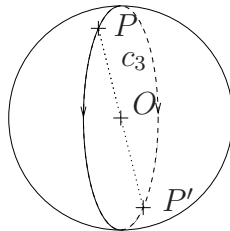


FIGURE 14.3 – La succession de deux lacets  $c_1$ , homotope à un lacet constant.

Ceci peut se visualiser par l'expérience suivante :

- ◇ partant d'une position avec la main paume tournée vers le haut, faites tourner votre main autour de votre bras tout en gardant la paume vers le haut. Au bout d'un tour ( $2\pi$ !), on parvient à une situation inconfortable qui est clairement distincte de la situation initiale. La main a une position identique à celle qu'elle avait au départ, mais pas votre corps. La succession des transformations effectuées correspond au lacet  $c_1$  (ou plutôt en pratique un lacet homotope à  $c_1$ , car sauf si vous êtes très souples, votre main n'a pas tout à fait tourné dans un seul plan, orthogonal à un axe  $PP'$ , de  $-\pi$  à  $\pi$ , mais comme on l'a vu ci-dessus, l'homotopie rattrape votre éventuel manque de souplesse...). L'inconfort matérialise le fait que  $c_1$  n'est pas contractile en l'identité!
- ◇ Mais vous aller retrouver votre confort si vous continuez de tourner dans le même sens, en répétant à nouveau la rotation de  $2\pi$  : le nœud de votre bras disparaît par miracle, après deux lacets  $c_1$ ...

Autre expérience du même type : tenez l'extrémité d'une ceinture dans une main, et effectuez une torsade de la ceinture sur elle-même, en faisant effectuer à l'autre extrémité un tour. Puis effectuer une second torsade dans le même sens. Vous constaterez qu'en faisant simplement passer l'extrémité qui vous a servi à torsader au-dessus de la ceinture elle-même, par simple translation, la double torsade disparaît !

Pour le lecteur curieux, on notera que le fait que  $SO(3)$  soit homéomorphe à l'espace projectif  $P^3(\mathbb{R})$  se retrouve bien ici. En effet  $P^3(\mathbb{R})$  est par définition l'ensemble des directions

des droites dans l'espace de dimension 4. Le chemin  $c_1$  est simplement une rotation d'angle  $\pi$  d'une droite dans cet espace, suivant un axe arbitraire orthogonal à cette droite. La direction obtenue après cette rotation est bien la même (puisque la droite s'est retournée sur elle-même), mais il n'y a pas moyen de contracter cette rotation en l'identité.

*Remarque :*

Ce résultat est général : le groupe fondamental de  $O(n)$  et de  $SO(n)$  est  $\mathbb{Z}_2$  pour  $n \leq 3$  (la situation est différente pour  $n = 2$  : c'est tout simplement  $\mathbb{Z}$  (le nombre de tour !) pour  $O(2)$  et  $SO(2)$ ).

## 14.3 Morphisme du groupe $SU(2)$ sur le groupe $SO(3)$

### 14.3.1 Morphisme

**Théorème 14.1 :**

Il existe un morphisme du groupe  $SU(2)$  sur le groupe des rotations  $SO(3)$  défini par

$$\boxed{\begin{array}{l} \pm U \mapsto R(U) \text{ tel que} \\ \forall \vec{x} \in \mathbb{R}^3, \vec{x} \mapsto \vec{x}' = R(U) \vec{x} \\ \text{avec } \vec{x}' \cdot \vec{\sigma} = U \vec{x} \cdot \vec{\sigma} U^{-1} \end{array}} \quad (14.26)$$

*Preuve :*

Il s'agit de montrer que pour  $U \in SU(2)$  donné, la transformation  $R(U)$  définie ci-dessus est bien une rotation.

◇ On remarque tout d'abord que  $U \vec{x} \cdot \vec{\sigma} U^{-1}$  peut s'écrire  $\vec{x}' \cdot \vec{\sigma}$  et non  $x_0 \mathbf{1} + \vec{x}' \cdot \vec{\sigma}$  puisque  $\text{Tr}[U \vec{x} \cdot \vec{\sigma} U^{-1}] = 0$  alors que  $\text{Tr} \mathbf{1} \neq 0$ . La définition précédente qui définit l'image de  $\vec{x}$  par  $R(U)$  a donc un sens.

◇ D'après l'éq. (14.23),  $\vec{x}'^2 = -\det[\vec{x}' \cdot \vec{\sigma}]$ . On déduit donc de la définition de  $\vec{x}'$  que

$$\vec{x}'^2 = -\det[U \vec{x} \cdot \vec{\sigma} U^{-1}] = -\det \vec{x} \cdot \vec{\sigma} = \vec{x}^2.$$

Ceci prouve que la transformation (3) conserve les longueurs, et donc qu'elle appartient à  $O(3)$ .

◇ Il reste à prouver qu'elle conserve aussi l'orientation des trièdres. On utilise pour cela la relation

$$\boxed{(\vec{x}, \vec{y}, \vec{z}) = \frac{1}{2i} \text{Tr}[(\vec{\sigma} \cdot \vec{x})(\vec{\sigma} \cdot \vec{y})(\vec{\sigma} \cdot \vec{z})]} \quad (14.27)$$

*Preuve :*

$\text{Tr}[(\vec{\sigma} \cdot \vec{x})(\vec{\sigma} \cdot \vec{y})(\vec{\sigma} \cdot \vec{z})] = x_i y_j z_k \text{Tr}[\sigma_i \sigma_j \sigma_k]$ . Or  $\text{Tr}[\sigma_i \sigma_j \sigma_k] = \text{Tr}[(\delta_{ij} + i\varepsilon_{ijk} \sigma_k) \sigma_k] = i\varepsilon_{ijk} \text{Tr}[\sigma_k \sigma_k]$ . Comme  $\text{Tr}[\sigma_k \sigma_k] = 2\delta_{kk}$ , on en déduit que  $\text{Tr}[(\vec{\sigma} \cdot \vec{x})(\vec{\sigma} \cdot \vec{y})(\vec{\sigma} \cdot \vec{z})] =$

$2i \varepsilon_{ijk} x_i y_j z_k$ , ce qui démontre le résultat.

On obtient alors immédiatement  $(\vec{x}', \vec{y}', \vec{z}') = (\vec{x}, \vec{y}, \vec{z})$  en utilisant l'invariance de la trace par permutation cyclique.

◇ Il reste à vérifier la structure de morphisme, c'est-à-dire que

$\forall U, V \in SU(2)$ ,  $R(UV) = R(U)R(V)$ , ce qui se démontre immédiatement en comparant l'action de  $R(UV)$  et de  $R(U)R(V)$  sur un vecteur arbitraire de  $\mathbb{R}^3$ .

### 14.3.2 Caractérisation

La paramétrisation  $U = \exp -i\theta \frac{\vec{n} \cdot \vec{\sigma}}{2}$  a une interprétation géométrique immédiate :

$$\boxed{\text{A l'élément } U = \cos \frac{\theta}{2} - i\vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} \text{ de } SU(2) \text{ correspond la rotation d'angle } \theta \text{ autour de l'axe } \vec{n}.}$$
 (14.28)

*Preuve :*

◇ Axe de la rotation :

De l'écriture de  $U$  on tire  $[U, \vec{\sigma} \cdot \vec{n}] = 0$  donc  $U \vec{\sigma} \cdot \vec{n} - \vec{\sigma} \cdot \vec{n} U = 0$  soit encore  $U \vec{\sigma} \cdot \vec{n} U^{-1} = \vec{\sigma} \cdot \vec{n}$ , ce qui montre que  $\vec{n}$  est conservé dans la rotation.

◇ Angle de la rotation :

Soit  $\vec{V} \perp \vec{n}$ . Alors  $\vec{V}$  se transforme sous l'action de  $R(U)$  en  $\vec{V}'$  tel que

$$\begin{aligned} \vec{\sigma} \cdot \vec{V}' &= \left( \cos \frac{\theta}{2} - i\vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} \right) \vec{\sigma} \cdot \vec{V} \left( \cos \frac{\theta}{2} + i\vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} \right) \\ &= \cos^2 \frac{\theta}{2} \vec{\sigma} \cdot \vec{V} + \frac{i}{2} [\vec{\sigma} \cdot \vec{V}, \vec{\sigma} \cdot \vec{n}] \sin \theta + (\vec{\sigma} \cdot \vec{n})(\vec{\sigma} \cdot \vec{V})(\vec{\sigma} \cdot \vec{n}) \sin^2 \frac{\theta}{2}. \end{aligned}$$

– Le second terme se simplifie grâce à la relation  $[\vec{\sigma} \cdot \vec{V}, \vec{\sigma} \cdot \vec{n}] = 2i \vec{V} \wedge \vec{n}$ .

– En utilisant la relation  $(\vec{\sigma} \cdot \vec{V})(\vec{\sigma} \cdot \vec{n}) = i\vec{\sigma} \cdot (\vec{V} \wedge \vec{n})$ , puisque  $\vec{V} \cdot \vec{n} = 0$ , le dernier terme peut se réorganiser grâce à l'identité  $i(\vec{\sigma} \cdot \vec{n})(\vec{\sigma} \cdot (\vec{V} \wedge \vec{n})) = -\vec{\sigma} \cdot (\vec{n} \wedge (\vec{V} \wedge \vec{n}))$ , que l'on développe en utilisant la relation bien connue

$$\vec{a} \wedge (\vec{b} \wedge \vec{c}) = (\vec{a} \cdot \vec{c}) \vec{b} - (\vec{a} \cdot \vec{b}) \vec{c},$$
 (14.29)

qui conduit à

$$(\vec{\sigma} \cdot \vec{n})(\vec{\sigma} \cdot \vec{V})(\vec{\sigma} \cdot \vec{n}) = \vec{\sigma} \cdot \vec{V}.$$

On obtient donc finalement

$$\vec{\sigma} \cdot \vec{V}' = \cos \theta \vec{\sigma} \cdot \vec{V} + \sin \theta \vec{\sigma} \cdot (\vec{n} \wedge \vec{V}),$$

d'où  $\vec{V}' = \cos \theta \vec{V} + \sin \theta \vec{n} \wedge \vec{V}$ , ce qui prouve que  $\vec{V}'$  se déduit de  $\vec{V}$  par une rotation d'angle  $\theta$  autour de  $\vec{n}$ .

### 14.3.3 Formules d'inversion

On pose

$$\tilde{x} = x_i \sigma_i. \quad (14.30)$$

Alors

$$\boxed{x_i = \frac{1}{2} \text{Tr} \sigma_i \tilde{x}} \quad (14.31)$$

et

$$\boxed{R(U)^i_j = \frac{1}{2} \text{Tr}(\sigma_i U \sigma_j U^\dagger)}. \quad (14.32)$$

*Preuves :*

- ◇ La relation (14.31) est immédiate en utilisant la relation (14.11).
- ◇ La relation (14.32) se démontre en calculant l'image  $R(U) \vec{x}$  de  $\vec{x}$  par la rotation  $R(U)$  et en vérifiant qu'elle est bien donnée par  $\vec{x}'$  défini en accord avec (14.26) :

$$R(U)^i_j x^j = \frac{1}{2} \text{Tr}(\sigma_i U \sigma_j U^\dagger) x^j = \frac{1}{2} \text{Tr}(\sigma_i U \tilde{x} U^\dagger) = \frac{1}{2} \text{Tr}(\sigma_i \tilde{x}') = \tilde{x}'^i.$$

### 14.3.4 Éléments conjugués

$\forall V \in SU(2)$ , les classes d'éléments conjugués s'écrivent

$$V U(\theta, \vec{n}) V^{-1} = U(\theta, \vec{n}')$$

où  $\vec{n}'$  se déduit de  $\vec{n}$  par la rotation  $R(V) : \vec{n}' = R(V) \vec{n}$ .

(14.33)

*Preuve :*

$$\begin{aligned} V U(\theta, \vec{n}) V^{-1} &= V \left( \cos \frac{\theta}{2} - i \vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} \right) V^{-1} \\ &= \cos \frac{\theta}{2} - i \sin \frac{\theta}{2} \vec{\sigma} \cdot \vec{n}' \end{aligned}$$

où  $\vec{\sigma} \cdot \vec{n}' = V \vec{\sigma} \cdot \vec{n} V^{-1}$ .

Les classes d'éléments conjugués sont donc caractérisées par le même angle. Cette propriété est très utile en pratique pour écrire l'élément de  $SU(2)$  qui code une rotation d'angle donné et d'axe quelconque en fonction d'un élément déjà connu de  $SU(2)$  codant une rotation de même angle autour d'un autre axe.

### 14.3.5 Forme infinitésimale

Cherchons la forme des matrices unitaires  $2 \times 2$  infiniment voisines de l'identité. A partir de la forme exponentielle (14.7) déjà obtenue plus haut, le résultat est immédiat. Il est cependant instructif d'obtenir directement le résultat.

Une matrice  $2 \times 2$  la plus générale est au voisinage de  $\mathbb{1}$  de la forme

$$U = \mathbb{1} + \varepsilon_0 + \vec{\varepsilon} \cdot \vec{\sigma} \quad \text{avec } \varepsilon_0, \varepsilon \in \mathbb{C}. \quad (14.34)$$

La contrainte  $\det U = 1$  conduit donc à  $\varepsilon_0 = 0$  à cause de la relation (12.31)

La contrainte d'unitarité s'écrit  $(\mathbb{1} + \vec{\varepsilon} \cdot \vec{\sigma})(\mathbb{1} + \vec{\varepsilon}^* \cdot \vec{\sigma}) = \mathbb{1}$  soit encore  $\vec{\varepsilon} + \vec{\varepsilon}^* = 0$ , d'où l'on déduit que  $\vec{\varepsilon}$  est imaginaire pur. Après redéfinition de  $\vec{\varepsilon}$ , on a donc

$$U = \mathbb{1} - i \vec{\varepsilon} \cdot \frac{\vec{\sigma}}{2} = \mathbb{1} - i \varepsilon \vec{n} \cdot \frac{\vec{\sigma}}{2} \quad (14.35)$$

que nous aurions bien sûr pu obtenir directement par développement au premier ordre de (14.7). Ce résultat correspond d'après la paramétrisation générale à une rotation infinitésimale d'angle  $\varepsilon$  autour de l'axe  $\vec{n}$ , qui s'écrit d'après (14.3)

$$\vec{V}' = R(U) \vec{V} = \vec{V} + \varepsilon \vec{n} \wedge \vec{V}$$

La relation (14.35) montre que les matrices de Pauli sont les générateurs de  $SU(2)$ . L'algèbre de Lie est donc définie par les relations de commutation

$$\left[ \frac{\sigma_i}{2}, \frac{\sigma_j}{2} \right] = i \varepsilon_{ijk} \frac{\sigma_k}{2} \quad (14.36)$$

Cette algèbre de Lie est isomorphe à l'algèbre de Lie du groupe  $SO(3)$ , comme on peut le voir par comparaison immédiate avec les relations de commutation (14.6) satisfaites par les générateurs de  $SO(3)$  (définis par  $R_{\vec{n}}(\omega) = 1 - i \vec{J} \cdot \vec{n} \omega$ ). Ce résultat ne doit pas nous étonner puisque nous avons montré dans la partie 14.3 qu'il existe un morphisme entre  $SU(2)$  et  $SO(3)$ , qui est bijectif dans un voisinage de l'identité (l'indétermination  $\pm U$  ne pose de problème que lorsque l'on s'éloigne suffisamment de l'identité).

Retenons donc que

$$\boxed{\mathfrak{so}(3) \cong \mathfrak{su}(2)}. \quad (14.37)$$

## 14.4 Représentations de $SU(2)$

### 14.4.1 Générateurs dans une représentation quelconque

A l'élément  $U = \mathbb{1} - i \vec{\varepsilon} \cdot \frac{\vec{\sigma}}{2}$  voisin de l'identité dans  $SU(2)$  correspond

$$D(U(\varepsilon_1, \varepsilon_2, \varepsilon_3)) = \mathbb{1} - i(\varepsilon_1 J_1 + \varepsilon_2 J_2 + \varepsilon_3 J_3) \quad (14.38)$$

dans la représentation  $D$ . Les  $J_1, J_2, J_3$  sont les générateurs de la représentation (ils sont hermitiques puisque  $U$  et donc  $D(U)$  sont unitaires). Pour une rotation infinitésimale autour de l'axe  $\vec{n}$  on a

$$D(U(\varepsilon, \vec{n})) = \mathbb{1} - i \varepsilon \vec{J} \cdot \vec{n}.$$

Considérons une autre matrice  $V \in SU(2)$ , on a alors

$$V U(\varepsilon, \vec{n}) V^{-1} = U(\varepsilon, \vec{n}') \quad \text{où} \quad \vec{n}' = R(V) \vec{n}.$$

Dans la représentation  $D$  ceci s'écrit

$$D(V) D(U(\varepsilon, \vec{n})) D(V^{-1}) = D(U(\varepsilon, \vec{n}')) \quad (14.39)$$

soit encore

$$D(V) (\mathbb{1} - i\varepsilon \vec{J} \cdot \vec{n}) D^{-1}(V) = \mathbb{1} - i\varepsilon \vec{J} \cdot \vec{n}', \quad (14.40)$$

*i.e.*

$$D(V) \vec{J} \cdot \vec{n} D^{-1}(V) = \vec{J} \cdot \vec{n}' = \vec{J} \cdot (R(V) \vec{n}) \quad (14.41)$$

et donc, en écrivant explicitement les composantes

$$D(V) J_i n_i D^{-1}(V) = J_i R_{ij} n_j \quad (14.42)$$

soit finalement, puisque le résultat est valable pour tout  $\vec{n}$ ,

$$\boxed{D(V) J_j D^{-1}(V) = R_{ij} J_i.} \quad (14.43)$$

Montrons réciproquement que la donnée de cette loi de transformation pour les opérateurs  $\vec{J}$  implique la structure d'algèbre de Lie. Soit  $V$  un élément de  $SU(2)$  codant une rotation infinitésimale d'angle  $\alpha$  autour de  $\vec{t}$ , *i.e.*

$$V = \mathbb{1} - i \frac{\alpha}{2} \vec{\sigma} \cdot \vec{t} \quad \text{et donc} \quad D(V) = \mathbb{1} - i \alpha \vec{J} \cdot \vec{t}, \quad (14.44)$$

où l'on note  $\vec{J} = D(\frac{\vec{\sigma}}{2})$ . La relation (14.43) s'écrit au premier ordre en  $\alpha$

$$\left( \mathbb{1} - i \alpha \vec{J} \cdot \vec{t} \right) J_j \left( \mathbb{1} + i \alpha \vec{J} \cdot \vec{t} \right) = R_{ij} J_i. \quad (14.45)$$

Les éléments de matrice de  $R$  s'obtiennent facilement en utilisant l'expression de la rotation infinitésimale d'un vecteur  $\vec{q}$

$$\vec{q}' = R(V) \vec{q} = \vec{q} + \alpha \vec{t} \wedge \vec{q}$$

soit

$$q'_i = q_i + \alpha \varepsilon_{ikj} t_k q_j = R_{ij} q_j,$$

d'où

$$R_{ij} = \delta_{ij} - \alpha \varepsilon_{ijk} t_k. \quad (14.46)$$

On en déduit, en égalant les termes du 1er ordre de la relation (14.45) et en utilisant le résultat (14.46), que

$$i \alpha [\vec{J} \cdot \vec{t}, J_j] = \alpha \varepsilon_{ijk} t_k J_i$$

soit encore

$$[J_k, J_j] = -i\varepsilon_{ijk} J_i = i\varepsilon_{kji} J_i.$$

On retrouve les règles de commutation bien connues du moment cinétique :

$$\boxed{[J_i, J_j] = i\varepsilon_{ijk} J_k.} \quad (14.47)$$

Les constantes de structures sont données par  $f_{ij}^k = -C_{ij}^k = \varepsilon_{ijk}$ . Ces règles sont valables quelle que soit la représentation considérée, par définition de la représentation d'une algèbre de Lie, comme nous l'avons vu au chap. 12, Eq. (13.88). En particulier, elles sont satisfaites par les générateurs de  $SU(2)$ , voir Eq. (14.6), et de  $SO(3)$ , voir Eq. (14.36), qui constituent deux représentations particulières de  $SU(2)$ , sur lesquelles nous allons revenir maintenant.

### 14.4.2 Exemples de représentation de $SU(2)$

#### Représentation spinorielle

L'espace de la représentation est à 2 dimensions : il est constitué de vecteurs dans  $\mathbb{C}^2$ , appelés *spinors*. Dans cette représentation,

$$\boxed{D(U) = U \quad \text{et} \quad \vec{J} = \frac{\vec{\sigma}}{2}.} \quad (14.48)$$

Nous étudierons en détail cette représentation et ses produits tensoriels dans le chap. 15.

#### Représentation adjointe, ou vectorielle

Dans cette représentation de dimension 3, chaque élément de  $SU(2)$  agit sur lui-même suivant la représentation adjointe, définie à l'aide de l'opérateur adjoint (12.11) par  $\text{ad } U(V) = UVU^{-1}$ . Cette structure a déjà été examinée dans la relation (14.33) : elle provient directement du morphisme de  $SU(2)$  sur  $SO(3)$ . Dans la définition de ce morphisme, la relation

$$\vec{x}' \cdot \vec{\sigma} = U \vec{x} \cdot \vec{\sigma} U^{-1}$$

peut encore s'écrire

$$\vec{x} \cdot \vec{\sigma}' = U \vec{x} \cdot \vec{\sigma} U^{-1} \quad \text{soit} \quad \vec{\sigma}' = U \vec{\sigma} U^{-1}$$

à condition de poser

$$\vec{\sigma}' = {}^t R(U) \vec{\sigma},$$

i.e. de considérer que les vecteurs de base  $\vec{\sigma}$  se transforment, de manière duale (c'est bien la propriété des vecteurs de base de façon générale!), comme des vecteurs de  $\mathbb{R}^3$  sous l'action de  $SO(3)$ .

Ainsi la représentation adjointe est donnée par

$$\boxed{D(U) = R(U),} \quad (14.49)$$

d'où le nom de représentation vectorielle. Ceci est possible puisque l'algèbre de Lie  $\mathfrak{su}(2)$  a même dimension que  $\mathbb{R}^3$ .

Remarque :

La relation (14.43) est une généralisation de la relation

$$U \vec{\sigma} U^{-1} = \vec{\sigma}' = {}^t R(U) \vec{\sigma}.$$

Comme nous en avons déjà fait la remarque page 197, la dimension de l'espace sur lequel les générateurs  $\vec{J}$  agissent ne joue aucun rôle dans la définition de la représentation adjointe du groupe (ici  $SU(2)$ ). L'équation (14.43) s'écrit toujours  $D(V) \vec{J} D^{-1}(V) = {}^t R(V) \vec{J}$ , quelque soit la forme explicite de  $\vec{J}$ . La représentation adjointe a été construite à partir de son action sur les générateurs de la représentation spinorielle (où  $\vec{J} = \vec{\sigma}/2$ ), mais choisir des générateurs dans une autre représentation ne change pas l'expression des générateurs de la représentation adjointe :  $R(V)$  n'est pas modifiée, de même bien sûr que la forme infinitésimale de  $R(V)$  qui définit les générateurs de la représentation adjointe.

Il n'est pas inutile de se convaincre de façon pédestre des résultats que nous venons d'obtenir. Pour une rotation infinitésimale on a

$$R_{ij} = \delta_{ij} - \alpha \varepsilon_{ijk} t_k$$

d'où l'on tire l'expression des générateurs  $\vec{J}$  de la représentation adjointe : par comparaison avec  $D(U) = 1 - i\alpha \vec{J} \cdot \vec{t}$  on obtient

$$\boxed{(J_k)_{ij} = -i \varepsilon_{ijk}}, \quad (14.50)$$

en accord avec la définition vue page 197 de la représentation adjointe :

$$\begin{aligned} (J_k)_{ij} &= D(-iX_k)^i_j = -iC_{kj}^i = if_{kj}^i = +i\varepsilon_{kji} \\ &= -i\varepsilon_{ijk}. \end{aligned}$$

Par changement de base, on obtient aisément une forme équivalente des générateurs de la représentation adjointe :

$$J_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (14.51)$$

### Moment orbital (représentation de dimension infinie)

Considérons l'espace des fonctions d'ondes de carré intégrable  $L^2(\mathbb{R}^3)$ . On définit la représentation

$${}^R \psi = D(R) \psi$$

telle que

$${}^R \psi(\vec{r}) = \psi(R^{-1}\vec{r}) \text{ (rotation active).}$$



Cette représentation est unitaire, puisque

$$\begin{aligned} \int |{}^R\psi(\vec{r})|^2 d^3\vec{r} &= \int |\psi(R^{-1}\vec{r})|^2 d^3\vec{r} \\ &= \int |\psi(\vec{r})|^2 d^3\vec{r}_1 \quad (\vec{r}_1 = R^{-1}\vec{r}), \end{aligned}$$

comme nous l'avons déjà vérifié page 145. Pour une rotation infinitésimale, nous pouvons écrire

$$\begin{aligned} {}^R\psi(\vec{r}) = \psi(\vec{r} - \alpha \vec{t} \wedge \vec{r}) &= \psi(\vec{r}) - \alpha (\vec{t} \wedge \vec{r}) \cdot \vec{\nabla} \psi(\vec{r}) \\ &= \psi(\vec{r}) - \alpha (\vec{r} \wedge \vec{\nabla} \psi) \cdot \vec{t} \end{aligned}$$

d'où

$$\boxed{\vec{J} = \frac{1}{i} \vec{r} \wedge \vec{\nabla} \quad \text{noté} \quad \vec{L} = \vec{r} \wedge \vec{p}.} \quad (14.52)$$

### Forme exponentielle

Pour finir, montrons que les opérateurs  $D(U)$  d'une représentation unitaire de  $SU(2)$  peuvent se calculer connaissant  $U$  et les générateurs de la représentation. On a

$$D(U(\varepsilon, \vec{n})) D(U(\theta, \vec{n})) = D(U(\varepsilon + \theta, \vec{n})) \quad (14.53)$$

donc

$$\begin{aligned} D(U(\varepsilon + \theta, \vec{n})) - D(U(\theta, \vec{n})) &= [D(U(\varepsilon, \vec{n})) - 1] D(U(\theta, \vec{n})) \\ &= -i\varepsilon \vec{J} \cdot \vec{n} D(U(\theta, \vec{n})) \end{aligned}$$

d'où

$$\frac{d}{d\theta} D(U(\theta, \vec{n})) = -i\vec{J} \cdot \vec{n} D(U(\theta, \vec{n}))$$

et

$$\boxed{D(U(\theta, \vec{n})) = \exp -i\theta \vec{J} \cdot \vec{n}.} \quad (14.54)$$

*Remarque :*

On n'aura pas l'équivalent de (14.53) pour le groupe de Lorentz restreint  $\mathcal{L}_+^\uparrow$ , à cause du fait qu'une transformation de Lorentz pure donnée n'engendre pas un sous-groupe à un paramètre (contrairement aux rotations), ce qui explique pourquoi on ne pourra écrire tout élément de  $\mathcal{L}_+^\uparrow$  comme l'exponentielle d'un élément de son algèbre de Lie. Le résultat présent est en accord avec la compacité de  $SO(3)$  (alors que  $\mathcal{L}_+^\uparrow$  n'est pas compact).

## 14.5 Représentations unitaires irréductibles de $SU(2)$

Le groupe  $SU(2)$  est compact, donc d'après le th. 13.45, les représentations de  $SU(2)$  sont équivalentes à des représentations unitaires, et toute représentation irréductible est de dimension finie.

### 14.5.1 Spectre

La structure d'algèbre de Lie

$$[J_i, J_j] = i\varepsilon_{ijk} J_k \quad (14.55)$$

implique que le Casimir  $J^2 \equiv \vec{J}^2 = J_1^2 + J_2^2 + J_3^2$  commute avec tous les générateurs  $J_i$ , comme on peut le vérifier directement. C'est une conséquence immédiate du résultat général démontré dans l'exercice 12.2. D'après le Lemme de Schur,  $J^2$  est donc un multiple de l'identité, que l'on écrira sous la forme  $J^2 = j(j+1)\mathbf{1}$  avec  $j(j+1)$  positif. On peut choisir de prendre  $j$  réel positif sans perte de généralité.

L'algèbre de Lie de  $SU(2)$  que nous avons considérée est l'algèbre de Lie *réelle* définie par les constantes de structure (14.55). Il est techniquement pratique d'introduire une autre algèbre de Lie réelle ayant la même extension complexe que celle de  $SU(2)$ . On pose donc

$$J_{\pm} \equiv J_1 \pm iJ_2. \quad (14.56)$$

L'algèbre de Lie réelle formée à partir des générateurs  $J^{\pm}$  et  $J_3$  a alors pour structure

$$[J_3, J_+] = J_+, \quad [J_3, J_-] = -J_- \quad \text{et} \quad [J_+, J_-] = 2J_3. \quad (14.57)$$

On vérifie immédiatement que

$$\begin{aligned} J_+ J_- &= J^2 - J_3(J_3 - 1) \\ J_- J_+ &= J^2 - J_3(J_3 + 1). \end{aligned} \quad (14.58)$$

Soit  $|jm\rangle$  un état propre de  $J^2$  et  $J_3$  de valeurs propres respectives  $j(j+1)$  et  $m$  (*remarque* : si  $J^2$  et  $J_3$  ne forment pas un Ensemble Complet d'Observables Commutantes,  $|jm\rangle$  est un état particulier choisi une fois pour toute dans le sous-espace de moment cinétique ( $jm$ )). On peut alors écrire

$$J_3|jm\rangle = m|jm\rangle, \quad (14.59)$$

et

$$\begin{aligned} J_- J_+ |jm\rangle &= J^2 |jm\rangle - J_3(J_3 + 1) |jm\rangle \\ &= [j(j+1) - m(m+1)] |jm\rangle = (j-m)(j+m+1) |jm\rangle \end{aligned} \quad (14.60)$$

d'où

$$\langle jm | J_- J_+ |jm\rangle = (j-m)(j+m+1) = \|J_+ |jm\rangle\|^2, \quad (14.61)$$

qui est par définition positif ou nul. On en déduit donc que

$(j-m)(j+m+1) \geq 0$ , d'où

$$-j-1 \leq m \leq j. \quad (14.62)$$

On peut reproduire le même raisonnement avec l'opérateur  $J_+ J_-$ , ce qui permet de montrer que

$$(j+m)(j-m+1) \geq 0,$$

d'où l'on déduit que

$$-j \leq m \leq j+1 \quad (14.63)$$

puisque  $\|J_-|jm\rangle\|^2 \geq 0$ . Les inégalités (14.62) et (14.63) donnent finalement

$$\boxed{-j \leq m \leq j.} \quad (14.64)$$

De plus, comme  $v = 0 \Leftrightarrow \|v\| = 0$ , on déduit de l'éq. (14.61) que

$$J_+|jm\rangle = 0 \Leftrightarrow (j - m)(j + m + 1) = 0 \quad \text{soit (puisque } m \in [-j, j]) \quad m = j \quad (14.65)$$

et

$$J_-|jm\rangle = 0 \Leftrightarrow (j + m)(j - m + 1) = 0 \quad \text{soit (puisque } m \in [-j, j]) \quad m = -j. \quad (14.66)$$

Dans le cas où  $m \neq j$ , le vecteur non nul  $J_+|jm\rangle$  est un vecteur de moment cinétique  $(j, m + 1)$ .

*Preuve :*

$J^2$  étant un Casimir,  $J^2 J_+|jm\rangle = J_+ J^2|jm\rangle = j(j + 1)J_+|jm\rangle$ . De plus, l'éq. (14.57) implique que  $J_3 J_+ = J_+(J_3 + 1)$  ce qui donne après action sur le ket  $|jm\rangle$ ,

$$J_3 J_+|jm\rangle = J_+(J_3 + 1)|jm\rangle = (m + 1)J_+|jm\rangle.$$

On montre de même que  $J_-|jm\rangle$  est un vecteur de moment cinétique  $(j, m - 1)$ .

◇ Considérons à présent la suite de vecteurs obtenus par application répétée de  $J_+$  sur  $|jm\rangle$  :

$$J_+|jm\rangle, J_+^2|jm\rangle, \dots, J_+^p|jm\rangle$$

pour  $-j \leq m \leq j$  :

- Si  $m = j$ ,  $J_+|jm\rangle = 0$
- Si  $m < j$ ,  $J_+|jm\rangle$  est un vecteur non nul de moment cinétique  $(j, m + 1)$  donc nécessairement  $m + 1 \leq j$
- Si  $m + 1 = j$ ,  $J_+^2|jm\rangle = 0$
- Si  $m + 1 < j$ ,  $J_+^2|jm\rangle \neq 0$ , de moment cinétique  $(j, m + 2)$
- etc.

Les vecteurs de cette suite doivent tous s'annuler à partir d'un certain rang : sinon on serait parvenu à former des vecteurs propres de  $J_3$  dont la valeur propre serait arbitrairement grande, et en particulier supérieure à  $j$ .

Donc  $\exists p \in \mathbb{N}$  tel que  $J_+^p|jm\rangle \neq 0$ , de moment cinétique  $(j, m + p)$  et vérifiant  $J_+^{p+1}|jm\rangle = 0$  (ce vecteur  $J_+^p|jm\rangle$  est appelé vecteur de plus haut poids). Ainsi  $m + p = j$  d'après l'Eq. (14.65), ce qui prouve que  $j - m \in \mathbb{N}$ .

◇ On peut de même agir de façon répétée sur  $|jm\rangle$  à l'aide de  $J_-$ . On en déduit alors que  $j + m = q \in \mathbb{N}$ .

On a donc obtenu les  $p + q + 1$  états

$$\begin{aligned}
 & J_-^q |jm\rangle, \quad J_-^{q-1} |jm\rangle, \quad \dots, \quad J_- |jm\rangle, \quad |jm\rangle, \quad J_+ |jm\rangle, \quad \dots, \quad J_+^p |jm\rangle \\
 & m - q = -j \quad -j + 1 \qquad \qquad \qquad m - 1 \quad m \quad m + 1 \qquad \qquad \qquad m + p = j
 \end{aligned} \tag{14.67}$$

où l'on a indiqué sous chaque vecteur la valeur propre de  $J_3$ .

La valeur propre de  $J^2$  pour chacun de ces états est  $j(j + 1)$ . D'autre part, comme  $p \in \mathbb{N}$  et  $q \in \mathbb{N}$ , on en déduit que  $p + q = j - m + j + m = 2j \in \mathbb{N}$  et donc que

$$\boxed{j \text{ est entier ou demi-entier .}}$$

### 14.5.2 Vecteurs propres de $J^2$ et $J_3$ et construction des sous-espaces invariants $\mathcal{E}^{(j)}$

Il nous reste à normaliser les  $2j + 1$  vecteurs précédents. Supposons que  $|jm\rangle$  soit de norme 1.

◇ Si  $m = j$ ,  $J_+ |jm\rangle = 0$

◇ Si  $m < j$ , notons  $|jm + 1\rangle$  le vecteur de norme 1 qui vérifie

$J_+ |jm\rangle = c_m |jm + 1\rangle$ . Comme  $\|J_+ |jm\rangle\|^2 = j(j + 1) - m(m + 1)$ , en choisissant  $c_m \in \mathbb{R}_+$  on aura

$$J_+ |jm\rangle = \sqrt{j(j + 1) - m(m + 1)} |jm + 1\rangle.$$

Si l'on fait agir  $J_-$  à gauche dans l'égalité précédente, on aura donc, en utilisant l'éq. (14.60),

$$J_- J_+ |jm\rangle = [j(j + 1) - m(m + 1)] |jm\rangle = \sqrt{j(j + 1) - m(m + 1)} J_- |jm + 1\rangle$$

d'où

$$J_- |jm + 1\rangle = \sqrt{j(j + 1) - m(m + 1)} |jm\rangle.$$

On peut répéter les manipulations précédentes sur  $|jm + 1\rangle$  :

◇ Si  $m + 1 = j$ ,  $J_+ |jm + 1\rangle = 0$ .

◇ Sinon on forme  $|jm + 2\rangle$  de norme 1,

◇ etc. jusqu'au vecteur de plus haut poids  $|jj\rangle$ .

De même par action répétée de  $J_-$ .

Les opérateurs  $J_+$  et  $J_-$  sont généralement appelés opérateurs d'échelle, puisqu'ils permettent respectivement de monter ou de descendre le long de l'échelle des états propres de  $J_3$ .

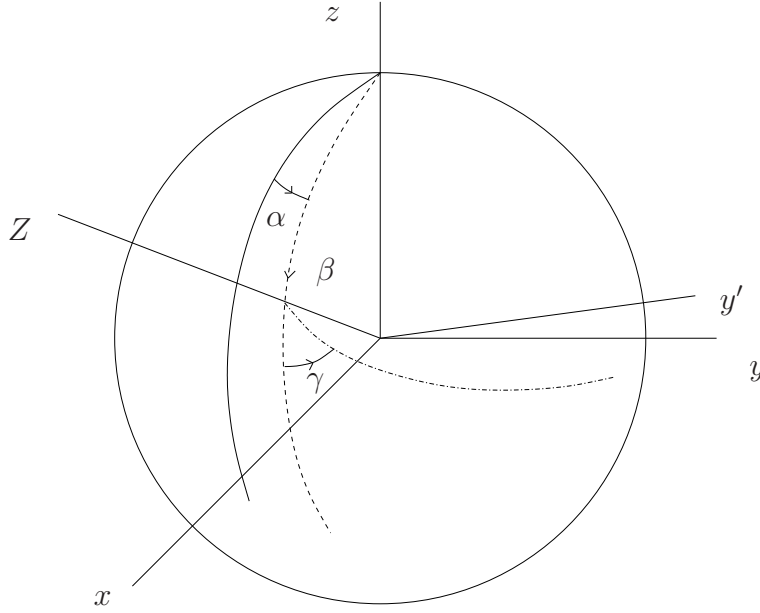


FIGURE 14.4 – Angles d'Euler.

On obtient ainsi une suite de  $2j + 1$  vecteurs orthonormés  $|j - j\rangle, \dots, |j j\rangle$  qui sous-tendent l'espace  $\mathcal{E}^{(j)}$  de la représentation irréductible  $D^j$  de dimension  $2j + 1$ . Ces vecteurs satisfont les relations

$$\begin{aligned}
 \langle j m | j m' \rangle &= \delta_{mm'} \\
 J_3 |j m\rangle &= m |j m\rangle \\
 J_+ |j m\rangle &= \sqrt{j(j+1) - m(m+1)} |j m+1\rangle \\
 J_- |j m\rangle &= \sqrt{j(j+1) - m(m-1)} |j m-1\rangle.
 \end{aligned}
 \tag{14.68}$$

## 14.6 Angles d'Euler

Les angles d'Euler permettent de paramétrer une rotation quelconque. C'est un outil classique de la mécanique du solide, qui permet de coder la position arbitraire d'un objet à l'aide de trois angles.

### 14.6.1 Définition et construction

Considérons une courbe dessinée en continu sur la fig. 14.4, que l'on cherche à amener dans la position finale tracée en traits pointillés-tirés, grâce à une transformation de  $SO(3)$  dont nous souhaitons trouver une paramétrisation. Nous allons pour cela utiliser 3 transformations successives :

- ◇ Rotation d'angle  $\alpha$  autour de  $Oz$  :  $R(U(\alpha, z))$
- ◇ Rotation d'angle  $\beta$  autour de  $Oy'$ ,  $Oy'$  étant le transformé de  $Oy$  par  $R(U(\alpha, z))$  :

$$R(V) \text{ avec } V = U(\alpha, z) U(\beta, y) U^{-1}(\alpha, z),$$

d'après la relation (14.33).

- ◇ Rotation d'angle  $\gamma$  autour de  $OZ$ ,  $OZ$  étant le transformé de  $Oz$  par la succession des rotations  $R(U(\alpha, z))$  puis  $R(U(\alpha, z)U(\beta, y)U^{-1}(\alpha, z))$ , *i.e.* le transformé de  $Oz$  par  $R(U(\alpha, z)U(\beta, y)) \equiv R(T)$ .

Cette rotation s'écrit donc  $R(W)$  avec  $W = T U(\gamma, z) T^{-1}$ , soit encore

$$W = U(\alpha, z) U(\beta, y) U(\gamma, z) U^{-1}(\beta, y) U^{-1}(\alpha, z). \quad (14.69)$$

Au total, la rotation complète s'écrit  $R = R(U)$  avec

$$\begin{aligned} U &= W V U(\alpha, z) \\ &= U(\alpha, z) U(\beta, y) U(\gamma, z) U^{-1}(\beta, y) U^{-1}(\alpha, z) U(\alpha, z) U(\beta, y) U^{-1}(\alpha, z) U(\alpha, z) \end{aligned}$$

soit encore

$$\boxed{U = U(\alpha, z) U(\beta, y) U(\gamma, z)}. \quad (14.70)$$

*Remarques :*

- ◇ La convention utilisée ici est différente de celle de Wigner (où  $\gamma \leftrightarrow \alpha$ ). Il n'y a de toute façon pas de convention universelle, donc prudence...
- ◇ L'axe  $OZ$ , transformé de  $Oz$ , s'obtient aussi par la rotation

$R(U(\alpha, z)U(\beta, y)U^{-1}(\alpha, z))$  (puisque la première rotation  $R(U(\alpha, z))$  n'affecte pas l'axe  $Oz$ ).

On aurait alors  $T' = U(\alpha, z)U(\beta, y)U^{-1}(\alpha, z)$  et  $W' = T' U(\gamma, z) T'^{-1}$  donc

$$\begin{aligned} U = W' V U(\alpha, z) &= U(\alpha, z) U(\beta, y) U^{-1}(\alpha, z) U(\gamma, z) U(\alpha, z) U^{-1}(\beta, y) U^{-1}(\alpha, z) \\ &\quad \times U(\alpha, z) U(\beta, y) U^{-1}(\alpha, z) U(\alpha, z) \\ &= U(\alpha, z) U(\beta, y) U^{-1}(\alpha, z) U(\gamma, z) U(\alpha, z). \end{aligned}$$

Comme  $[U(\alpha, z), U(\gamma, z)] = 0$ , on a encore  $U = U(\alpha, z)U(\beta, y)U(\gamma, z)$ , ce qui est donc bien identique à (14.70).

- ◇ Domaine de variation des paramètres :

$$SO(3) : \begin{cases} \alpha \in [0, 2\pi] \\ \beta \in [0, \pi] \\ \gamma \in [0, 2\pi] \end{cases} \quad \text{et} \quad SU(2) : \begin{cases} \alpha \in [0, 2\pi] \\ \beta \in [0, \pi] \\ \gamma \in [0, 4\pi] \end{cases}.$$

- ◇ Autre point de vue, bien sûr équivalent :

$R(U)$  avec  $U = U(\alpha, z)U(\beta, y)U(\gamma, z)$  peut aussi être interprétée littéralement : rotation d'angle  $\gamma$  par rapport à l'axe  $z$ , puis rotation d'angle  $\beta$  par rapport à l'axe  $y$ , puis rotation d'angle  $\alpha$  par rapport à l'axe  $z$ , tous ces axes étant ici les axes du repère *fixe* initial.

### 14.6.2 Expression explicite

Partant des expressions

$$\begin{aligned} U(\alpha, z) &= \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix}, \\ U(\beta, y) &= \begin{pmatrix} \cos \beta/2 & \sin \beta/2 \\ -\sin \beta/2 & \cos \beta/2 \end{pmatrix}, \\ U(\gamma, z) &= \begin{pmatrix} e^{-i\gamma/2} & 0 \\ 0 & e^{i\gamma/2} \end{pmatrix} \end{aligned}$$

où l'on a utilisé la paramétrisation (14.22), on déduit aisément que

$$U = U(\alpha, z)U(\beta, y)U(\gamma, z) = \begin{pmatrix} e^{-i\alpha/2} \cos \frac{\beta}{2} e^{-i\gamma/2} & -e^{-i\alpha/2} \sin \frac{\beta}{2} e^{i\gamma/2} \\ e^{i\alpha/2} \sin \frac{\beta}{2} e^{-i\gamma/2} & e^{i\alpha/2} \cos \frac{\beta}{2} e^{i\gamma/2} \end{pmatrix}. \quad (14.71)$$

### 14.6.3 Propriétés de symétrie de la paramétrisation d'Euler

#### Transposition

$$U \leftrightarrow U^T \Leftrightarrow \begin{pmatrix} \alpha \leftrightarrow \gamma \\ \beta \leftrightarrow -\beta \end{pmatrix} \quad (14.72)$$

car  $U = U(\alpha, z)U(\beta, y)U(\gamma, z)$  et

$$U^T = U^T(\gamma, z)U^T(\beta, y)U^T(\alpha, z) = U(\gamma, z)U(-\beta, y)U(\alpha, z)$$

#### Inversion

$$U \leftrightarrow U^\dagger = U^{-1} \Leftrightarrow \begin{pmatrix} \alpha \leftrightarrow -\gamma \\ \beta \leftrightarrow -\beta \end{pmatrix} \quad (14.73)$$

car  $U^{-1} = U^{-1}(\gamma, z)U^{-1}(\beta, y)U^{-1}(\alpha, z) = U(-\gamma, z)U(-\beta, y)U(-\alpha, z)$ .

## 14.7 Définition des matrices de rotation

Considérons un élément de  $SU(2)$  codé par les angles d'Euler (14.70) et sa représentation correspondante dans  $\mathcal{E}^{(j)}$  qui s'écrit, d'après (14.54),

$$\boxed{\mathcal{D}(U) = \exp(-i\alpha J_3) \exp(-i\beta J_2) \exp(-i\gamma J_3)}. \quad (14.74)$$

Par définition, les matrices de rotation sont définies par

$$\boxed{\mathcal{D}_{mm'}^j(U) = \langle jm | \mathcal{D}(U) | jm' \rangle}. \quad (14.75)$$

On a donc

$$\mathcal{D}_{mm'}^j(U) = e^{-im\alpha} \langle jm | \exp(-i\beta J_2) | jm' \rangle e^{-im'\gamma}.$$

On pose par définition

$$d_{mm'}^j(\beta) = \langle jm | \exp -i\beta J_2 | jm' \rangle = \mathcal{D}_{mm'}^j(U(0, \beta, 0)), \quad (14.76)$$

d'où

$$\boxed{\mathcal{D}_{mm'}^j(U) = e^{-im\alpha} d_{mm'}^j(\beta) e^{-im'\gamma}.} \quad (14.77)$$

Comme  $iJ_2 = \frac{1}{2}(J_+ - J_-)$ , les  $d_{mm'}^j$  sont réels. Notons que

$$\mathcal{D}(U) |jm\rangle = \sum_{m'} |jm'\rangle \mathcal{D}_{m'm}^j(U) \quad (14.78)$$

en accord avec

$$\langle jm'' | \mathcal{D}(U) |jm\rangle = \sum_{m'} \delta_{m'm''} \mathcal{D}_{m'm}^j(U) = \mathcal{D}_{m''m}^j(U).$$

On notera la position des indices : les kets  $|jm'\rangle$  se transforment comme des vecteurs de base, et donc de manière duale sous l'action de  $\mathcal{D}(U)$ .

*Exemples :*

◇ Représentation fondamentale ( $j = \frac{1}{2}$ ) :

$$d_{mm'}^{\frac{1}{2}}(\beta) = \left[ \exp \left( -i\beta \frac{\sigma_2}{2} \right) \right]_{mm'} = \begin{pmatrix} \cos \frac{\beta}{2} & -\sin \frac{\beta}{2} \\ \sin \frac{\beta}{2} & \cos \frac{\beta}{2} \end{pmatrix} = \begin{pmatrix} d_{1/2 \ 1/2} & d_{1/2 \ -1/2} \\ d_{-1/2 \ 1/2} & d_{-1/2 \ -1/2} \end{pmatrix}.$$

◇ Représentation vectorielle ( $j = 1$ ) :

$$R(\beta, y) = \begin{pmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{pmatrix} = \exp(-i\beta J_2), \quad (14.79)$$

qui s'obtient soit de façon directe en écrivant explicitement la matrice de la rotation, soit en calculant  $\exp(-i\beta J_2)$  par son développement en série (le seul calcul matriciel à faire étant de vérifier que  $J_2 \cdot J_2 = \text{Diag}(-1, 0, -1)$  et  $J_2 \cdot J_2 \cdot J_2 = -J_2$  dans la base ci-dessous).

Dans la base cartésienne,

$$J_1 = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = i \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad J_3 = i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (14.80)$$

en accord avec l'éq. (14.5). Les états propres de  $J_3$  dans cette base cartésienne sont

$$|1 \ 0\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad |1 \ 1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ -i \\ 0 \end{pmatrix} = \frac{J_1 + iJ_2}{\sqrt{2}} |1 \ 0\rangle \quad (14.81)$$

$$\text{et } |1 \ -1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} = \frac{J_1 - iJ_2}{\sqrt{2}} |1 \ 0\rangle. \quad (14.82)$$



Les vecteurs  $|1 - 1\rangle$ ,  $|1 0\rangle$  et  $|1 1\rangle$ , définissent une nouvelle base dite standard, dans laquelle les générateurs  $J_i$  ont la forme déjà mentionnée en (14.51).

### Exercice 14.1

Vérifier que dans cette base standard, les valeurs de  $d_{mm'}^1(\beta)$  sont données par

$m'$	1	0	-1
$m$			
1	$\frac{1}{2}(1 + \cos \beta)$	$-\frac{\sin \beta}{\sqrt{2}}$	$\frac{1}{2}(1 - \cos \beta)$
0	$\frac{\sin \beta}{\sqrt{2}}$	$\cos \beta$	$-\frac{\sin \beta}{\sqrt{2}}$
-1	$\frac{1}{2}(1 - \cos \beta)$	$\frac{\sin \beta}{\sqrt{2}}$	$\frac{1}{2}(1 + \cos \beta)$

## 14.8 Transformation des états et des observables sous une rotation

### 14.8.1 Etats et observables

Nous nous proposons d'interpréter physiquement l'équation

$$D(U)|jm\rangle = \sum_{m'} D_{m'm}^j(U)|jm'\rangle. \quad (14.83)$$

On a vu dans la section 14.4 que

$$D(U) \vec{J} \cdot \vec{k} D^{-1}(U) = \vec{J} \cdot (R(U) \vec{k}), \quad (14.84)$$

donc

$$D(U) J_z D(U)^{-1} = J_z, \quad (14.85)$$

où  $OZ$  est l'axe obtenu après la rotation  $R(U)$  de l'axe  $Oz$ . Cette formule exprime comment se transforme l'observable  $J_z$  sous une rotation : si les états se transforment selon

$$|^R\psi\rangle = D(U) |\psi\rangle$$

alors les observables se transforment selon (14.85).

*Conséquences :*

- ◇ Les valeurs moyennes sont invariantes.

*Preuve :*

$$\langle ^R\psi | J_z | ^R\psi \rangle = \langle \psi | D^\dagger(U) D(U) J_z D^{-1}(U) D(U) | \psi \rangle = \langle \psi | J_z | \psi \rangle.$$

- ◇ Si  $|j m\rangle$  est état propre de  $J_z$ , alors de façon équivalente,  $D(U)|j m\rangle$  est état propre de  $J_Z$ , tout deux avec la valeur propre  $m$ .

*Preuve :*

$$D(U) J_z D^{-1}(U) = J_Z, \text{ i.e. } D(U) J_z = J_Z D(U)$$

soit, en agissant sur l'état  $|j m\rangle$ ,  $D(U) J_z |j m\rangle = J_Z D(U) |j m\rangle$ , d'où

$$m D(U) |j m\rangle = J_Z D(U) |j m\rangle,$$

ce qui montre que  $D(U) |j m\rangle$  est état propre de  $J_Z$ , de valeur propre  $m$ .

Ces deux résultats traduisent le fait que si l'on tourne à la fois les états et les axes, il ne se passe évidemment rien : pour matérialiser une rotation, il faut soit faire tourner un état en ne touchant pas au repère (point de vue actif), soit faire tourner le repère sans toucher au système (point de vue passif), mais certainement pas faire les deux opérations en même temps...!

La relation (14.83) exprime donc comment se transforment les vecteurs d'états quand on change l'axe de quantification.

### 14.8.2 Notion d'opérateur tensoriel

La relation (14.85) donne un autre point de vue sur l'équation (14.43) : le membre de gauche traite  $\vec{J}$  comme un *opérateur* se transformant sous l'action d'une rotation, tandis que le membre de droite traduit le fait que cet opérateur possède 3 composantes se transformant comme un *vecteur* sous l'action du groupe des rotations. On appelle  $\vec{J}$  un *opérateur vectoriel*. En termes des composantes standards ( $J_-, J_0 = J_3, J_+$ ), on écrira la relation (14.43) sous la forme

$$D(U) J_m D^{-1}(U) = \sum_{m'} J_{m'}^j \mathcal{D}_{m'm}^j(U), \quad (14.86)$$

qui est une réalisation particulière (correspondant à  $j = 1$ ) de la notion plus générale d'*opérateur tensoriel irréductible de spin  $j$*  (dans le cas général, un tel opérateur possède  $2j+1$  composantes).

### 14.8.3 Exemples

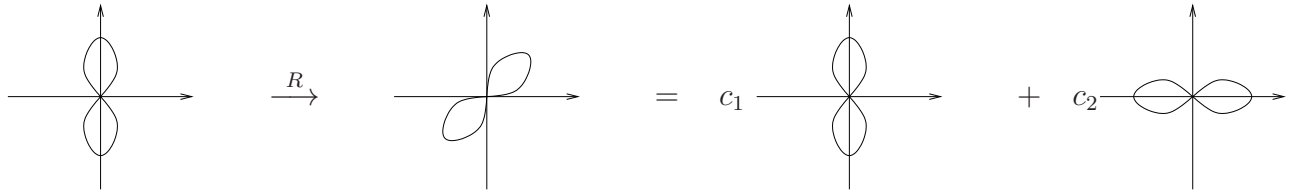
#### Rotation agissant sur une harmonique sphérique

L'état  $\langle \vec{r} | \ell m \rangle$  est un état propre de  $L^2$ ,  $L_z$ . Cet état  $\psi_{\ell m}(\vec{r})$  est donné pour sa partie angulaire par l'harmonique sphérique  $Y_{\ell m}(\theta, \varphi)$  où  $\theta, \varphi$  sont les angles sphériques du vecteur  $\vec{n} = \frac{\vec{r}}{r}$ . On posera par définition

$$\langle \vec{n} | \ell m \rangle = Y_{\ell m}(\theta, \varphi) = Y_{\ell m}(\vec{n}). \quad (14.87)$$

Étudions la loi de transformation des harmoniques sphériques sous l'action du groupe des rotations. Partons de la loi de transformation des états

$$D(U) | \ell m \rangle = \sum_{m'} | \ell m' \rangle D_{m'm}^{\ell}(U).$$


 FIGURE 14.5 – Rotation de l'orbitale  $p$  d'un atome.

Il vient

$$\langle \vec{n} | D(U) | \ell m \rangle = \sum_{m'} Y_{\ell m'}(\vec{n}) D_{m' m}^{\ell}(U).$$

Déterminons  $\langle \vec{n} | D(U)$ . La relation

$${}^R\psi(\vec{r}) = \psi(R^{-1}\vec{r}) \quad (14.88)$$

est équivalente à

$$\forall \psi, \langle \vec{r} | {}^R\psi \rangle = \langle R^{-1}\vec{r} | \psi \rangle, \text{ i.e. } \langle \vec{r} | D(U) | \psi \rangle = \langle R^{-1}\vec{r} | \psi \rangle,$$

ce qui montre que

$$\langle \vec{n} | D(U) = \langle R^{-1}\vec{n} |.$$

Ainsi  $\langle R^{-1}\vec{n} | \ell m \rangle = \sum_{m'} Y_{\ell m'}(\vec{n}) D_{m' m}^{\ell}(U)$ , soit

$$\boxed{Y_{\ell m}(R^{-1}\vec{n}) = \sum_{m'} D_{m' m}^{\ell}(U) Y_{\ell m'}(\vec{n})}. \quad (14.89)$$

Physiquement cette relation exprime par exemple comment se transforment les orbitales quand on fait tourner un atome (par exemple une orbitale  $p$ ), comme illustré dans la fig. 14.5. Seuls les états ayant un moment angulaire donné interviennent dans la décomposition.

### Rotation agissant sur un spineur

Pour une fonction d'onde scalaire, l'opérateur de rotation correspondant à une rotation active d'angle  $\theta$  et d'axe  $\vec{n}$  s'écrit

$$D(U) = \exp(-i\theta \vec{L} \cdot \vec{n}).$$

Certains états physiques, décrivant des systèmes possédant un spin intrinsèque  $\frac{1}{2}$ , se transforment selon la représentation de générateurs<sup>1</sup>

$$\vec{J} = \vec{L} + \frac{\vec{\sigma}}{2} \quad \text{agissant dans } \mathcal{H}_r \otimes \mathcal{H}_s. \quad (14.90)$$

1. Cette écriture est symbolique et doit être comprise comme  $\vec{J} = \vec{L} \otimes \mathbf{1} + \mathbf{1} \otimes \frac{\vec{\sigma}}{2}$ , puisque cet opérateur agit dans  $\mathcal{H}_r \otimes \mathcal{H}_s$ .

L'état transformé par rotation s'écrira donc

$$|^R\psi\rangle = \exp\left[-i\theta\left(\vec{L} + \frac{\vec{\sigma}}{2}\right) \cdot \vec{n}\right] |\psi\rangle.$$

$\vec{L}$  et  $\vec{\sigma}$  opérant dans deux espaces différents,  $|\psi\rangle$  s'écrira

$$|\psi\rangle = \sum_{\sigma} \left|\frac{1}{2}, \sigma\right\rangle |\psi_{\sigma}\rangle \in \mathcal{H}_s \otimes \mathcal{H}_r$$

où les états  $|\frac{1}{2}, \sigma\rangle$  sous-tendent la représentation spinorielle. On définit les fonctions d'onde  $\psi_{\sigma}(r)$  par la relation

$$\begin{aligned} \langle \vec{r} \left| \frac{1}{2} \sigma' \right| \psi \rangle &= \sum_{\sigma} \langle \frac{1}{2} \sigma' \left| \frac{1}{2} \sigma \right\rangle \langle \vec{r} | \psi_{\sigma} \rangle \\ &= \sum_{\sigma} \delta_{\sigma\sigma'} \psi_{\sigma}(\vec{r}) = \psi_{\sigma'}(\vec{r}). \end{aligned}$$

Déterminons la loi de transformation des fonctions d'onde sous l'action du groupe des rotations :

$$\begin{aligned} \langle \vec{r} \left| \frac{1}{2} \sigma' \right| ^R\psi \rangle &= {}^R\psi_{\sigma'}(\vec{r}) = \langle \vec{r} \left| \frac{1}{2} \sigma' \right| \exp\left[-i\theta\left(\vec{L} + \frac{\vec{\sigma}}{2}\right) \cdot \vec{n}\right] |\psi\rangle \\ &= \sum_{\sigma} \langle \vec{r} \left| \frac{1}{2} \sigma' \right| \exp\left[-i\theta\left(\vec{L} + \frac{\vec{\sigma}}{2}\right) \cdot \vec{n}\right] \left| \frac{1}{2} \sigma \right\rangle |\psi_{\sigma}\rangle \\ &= \sum_{\sigma} \langle \frac{1}{2} \sigma' \left| \exp\left(-i\theta \frac{\vec{\sigma}}{2} \cdot \vec{n}\right) \right| \frac{1}{2} \sigma \right\rangle \langle \vec{r} | \exp(-i\theta \vec{L} \cdot \vec{n}) | \psi_{\sigma} \rangle \end{aligned}$$

d'où

$$\boxed{{}^R\psi_{\sigma'}(\vec{r}) = \sum_{\sigma} D_{\sigma'\sigma}^{1/2}(U) \psi_{\sigma}(R^{-1}\vec{r})}. \quad (14.91)$$

### Généralisation

Une particule de spin  $s$  se transforme selon la représentation  $\mathcal{D}_{m'm}^s(U)$ . Elle est décrite par une fonction d'onde à  $2s + 1$  composantes.

*Exemple :*

Une particule de spin 1 est un champ de vecteur dont les 3 composantes standards  $\psi_{10}(\vec{r})$ ,  $\psi_{11}(\vec{r})$  et  $\psi_{1-1}(\vec{r})$  correspondent aux 3 composantes d'un champ de vecteur  $V_i(\vec{r})$  ( $i = 1, 2, 3$ ) dont on sait qu'il se transforme selon

$${}^R V_i(\vec{r}) = R_{ij} V_j(R^{-1}\vec{r}).$$

## 14.9 Représentation à une phase près de $SO(3)$

Nous avons vu dans la section 14.3 qu'aux deux éléments  $U$  et  $-U$  de  $SU(2)$  correspondent la même rotation par morphisme. Cela a une conséquence immédiate sur les représentations de  $SU(2)$  et de  $SO(3)$ . Montrons que

$$\boxed{\mathcal{D}^j(-U) = (-1)^{2j} \mathcal{D}^j(U) .} \quad (14.92)$$

*Preuve :*

En utilisant la paramétrisation (14.71) de  $SU(2)$  en terme des angles d'Euler, on constate que la transformation  $U \rightarrow -U$  peut s'obtenir par le remplacement  $\gamma \rightarrow \gamma + 2\pi$ . En utilisant la relation (14.77), on a donc

$$\mathcal{D}_{mm'}^j(U) = e^{-im\alpha - im'\gamma} d_{mm'}^j(\beta)$$

qui se transforme en

$$\mathcal{D}_{mm'}^j(-U) = e^{-im\alpha} e^{-im'\gamma} d_{mm'}^j(\beta) e^{-im'2\pi} .$$

Or d'après l'analyse du spectre des représentations irréductibles  $D^j$  de  $SU(2)$  (voir 14.5.1),  $2m$  a même parité que  $2j$ , d'où le résultat.

*Conséquences :*

- ◇ Pour  $j$  entier,  $\mathcal{D}^j(U) = \mathcal{D}^j(-U)$ . Donc  $\mathcal{D}^j(U)$  forme une représentation de  $SO(3)$ .
- ◇ Pour  $j$  demi-entier  $\mathcal{D}^j(U) = -\mathcal{D}^j(-U)$ . Donc  $\mathcal{D}^j(U)$  ne forme pas une représentation de  $SO(3)$ .

Toutefois, dans le cas où  $j$  est demi-entier, en restreignant l'intervalle de variation des paramètres, on peut construire une représentation à une phase près de  $SO(3)$  : nous allons le montrer explicitement dans le cas  $j = \frac{1}{2}$ , pour lequel

$$\mathcal{D}^j(U) = U .$$

Soit une rotation d'angle  $\varphi$  ( $0 \leq \varphi \leq \pi$ ) autour de l'axe  $\vec{k}$ . On lui associe

$$\mathcal{D}(U) = U_{\vec{k}}(\varphi) = \cos \frac{\varphi}{2} - i\vec{\sigma} \cdot \vec{k} \sin \frac{\varphi}{2} = \begin{pmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix}$$

(donc  $-U$  ne peut apparaître puisque  $0 \leq \varphi \leq \pi$ ).

Effectuons deux rotations successives d'angles  $\varphi_1$ ,  $\varphi_2$  autour de l'axe  $\vec{k}$ .

- ◇ Si  $0 \leq \varphi_1 + \varphi_2 \leq \pi$ , à la rotation  $R_1 R_2$  sera associé

$$U_{\vec{k}}(\varphi_1 + \varphi_2) = \begin{pmatrix} e^{-\frac{i}{2}(\varphi_1 + \varphi_2)} & 0 \\ 0 & e^{\frac{i}{2}(\varphi_1 + \varphi_2)} \end{pmatrix} = U(\varphi_1)U(\varphi_2)$$

et donc

$$U(R_1 R_2) = U(R_1)U(R_2) .$$

- ◇ Si  $\varphi_1 + \varphi_2 > \pi$ , à la rotation  $R_1 R_2$  (qui est en fait une rotation d'angle  $2\pi - (\varphi_1 + \varphi_2)$  autour de l'axe  $-\vec{k}$ ) on associe

$$\begin{aligned} U_{-\vec{k}}(2\pi - (\varphi_1 + \varphi_2)) &= \cos\left(\pi - \frac{\varphi_1 + \varphi_2}{2}\right) + i\vec{\sigma} \cdot \vec{k} \sin\left(\pi - \frac{\varphi_1 + \varphi_2}{2}\right) \\ &= -\left[\cos\frac{\varphi_1 + \varphi_2}{2} - i\vec{\sigma} \cdot \vec{k} \sin\frac{\varphi_1 + \varphi_2}{2}\right] \\ &= -U_{\vec{k}}(\varphi_1)U_{\vec{k}}(\varphi_2) \end{aligned}$$

donc

$$U(R_1 R_2) = -U(R_1)U(R_2).$$

La généralisation au cas de deux rotations d'axes quelconques (non identiques *a priori*) est aisée en utilisant la paramétrisation d'Euler. On choisit pour axe  $Oz$  l'axe de la première rotation d'angle  $\varphi_1$ . On a alors  $U(R_1) = U_{\vec{k}}(\varphi_1)$  et  $U(R_2) = U(\alpha, z) U(\beta, y) U(\gamma, z)$ . En utilisant le fait que la paramétrisation d'Euler est invariante sous la transformation

$$(\alpha \rightarrow \alpha + \pi, \beta \rightarrow -\beta, \gamma \rightarrow \gamma - \pi),$$

on peut restreindre le domaine de variation de  $\gamma$  à  $[0, \pi]$ . On définit alors  $U(\gamma, z) U(\varphi_1, z)$  en utilisant la règle définie ci-dessus, d'où l'on déduit que l'on a construit une représentation de  $SO(3)$  à une phase près (qui est  $\pm 1$  ici).

Le fait de choisir une rotation d'axe  $\vec{k}$  pour la première rotation ne joue bien sûr aucun rôle. On pourra s'en convaincre en utilisant la propriété (14.33) sous la forme

$$U(\varphi, \vec{n}_1) = V U(\varphi, \vec{k}) V^{-1}$$

avec  $\vec{n}_1 = R(V)\vec{k}$ , et en écrivant la seconde rotation sous la forme

$$U(R_2) = V U(\alpha', z) U(\beta', y) U(\gamma', z) V^{-1},$$

puisque  $V^{-1} U(R_2) V$  est un élément de  $SU(2)$  qui peut donc s'écrire dans la représentation d'Euler, ce qui permet ainsi de ramener la discussion au cas précédent.

# Chapitre 15

## Représentations spinorielles de $SU(2)$

La méthode tensorielle est l'une des méthodes usuelles pour construire les représentations d'un groupe. Elle consiste à partir de la (les) représentation(s) fondamentale(s) du groupe, puis à construire des représentations de dimensions plus élevées par produit tensoriel, dont il s'agit ensuite d'extraire les représentations irréductibles. Dans ce chapitre, nous allons illustrer cette méthode sur le cas très important physiquement du groupe  $SU(2)$ .

### 15.1 Spineurs

Comme nous l'avons vu au chap. 14, un spineur est par définition un vecteur de  $\mathbb{C}^2$ , espace sur lequel agit la représentation fondamentale (appelée aussi spinorielle) de  $SU(2)$ .

#### 15.1.1 Action contravariante

Pour tout  $U$  dans  $SU(2)$  on note  $U^\alpha_\beta$  les éléments de matrice de  $U$ , de sorte que  $D^{1/2}(U) = U$  agit sur les spineurs  $\xi = \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix}$ ,  $\xi \in \mathbb{C}^2$  selon la relation matricielle usuelle

$$\xi^\alpha \rightarrow U^\alpha_\beta \xi^\beta = \xi'^\alpha \quad (15.1)$$

qui décrit l'action de  $U$  sur le spineur *contravariant*  $\xi^\alpha$ .

#### 15.1.2 Métrique

La matrice

$$C_{\alpha\beta} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in SU(2), \quad (15.2)$$

dont nous avons déjà étudié les propriétés en (14.13), permet de définir le produit scalaire de 2 spineurs :

$$\langle \phi, \xi \rangle \equiv \phi^T C \xi = \phi^\alpha C_{\alpha\beta} \xi^\beta = \phi^1 \xi^2 - \phi^2 \xi^1 = \begin{vmatrix} \phi^1 & \xi^1 \\ \phi^2 & \xi^2 \end{vmatrix}.$$

Ce produit scalaire est invariant sous les rotations  $R(U)$ .

*Preuve :*

$$\langle \phi, \xi \rangle = \begin{vmatrix} \phi^1 & \xi^1 \\ \phi^2 & \xi^2 \end{vmatrix} \xrightarrow{U} \left| U \begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} U \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} \right| = \det U \begin{vmatrix} \phi^1 & \xi^1 \\ \phi^2 & \xi^2 \end{vmatrix}. \quad (15.3)$$

Comme  $\det U = 1$ ,  $\langle \phi, \xi \rangle$  est invariant (noter l'invariance sous  $SL(2, \mathbb{C})$ , donc plus forte que simplement sous  $SU(2, \mathbb{C})$ ).

Cette invariance ne doit pas nous surprendre :

$$\langle \phi, \xi \rangle = \phi^T C \xi \xrightarrow{U} \phi^T U^T C U \xi.$$

Or d'après la relation (14.16),

$$U^T C U = C,$$

d'où l'invariance immédiate de ce produit scalaire,  $\forall U \in SL(2, \mathbb{C})$ .

A une constante multiplicative près,  $C$  est le seul tenseur  $2 \times 2$  invariant sous  $SL(2, \mathbb{C})$ .

*Preuve :*

Il s'agit de résoudre l'équation en  $G$

$$\forall U \in SL(2, \mathbb{C}), U^T G U = G. \quad (15.4)$$

Comme  $SU(2) \subset SL(2, \mathbb{C})$ , on peut se limiter à résoudre l'équation pour  $U \in SU(2)$ . Il suffira alors de vérifier que la solution obtenue est valable sur tout  $SL(2, \mathbb{C})$ .

$$\forall U \in SU(2), U^\dagger U = 1 \quad \text{i.e.} \quad U^T = \bar{U}^{-1},$$

donc (15.4) s'écrit encore  $\bar{U}^{-1} G U = G$ , soit

$$G U = \bar{U} G. \quad (15.5)$$

Soit  $U$  un élément de  $SU(2)$ , paramétré suivant (14.7) sous la forme

$$U = \cos \frac{\theta}{2} - i \vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2}.$$

D'après l'expression des matrices de Pauli (14.8),

$$\bar{\sigma}_1 = \sigma_1, \quad \bar{\sigma}_2 = -\sigma_2, \quad \text{et} \quad \bar{\sigma}_3 = \sigma_3,$$

donc si  $\vec{n} = (n_1, n_2, n_3)$ , en notant  $\tilde{\vec{n}} = (n_1, -n_2, n_3)$ , (notation non orthodoxe qui ne sera pas utilisée en dehors de cette preuve), on aura

$$\bar{U} = \cos \frac{\theta}{2} + i \vec{\sigma} \cdot \tilde{\vec{n}} \sin \frac{\theta}{2}.$$

Cherchons  $G$  sous la forme

$$G = a_0 \mathbf{1} + \vec{a} \cdot \vec{\sigma}.$$



Alors (15.5) s'écrit encore

$$\left( \cos \frac{\theta}{2} + i \vec{\sigma} \cdot \tilde{\vec{n}} \sin \frac{\theta}{2} \right) (a_0 \mathbb{1} + \vec{a} \cdot \vec{\sigma}) = (a_0 \mathbb{1} + \vec{a} \cdot \vec{\sigma}) \left( \cos \frac{\theta}{2} - i \vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} \right)$$

soit explicitement

$$\begin{aligned} & a_0 \cos \frac{\theta}{2} + i a_0 \vec{\sigma} \cdot \tilde{\vec{n}} \sin \frac{\theta}{2} + \vec{a} \cdot \vec{\sigma} \cos \frac{\theta}{2} + i \tilde{\vec{n}} \cdot \vec{a} \sin \frac{\theta}{2} - (\tilde{\vec{n}} \wedge \vec{a}) \cdot \vec{\sigma} \sin \frac{\theta}{2} \\ = & a_0 \cos \frac{\theta}{2} - i a_0 \vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} + \vec{a} \cdot \vec{\sigma} \cos \frac{\theta}{2} - i \vec{n} \cdot \vec{a} \sin \frac{\theta}{2} - (\vec{n} \wedge \vec{a}) \cdot \vec{\sigma} \sin \frac{\theta}{2}. \end{aligned}$$

On en déduit par identification des termes proportionnels à  $\mathbb{1}$  que  $\tilde{\vec{n}} \cdot \vec{a} = -\vec{n} \cdot \vec{a}$  pour tout  $\vec{n}$ , i.e. que  $a_1 = a_3 = 0$ . Ceci implique que  $\tilde{\vec{n}} \wedge \vec{a} = \vec{n} \wedge \vec{a}$ , d'où l'on déduit par identification des termes proportionnels à  $\sigma_i$  ( $i = 1, 2, 3$ ) que  $a_0 = 0$ , ce qui montre que  $G$  est multiple de  $C$ .

D'après la relation (14.17), la solution obtenue pour  $U \in SU(2)$  est bien solution de l'équation (15.4) pour tout  $U \in SL(2, \mathbb{C})$ , ce qui achève la preuve.

*Remarques :*

- ◇ Les éléments d'une matrice sont notés  $U^\alpha_\beta$  à l'exception de la matrice  $C$  qui définit la métrique, dont les éléments de matrice sont notés  $C_{\alpha\beta}$ , ce qui va permettre de définir un calcul covariant cohérent. Les éléments de matrices de  $U^T$  sont notés  $(U^T)^\alpha_\beta = U^\beta_\alpha$ .
- ◇  $\langle \phi, \phi \rangle = 0$  donc on ne peut utiliser ce produit scalaire pour définir une norme sur les spineurs. On verra plus loin comment construire une norme adaptée.
- ◇ Ces spineurs interviennent pour décrire des particules de spin  $\frac{1}{2}$  non relativistes : les fonctions d'onde solution de l'équation de Pauli, qui sera étudiée au chapitre 21, se transforment comme ces spineurs.

## 15.2 Calcul tensoriel

Un calcul tensoriel peut être défini grâce à la métrique  $C$ .

### 15.2.1 Spineur covariant

Étant donné un spineur contravariant  $\xi^\alpha$  se transformant suivant (15.1), on définit le spineur covariant  $\xi_\alpha$  par

$$\boxed{\xi_\alpha = C_{\alpha\beta} \xi^\beta}. \quad (15.6)$$

Par définition

$$C^{\alpha\beta} \equiv (C^{-1})_{\alpha\beta} \quad \text{i.e.} \quad C^{\alpha\beta} C_{\beta\alpha'} = \delta^\alpha_{\alpha'} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

D'après les propriétés de la matrice  $C$  (voir (14.13)),  $(C^{-1})_{\alpha\beta} = -C_{\alpha\beta}$  donc  $C^{\alpha\beta} = (C^{-1})_{\alpha\beta} = -C_{\alpha\beta}$ . Par construction,

$$\boxed{\xi^\alpha = C^{\alpha\beta} \xi_\beta.} \quad (15.7)$$

*Preuve :*

$$C^{\beta'\alpha} \xi_\alpha = C^{\beta'\alpha} C_{\alpha\beta} \xi^\beta = \delta^{\beta'}_\beta \xi^\beta = \xi^{\beta'}.$$

Ainsi, on monte ou on descend les indices à l'aide de  $C^{\alpha\beta}$  ou  $C_{\alpha\beta}$ . En particulier,

$$\boxed{\langle \phi, \xi \rangle = \phi^\alpha \xi_\alpha = \phi_\beta \xi^\beta.} \quad (15.8)$$

*Preuve :*

Puisque  $C_{\alpha\beta} = -C_{\beta\alpha}$ , on a donc

$$\langle \phi, \xi \rangle = \phi^\alpha C_{\alpha\beta} \xi^\beta = \phi^\alpha \xi_\alpha = -\phi^\alpha C_{\beta\alpha} \xi^\beta = -C_{\beta\alpha} \phi^\alpha \xi^\beta = -\phi_\beta \xi^\beta. \quad (15.9)$$

### 15.2.2 Action d'une rotation sur un spineur covariant

$$\boxed{\text{Sous l'action de } SU(2), \xi_\alpha \text{ se transforme comme } \bar{\xi}^\alpha.} \quad (15.10)$$

*Preuve :*

Le spineur contravariant  $\xi^\alpha$  se transforme suivant la relation

$$\xi'^\alpha = U^\alpha_\beta \xi^\beta = U^\alpha_\beta C^{\beta\alpha_1} \xi_{\alpha_1}$$

donc

$$\xi'_{\alpha_2} = C_{\alpha_2\alpha} U^\alpha_\beta C^{-1}_{\beta\alpha_1} \xi_{\alpha_1} = U_{\alpha_2}^{\alpha_1} \xi_{\alpha_1}.$$

D'après les propriétés vues en (14.13),  $CU C^{-1} = (U^T)^{-1}$ . Cette relation s'écrit

$$C_{\alpha_2\alpha} U^\alpha_\beta (C^{-1})_{\beta\alpha_1} = ((U^T)^{-1})^{\alpha_2}_{\alpha_1}, \quad \text{i.e. } U_{\alpha_2}^{\alpha_1} = (U^{-1})^{\alpha_1}_{\alpha_2}.$$

Pour  $U \in SU(2)$ ,  $U^{-1} = U^\dagger$  donc  $(U^T)^{-1} = \bar{U}$  et  $CU C^{-1} = \bar{U}$ , ce qui donne ici

$$C_{\alpha_2\alpha} U^\alpha_\beta C^{-1}_{\beta\alpha_1} = U_{\alpha_2}^{\alpha_1} = \bar{U}^{\alpha_2}_{\alpha_1},$$

d'où  $\xi'_{\alpha_2} = \bar{U}^{\alpha_2}_{\alpha_1} \xi_{\alpha_1}$ . Les  $\bar{U}^{\alpha_2}_{\alpha_1}$  sont les éléments de matrice de  $\bar{U}$ , qui est bien la matrice de transformation de  $\bar{\xi}$ , puisque  $\bar{\xi}'^{\alpha_2} = \bar{U}^{\alpha_2}_{\alpha_1} \bar{\xi}^{\alpha_1}$ .

### 15.2.3 Equivalence des représentations $U$ et $\bar{U}$

Pour tout  $U \in SU(2)$ ,  $CUC^{-1} = \bar{U}$  donc les représentations données par  $U$  et  $\bar{U}$  sont équivalentes. Ainsi

La représentation  $U$  agissant sur les spineurs contravariants  
et  
la représentation  $\bar{U}$  agissant sur les spineurs covariants  
sont équivalentes.

Ceci est très spécifique au groupe  $SU(2)$ . Ce n'est pas vrai en général pour  $SU(N)$ , dès que  $N > 2$ . Par exemple, pour  $SU(3)$  utilisé comme groupe de jauge de QCD, les représentations fondamentales  $3$  et  $\bar{3}$  correspondants aux quarks et aux antiquarks ne sont pas équivalentes. De même, comme on le verra au chapitre ??, ce n'est pas vrai sur  $SL(2, \mathbb{C})$ .

*Remarque :*

On a vu plus haut que  $C$  est la seule métrique invariante sous  $SL(2, \mathbb{C})$ . D'autre part, si  $G$  est la métrique qui fait passer des spineurs contravariants aux spineurs covariants, et que l'on souhaite imposer que les spineurs covariants se transforment sous la représentation complexe conjuguée  $\bar{U}$ , alors si  $\xi' = U \xi$  pour  $\xi$  contravariant,

$$\underbrace{G \xi'}_{\text{spineur covariant}} = \underbrace{G U G^{-1}}_{\bar{U}} \underbrace{G \xi}_{\text{spineur covariant}}$$

donc  $G$  doit satisfaire à la condition  $\bar{U} = G U G^{-1}$ . On retrouve donc la même équation qui nous a servi à définir une métrique invariante sous  $SU(2)$  et l'on en déduit que  $G$  doit être proportionnel à  $C$ .

### 15.2.4 Norme sur les spineurs

La remarque précédente permet de déduire que la seule norme (à une constante multiplicative près) invariante sous  $SU(2)$  et définie positive est

$$\|\xi\|^2 = \xi^T \bar{\xi} = \xi^1 \bar{\xi}^1 + \xi^2 \bar{\xi}^2. \quad (15.11)$$

Elle sert de définition pour la densité de probabilité des particules non relativistes de spin 1/2.

*Preuve (pour le lecteur admirateur de St Thomas) :*

◇ De façon pédestre :

$$\|\xi\|^2 = \xi^T \bar{\xi} = (\xi^T)_\alpha \bar{\xi}^\alpha = \xi^\alpha \bar{\xi}^\alpha,$$

qui se transforme sous  $SU(2)$  suivant

$$\xi'^\alpha \bar{\xi}'^\alpha = U^\alpha_{\alpha'} \xi^{\alpha'} \bar{U}^\alpha_{\alpha_1} \bar{\xi}^{\alpha_1}.$$

$$\text{Or } U^\alpha_{\alpha'} \bar{U}^\alpha_{\alpha_1} = U^\alpha_{\alpha'} (\bar{U}^T)^{\alpha_1}_\alpha = (\bar{U}^T)^{\alpha_1}_\alpha U^\alpha_{\alpha'} = \delta^{\alpha_1}_{\alpha'} \quad \text{donc} \quad \|\xi\|^2 \rightarrow \|\xi'\|^2 = \xi^\alpha \bar{\xi}^\alpha = \|\xi\|^2.$$

- ◇ L'invariance de ce produit scalaire est bien sûr directement liée au fait que  $\xi_\alpha$  se transforme comme  $\bar{\xi}^\alpha$  :

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} = \begin{pmatrix} +\xi^2 \\ -\xi^1 \end{pmatrix}$$

donc

$$\begin{aligned} \|\xi\|^2 &= \xi^1 \bar{\xi}^1 + \xi^2 \bar{\xi}^2 = -\xi_2 \bar{\xi}^1 + \xi_1 \bar{\xi}^2 = -\det \begin{pmatrix} \bar{\xi}^1 & \xi_1 \\ \bar{\xi}^2 & \xi_2 \end{pmatrix} \\ &\xrightarrow{SU(2)} -\det \left[ \bar{U} \begin{pmatrix} \bar{\xi}^1 & \xi_1 \\ \bar{\xi}^2 & \xi_2 \end{pmatrix} \right] = -\det \begin{pmatrix} \bar{\xi}^1 & \xi_1 \\ \bar{\xi}^2 & \xi_2 \end{pmatrix} \end{aligned} \quad (15.12)$$

car  $\det \bar{U} = 1$ .

Le lecteur avisé aura sans doute remarqué que la preuve de l'invariance de  $\|\xi\|^2$  que nous venons de donner est exactement la même que celle présentée en (15.3). En effet, d'après (15.9),

$$\langle \phi, \xi \rangle = \det(\phi^\alpha, \xi^\alpha) = \det(\phi_\alpha, \xi_\alpha)$$

et comme  $\phi_\alpha$  se transforme comme  $\bar{\phi}^\alpha$ , le résultat est immédiat.

## 15.3 Tenseurs spinoriels

### 15.3.1 Tenseurs de rang quelconque

Nous allons à présent généraliser la notion de spineur covariant et contravariant. Nous avons pour le moment manipulé les objets suivants :

$$\begin{aligned} \xi^\alpha &= \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} && \text{spineur contravariant} \\ \xi_\alpha &= \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} && \text{spineur covariant.} \end{aligned}$$

Soit  $\psi^{\alpha\beta}$  un tenseur spinoriel (ou spineur) contravariant de rang 2 (i.e. qui possède 2 indices de type contravariant). Par définition, il se transforme comme  $\xi^\alpha \phi^\beta$ .

Avec  $C_{\alpha\beta}$  on fabrique des spineurs covariants de rang 2 :

$$\psi_{\alpha\beta} = C_{\alpha\alpha_1} C_{\beta\beta_1} \psi^{\alpha_1\beta_1}$$

ou mixte :

$$\psi^\alpha{}_\beta = C_{\beta\beta_1} \psi^{\alpha\beta_1}.$$

La généralisation à des tenseurs de rang quelconque est immédiate : par produits de spineurs contravariants et (ou) covariants on fabrique des tenseurs spinoriels (ou spineurs) de rang quelconque, se transformant suivant les types d'indices qu'il portent (en respectant les positions relatives horizontales et verticales des indices) :

$$\Psi^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n} = U^{\alpha_1}_{\alpha'_1} \dots U^{\alpha_m}_{\alpha'_m} U^{\beta'_1}_{\beta_1} \dots U^{\beta'_n}_{\beta_n} \Psi^{\alpha'_1 \dots \alpha'_m}_{\beta'_1 \dots \beta'_n}. \quad (15.13)$$

### 15.3.2 Tenseurs antisymétriques

Comme  $\alpha = 1, 2$ , les seuls tenseurs antisymétriques sont de rang 2. Ce sont les produits d'une constante par  $C_{\alpha\beta}$ . Or  $C^{\alpha\beta}$  a été défini par  $C^{\alpha\beta} = (C^{-1})_{\alpha\beta}$  de sorte que  $C_{\alpha\beta}C^{\beta\gamma} = \delta_\alpha^\gamma$  est bien le spineur unité, i.e. le spineur symétrique de rang 2 tel que  $\delta_1^1 = \delta_2^2 = 1$  et  $\delta_2^1 = \delta_1^2 = 0$ .

### 15.3.3 Norme sur les tenseurs spinoriels

Le produit scalaire naturel sur les tenseurs spinoriels se définit suivant

$$\|\psi\|^2 = \sum_{\alpha\beta\cdots} |\psi^{\alpha\beta\cdots}|^2 = \psi^T \bar{\psi}$$

par extension immédiate de la définition (15.11).

*Preuve :*

- ◇ Il suffit d'utiliser le fait que  $\bar{\Psi}^{\alpha\beta\cdots}$  se transforme comme  $\Psi_{\alpha\beta\cdots}$ . En écrivant  $\Psi_{\alpha\beta\cdots} = \xi_\alpha \Phi_\beta \cdots$  on a

$$\|\psi\|^2 = (-1)^{\text{rang } \Psi} \det(\bar{\xi}^\alpha, \xi_\alpha) \det(\bar{\Phi}^\alpha, \Phi_\alpha) \cdots$$

d'où le résultat par extension de la preuve (15.12).

- ◇ De façon pédestre,

$$\psi^{\alpha\beta\cdots} \rightarrow U^\alpha_{\alpha_1} U^\beta_{\beta_1} \cdots \psi^{\alpha_1\beta_1\cdots} \quad \text{et} \quad (\psi^T)_{\alpha\beta\cdots} = \psi^{\alpha\beta\cdots}.$$

De plus

$$\bar{\psi}^{\alpha\beta\cdots} \rightarrow \bar{U}^\alpha_{\alpha_1} \bar{U}^\beta_{\beta_1} \cdots \bar{\psi}^{\alpha_1\beta_1\cdots}$$

donc

$$\psi^T \bar{\psi} = \psi^{\alpha\beta\cdots} \bar{\psi}^{\alpha\beta\cdots} \rightarrow \psi'^{\alpha\beta\cdots} \bar{\psi}'^{\alpha\beta\cdots} = U^\alpha_{\alpha_1} U^\beta_{\beta_1} \cdots \psi^{\alpha_1\beta_1\cdots} \bar{U}^\alpha_{\alpha'_1} \bar{U}^\beta_{\beta'_1} \cdots \bar{\psi}^{\alpha'_1\beta'_1\cdots}$$

or  $U^\alpha_{\alpha_1} \bar{U}^\alpha_{\alpha'_1} = U^\alpha_{\alpha_1} U^{\dagger\alpha'_1}_{\alpha} = (U^\dagger U)_{\alpha_1}^{\alpha'_1} = \delta_{\alpha_1}^{\alpha'_1}$ , d'où le résultat.

### 15.3.4 Opérations sur les tenseurs spinoriels

On peut définir deux types d'opérations covariantes sur les tenseurs spinoriels :

- ◇ Multiplication de deux tenseurs de rangs  $m$  et  $n$ , ce qui donne un tenseur de rang  $m+n$ .
- ◇ Contraction de deux indices du même type par le tenseur métrique, ce qui abaisse de rang du tenseur de 2 unités.

*Exemples :*

- ◇ Le tenseur  $C_{\nu\mu} \psi_\lambda^{\nu\rho\sigma\kappa} = \psi_\lambda^{\nu\rho\sigma\kappa}{}_\nu$  est un tenseur de rang 4.

- ◇ Le tenseur  $C_{\nu\mu} \psi^{\mu\nu} = \psi_\nu{}^\nu = \psi^\nu{}_\nu = \psi^\nu_\nu$  est un scalaire (spineur de rang 0).

Comme  $C$  est antisymétrique, la contraction d'un tenseur sur 2 indices symétriques donne le tenseur nul. L'opération de contraction est covariante. C'est la seule qui permette d'abaisser de façon covariante le rang d'un tenseur.

## 15.4 Représentations irréductibles de $SU(2)$

### 15.4.1 Lien avec les tenseurs spinoriels

**Théorème 15.1 :**

Les spineurs complètement symétriques sur tous les indices forment les espaces de représentations irréductibles de  $SU(2)$ .

*Preuve :*

Soit  $u$  un tenseur spinoriel de rang  $n$ , i.e.  $u \in (\mathbb{C}^2)^{\otimes n}$ . Le groupe  $SU(2)$  agit alors de la façon suivante sur ce tenseur spinoriel :

$$\begin{aligned} (D^{1/2})^{\otimes n} : (\mathbb{C}^2)^{\otimes n} &\longrightarrow (\mathbb{C}^2)^{\otimes n} \\ u &\longmapsto u' \\ \text{avec } u'^{\alpha_1 \dots \alpha_n} &= U_{\beta_1}^{\alpha_1} \dots U_{\beta_n}^{\alpha_n} u^{\beta_1 \dots \beta_n} \end{aligned} \quad (15.14)$$

Si  $n > 1$ , cette représentation est réductible. Contractons en effet deux indices arbitraire (par exemple les deux premiers), et considérons le tenseur  $v$  défini par

$$\begin{aligned} v^{\alpha_3 \dots \alpha_n} &= C_{\alpha_1 \alpha_2} u^{\alpha_1 \dots \alpha_n} \\ &= u^{12 \alpha_3 \dots \alpha_n} - u^{21 \alpha_3 \dots \alpha_n} \\ &\sim \underbrace{(\xi^1 \phi^2 - \xi^2 \phi^1)}_{\text{scalaire invariant}} \eta^{\alpha_3} \dots \theta^{\alpha_n}. \end{aligned}$$

L'espace des  $v$  est invariant par  $(D^{1/2})^{\otimes n}$ , donc  $(D^{1/2})^{\otimes n}$  est réductible. Seuls les espaces de tenseurs complètement symétriques n'ont pas de sous-espaces invariants (puisque la contraction d'indices symétriques donne le tenseur nul).

**Théorème 15.2 :**

La dimension de l'espace de représentation des spineurs complètement symétriques d'ordre  $n$  est  $n + 1$ . On symbolise ces tenseurs par la notation  $\Psi^{\{n\}}$ .

*Preuve :*

Considérons une composante donnée du tenseur, par exemple  $\psi^{121}$  pour un tenseur symétrique de rang 3. Seul compte (à cause de la symétrie complète) le nombre d'indices 1 d'une part, 2 d'autre part. Dans cet exemple  $\psi^{121} = \psi^{112} = \psi^{211}$  est la composante où 1 apparaît deux fois et 2 une fois. Plus généralement, toute composante du tenseur  $u^{\alpha_1 \dots \alpha_n}$  se ramène à l'une des  $n + 1$  composantes suivantes :

$$u^{\overbrace{11 \dots 1}^n}, \dots, u^{\overbrace{1 \dots 1}^{n-k} \overbrace{2 \dots 2}^k}, \dots, u^{\overbrace{2 \dots 2}^n},$$

ce qui prouve le résultat.

Méthode pratique pour décomposer n'importe quel tenseur spinoriel de rang  $n$  sur les espaces de représentations irréductibles (qui sont constitués des spineurs complètement symétriques de rang  $n, n-2, n-4, \dots$ ) :

- ◇ On symétrise  $\psi^{\alpha\beta\gamma\dots}$  complètement  $\rightarrow$  tenseur symétrique de rang  $n$ .
- ◇ On contracte le tenseur initial sur un couple d'indices, de toute les façons possibles parmi les divers couples d'indices possibles, et l'on symétrise complètement chacun des tenseurs obtenus  $\rightarrow$  tenseurs symétriques de rang  $n-2$ .
- ◇ On contracte le tenseur initial sur 2 couples d'indices, on symétrise, etc.

On peut symboliser cette procédure par

$$\psi^n = \psi^{\{n\}} + C\psi^{\{n-2\}} + CC\psi^{\{n-4\}} + \dots$$

*Exemple :*

Considérons un tenseur  $\psi^{\alpha\beta}$  (donc  $n=2$ ).

- ◇ On symétrise sur  $\alpha\beta$  :

$$\psi^{\{2\}} = \frac{1}{2}(\psi^{\alpha\beta} + \psi^{\beta\alpha})$$

- ◇ On contracte  $\psi^{\alpha\beta}$  sur 2 indices (ici on n'a pas le choix!). Le tenseur  $C_{\alpha\beta}\psi^{\alpha\beta}$  est automatiquement symétrisé ici car c'est un scalaire. On obtient donc la décomposition

$$\psi^{\alpha\beta} = \underbrace{\frac{1}{2}(\psi^{\alpha\beta} + \psi^{\beta\alpha})}_{\psi^{\{2\}}} + \underbrace{\frac{1}{2}(\psi^{\alpha\beta} - \psi^{\beta\alpha})}_{C^{\alpha\beta} \times \underbrace{\text{scalaire}}_{\psi^{\{0\}}}}.$$

Or  $C^{\alpha\beta}C_{\beta\alpha'} = \delta_{\alpha'}^{\alpha}$  donc  $\psi^{\{0\}}$  s'obtient par contraction de  $\psi^{\alpha\beta} - \psi^{\beta\alpha}$  (ou de  $\psi^{\alpha\beta}$  puisque  $\psi^{\{2\}}$  est symétrique), d'où

$$\psi^{\{0\}} = \frac{1}{2}(\psi^{\alpha'\beta'} - \psi^{\beta'\alpha'}) C_{\alpha'\beta'} = C_{\alpha\beta}\psi^{\alpha\beta}. \quad (15.15)$$

Finalement

$$\psi^{\alpha\beta} = \underbrace{\frac{1}{2}(\psi^{\alpha\beta} + \psi^{\beta\alpha})}_{\psi^{\{2\}}} + \underbrace{C^{\alpha\beta}\psi^{\alpha'\beta'}C_{\alpha'\beta'}}_{\psi^{\{0\}}}. \quad (15.16)$$

Le spineur  $\psi^{\{2\}}$  correspond à la représentation vectorielle tandis que  $\psi^{\{0\}}$  correspond à la représentation scalaire.

### 15.4.2 Lien avec les représentations $D^j$

Les tenseurs symétriques de rang  $2j$  étant obtenus par produit tensoriel de  $2j$  spineurs puis symétrisation, les générateurs  $J_i$  de  $SU(2)$  agissant sur ces tenseurs se définissent également par produit tensoriel, sous la forme :

$$J_i = J_i^{1/2} \otimes 1 \otimes \cdots \otimes 1 + 1 \otimes J_i^{1/2} \otimes 1 \cdots \otimes 1 + \cdots + 1 \otimes \cdots \otimes 1 \otimes J_i^{1/2} \quad (15.17)$$

où  $J_i^{1/2}$  agit sur l'espace  $\mathcal{H}_{1/2}$  de la façon usuelle :

$$\begin{aligned} J_+^{1/2} \xi^1 &= 0, & J_-^{1/2} \xi^1 &= \xi^2, & J_3^{1/2} \xi^1 &= \frac{1}{2} \xi^1, \\ J_+^{1/2} \xi^2 &= \xi^1, & J_-^{1/2} \xi^2 &= 0, & J_3^{1/2} \xi^2 &= -\frac{1}{2} \xi^2. \end{aligned} \quad (15.18)$$

Les générateurs  $J_i$  satisfont les relations de commutation de l'algèbre de Lie de  $SU(2)$  de manière immédiate, par application des relations de commutation des  $J_i^{1/2}$ .

Nous allons maintenant prouver le résultat important suivant, qui fait le lien entre les tenseurs spinoriels et les représentations  $\mathcal{D}(j)$  de  $SU(2)$  :

#### Théorème 15.3 :

Les tenseurs symétriques de rang  $2j$  sous-tendent la représentation de dimension  $2j + 1$  du groupe  $SU(2)$ . Les vecteurs

$$\Phi_{jm} = \sqrt{\frac{(2j)!}{(j+m)!(j-m)!}} \psi_{\underbrace{1 \cdots 1}_{j+m} \underbrace{2 \cdots 2}_{j-m}}$$

forment une base sous-tendant la représentation  $\mathcal{D}(j)$  de  $SU(2)$ .

*Preuve :*

En utilisant les équations (15.18) on obtient

$$J_+ \psi_{\underbrace{1 \cdots 1}_{j+m} \underbrace{2 \cdots 2}_{j-m}} = (j-m) \psi_{\underbrace{1 \cdots 1}_{j+m+1} \underbrace{2 \cdots 2}_{j-m-1}},$$

$$J_- \psi_{\underbrace{1 \cdots 1}_{j+m} \underbrace{2 \cdots 2}_{j-m}} = (j+m) \psi_{\underbrace{1 \cdots 1}_{j+m-1} \underbrace{2 \cdots 2}_{j-m+1}},$$

$$J_3 \psi_{\underbrace{1 \cdots 1}_{j+m} \underbrace{2 \cdots 2}_{j-m}} = [(j+m)\frac{1}{2} - (j-m)\frac{1}{2}] \psi_{\underbrace{1 \cdots 1}_{j+m} \underbrace{2 \cdots 2}_{j-m}} = m \psi_{\underbrace{1 \cdots 1}_{j+m} \underbrace{2 \cdots 2}_{j-m}}.$$

D'après cette dernière relation, on peut poser



$$\Phi_{jm} = c_m \underbrace{\psi_1 \dots 1}_{j+m} \underbrace{2 \dots 2}_{j-m}.$$

Alors

$$J_+ \Phi_{jm} = \sqrt{j(j+1) - m(m+1)} \Phi_{j, m+1} = (j-m) \frac{c_m}{c_{m+1}} \Phi_{j, m+1}$$

et

$$J_- \Phi_{jm} = \sqrt{j(j+1) - m(m-1)} \Phi_{j, m-1} = (j+m) \frac{c_m}{c_{m-1}} \Phi_{j, m-1}$$

qui donne donc les deux relations équivalentes

$$c_m = \frac{\sqrt{j(j+1) - m(m+1)}}{j-m} c_{m+1} = \sqrt{\frac{j+m+1}{j-m}} c_{m+1},$$

$$c_m = \frac{\sqrt{j(j+1) - m(m-1)}}{j+m} c_{m-1} = \sqrt{\frac{j-m+1}{j+m}} c_{m-1}.$$

Si l'on suppose que  $\psi_{2\dots 2}$  est normalisé à 1, alors  $c_{-j} = 1$ . On en déduit que

$$c_m = \sqrt{\frac{(2j)!}{(j+m)!(j-m)!}},$$

d'où le résultat.

### 15.4.3 Retour sur le morphisme $SU(2) \rightarrow SO(3)$

Nous allons réexaminer le morphisme (1) à la lumière de ce que nous avons vu dans ce chapitre. La loi de transformation sous  $SU(2)$  de  $\tilde{x}$  défini par (14.30) s'écrit  $\tilde{x}' = U \tilde{x} U^{-1}$  soit encore

$$\tilde{x}'^{\alpha}_{\beta} = U^{\alpha}_{\gamma} \tilde{x}^{\gamma}_{\delta} (U^{-1})^{\delta}_{\beta} = U^{\alpha}_{\gamma} \bar{U}^{\beta}_{\delta} \tilde{x}^{\gamma}_{\delta} \quad (15.19)$$

car  $U^{-1} = U^+$  donc

$$\tilde{x}'^{\alpha\beta} = C^{\beta\beta'} \tilde{x}'^{\alpha}_{\beta'} \quad (15.20)$$

$$= C^{\beta\beta'} U^{\alpha}_{\gamma} \tilde{x}^{\gamma}_{\delta} (U^{-1})^{\delta}_{\beta'} \quad (15.21)$$

$$= C^{\beta\beta'} U^{\alpha}_{\gamma} C_{\delta\delta'} \tilde{x}^{\gamma\delta'} (U^{-1})^{\delta}_{\beta'} \quad (15.22)$$

or

$$C^{\beta\beta'} C_{\delta\delta'} (U^{-1})^{\delta}_{\beta'} = C_{\delta'\delta} (U^{-1})^{\delta}_{\beta'} C^{\beta'\beta} \quad (15.23)$$

où l'on utilise le fait que ce sont des  $\mathbb{C}$ -nombres et d'autre part que  $C_{\delta'\delta} = -C_{\delta\delta'}$  et  $C^{\beta\beta'} = -C^{\beta'\beta}$ . On peut donc écrire, par définition de  $C^{\beta'\beta}$ , que

$$C^{\beta\beta'} C_{\delta\delta'} (U^{-1})^{\delta}_{\beta'} = C_{\delta'\delta} (U^{-1})^{\delta}_{\beta'} (C^{-1})_{\beta'\beta} = (U^T)^{\delta'}_{\beta} \quad (15.24)$$

et donc comme attendu que

$$\tilde{x}'^{\alpha\beta} = U^\alpha_\gamma \tilde{x}^{\gamma\delta'} (U^T)^{\delta'}_\beta = U^\alpha_\gamma U^\beta_{\delta'} \tilde{x}^{\gamma\delta'}. \quad (15.25)$$

Cette relation de transformation est un exemple de transformation d'un tenseur contravariant de rang 2. D'autre part, elle met en évidence la relation

$$D^{1/2} \otimes D^{1/2} = D^0 \oplus D^1. \quad (15.26)$$

En effet, l'éq. (15.25) s'écrit encore

$$\tilde{x}'^{\alpha\beta} = U^\alpha_\gamma U^\beta_\delta \left[ \underbrace{\frac{1}{2} (\tilde{x}^{\gamma\delta} + \tilde{x}^{\delta\gamma})}_{\tilde{x}'^{\{\gamma\delta\}}} + \underbrace{\frac{1}{2} (\tilde{x}^{\gamma\delta} - \tilde{x}^{\delta\gamma})}_{C^{\gamma\delta} \times \text{scalaire}} \right] \quad (15.27)$$

qui est une illustration de la décomposition (15.16) étudiée plus haut.

Il est intéressant de faire deux remarques, qui nous serviront lors de l'étude du groupe de Lorentz :

◇  $\bar{\xi}^T \xi$  est un scalaire.

◇  $\vec{V} = \bar{\xi}^T \vec{\sigma} \xi$  est un vecteur :  $V'_i = (\bar{\xi}^T \vec{\sigma} \xi)'_i = \bar{\xi}^T U^\dagger \sigma_i U \xi$

or  $U^{-1} \sigma_i U = R_{ij} \sigma_j = U^\dagger \sigma_i U$  donc  $V'_i = \bar{\xi}^T R_{ij} \sigma_j \xi = R_{ij} V_j$ .

Ce second résultat peut s'obtenir à l'aide de l'analyse tensorielle précédente. En effet,

$$V_i = \bar{\xi}^T \sigma_i \xi = \bar{\xi}^T (\sigma_i)^\alpha_\beta \xi^\beta = \bar{\xi}^\alpha (\sigma_i)^\alpha_\beta \xi^\beta. \quad (15.28)$$

Or  $\bar{\xi}^\alpha$  se transforme comme  $\phi_\alpha$  sous  $SU(2)$ , donc

$$V_i \sim \phi_\alpha (\sigma_i)^\alpha_\beta \xi^\beta \sim (\sigma_i)^{\alpha\beta} \zeta_{\beta\alpha}. \quad (15.29)$$

Or

$$(\sigma_i)^{\alpha\beta} C_{\beta\alpha} = (\sigma_i)^\alpha_\alpha = \text{Tr} \sigma_i = 0, \quad (15.30)$$

donc seule la partie *symétrique* de  $\zeta_{\beta\alpha}$  contribue, *q.e.d.*

# Chapitre 16

## Notion de groupe de recouvrement

### 16.1 Homotopie

L'homotopie est une notion fondamentale de topologie algébrique. Elle permet de formaliser la notion de déformation continue lorsque l'on passe d'un objet à un autre. En particulier, deux *lacets* seront dits homotopes s'il est possible de passer continûment de l'un à l'autre.

#### 16.1.1 Chemins

**Définition 16.1 :** *Chemin*

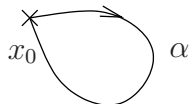
Soit  $X$  un espace topologique.

Une application continue de  $[0, 1]$  dans  $X$  est appelée chemin de point de départ  $x_0$  et de point final  $x_1$  si  $\alpha(0) = x_0$  et  $\alpha(1) = x_1$ .



**Définition 16.2 :** *Boucle*

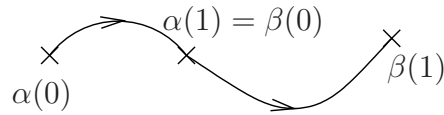
Si  $\alpha(0) = \alpha(1) = x_0$ , le chemin est appelé boucle de point de base  $x_0$ .



**Définition 16.3 :** *Produit de chemins*

Soient  $\alpha, \beta : [0, 1] \rightarrow X$ , des chemins tels que  $\alpha(1) = \beta(0)$ . Le produit de  $\alpha$  et  $\beta$ , noté  $\alpha * \beta$ , est le chemin défini par

$$\alpha * \beta(s) = \begin{cases} \alpha(2s) & 0 \leq s \leq \frac{1}{2} \\ \beta(2s - 1) & \frac{1}{2} \leq s \leq 1 \end{cases}$$



**Définition 16.4 :** *Chemin inverse*

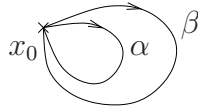
Au chemin  $\alpha : [0, 1] \rightarrow X$  on associe le chemin inverse  $\alpha^{-1}$  tel que  $\alpha^{-1}(s) = \alpha(1 - s)$ .

**Définition 16.5 :** *Homotopie*

Soient  $\alpha, \beta : [0, 1] \rightarrow X$  des boucles de base  $x_0$ . Elles sont dites homotopes (noté  $\alpha \sim \beta$ ) s'il existe une application continue  $f : [0, 1] \times [0, 1] \rightarrow X$  t.q.

$$\diamond f(s, 0) = \alpha(s) \quad ; \quad f(s, 1) = \beta(s) \quad \forall s \in [0, 1]$$

$$\diamond f(0, t) = f(1, t) = x_0 \quad \forall t \in [0, 1]$$



**Proposition 16.6 :**

- $\diamond$  L'homotopie  $\alpha \sim \beta$  est une relation d'équivalence. On note  $[\alpha]$  la classe d'homotopie de la boucle  $\alpha$ .
- $\diamond$  le produit  $*$  entre chemins définit alors naturellement le produit entre classes :  $[\alpha] * [\beta] = [\alpha * \beta]$ .

### 16.1.2 Groupe d'homotopie (ou groupe fondamental)

**Théorème 16.7 :**

Soit  $X$  un espace topologique.

L'ensemble des classes d'homotopie des boucles de point de base  $x_0 \in X$ , noté  $\pi_1(X, x_0)$ , est un groupe. C'est le groupe fondamental (ou premier groupe d'homotopie) de  $X$  en  $x_0$ .

*Exemples :*

$$\diamond \text{ cercle } \mathbb{S}^1 \quad \pi_1(\mathbb{S}^1) \cong \mathbb{Z}$$

$$\diamond \text{ tore } \mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1 \quad \pi_1(\mathbb{T}^2) \cong \mathbb{Z} \oplus \mathbb{Z}$$

◇ plus généralement, en pour le tore en dimension  $n$ ,  $T^n = \underbrace{\mathbb{S}^1 \times \mathbb{S}^1 \times \cdots \times \mathbb{S}^1}_n \quad \pi_1(T^n) \cong \underbrace{\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}}_n$

◇ cylindre  $X = \mathbb{S}^1 \times \mathbb{R} \quad \pi_1(X) \cong \mathbb{Z} \oplus \{e\} \cong \mathbb{Z}$ .

◇ plan projectif réel :  $\pi_1(P^2(\mathbb{R})) \cong \mathbb{Z}/2\mathbb{Z}$ .

**Th. :** soit  $X$  un espace connexe par arc, et  $x_0, x_1 \in X$ .  
Alors  $\pi_1(X, x_0)$  et  $\pi_1(X, x_1)$  sont isomorphes.

**Définition 16.8 :** *Simple connexité*

Un espace  $X$  connexe par arc possédant un groupe fondamental trivial est dit simplement connexe.

### 16.1.3 Topologie des groupes $SU(2)$ et $SO(3)$

#### Groupe $SU(2)$

Soit

$$U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2). \quad (16.1)$$

Alors, puisque  $\alpha\delta - \beta\gamma = 1$ ,

$$U^+ = \begin{pmatrix} \bar{\alpha} & \bar{\gamma} \\ \bar{\beta} & \bar{\delta} \end{pmatrix} \quad U^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \quad (16.2)$$

et donc la condition  $U^+ = U^{-1}$  est équivalente à  $\delta = \bar{\alpha}$  et  $\gamma = -\bar{\beta}$ , d'où l'écriture générique des matrices de  $SU(2)$

$$U = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}. \quad (16.3)$$

De plus, comme  $\det U = 1$ ,  $|\alpha|^2 + |\beta|^2 = 1$ .

Chaque élément de  $SU(2)$  est donc défini par la donnée de deux nombres complexes  $(\alpha, \beta)$ , tels que  $|\alpha|^2 + |\beta|^2 = 1$ , et réciproquement.

Posons  $\begin{cases} \alpha = \alpha_1 + i\alpha_2 \\ \beta = \beta_1 + i\beta_2 \end{cases}$  ; alors  $|\alpha|^2 + |\beta|^2 = 1$  s'écrit  $|\alpha_1|^2 + |\alpha_2|^2 + |\beta_1|^2 + |\beta_2|^2 = 1$ .

Donc  $SU(2)$ , comme espace topologique, est homéomorphe à la sphère unité  $\mathbb{S}^3$  de  $\mathbb{R}^4$ .

La connexité par arc et la simple connexité de  $\mathbb{S}^3$  impliquent donc celle de  $SU(2)$ .

#### Groupe $SO(3)$

A chaque point de  $S^3$  correspond toujours une rotation. Cependant aux points  $(1, 0, 0, 0)$  et  $(-1, 0, 0, 0)$  correspondent l'identité de  $SO(3)$ . Donc une courbe sur la sphère  $S^3$  commençant en  $(1, 0, 0, 0)$  et finissant en  $(-1, 0, 0, 0)$  est en fait une boucle de base  $\mathbf{1}$  dans  $SO(3)$ . Il n'y

a clairement aucune façon de déformer continûment cette boucle en le point  $(1, 0, 0, 0)$ , donc  $SO(3)$  n'est pas simplement connexe. On montre qu'il est doublement connexe :  $\pi_1(SO(3)) \cong \mathbb{Z}/2\mathbb{Z} = \{1, -1\}$ , la rotation d'angle  $2\pi$  étant un représentant de la classe non topologiquement triviale, tandis que la rotation d'angle  $4\pi$  est un représentant de la classe triviale.

## 16.2 Groupe de recouvrement

Avant de parvenir au théorème important concernant le groupe de recouvrement universel d'un groupe, les sous-groupes du centre d'un groupe  $G$  jouent un rôle essentiels. Nous avons vu au paragraphe 4.4.4 que centre d'un groupe, sous-groupe de  $G$  qui contient tous les éléments commutant avec chacun des éléments de  $G$ , est abélien et distingué dans  $G$ . De façon immédiate, on en déduit la propriété suivante.

**Proposition 16.9 :** *Sous-groupe invariant central*

Tout sous-groupe de  $Z$  est abélien et constitue un sous-groupe invariant de  $G$ . Un tel sous-groupe est appelé sous-groupe invariant central de  $G$ .

**Théorème 16.10 :**

Si  $G$  est un groupe de Lie connexe, alors il existe un groupe de Lie simplement connexe  $\tilde{G}$  (qui est unique à isomorphisme près) tel que

a)  $G$  est analytiquement isomorphe à un groupe quotient  $\tilde{G}/K$ , où  $K$  est un sous-groupe invariant central discret de  $\tilde{G}$ , isomorphe au groupe fondamental  $\pi_1(G)$  de la variété sous-jacente à  $G$

b) si  $G$  est lui-même simplement connexe alors  $G$  est isomorphe à  $\tilde{G}$

c) les algèbres de Lie de  $G$  et  $\tilde{G}$  sont isomorphes

d) chaque représentation de l'algèbre de Lie de  $\tilde{G}$  est associée à une représentation du groupe  $\tilde{G}$  suivant la relation

$$\forall a \in \mathfrak{g}, \quad \Gamma_g(a) = \left[ \frac{d}{dt} \Gamma_G(\exp(ta)) \right]_{t=0}$$

$\tilde{G}$  est appelé groupe de recouvrement universel de  $G$ .

Chaque représentation  $\Gamma_{\tilde{G}}$  de  $\tilde{G}$  (cf d)) fournit une représentation  $\Gamma_G$  de  $G$  à l'aide de  $\Gamma_G(TK) = \Gamma_{\tilde{G}}(T)$  ( $\forall T \in G$ ) si et seulement si  $\Gamma_{\tilde{G}}(T) = 1 \forall T \in K$ .

Sinon on obtient une représentation projective de  $G$ .

**Théorème 16.11 :**

Si les algèbres de Lie  $\mathfrak{g}_1$  et  $\mathfrak{g}_2$  de deux groupes de Lie simplement connexes  $\tilde{G}_1$  et  $\tilde{G}_2$  sont isomorphes, alors  $\tilde{G}_1$  et  $\tilde{G}_2$  sont analytiquement isomorphes.

Conséquence de ces deux théorèmes :

Pour chaque algèbre de Lie réelle  $\mathfrak{g}$  il existe un groupe de Lie simplement connexe  $\tilde{G}$ , qui est unique (à isomorphisme près), tel que chaque groupe de Lie  $G$  ayant pour algèbre de Lie

isomorphe  $g$  est isomorphe à  $\tilde{G}/K$ , où  $K$  est un sous-groupe invariant central discret de  $\tilde{G}$ .  $\tilde{G}$  est alors appelé groupe de recouvrement universel de  $g$ .

**Théorème 16.12 :**

Si  $g_1$  et  $g_2$  sont deux algèbres de Lie réelles de groupe de recouvrement universels  $\tilde{G}_1$  et  $\tilde{G}_2$ , alors le groupe de recouvrement universel de  $g_1 \oplus g_2$  est isomorphe à  $\tilde{G}_1 \otimes \tilde{G}_2$ .

*Ex.* : application du th. 2 à  $SU(N)$

$SU(N)$  ( $N = 2, 3, 4, \dots$ ) est simplement connexe

Déterminons le centre de  $SU(N)$  noté  $Z$

Comme l'ensemble des matrices  $N \times N$  unitaires de déterminant 1 forme une représentation unitaire irréductible de  $SU(N)$ , d'après le lemme de Schur  $u \in Z \Rightarrow u = \alpha \mathbf{1}_N$ .  $\det u = 1$  donc  $\alpha^N = 1$ . Donc  $Z$  est le groupe fini d'ordre  $N$  d'éléments  $e^{2\pi \frac{ip}{N}} \mathbf{1}_N$ ,  $p = 0, \dots, N-1$

$N = 2$  :  $Z = \{\mathbf{1}_2, -\mathbf{1}_2\}$  groupe de Lie compact d'algèbre de Lie isomorphe à

$SU(2)$  :  $SU(2)$  et  $SU(2)/Z \simeq SO(3)$

$N = 3$  :  $Z = \{\mathbf{1}_3, e^{2\pi i/3} \mathbf{1}_3, e^{4\pi i/3} \mathbf{1}_3\}$   $SU(3)$  et  $SU(3)/Z$

$N = 4$  :  $Z = \{\mathbf{1}_4, i\mathbf{1}_4, -\mathbf{1}_4, -i\mathbf{1}_4\}$   $SU(4)$   $SU(4)/Z$   $SU(4)/K$  avec  $K = \{\mathbf{1}_4, -\mathbf{1}_4\}$

**Application à la mécanique quantique**

Toute représentation à une phase près continue unitaire de  $SO(3)$  est une représentation vraie, continue, unitaire de son groupe de recouvrement universel  $SU(2)$ .

Par conséquent pour tout système qui admet les rotations comme groupe de symétrie, les vecteurs de l'espace de Hilbert se transforment non pas selon une représentation unitaire de  $SO(3)$  mais selon une représentation unitaire de  $SU(2)$ .

rep. unitaire	$SO(3) \rightarrow$ spins entiers
	$SU(2) \rightarrow$ spins entiers et demi-entiers

Ainsi l'existence de spins demi-entiers est-il conséquence directe du théorème de Wigner.





Cinquième partie  
Symétries relativistes



# Chapitre 17

## Le groupe de Lorentz

Les équations de Maxwell qui rendent compte des phénomènes électriques et magnétiques ne sont pas invariantes sous le groupe de Galilée. Donc contrairement à ce que stipule le principe de relativité galiléenne, il devrait être possible de mettre en évidence un référentiel absolu (éther). Les expériences de Michelson et Morley (1881) ont infirmé cette hypothèse : la vitesse de la lumière est absolument indépendante de la direction de propagation. Le temps ne peut plus être considéré comme absolu : la simultanéité de deux phénomènes dans un référentiel n'implique pas cette simultanéité dans un autre référentiel, contrairement à la mécanique classique.

### 17.1 Intervalle d'espace-temps

Nous allons brièvement rappeler ici la notion d'intervalle d'espace-temps en relativité restreinte. Considérons tout d'abord un événement décrit par un observateur lié à un référentiel inertiel par ses coordonnées d'espace-temps  $(t_1, \vec{x}_1)$ .

*Exemple* : émission d'un photon dans la désintégration

$$\pi^0 \rightarrow \gamma\gamma$$

Le signal se propage à la vitesse de la lumière (photon). Il est détecté en  $(t_2, \vec{x}_2)$ .

Par définition l'intervalle entre ces événements est

$$(\Delta s)^2 = c^2 (t_2 - t_1)^2 - (\vec{x}_2 - \vec{x}_1)^2 ; \quad \text{ici } (\Delta s)^2 = 0.$$

$c$  ne dépend pas du référentiel donc  $(\Delta s)^2 = 0$  dans **tous** les référentiels inertiels par rapport auxquels on peut décrire cette expérience.

L'intervalle infinitésimal s'écrit

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2. \quad (17.1)$$

Si  $ds^2 = 0$  dans un référentiel d'inertie  $K$ , alors  $ds'^2 = 0$  dans tout autre référentiel d'inertie  $K'$ .

Donc  $ds^2 = a ds'^2$

En utilisant l'homogénéité du temps et de l'espace et l'isotropie de l'espace, on montre que  $a = \text{constante} = 1$ .

### exercice 17.1

le montrer (cf. Landau T.C. p. 12).

Un point de l'espace de Minkowski  $M$  est repéré par ses coordonnées dites **contravariantes**  $x^\mu = (ct, \vec{x})$ .

**Métrie** : elle est donnée par la matrice

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \quad (17.2)$$

donc

$$(\Delta s)^2 = g_{\mu\nu} (x_2^\mu - x_1^\mu) (x_2^\nu - x_1^\nu),$$

qui est bien identique, avec cette définition de la métrique, à la notion d'intervalle introduite plus haut.

On définit 3 types d'intervalles selon le signe de  $(\Delta s)^2$  :

$(\Delta s)^2 > 0$  intervalle de genre temps

$(\Delta s)^2 < 0$  intervalle de genre espace

$(\Delta s)^2 = 0$  intervalle de genre lumière

### exercice 17.2

Montrer, en reprenant l'exemple de l'émission d'une particule dans la désintégration  $\beta^-$  :  $m \rightarrow p + e^- + \bar{\nu}$  que l'intervalle séparant l'émission et la détection d'une particule est du genre temps ou lumière.

## 17.2 Groupe de Poincaré et groupe de Lorentz

Ayant postulé que la vitesse de la lumière est indépendante du référentiel, nous en avons déduit que l'intervalle  $ds^2$  est identique dans tous les référentiels inertiels. Il nous faut à présent caractériser ces référentiels d'inertie, en déterminant les transformations relativistes reliant ces référentiels entre eux. Ces transformations possèdent une structure de groupe de symétrie d'espace-temps relativiste que nous allons étudier dans ce chapitre et dans les deux suivants.

### 17.2.1 Caractérisation de $\mathcal{P}$

Soient  $x$  les coordonnées d'un point de  $M$ , mesurées dans le référentiel  $K$ , et  $x'$  les coordonnées du même point mesurées dans le référentiel  $K'$ .

La condition d'invariance de l'intervalle infinitésimal

$$ds'^2 = ds^2 \tag{17.3}$$

implique que

translation (4 paramètres réels)

$x'^\alpha = \Lambda^\alpha_\beta x^\beta + a^\alpha$  : groupe de Poincaré (10 paramètres réels)

groupe de Lorentz (6 paramètres réels)

(17.4)

*Preuve :*

La condition (17.3) s'écrit

$$g_{\alpha\beta} dx'^\alpha dx'^\beta = g_{\mu\nu} dx^\mu dx^\nu,$$

i.e.

$$g_{\alpha\beta} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu$$

donc

$$g_{\alpha\beta} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} = g_{\mu\nu}. \tag{17.5}$$

En prenant le déterminant membre à membre, on en tire donc

$$\det g \left( \det \frac{\partial x'}{\partial x} \right)^2 = \det g.$$

$g$  étant régulière, on en déduit que  $\det \frac{\partial x'}{\partial x} \neq 0$  : la transformation de Poincaré est donc inversible.

En faisant agir  $\frac{\partial}{\partial x^\rho}$  sur (17.5) on obtient

$$g_{\alpha\beta} \frac{\partial^2 x'^\alpha}{\partial x^\mu \partial x^\rho} \frac{\partial x'^\beta}{\partial x^\nu} + g_{\alpha\beta} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial^2 x'^\beta}{\partial x^\nu \partial x^\rho} = 0$$

qui s'écrit symboliquement

$$A_{(\mu\rho)\nu} + A_{(\nu\rho)\mu} = 0, \quad (17.6)$$

où  $A$  est symétrique sur les indices entre parenthèses.

$$\mu \leftrightarrow \rho \quad \text{donne} \quad A_{(\mu\rho)\nu} + A_{(\mu\nu)\rho} = 0 \quad (17.7)$$

$$\nu \leftrightarrow \rho \quad \text{donne} \quad A_{(\mu\nu)\rho} + A_{(\nu\rho)\mu} = 0. \quad (17.8)$$

La combinaison (17.6) + (17.7) -(17.8) donne donc

$$g_{\alpha\beta} \frac{\partial^2 x'^\alpha}{\partial x^\mu \partial x^\rho} \frac{\partial x'^\beta}{\partial x^\nu} = 0$$

$$\text{d'où (puisque } \det \frac{\partial x'}{\partial x} \neq 0) : \frac{\partial^2 x'^\alpha}{\partial x^\mu \partial x^\rho} = 0$$

ce qui prouve que les  $x'$  sont des fonctions linéaires des  $x$ .

On utilise les mêmes notations matricielles que celle vue dans le chapitre ?? pour les matrices de  $SU(2)$ , i.e.

$$\Lambda^\alpha_\beta \quad \begin{array}{l} \swarrow \text{indice de ligne} \\ \nwarrow \text{indice de colonne} \end{array}$$

On déduit de (17.5) et de (17.4) que  $\Lambda$  est une matrice réelle vérifiant

$$\boxed{g_{\alpha\beta} \Lambda^\alpha_\mu \Lambda^\beta_\nu = g_{\mu\nu}} \quad (17.9)$$

Le groupe de Lorentz est bien un groupe :

*Preuve :*

- c'est immédiat d'après sa définition : le produit de 2 transformations homogènes qui préservent  $ds^2$  est une transformation homogène qui préserve  $ds^2$ ; d'autre part, l'inverse d'une transformation homogène préservant  $ds^2$  est une transformation homogène préservant  $ds^2$ .

- algébriquement, si  $\Lambda$  et  $\Lambda'$  sont deux éléments de  $L$ , alors

$$\begin{cases} g_{\alpha\beta} \Lambda^\alpha_\mu \Lambda^\beta_\nu = g_{\mu\nu} \\ g_{\alpha\beta} \Lambda'^\alpha_\mu \Lambda'^\beta_\nu = g_{\mu\nu}. \end{cases}$$

On en déduit que

$$g_{\alpha\beta} \Lambda^\alpha_\mu \Lambda'^\mu_{\mu'} \Lambda'^\beta_{\nu'} \Lambda'^{\nu'}_{\nu} = g_{\mu\nu} \Lambda'^\mu_{\mu'} \Lambda'^{\nu'}_{\nu} = g_{\mu'\nu'},$$

ce qui montre que  $\Lambda \Lambda' \in L$ .

La première des preuves précédentes s'étend immédiatement au cas du groupe de Lorentz inhomogène (ou groupe de Poincaré), ce qui montre que  $\mathcal{P}$  est un groupe. La seconde preuve

exige de connaître explicitement la loi de groupe de  $\mathcal{P}$ , que nous examinerons au chapitre 19.

### Nombre de paramètres :

$$\boxed{L \text{ dépend de 6 paramètres réels.}} \quad (17.10)$$

*Preuve :*

Considérons une transformation infinitésimale

$$\Lambda^{\mu\nu} = g^{\mu\nu} + \omega^{\mu\nu}, \text{ avec } \|\omega^{\mu\nu}\| \ll 1.$$

La contrainte  $g_{\mu\nu}\Lambda^{\mu\rho}\Lambda^{\nu\sigma} = g^{\rho\sigma}$  donne

$$g_{\mu\nu}(g^{\mu\rho} + \omega^{\mu\rho})(g^{\nu\sigma} + \omega^{\nu\sigma}) = g^{\rho\sigma}$$

soit au premier ordre en  $\omega$ ,

$$g_{\mu\nu}g^{\mu\rho}g^{\nu\sigma} + g_{\mu\nu}g^{\mu\rho}\omega^{\nu\sigma} + g_{\mu\nu}\omega^{\mu\rho}g^{\nu\sigma} = g^{\rho\sigma}$$

d'où

$$\delta^\rho_\nu g^{\nu\sigma} + \delta_\nu^\rho \omega^{\nu\sigma} + \delta_\mu^\sigma \omega^{\mu\rho} = g^{\rho\sigma} + \omega^{\rho\sigma} + \omega^{\sigma\rho} = g^{\rho\sigma}$$

et donc

$$\omega^{\rho\sigma} + \omega^{\sigma\rho} = 0. \quad (17.11)$$

$\omega^{\mu\nu}$  est donc un tenseur réel  $4 \times 4$  antisymétrique : il dépend de 6 paramètres réels.

$$\boxed{\mathcal{P} \text{ dépend de 10 paramètres réels :}} \quad (17.12)$$

il dépend des 6 paramètres réels décrivant  $L$  auxquels il faut ajouter les 4 translations d'espace-temps.

### 17.2.2 Structure du groupe de Lorentz

De (17.9) on déduit que  $\det \Lambda = \pm 1$  :

$$\left| \begin{array}{l} \det \Lambda = +1 : \text{ transformations propres} \\ \det \Lambda = -1 : \text{ transformations impropres} \end{array} \right.$$

En considérant à nouveau (17.9) avec  $\mu = \nu = 0$ , on obtient  $(\Lambda^0_0)^2 - \sum_{i=1}^3 (\Lambda^i_0)^2 = 1$  :

$$\left| \begin{array}{ll} \Lambda^0_0 \geq 1 : & \text{pas de renversement du temps (orthochrone)} \\ \Lambda^0_0 \leq -1 : & \text{avec renversement du temps} \end{array} \right.$$

On démontre (voir Gelfand p165) que

$\boxed{\text{Le groupe de Lorentz } O(3,1) \text{ possède 4 composantes connexes. Chacune est doublement connexe. mais non simplement connexe.}}$

Ces différentes composantes connexes sont reliées par des transformations discrètes :

- transformation de réflexion (ou parité) :

$$P = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

- transformation de renversement du temps :

$$T = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

figure

$L$  laisse invariante les 3 régions  $x^2 > 0$ ,  $x^2 = 0$  (cône de lumière) et  $x^2 < 0$

$L_+^\uparrow$  laisse invariante les 4 régions  $\left| \begin{array}{l} x^2 > 0 \\ x_0 > 0 \end{array} \right|$ ,  $\left| \begin{array}{l} x^2 > 0 \\ x_0 < 0 \end{array} \right|$ ,  $x^2 = 0$ ,  $x^2 < 0$

figure

$L_+^\uparrow$  est le groupe de Lorentz **restreint**

$L_+^\uparrow$  et  $L_-^\uparrow$  forment le groupe de Lorentz **complet**



$L_+^\uparrow$  et  $L_+^\downarrow$  forment le groupe de Lorentz **propre**

### exercice 17.3

Vérifier que le groupe de Lorentz restreint, le groupe de Lorentz complet et le groupe de Lorentz propre sont bien des sous-groupes du groupe de Lorentz  $L$ .

### Compacité :

Le groupe de Lorentz restreint  $L_+^\uparrow$  est un groupe non compact.

Ceci est dû à l'existence des boost, ou transformation de Lorentz spéciales, que nous examinerons plus loin, qui sont codées par une rapidité variant de  $-\infty$  à  $+\infty$ .

### 17.2.3 Vecteurs contravariants et covariants

On peut rapporter l'espace de Minkowski  $M$  à une base  $(\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4)$  notée  $\underline{e}_\mu$  ( $\mu = 0, \dots, 3$ ).

Un quadrivecteur  $\underline{A}$  arbitraire peut alors se décomposer dans cette base sous la forme :

$$\underline{A} = A^\mu \underline{e}_\mu \quad (\text{convention d'Einstein})$$

Par définition le tenseur métrique (ou métrique) est relié aux vecteurs de base par

$$\underline{e}_\mu \cdot \underline{e}_\nu = g_{\mu\nu}$$

avec  $g_{\mu\nu}$  donné par (17.2). On a alors

$$ds^2 = d\underline{x} \cdot d\underline{x} = dx^\mu dx^\nu \underline{e}_\mu \cdot \underline{e}_\nu = dx^\mu dx^\nu g_{\mu\nu},$$

en accord avec la définition donnée plus haut. Dans toute la suite, sauf mention contraire, on fera  $c = 1$ , et donc  $x^0 = t$ ,  $x^1 = x$ ,  $x^2 = y$ ,  $x^3 = z$ .

### Coordonnées co- et contra-variantes :

$$\underline{A} = A^\mu \underline{e}_\mu \quad \swarrow \text{composante contra-variante} \qquad A_\mu = \underline{A} \cdot \underline{e}_\mu \quad \swarrow \text{composante co-variante}$$

en général  $A^\mu \neq A_\mu$

*Exemple :* plan euclidien

figure

*Remarque* : les  $\underline{e}_\mu$  ne sont pas des quadrivecteurs puisqu'ils ne sont pas invariants lorsque l'on change les axes (alors que par définition un quadrivecteur est un objet à 4 composantes qui est invariant par changement d'axe).

**Lien entre composantes covariantes et contravariantes :**

$$A_\mu = \underline{A} \cdot \underline{e}_\mu = A^\nu \underline{e}_\nu \cdot \underline{e}_\mu = g_{\mu\nu} A^\nu$$

De la même manière que pour la métrique  $C$  de  $SU(2)$ , (voir chapitre ??), on définit  $g^{\mu\nu}$  comme l'inverse de  $g_{\mu\nu}$  :  $g^{\mu\nu} g_{\nu\rho} = \delta_\rho^\mu$ .

Comme  $g_{\mu\nu}$  est son propre inverse,  $g_{\mu\nu} = g^{\mu\nu}$ .

Ainsi

$$\boxed{\begin{aligned} A_\mu &= g_{\mu\nu} A^\nu \\ A^\mu &= g^{\mu\nu} A_\nu. \end{aligned}} \quad (17.13)$$

On peut donc écrire  $ds^2 = g_{\mu\nu} dx^\mu dx^\nu = dx^\mu dx_\mu = g^{\mu\nu} dx_\mu dx_\nu$   
et  $\underline{A} \cdot \underline{B} = A^\mu B^\nu g_{\mu\nu} = A^\mu B_\mu = A_\mu B^\mu$ .

### 17.2.4 Transformation d'un vecteur covariant

Examinons à présent la transformation par  $L$  d'un vecteur covariant :  
de

$$g_{\alpha\beta} \Lambda_\mu^\alpha \Lambda_\nu^\beta = g_{\mu\nu},$$

on tire

$$g_{\alpha\beta} \Lambda_\mu^\alpha = g_{\mu\nu} (\Lambda^{-1})^\nu_\beta$$

donc

$$(\Lambda^{-1})^\nu_\beta = g_{\beta\alpha} \Lambda_\mu^\alpha g^{\mu\nu}.$$

Posons

$$\Lambda_\beta^\nu = g_{\beta\alpha} \Lambda_\mu^\alpha g^{\mu\nu}$$

ce qui revient à considérer  $\Lambda_\mu^\alpha$  comme un tenseur d'ordre 2 qui se transforme comme  $A^\alpha A_\mu$ .

On a donc  $\Lambda_\beta^\nu = (\Lambda^{-1})^\nu_\beta = ({}^t \Lambda^{-1})^\beta_\nu$  et l'on peut vérifier explicitement que  $\Lambda_\beta^\nu \Lambda_\nu^\rho = (\Lambda^{-1})^\nu_\beta \Lambda_\nu^\rho = \Lambda_\nu^\rho (\Lambda^{-1})^\nu_\beta = \delta_\beta^\rho$ .

*Remarques :*

- ceci est l'analogie de  $CUC^{-1} = {}^t U^{-1}$  pour  $SL(2, C)$ , qui s'écrit  $C_{\alpha'\alpha} U_\beta^\alpha C^{\beta\beta'} = U^{-1\beta'}_{\alpha'}$ .

- avec ces notations, puisque  $\mathbb{1} = \mathbb{1}^{-1} = {}^t \mathbb{1}$ , on pourra écrire plus simplement

$$\delta^\mu_\nu = \delta_\nu^\mu = \delta^\nu_\mu = \delta_\mu^\nu = \delta_\nu^\mu.$$

Ainsi l'inverse d'une matrice se calcule simplement en montant et en descendant les indices grâce à la métrique, puis en transposant.

Loi de transformation des composantes covariantes sous l'action de  $L$  :

$$\boxed{x'_\mu = \Lambda_\mu^\nu x_\nu} \quad (17.14)$$

*Preuve :*

$$\begin{aligned} \underline{x} &= x^\mu \underline{e}_\mu = x'^\nu \underline{e}'_\nu = \Lambda_\mu^\nu x^\mu \underline{e}'_\nu \\ \text{donc } \underline{e}_\mu &= \Lambda_\mu^\nu \underline{e}'_\nu \quad \text{et} \quad \underline{e}'_\mu = \Lambda_\mu^\nu \underline{e}_\nu \quad \text{d'après ce qui précède.} \\ \text{Ainsi } x'_\mu &= \underline{x} \cdot \underline{e}'_\mu = \underline{x} \cdot \Lambda_\mu^\nu \underline{e}_\nu = \Lambda_\mu^\nu x_\nu. \end{aligned}$$

*Remarque :* le fait que les composantes covariantes se transforment par  ${}^t\Lambda^{-1}$ , qui est la transposée inverse de la matrice de transformation des composantes contravariantes, est dû à l'invariance de  $ds^2 = dx^\mu dx_\mu$  qui implique que  ${}^t\Lambda g \Lambda = g$ .

En effet, ceci s'écrit encore  ${}^t\Lambda^{-1} = g \Lambda g^{-1}$ . Partant de  $v' = \Lambda v$  pour  $v$  contravariant, on en déduit :

$$\underbrace{g v'}_{\text{vecteur covariant}} = \underbrace{g \Lambda g^{-1}}_{{}^t\Lambda^{-1}} \underbrace{g \xi}_{\text{vecteur covariant}}$$

C'est l'analogie de la preuve donnée page 231 dans le cas de  $SU(2)$ .

### 17.2.5 Calcul tensoriel

Nous pouvons maintenant de façon générale parler de covariance relativiste : les grandeurs physiques qui décrivent l'état d'un système ne sont en général pas les mêmes pour différents observateurs :

- les quantités scalaires ne dépendent pas du choix d'un référentiel particulier : ce sont des fonctions  $\phi$  définies sur  $M$  t.q.  $\phi'(x') = \phi(x)$
- les composantes contravariantes  $v^\mu$  d'un vecteur se transforment comme les coordonnées (les translations n'ont d'effet que sur les coordonnées) :

$$v'^\mu = \Lambda^\mu_\nu v^\nu \left( = \frac{\partial x'^\mu}{\partial x^\nu} v^\nu \right)$$

- les composantes covariantes  $v_\mu$  d'un vecteur se transforment comme  $x_\mu$  :

$$v'_\mu = \Lambda_\mu^\nu x_\nu$$

- tenseur de rang  $(\underbrace{m}, \underbrace{n})$

co- contra-

$$F'_{\alpha_1 \dots \alpha_m}{}^{\mu_1 \dots \mu_n} = \Lambda_{\alpha_1}{}^{\beta_1} \dots \Lambda_{\alpha_m}{}^{\beta_m} \Lambda^{\mu_1}{}_{\rho_1} \dots \Lambda^{\mu_n}{}_{\rho_n} F_{\beta_1 \dots \beta_m}{}^{\rho_1 \dots \rho_n}$$

★ dérivation :

$$\frac{\partial \phi}{\partial x'^{\mu}} = \frac{\partial x^{\nu}}{\partial x'^{\mu}} \frac{\partial \phi}{\partial x^{\nu}} \quad \text{avec} \quad \Lambda_{\mu}{}^{\nu} = \frac{\partial x^{\nu}}{\partial x'^{\mu}},$$

donc  $\frac{\partial \phi}{\partial x^{\nu}}$  se transforme comme un vecteur covariant, noté  $\partial_{\nu} \phi$

De la définition des coordonnées contravariantes  $x^{\mu} = (x^0, \vec{x})$  on tire

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}} = \left( \frac{\partial}{\partial x^0}, \vec{\nabla} \right)$$

$$\partial^{\mu} = \frac{\partial}{\partial x_{\mu}} = \left( \frac{\partial}{\partial x_0}, -\vec{\nabla} \right)$$

attention aux signes :  $\vec{x}$  est contravariant mais  $\vec{\nabla}$  est covariant.

La quadri-divergence d'un quadri-vecteur est le scalaire

$$\partial^{\mu} A_{\mu} = \partial_{\mu} A^{\mu} = \frac{\partial A^0}{\partial x^0} + \vec{\nabla} \cdot \vec{A}$$

opérateur d'Alembertien :  $\square = \partial_{\mu} \partial^{\mu} = \frac{\partial^2}{\partial x_0^2} - \Delta$

Retenons :

- que l'on passe de l'une à l'autre forme des tenseurs en montant et en descendant les indices comme pour les quadri-vecteurs
- que la transformation se fait en mettant la matrice  $\Lambda$  à **gauche** du quadri-vecteur et en sommant sur les indices voisins (l'un en haut et l'autre en bas).

## 17.3 Algèbre de Lie du groupe de Lorentz restreint $L_{+}^{\uparrow}$

### 17.3.1 Forme non covariante de générateurs

De la contrainte

$$\omega_{\beta\alpha} = -\omega_{\alpha\beta}$$

on tire la relation matricielle  $\omega^{\beta}_{\alpha} = -\omega_{\alpha}{}^{\beta} = -g_{\alpha\alpha'} g^{\beta\beta'} \omega^{\alpha'}_{\beta'}$ .

En utilisant la forme explicite de la métrique  $g^{\mu\nu}$ , qui s'écrit

$$g_{i\alpha'} = -\delta_i^{\alpha'}, \quad g^{0\beta'} = \delta_0^{\beta'},$$

on en déduit l'écriture matricielle

$$\begin{cases} \omega^0_i = \omega^i_0 \\ \omega^i_j = -\omega^j_i \end{cases}$$

soit

$$\omega_\nu^\mu = \begin{pmatrix} 0 & d\phi_1 & d\phi_2 & d\phi_3 \\ d\phi_1 & | & 0 & -d\theta_3 & +d\theta_2 \\ d\phi_2 & | & +d\theta_3 & 0 & -d\theta_1 \\ d\phi_3 & | & -d\theta_2 & +d\theta_1 & 0 \end{pmatrix} \quad (17.15)$$

où le bloc  $3 \times 3$  correspond à la matrice  $3 \times 3$  d'une rotation infinitésimale dans le sous-espace des rotations spatiales.

• Introduisons les générateurs  $J^i$  (ce sont les générateurs des rotations; nous le justifierons explicitement dans le paragraphe 17.3.4) de la forme

$$J^i = \begin{pmatrix} 0 & \\ & J_{3 \times 3}^i \end{pmatrix} \quad (17.16)$$

où les matrices  $3 \times 3$   $J_{3 \times 3}^i$  sont les générateurs des rotations à 3 dimensions, qui s'écrivent  $(J_{3 \times 3}^i)_{kj} = i\epsilon_{ijk}$ . On a  $J^+ = J$ . Explicitement,

$$J_{3 \times 3}^1 = i \begin{pmatrix} & -1 \\ & \\ 1 & \end{pmatrix} \quad J_{3 \times 3}^2 = i \begin{pmatrix} & -1 \\ & \\ & 1 \end{pmatrix} \quad J_{3 \times 3}^3 = i \begin{pmatrix} -1 & & \\ & & \\ & & 1 \end{pmatrix}. \quad (17.17)$$

• De même, on introduit les générateurs  $K^i$  (qui sont les générateurs des boosts : voir paragraphe 17.3.3), sous de la forme (en ne représentant que les éléments non nuls)

$$K^1 = -i \begin{pmatrix} 0 & 1 \\ & \\ 1 & \end{pmatrix} \quad K^2 = -i \begin{pmatrix} 0 & 0 & 1 \\ 0 & & \\ 1 & & \end{pmatrix} \quad K^3 = -i \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & & & \\ 0 & & & \\ 1 & & & \end{pmatrix} \quad (17.18)$$

On a  $K^+ \neq K$ .

Nous pouvons alors écrire (17.15) sous la forme

$$\omega = id\vec{\phi} \cdot \vec{K} - id\vec{\theta} \cdot \vec{J}. \quad (17.19)$$

Dans le cas d'une transformation finie, nous aurons le résultat suivant, justifié aux paragraphes 17.3.3 et 17.3.4,

Boost pur	$\Lambda = e^{i\vec{\phi} \cdot \vec{K}}$
Rotation pure	$\Lambda = e^{-i\vec{\theta} \cdot \vec{J}}$

Un boost pur est non unitaire car  $K_i^\dagger \neq K_i$ .

Dans le cas général, un élément quelconque de  $L_+^\uparrow$  pourra s'écrire  $\Lambda = \Lambda_1 \Lambda_2$  où  $\Lambda_1$  est une rotation et  $\Lambda_2 = e^{i\vec{\phi} \cdot \vec{K}}$  est un boost (voir chapitre ??), mais  $\Lambda$  ne peut s'écrire comme l'exponentielle d'un élément de son algèbre de Lie (sauf au voisinage de l'identité), en raison de la non compacité de la composante connexe  $L_+^\uparrow$  de  $L$ .

### 17.3.2 Algèbre de Lie

Des formes explicites (17.16) et (17.18) des générateurs on tire immédiatement les relations de commutation

$$[J_i, J_j] = i\varepsilon_{ijk}J_k \quad (17.20)$$

$$[K_i, K_j] = -i\varepsilon_{ijk}J_k \quad (17.21)$$

$$[J_i, K_j] = i\varepsilon_{ijk}K_k \quad (17.22)$$

#### exercice 17.4

le vérifier en utilisant les expressions intrinsèques

$$(J_i)^k{}_j = i\varepsilon_{ijk} \quad (17.23)$$

et

$$(K_i)^\mu{}_\nu = -i[\delta^\mu{}_\nu\delta^i{}_i + \delta^\mu{}_0\delta^i{}_\nu]. \quad (17.24)$$

#### Conséquences importantes :

- la relation (17.20) prouve que le groupe des rotations est un sous-groupe de  $L_+^\uparrow$ . En revanche, on constate d'après (17.22) que les boosts ne constituent pas un sous-groupe de  $L_+^\uparrow$ .

- $[J_1, K_1] = 0$  donc une rotation effectuée autour de la direction de la vitesse d'une transformation de Lorentz ne modifie pas cette transformation :

$$e^{-i\theta_1 J_1} e^{i\phi_1 K_1} e^{i\theta_1 J_1} = e^{i\phi_1 K_1}$$

puisque la loi de transformation d'un opérateur  $T$  s'écrit  $T' = D(\Lambda)TD^+(\Lambda)$ .

- l'équation (17.22) exprime que  $\vec{K} = (K_1, K_2, K_3)$  se transforme par rotation comme les composantes d'un vecteur (voir la représentation adjointe de  $SO(3)$  discutée au chapitre ??) : c'est un opérateur tensoriel d'ordre 1, ou opérateur vectoriel.

- l'équation (17.21)  $[K_i, K_j] = -i\varepsilon_{ijk}J_k$  exprime le fait que deux transformations de Lorentz pures  $\Lambda_i$  et  $\Lambda_j$  le long des axes  $i$  et  $j$  engendrent une rotation lorsqu'on effectue l'opération  $\Lambda_i\Lambda_j\Lambda_i^{-1}$ . Ce résultat est à l'origine du phénomène de "précession de Thomas" : une succession de transformations de Lorentz pures appliquées à une particule à spin peuvent ramener l'impulsion à sa valeur initiale (représentation vectorielle) alors que le spin a tourné (représentation spinorielle).

- attention : comme  $[J, K] \neq 0$ ,

$$\Lambda = \exp(i\vec{\phi} \cdot \vec{K}) \exp(-i\vec{\theta} \cdot \vec{J}) \neq \exp(-i\vec{\theta} \cdot \vec{J}) \exp(i\vec{\phi} \cdot \vec{K}).$$

### 17.3.3 Transformation de Lorentz pure

Considérons une transformation de Lorentz pure d'axe  $\vec{x}$  engendré par  $K_1$ , de paramètre  $\phi_1$  (transformation active)

$$e^{iK_1\phi_1} = \sum_{k=0}^{\infty} \frac{(iK_1\phi_1)^{2k}}{(2k)!} + iK_1 \sum_{k=0}^{\infty} \frac{(iK_1\phi_1)^{2k+1}}{(2k+1)!}$$

$$(iK_1) = \begin{pmatrix} \sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{donc} \quad (iK_1)^2 = \begin{pmatrix} \mathbb{1}_{2 \times 2} & 0 \\ 0 & 0 \end{pmatrix}$$

puisque  $\sigma_1^2 = \mathbb{1}_{2 \times 2}$ , donc

$$\begin{aligned} e^{iK_1\phi_1} &= \sum_{k=0}^{\infty} \frac{(\phi_1)^{2k}}{(2k)!} + iK_1 \sum_{k=0}^{\infty} \frac{\phi_1^{2k+1}}{(2k+1)!} \\ &= \text{ch } \phi_1 + iK_1 \text{sh } \phi_1 \\ &= \begin{pmatrix} \text{ch } \phi_1 & \text{sh } \phi_1 & 0 & 0 \\ \text{sh } \phi_1 & \text{ch } \phi_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

La relation  $\vec{x}' = e^{iK_1\phi_1} \vec{x}$  s'écrit donc

$$\begin{cases} x'^0 = \text{ch } \phi_1 x^0 + \text{sh } \phi_1 x^1 \\ x'^1 = \text{sh } \phi_1 x^0 + \text{ch } \phi_1 x^1 \end{cases} \Leftrightarrow \begin{cases} x^0 = \text{ch } \phi_1 x'^0 - \text{sh } \phi_1 x'^1 \\ x^1 = -\text{sh } \phi_1 x'^0 + \text{ch } \phi_1 x'^1. \end{cases}$$

En considérant un point particulier tel que  $x^1 = \text{constante}$ , on tire immédiatement  $v = \frac{dx'^1}{dx'^0} = \text{th } \phi$  : c'est la vitesse du point après boost, mesurée par rapport au référentiel par rapport auquel il était fixe avant boost.

$$\boxed{\phi = \text{Argth } v \text{ est appelée rapidité.}}$$

On déduit de ce qui précède la paramétrisation

$$\begin{cases} \text{ch } \phi_1 = \frac{1}{\sqrt{1-v^2}} \\ \text{sh } \phi_1 = \frac{v}{\sqrt{1-v^2}} \end{cases}$$

et

$$\boxed{\begin{aligned} x'^0 &= \frac{x^0}{\sqrt{1-v^2}} + \frac{v}{\sqrt{1-v^2}} x^1 \\ x'^1 &= \frac{v x^0}{\sqrt{1-v^2}} + \frac{x^1}{\sqrt{1-v^2}} \end{aligned}} \quad (17.25)$$

qui est bien la forme usuelle d'une transformation spéciale de Lorentz.

Cas général :

Considérons une transformation de Lorentz pure d'axe quelconque  $\vec{n}$  ( $\vec{n}^2 = 1$ ) et de vitesse  $v$  (point de vue actif) en terme de matrices,  $(K_i)^\mu_\nu = -i[\delta_\nu^0 \delta_i^\mu + \delta_0^\mu \delta_\nu^i]$

Montrons que  $\vec{x}' = e^{i\phi \vec{n} \cdot \vec{k}} \vec{x}$  est donné par

$$\begin{cases} x'^0 &= x^0 \text{ch } \phi + (\vec{n} \cdot \vec{x}) \text{sh } \phi \\ \vec{x}' &= \vec{x} - \vec{n} (\vec{x} \cdot \vec{n}) + (x^0 \text{sh } \phi + (\vec{n} \cdot \vec{x}) \text{ch } \phi) \vec{n} \end{cases} \quad (17.26)$$

*Preuve :*

$$\begin{aligned} (i\vec{K} \cdot \vec{n})^\mu_\nu &= +[\delta_\nu^0 n^\mu - \delta_0^\mu n_\nu] \\ &\swarrow \text{à cause de } \vec{n} \cdot \vec{K} = n^i K^i (= n^i K_i \text{ car dans nos conventions } K_i = K^i!) \end{aligned}$$

$$\begin{aligned} \left[ (i\vec{K} \cdot \vec{n})^2 \right]^\mu_\nu &= \left[ i\vec{K} \cdot \vec{n} \right]^\mu_{\nu'} \left[ i\vec{K} \cdot \vec{n} \right]^{\nu'}_\nu \\ &= +[\delta_\nu^0 n^\mu - \delta_0^\mu n_{\nu'}] [\delta_\nu^0 n^{\nu'} - \delta_0^{\nu'} n_\nu] \\ &= + (n^0 n^\mu \delta_\nu^0 - n^\mu n_\nu - \delta_0^\mu \delta_\nu^0 n^2 + n_0 \delta_0^\mu n_0) \\ &= -n^\mu n_\nu + \delta_0^\mu \delta_\nu^0 \end{aligned}$$

puisque  $n^2 = -1$ . Enfin

$$\begin{aligned} \left[ (i\vec{K} \cdot \vec{n})^3 \right]^\mu_\nu &= +[\delta_\nu^0 n^\mu - \delta_0^\mu n_{\nu'}] [-n^{\nu'} n_\nu + \delta_0^{\nu'} \delta_\nu^0] \\ &= +n^\mu \delta_\nu^0 - \delta_0^\mu n_\nu = (i\vec{K} \cdot \vec{n})^\mu_\nu. \end{aligned}$$

La série s'écrit donc

$$\begin{aligned} e^{i\phi \vec{n} \cdot \vec{k}} &= \delta_\nu^\mu + [\delta_\nu^0 n^\mu - \delta_0^\mu n_\nu] \sum_{k=0}^{\infty} \frac{\phi^{2k+1}}{(2k+1)!} \\ &\quad + [-n^\mu n_\nu + \delta_0^\mu \delta_\nu^0] \sum_{k=0}^{\infty} \frac{\phi^{2k}}{(2k+1)!} + n^\mu n_\nu - \delta_0^\mu \delta_\nu^0 \\ &= \delta_\nu^\mu + n^\mu n_\nu - \delta_0^\mu \delta_\nu^0 + [\delta_\nu^0 n^\mu - \delta_0^\mu n_\nu] \text{sh } \phi + [-n^\mu n_\nu + \delta_0^\mu \delta_\nu^0] \text{ch } \phi. \end{aligned}$$

Ainsi

$$x'^\mu = x^\mu + n^\mu (x \cdot n) - \delta_0^\mu x^0 + [x^0 n^\mu - \delta_0^\mu (n \cdot x)] \text{sh } \phi + [-n^\mu (n \cdot x) + \delta_0^\mu x^0] \text{ch } \phi$$

ce qui prouve le résultat.

*Remarque :*

une transformation passive (changement de référentiel) du référentiel  $K$  vers le référentiel  $K'$  (animé d'une vitesse  $v$  par rapport à  $K$ ) est une transformation active (référentiel fixe, les points sont transformés) de vitesse  $-v$ .



### 17.3.4 Rotation

A titre d'exercice, nous pouvons également déterminer la forme explicite d'une rotation, que nous avons obtenue directement en (14.3).

Pour une rotation active d'axe  $\vec{z}$  et d'angle  $\theta_3$ ,

$$\begin{aligned} e^{-i\vec{\theta}\cdot\vec{J}} = e^{-i\theta_3\cdot J_3} &= \sum_{k=0}^{\infty} \frac{(J_3)^{2k}}{(2k)!} (-i\theta_3)^{2k} - iJ_3 \sum_{k=0}^{\infty} \frac{(J_3)^{2k}}{(2k+1)!} \theta_3^{2k+1} (-i)^{2k} \\ &= \underbrace{\sum_{k=0}^{\infty} \frac{(+\theta_3)^{2k}}{(2k)!} (-1)^k}_{\cos \theta_3} + \underbrace{\sum_{k=0}^{\infty} \frac{\theta_3^{2k+1}}{(2k+1)!} (-1)^k (-iJ_3)}_{\sin \theta_3} \end{aligned}$$

En effet,

$$J_3 = i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & & & \\ & \sigma_2 & & \\ & & & \\ & & & 0 \end{pmatrix}$$

donc

$$J_3^2 = \begin{pmatrix} 0 & & & \\ & \mathbb{1}_{2\times 2} & & \\ & & & \\ & & & 0 \end{pmatrix}$$

puisque  $\sigma_2^2 = \mathbb{1}_{2\times 2}$ . Ainsi

$$e^{-i\theta_3 J_3} = \begin{pmatrix} 1 & & & \\ & \cos \theta_3 & -\sin \theta_3 & \\ & \sin \theta_3 & \cos \theta_3 & \\ & & & 1 \end{pmatrix}$$

La rotation passive (rotation des axes) est obtenue en faisant  $\theta \rightarrow -\theta$ .

#### Cas général :

Dans le cas d'une rotation quelconque d'axe  $\vec{n}$  et d'angle  $\theta$  (point de vue actif), montrons que

$$\boxed{\vec{x}' = e^{-i\theta\vec{n}\cdot\vec{J}}\vec{x} = (\vec{x} - \vec{n}(\vec{n}\cdot\vec{x})) \cos \theta + \vec{n} \wedge \vec{x} \sin \theta + \vec{n}(\vec{n}\cdot\vec{x})}$$

en accord avec la relation (14.3).

*Preuve :*

Calculons

$$e^{-i\theta\vec{n}\cdot\vec{J}} = \sum_{k=0}^{\infty} \frac{(-i\vec{n}\cdot\vec{J})^k}{k!} \theta^k.$$

$\left(-i\vec{J} \cdot \vec{m}\right)_j^k = -i(J^i)^k_j n^i = \varepsilon_{ijk} n^i$  (dans ce calcul, il n'est pas utile de distinguer les composantes co- et contra-variantes)

$$\text{donc } \left(-i\vec{J} \cdot \vec{n}\right)^k = \varepsilon_{ijk} n^i x^j = (\vec{n} \wedge \vec{x})^k$$

$$\left(-i\vec{J} \cdot \vec{n}\right)_{kk'}^2 = \left(-i\vec{J} \cdot \vec{n}\right)_{kj} \left(-i\vec{J} \cdot \vec{n}\right)_{jk'} = \varepsilon_{ijk} n^i \varepsilon_{i'k'j} n^{i'}$$

qui s'écrit encore  $-\delta_{k'}^k + n^k n^{k'}$  (car  $\vec{n}^2 = 1$ )

$$\begin{aligned} \left(-i\vec{J} \cdot \vec{n}\right)_{kk'}^3 &= \left(-i\vec{J} \cdot \vec{n}\right)_{kj} \left(-i\vec{J} \cdot \vec{n}\right)_{jk'}^2 \\ &= \varepsilon_{ijk} n^i \left(-\delta_{k'}^j + n^j n^{k'}\right) = -\varepsilon_{ik'k} n^i + \underbrace{\varepsilon_{ijk} n^i n^j n^{k'}}_0 \\ &\quad \text{(contraction d'un tenseur antisymétrique par} \\ &\quad \text{un tenseur symétrique)} \\ &= -\left(-i\vec{J} \cdot \vec{n}\right)_{kk'} \end{aligned}$$

donc

$$\begin{aligned} \left(e^{-i\theta\vec{n} \cdot \vec{J}}\right)_{kj} &= \delta_j^k + \theta \varepsilon_{ijk} n^i - \frac{\theta^2}{2} (\delta_j^k - n^k n^j) - \frac{\theta^3}{3!} \varepsilon_{ijk} n^i + \frac{\theta^4}{4!} (\delta_j^k - n^k n^j) + \dots \\ &= (\delta_j^k - n^k n^j) \left(1 - \frac{\theta^2}{2} + \frac{\theta^4}{4} + \dots\right) + n^k n^j + \varepsilon_{ijk} n^i \left(\theta - \frac{\theta^3}{3!} + \dots\right) \end{aligned}$$

soit

$$\left(e^{-i\theta\vec{n} \cdot \vec{J}}\right)_{kj} = (\delta_j^k - n^k n^j) \cos \theta + \varepsilon_{ijk} n^i \sin \theta + n^k n^j$$

d'où finalement

$$e^{-i\theta\vec{n} \cdot \vec{J}} \vec{x} = \vec{x}' = (\vec{x} - \vec{n}(\vec{n} \cdot \vec{x})) \cos \theta + \vec{n} \wedge \vec{x} \sin \theta + \vec{n}(\vec{n} \cdot \vec{x})$$

puisque  $\varepsilon_{ijk} n^i x^j = (\vec{n} \wedge \vec{x})_k$ .

### 17.3.5 Forme covariante de l'algèbre de Lie de $L_+^\uparrow$

Nous avons écrit en 17.2.2 une transformation infinitésimale sous la forme

$$\Lambda = \mathbb{1} + \omega$$

ou encore

$$\Lambda^\mu_\nu = g^\mu_\nu + \omega^\mu_\nu,$$

et montré que  $\omega_{\mu\nu}$  est antisymétrique, i.e. que  $\omega_{\mu\nu}$  dépend de 6 paramètres, ce qui définit la dimension de l'algèbre de Lie.

Plutôt que d'utiliser une paramétrisation non covariante de cette transformation infinitésimale à l'aide de  $(d\vec{\theta}, d\vec{\phi})$ , comme nous l'avons fait dans le paragraphe 17.3.1, on peut utiliser  $\omega_{\mu\nu}$  comme collection de paramètres. Ceci nous conduit à introduire la définition suivante, covariante, des générateurs de l'algèbre de Lie :

$$\Lambda = \mathbb{1} - \frac{i}{2} \omega^{\alpha\beta} J_{\alpha\beta}. \quad (17.27)$$

Dans cette définition, les indices  $(\alpha, \beta)$  jouent le rôle d'étiquetage des générateurs. Ceux-ci sont eux-mêmes des tenseurs, et portent donc également des indices. On peut imposer  $J_{\alpha\beta} = -J_{\beta\alpha}$  puisque les  $\omega^{\alpha\beta}$  sont antisymétriques, d'où le facteur conventionnel  $\frac{1}{2}$  pour éviter le double comptage Tensoriellement, la relation (17.27) s'écrit :

$$\Lambda_{\mu\nu} = g_{\mu\nu} - \frac{i}{2} \omega^{\alpha\beta} (J_{\alpha\beta})_{\mu\nu}.$$

Donc

$$\omega_{\mu\nu} = -\frac{i}{2} \omega^{\alpha\beta} (J_{\alpha\beta})_{\mu\nu}. \quad (17.28)$$

On en déduit que les  $J_{\alpha\beta}$  sont des tenseurs antisymétriques (i.e. que  $(J_{\alpha\beta})_{\mu\nu}$  est antisymétrique par rapport à  $(\mu\nu)$ ).

Introduisons la base

$$(a_{\mu'\nu'})^{\mu\nu} = \delta_{\mu'}^\mu \delta_{\nu'}^\nu - \delta_{\mu'}^\nu \delta_{\nu'}^\mu. \quad (17.29)$$

C'est une base, dite canonique, des tenseurs antisymétriques, qui s'écrit encore

$$a_{\mu'\nu'} = e_{\mu'\nu'} - e_{\nu'\mu'}$$

où

$$(e_{\mu'\nu'})^{\mu\nu} = \delta_{\mu'}^\mu \delta_{\nu'}^\nu$$

est la base canonique des tenseurs de rang 2. Dans cette base, (17.28) s'écrit

$$-\frac{i}{2} \left( \delta_{\mu'}^\alpha \delta_{\nu'}^\beta - \delta_{\mu'}^\beta \delta_{\nu'}^\alpha \right) (J_{\alpha\beta})^\mu{}_\nu = \delta_{\mu'}^\mu g_{\nu\nu'} - g_{\nu\mu'} \delta_{\nu'}^\mu.$$

Le premier membre de cette égalité s'écrit encore  $-\frac{i}{2} [(J_{\mu'\nu'})^\mu{}_\nu - (J_{\nu'\mu'})^\mu{}_\nu] = -i (J_{\mu'\nu'})^\mu{}_\nu$  par antisymétrie sur  $(\mu'\nu')$ .

Donc

$$\boxed{(J_{\mu'\nu'})^\mu{}_\nu = i (g_{\mu'}^\mu g_{\nu\nu'} - g_{\nu\mu'} g_{\nu'}^\mu),}$$

ou encore

$$\boxed{(J_{\mu'\nu'})_{\mu\nu} = i (a_{\mu'\nu'})_{\mu\nu}.$$

L'algèbre de Lie s'en déduit immédiatement :

$$\boxed{[J_{\mu\nu}, J_{\rho\sigma}] = i (g_{\nu\rho} J_{\mu\sigma} - g_{\mu\rho} J_{\nu\sigma} + g_{\mu\sigma} J_{\nu\rho} - g_{\nu\sigma} J_{\mu\rho}).} \quad (17.30)$$

*Preuve :*

- directement, en utilisant la forme matricielle :

$$\begin{aligned}
[J_{\mu\nu}, J_{\rho\sigma}]^\alpha_\beta &= (J_{\mu\nu})^\alpha_{\beta'} (J_{\rho\sigma})^{\beta'}_\beta - (J_{\rho\sigma})^\alpha_{\beta'} (J_{\mu\nu})^{\beta'}_\beta \\
&= -(\delta_\mu^\alpha g_{\nu\beta'} - g_{\beta'\mu} \delta_\nu^\alpha) (\delta_\rho^{\beta'} g_{\sigma\beta} - g_{\beta\rho} \delta_\sigma^{\beta'}) \\
&\quad + (\delta_\rho^\alpha g_{\sigma\beta'} - g_{\beta'\rho} \delta_\sigma^\alpha) (\delta_\mu^{\beta'} g_{\nu\beta} - g_{\beta\mu} \delta_\nu^{\beta'}) \\
&= -g_{\nu\rho} (\delta_\mu^\alpha g_{\sigma\beta} - \delta_\sigma^\alpha g_{\beta'\mu}) + g_{\mu\rho} (-\delta_\sigma^\alpha g_{\nu\beta} + \delta_\nu^\alpha g_{\sigma\beta}) \\
&\quad - g_{\mu\sigma} (-\delta_\rho^\alpha g_{\nu\beta} + \delta_\nu^\alpha g_{\beta\rho}) + g_{\nu\sigma} (\delta_\mu^\alpha g_{\rho\beta} - \delta_\rho^\alpha g_{\beta\sigma}) \\
&= i (g_{\nu\rho} J_{\mu\sigma} - g_{\mu\rho} J_{\nu\sigma} + g_{\mu\sigma} J_{\nu\rho} - g_{\nu\sigma} J_{\mu\rho})^\alpha_\beta
\end{aligned}$$

• en écrivant les générateurs sous la forme manifestement antisymétrique  $J_{\mu\nu} = i(e_{\mu\nu} - e_{\nu\mu}) = i(|\mu\rangle\langle\nu| - |\nu\rangle\langle\mu|)$ , avec le produit scalaire minkowskien  $\langle\alpha|\mu\rangle = g_{\alpha\mu}$ , on peut aisément retrouver l'équation précédente de manière intrinsèque. Le lecteur non convaincu vérifiera préalablement que

$$\begin{aligned}
(J_{\mu\nu})_{\alpha\beta} &= i (\langle\alpha|\mu\rangle\langle\nu|\beta\rangle - \langle\alpha|\nu\rangle\langle\mu|\beta\rangle) \\
&= i (g_{\alpha\mu} g_{\nu\beta} - g_{\alpha\nu} g_{\mu\beta})
\end{aligned}$$

en accord avec la forme obtenue plus haut. L'obtention des relations de commutation de l'algèbre de Lie est alors immédiate :

$$\begin{aligned}
[J_{\mu\nu}, J_{\rho\sigma}] &= -(|\mu\rangle\langle\nu| - |\nu\rangle\langle\mu|) (|\rho\rangle\langle\sigma| - |\sigma\rangle\langle\rho|) \\
&\quad + (|\rho\rangle\langle\sigma| - |\sigma\rangle\langle\rho|) (|\mu\rangle\langle\nu| - |\nu\rangle\langle\mu|) = -g_{\nu\rho} (|\mu\rangle\langle\sigma| - |\sigma\rangle\langle\mu|) \\
&\quad + g_{\mu\rho} (|\nu\rangle\langle\sigma| - |\sigma\rangle\langle\nu|) - g_{\mu\sigma} (|\nu\rangle\langle\rho| - |\rho\rangle\langle\nu|) + g_{\nu\sigma} (|\mu\rangle\langle\rho| - |\rho\rangle\langle\mu|) \\
&= i (g_{\nu\rho} J_{\mu\sigma} - g_{\mu\rho} J_{\nu\sigma} + g_{\mu\sigma} J_{\nu\rho} - g_{\nu\sigma} J_{\mu\rho}) .
\end{aligned}$$

On notera que dans ce calcul, la forme explicite des générateurs ne joue aucun rôle. Seul compte leur propriété d'antisymétrie par rapport à leur index d'étiquetage, qui provient elle-même de l'antisymétrie des  $\omega_{\mu\nu}$ , i.e. de la définition de  $L_+^\uparrow$ . Ceci ne doit pas nous surprendre : l'algèbre de Lie est la même quelque soit la représentation choisie. Nous allons le vérifier explicitement sur la représentation régulière discutée dans le paragraphe qui suit.

### 17.3.6 Représentation sur les fonctions scalaires (représentation appelée régulière)

$\forall \Lambda \in L_+^\uparrow$  on associe la transformation  $T_\Lambda$  agissant sur les fonctions  $f(x)$  :

$$f(x) \xrightarrow{T_\Lambda} (T_\Lambda f)(x) = f(\Lambda^{-1}x) .$$

T est une représentation de  $L_+^\uparrow$

Preuve :

$$\begin{aligned} (T_\Lambda T_{\Lambda'} f)(x) &= T_{\Lambda'} f(\Lambda^{-1}x) = f(\Lambda'^{-1}\Lambda^{-1}x) \\ &= f((\Lambda\Lambda')^{-1}x) \\ &= (T_{\Lambda\Lambda'} f)(x) \end{aligned}$$

ce qui prouve que  $T_\Lambda T_{\Lambda'} = T_{\Lambda\Lambda'}$ .

### Générateurs de cette représentation :

A la transformation infinitésimale de Lorentz

$$x'^\mu = x^\mu + \omega^{\mu\nu} x_\nu$$

correspond

$$(T_\Lambda f)(x) = \left(1 - \frac{i}{2}\omega^{\mu\nu} J_{\mu\nu}\right) f(x) \quad (\text{facteur 2 à cause du double comptage})$$

où les  $J_{\mu\nu}$  sont des opérateurs différentiels. En utilisant la définition de  $T$ , on a donc

$$\begin{aligned} \left(1 - \frac{i}{2}\omega^{\mu\nu} J_{\mu\nu}\right) f(x) &= f(\Lambda^{-1}x) = f(x^\mu - \omega^{\mu\nu} x_\nu) \\ &= f(x) - \omega^{\mu\nu} x_\nu \partial_\mu f \end{aligned}$$

Donc

$$\boxed{J_{\mu\nu} = i(x_\mu \partial_\nu - x_\nu \partial_\mu)}. \quad (17.31)$$

On vérifiera aisément que ces générateurs satisfont l'algèbre de Lie (17.30).

Cette représentation est **unitaire**. La preuve est la même que celle déjà donnée pages 145 et 212 pour la représentation de moment orbital du groupe des rotations. Notons que  $L_+^\uparrow$  n'est pas compact, et que la représentation unitaire de  $L_+^\uparrow$  que nous venons de construire est de dimension **infinie**, en accord avec les résultats généraux énoncés au paragraphe ??.

### 17.3.7 Forme réelle compacte

Considérons les relations de commutation

$$[J_{\mu\nu}, J_{\rho\sigma}] = ig_{\nu\rho} J_{\mu\sigma} - ig_{\mu\rho} J_{\nu\sigma} - ig_{\nu\sigma} J_{\mu\rho} + ig_{\mu\sigma} J_{\nu\rho}$$

En posant

$$\left| \begin{array}{l} J_i = \frac{1}{2}\varepsilon_{ijk} J_{jk} (= J^i) \quad (\Leftrightarrow J_{jk} = \varepsilon_{ijk} J_i) \\ K_i = J_{0i} \quad (= K^i = -J^{0i}) \end{array} \right.$$

on obtient

$$[J_i, J_j] = i\varepsilon_{ijk}J_k \quad [K_i, K_j] = -i\varepsilon_{ijk}J_k \quad [J_j, K_j] = i\varepsilon_{ijk}K_k$$

qui est bien identique aux relations obtenues plus haut. Le lien entre générateurs covariants  $J^\mu_\nu$  et générateurs  $(\vec{K}, \vec{J})$  se résume sous la forme

$$J = \begin{pmatrix} 0 & K_1 & K_2 & K_3 \\ -K_1 & 0 & J_3 & -J_2 \\ -K_2 & -J_3 & 0 & J_1 \\ -K_3 & J_2 & -J_1 & 0 \end{pmatrix}.$$

### Lien entre $\omega_{\alpha\beta}$ et $\vec{\theta}, \vec{\phi}$ :

en comparant

$$1 - \frac{i}{2}\omega_{\alpha\beta}J^{\alpha\beta} + o(\omega) = (1 - i\vec{\theta} \cdot \vec{J})(1 + i\vec{\phi} \cdot \vec{K}) + o(\theta, \phi)$$

on a

$$-\frac{i}{2}\omega_{\alpha\beta}J^{\alpha\beta} = i\vec{\phi} \cdot \vec{K} - i\vec{\theta} \cdot \vec{J} = i\phi^i K^i - i\theta^i J^i$$

(car  $K^i = K_i$  et  $J^i = J_i$ , puisque dans les conventions utilisées (cf. celles pour  $\vec{\sigma}$ ),  $K_i = K^i$  et  $J_i = J^i$  désignent **toujours** les composantes contravariantes). Par antisymétrie de  $J^{\mu\nu}$  on a donc

$$-\frac{1}{2}i\omega_{\alpha\beta}J^{\alpha\beta} = -i\omega_{0i}J^{0i} - \frac{i}{2}\omega_{jk}J^{jk} = -i\phi^i J^{0i} + \frac{i}{2}\theta_i \varepsilon_{ijk}J^{jk}$$

d'où

$$\begin{cases} \omega_{0i} = \phi^i \\ \omega_{jk} = -\varepsilon_{ijk}\theta_i \end{cases} \quad \text{ou encore} \quad \begin{cases} \omega^{0i} = -\phi^i \\ \omega^{jk} = \varepsilon_{ijk}\theta^i. \end{cases}$$

### Compactification

Il est possible de rendre l'algèbre de Lie compacte, en utilisant la méthode présentée page 165.

Posons

$$\begin{cases} M_i = \frac{1}{2}(J_i + iK_i) \\ N_i = \frac{1}{2}(J_i - iK_i) \end{cases}. \quad (17.32)$$

On obtient alors

$$[M_i, M_j] = i\varepsilon_{ijk}M_k \quad [N_i, N_j] = i\varepsilon_{ijk}N_k \quad [M_i, N_j] = 0.$$

Cette algèbre de Lie est donc identique à celle du groupe  $SU(2) \times SU(2)$ .

L'algèbre de Lie ainsi obtenue est une forme réelle compacte de même extension complexe que  $so(3, 1)$ .  $\vec{M}^2$  et  $\vec{N}^2$  sont deux opérateurs quadratiques invariants. Ces Casimirs commutent avec  $\vec{M}$  et  $\vec{N}$ , donc avec toute l'algèbre, et donc avec tout élément du groupe de Lorentz restreint  $L_+^\uparrow$ . Dans le cas d'une représentation irréductible, d'après le lemme de Schur, on a

$$\begin{aligned} \vec{M}^2 &= j_1(j_1 + 1)\mathbf{1} \\ \vec{N}^2 &= j_2(j_2 + 1)\mathbf{1} \end{aligned} \quad (17.33)$$

avec  $j_1, j_2$  entiers ou demi-entiers





# Chapitre 18

## Représentations spinorielles du groupe de Lorentz

### 18.1 Introduction

Soit un système physique qui possède le groupe de Lorentz restreint comme groupe de symétrie :

$$| \langle \psi | \mathcal{X} \rangle |^2 = | \langle \psi_\Lambda | \mathcal{X}_\Lambda \rangle |^2$$

Alors le théorème de Wigner implique qu'il existe un opérateur  $U(\Lambda)$  défini à une phase près tel que  $|\psi_n \rangle = U(\Lambda)|\psi \rangle$ .

Dans le cas du groupe de Lorentz restreint, la connexité du groupe implique que ces opérateurs sont unitaires. Ils constituent une représentation du groupe de Lorentz restreint à une phase près :  $U(\Lambda_1)U(\Lambda_2) = \pm U(\Lambda_1\Lambda_2)$ .

L'espace de Hilbert qui décrit un système quantique ayant pour groupe de symétrie le groupe de Lorentz restreint est un espace de représentation unitaire du groupe de recouvrement universel de  $L_+^\uparrow$ , qui est le groupe  $SL(2, \mathbb{C})$ .

rappel :  $SL(2, \mathbb{C})$  est le groupe des matrices complexes  $A$  unidomulaires ( $\det A = +1$ ) à deux dimensions.

L'introduction des spineurs quadridimensionnels est tout à fait analogue à la démarche suivie pour les rotations (voir chapitres ?? et ??). Cette généralisation date de la fin des années 20 (Van der Warden, Uhlenbeck et Laporte).

Les spineurs quadridimensionnels sont introduits comme espace de représentation de dimension minimale de  $SL(2, \mathbb{C})$ . La différence importante entre spineurs 3d et 4d est que dans ce dernier cas, deux représentations conjuguées l'une de l'autre ne sont pas équivalentes.

### 18.2 Homomorphismes du groupe $SL(2, \mathbb{C})$ sur le groupe $L_+^\uparrow$

#### 18.2.1 Préliminaires

- Toute matrice  $2 \times 2$  à coefficients dans  $\mathbb{C}$  peut s'écrire comme combinaison linéaire à coefficients dans  $\mathbb{C}$  de la matrice unité  $\sigma_0 = \mathbb{1}$  et des trois matrices de Pauli.

• Toute matrice  $2 \times 2$  hermitienne peut s'écrire comme combinaison linéaire à coefficient réels de ces 4 matrices. Nous avons utilisé cette base pour paramétrer le groupe  $SU(2)$  page 202.

Posons

$$\begin{cases} \sigma_\mu = (\mathbf{1}, \vec{\sigma}) = \underline{\sigma}^\mu \\ \sigma^\mu = (\mathbf{1}, -\vec{\sigma}) = \underline{\sigma}_\mu \end{cases}, \quad (18.1)$$

avec  $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ .

A tout quadrivecteur  $x^\mu$  on associe les matrices

$$\tilde{x} = x^\mu \sigma_\mu = x_0 + \vec{\sigma} \cdot \vec{x} = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \quad (18.2)$$

et

$$\tilde{x} = x^\mu \underline{\sigma}_\mu = x_0 - \vec{\sigma} \cdot \vec{x} = \begin{pmatrix} x^0 - x^3 & x^1 + ix^2 \\ x^1 - ix^2 & x^0 + x^3 \end{pmatrix}. \quad (18.3)$$

On déduit immédiatement de ces relations que  $\tilde{x}$  est hermitique  $\Leftrightarrow \underline{x}$  est hermitique  $\Leftrightarrow x$  réel

$$\begin{aligned} \sigma_\mu &= (\mathbf{1}, \vec{\sigma}) \\ \sigma^\mu &= (\mathbf{1}, -\vec{\sigma}) \end{aligned}$$

### Propriétés

1)  $\det \tilde{x} = x^2$

En effet  $\det \tilde{x} = (x^0 + x^3)(x^0 - x^3) - (x^1 + ix^2)(x^1 - ix^2) = x^{0^2} - x^{1^2} - x^{2^2} - x^{3^2} = x^2$

2)  $\frac{1}{2} Tr \sigma_\mu \sigma_\nu = \delta^\mu_\nu$  (correspond à la normalisation habituelle des générateurs de l'algèbre de Lie)

En effet :  $x \sigma_i \sigma_j = \delta_{ij} + i \varepsilon_{ijk} \sigma_k$  donc  $Tr \sigma_i \sigma_j = 2 \delta_{ij}$

•  $\sigma_i \sigma_0 = \sigma_i$  donc  $Tr \sigma_i \sigma_0 = Tr \sigma_i = 0$

•  $\sigma_0 \sigma_0 = \sigma_0$   $Tr \sigma_0 \sigma_0 = Tr \sigma_0 = 2$

3)  $x^\mu = \frac{1}{2} Tr \sigma_\mu \tilde{x}$  puisque  $\frac{1}{2} Tr \sigma_\mu \tilde{x} = \frac{1}{2} Tr \sigma_\mu x^\nu \sigma_\nu = \delta^\mu_\nu x^\nu = x^\mu$

### 18.2.2 Théorème

Il existe un homomorphisme du groupe  $SL(2, \mathbb{C})$  sur le groupe de Lorentz restreint  $L_+^\uparrow$  défini par

$$\pm A \rightarrow \Lambda(A)$$

$$x' = \Lambda(A)x$$

$$\tilde{x}' = A \tilde{x} A^+$$

**Preuve :**  $A \in SL(2, \mathbb{C})$  donc  $\det \tilde{x}' = |\det A|^2 \det \tilde{x} = \det \tilde{x}$

$$\text{d'où} \quad x'^2 = x^2$$

La transformation linéaire  $x \rightarrow x'$  conserve les longueurs relativistes. C'est donc une transformation de Lorentz.

Calcul des éléments de matrice  $\Lambda(A)^\mu_\nu$

$$x'^\mu \sigma_\mu = Ax^\rho \sigma_\rho A^+ = A \sigma_\rho A^+ x^\rho = \Lambda^\mu_\rho x^\rho \sigma_\mu$$

$$\uparrow x' = \Lambda(A)x$$

donc  $\Lambda^\mu_\rho \sigma_\mu = A \sigma_\rho A^+$   
 $\Lambda^\mu_\rho \sigma_\mu \sigma_\nu = A \sigma_\rho A^+ \sigma_\nu$

comme  $\frac{1}{2} \text{Tr} \sigma_\mu \sigma_\nu = \delta_\nu^\mu$  ,  $\frac{1}{2} \Lambda^\mu_\rho \text{Tr} \sigma_\mu \sigma_\nu = \boxed{\Lambda^\nu_\rho = \frac{1}{2} \text{Tr} A \sigma_\rho A^+ \sigma_\nu}$

• Montrons que les éléments de matrice sont réels :

$$(\Lambda^\nu_\rho)^* = \frac{1}{2} \text{Tr} (A \sigma_\rho A^+ \sigma_\nu)^+ = \frac{1}{2} \text{Tr} \sigma_\nu A \sigma_\rho A^+ = \frac{1}{2} \text{Tr} A \sigma_\rho A^+ \sigma_\nu = \Lambda^\nu_\rho$$

• La transformation est orthochrone :

$$A^0_0 = \frac{1}{2} \text{Tr} A \sigma_0 A^+ \sigma_0 = \frac{1}{2} \text{Tr} A A^+ > 0$$

•  $\det \Lambda = +1$  :

$\det \Lambda(A)$  est une fonction continue de  $A$ . Comme  $SL(2, \mathbb{C})$  est connexe,  $\det \Lambda(A)$  ne peut prendre qu'une seule des deux valeurs  $+1$  ou  $-1$  ( $x^1 = \Lambda(A)x$  et  $x^2 = x'^2$  donc  $\det \Lambda(A) = \pm 1$ ).

Considérons un élément particulier :  $A = \mathbb{1}$

$$\Lambda(\mathbb{1})^\nu_\rho = \frac{1}{2} \text{Tr} \sigma_\rho \sigma_\nu = \delta_{\rho\nu} \quad \text{donc} \quad \Lambda(\mathbb{1}) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad \text{et} \quad \det \Lambda = 1 \quad \mathbf{c.f.q.d.}$$

**Ainsi**  $\Lambda(A) \in \mathcal{L}_+^\uparrow$ . Il est immédiat que  $\Lambda(-A) = \Lambda(A)$ .

• le noyau de ce morphisme se réduit à  $Z_2$  : soit  $A \in SL(2, \mathbb{C})$  t.q.  $\Lambda(A)^\mu_\nu = \delta^\mu_\nu$

Alors  $\forall x, \tilde{x} = A \tilde{x} A^+$ .

Pour  $x = (1, 0, 0, 0)$ ,  $\tilde{x} = 1$  donc  $AA^+ = 1$ , d'où  $A$  unitaire.

On a ainsi  $[\tilde{x}, A] = 0, \forall \tilde{x}$ .

Ainsi  $A = \lambda \mathbb{1}$

Comme  $\det A = 1$ ,  $A = \pm \mathbb{1}$  **c.q.f.d.**

On peut construire un second homomorphisme de  $SL(2, \mathbb{C})$  sur  $L_+^\uparrow$  en considérant

$$\underset{\sim}{x} = x^0 - \vec{\sigma} \cdot \vec{x} = x_\mu \underset{\sim}{\sigma}^\mu \quad \underset{\sim}{\sigma}^\mu = (1, \vec{\sigma}) \quad \underset{\sim}{\sigma}_\mu = (1, -\vec{\sigma})$$

$\pm B \rightarrow \Lambda(B)$

$x' = \Lambda(B)x$

$\underset{\sim}{x}' = B \underset{\sim}{x} B^+$

$$\underset{\sim}{x} = \begin{pmatrix} x^0 - x^3 & x^1 + ix^2 \\ x^1 - ix^2 & x^0 + x^3 \end{pmatrix} \quad \det \underset{\sim}{x} = x^2$$

La preuve est identique à la preuve précédente. En particulier l'expression de  $\Lambda(B)$  est obtenue par :

$$\begin{aligned} x'_\mu \underset{\sim}{\sigma}^\mu &= B x_\rho \underset{\sim}{\sigma}^\rho B^+ = B \underset{\sim}{\sigma}^\rho B^+ x_\rho & (\underset{\sim}{\sigma}^\mu = \sigma_\mu) \quad \underset{\sim}{\sigma}^\mu &= (1, \vec{\sigma}) \\ &= \Lambda_\mu^\rho x_\rho \underset{\sim}{\sigma}^\mu \end{aligned}$$

$$\text{donc } \Lambda_\mu^\rho \underset{\sim}{\sigma}^\mu = B \underset{\sim}{\sigma}^\rho B^+$$

$$\Lambda_\mu^\rho \underset{\sim}{\sigma}^\mu \underset{\sim}{\sigma}^\nu = B \underset{\sim}{\sigma}^\rho B^+ \underset{\sim}{\sigma}^\nu$$

$$\begin{aligned} \text{d'où } \Lambda_\nu^\rho &= \frac{1}{2} \text{Tr } B \underset{\sim}{\sigma}^\rho B^+ \underset{\sim}{\sigma}^\nu \\ &= \Lambda_\nu^{-1\rho} = \left( {}^t \Lambda_{(B)}^{-1} \right)_\rho^\nu \end{aligned}$$

$$\text{or } x^1 = B x B^+ \Leftrightarrow x = B^{-1} x' B^{-1+} \quad \text{donc } \Lambda^{-1}(B) = \Lambda(B^{-1})$$

$$\begin{aligned} \text{soit } {}^t \Lambda(B)_\rho^\nu &= \frac{1}{2} \text{Tr } B^{-1} \underset{\sim}{\sigma}^\rho B^{-1+} \underset{\sim}{\sigma}^\nu \\ \parallel \Lambda(B)_\rho^\nu &= \frac{1}{2} \text{Tr } B^{-1} \underset{\sim}{\sigma}^\nu B^{-1+} \underset{\sim}{\sigma}^\rho \end{aligned}$$

En posant  $B = (A^+)^{-1}$  on décrit la même transformation de Lorentz que celle définie par  $\tilde{x} \rightarrow \tilde{x}' = A \tilde{x} A^+$  :

$$\begin{aligned} \Lambda(A^{+-1})^\nu e &= \frac{1}{2} \text{Tr } A^+ \underset{\sim}{\sigma}^\nu A \underset{\sim}{\sigma}^\rho \\ &= \frac{1}{2} \text{Tr } A \underset{\sim}{\sigma}^\rho A^+ \underset{\sim}{\sigma}^\nu \quad (\text{par permutation cyclique}) \\ &= \tilde{\Lambda}(A)_\rho^\nu \quad \mathbf{c.q.f.d.} \end{aligned}$$

**Théorème :** Tout élément  $\Lambda$  de  $L_+^\uparrow$  se décompose de façon unique en un produit  $\Lambda_1 \Lambda_2$ , où  $\Lambda_1$  est une rotation et  $\Lambda_2$  une transformation de Lorentz spéciale.

**Preuve :** il suffit, d'après ce qui précède, de montrer que tout élément  $A$  de  $SL(2, \mathbb{C})$  se décompose de façon unique en produit  $A = UH$  avec  $\det H = \det U = 1$  et  $\begin{cases} UU^+ = U^+U = 1 & \rightarrow \text{rotation} \\ H^+ = H & \rightarrow \text{boost} \end{cases}$

Or  $A^+A = H^2$ , donc  $H = (A^+A)^{1/2}$  est la racine carrée définie positive unique (qui existe toujours) de  $A^+A$ , et  $U = A(A^+A)^{-1/2}$ . On a bien  $\det H = 1$  et  $\det U = 1$  **c.q.f.d.**

## 18.3 Représentations spinorielles de $L_+^\uparrow$

### 18.3.1 Identification des représentations spinorielles

D'après le paragraphe précédent, les matrices  $A$  engendrent une représentation de  $L_+^\uparrow$ . Identifions-la

$$A = 1 + \varepsilon_0 + \vec{\varepsilon} \cdot \vec{\sigma} + O(\varepsilon^2) \quad \varepsilon_0, \vec{\varepsilon} \in \mathbb{C}$$

$\det A = 1$  donc  $\varepsilon_0 = 0$

Donc toute matrice de  $SL(2, \mathbb{C})$  proche de l'identité peut s'écrire

$$A = 1 + \vec{\varepsilon} \vec{\sigma} = 1_{\frac{1}{2}} i \left( -\vec{\alpha} - i\vec{\beta} \right) \cdot \vec{\sigma} \quad (1)$$

Déterminons les générateurs de la représentation correspondante de  $L_+^\uparrow$ , notés  $J_{\mu\nu}$ , qui vérifient

$$A = 1 - \frac{i}{2} \omega^{\mu\nu} J_{\mu\nu}$$

$$\text{avec } \begin{cases} x'^\mu = x^\mu + \omega^{\mu\nu} x_\nu = \Lambda^{\mu\nu} x_\nu \\ x'^0 + \vec{\sigma} \cdot \vec{x}' = (1 + \vec{\varepsilon} \cdot \vec{\sigma})(x^0 + \vec{\sigma} \cdot \vec{x})(1 + \vec{\varepsilon} \cdot \vec{\sigma})^+ \end{cases}$$

$$\text{d'où } x'^0 + \vec{\sigma} \cdot \vec{x}' = x^0 + \vec{\sigma} \cdot \vec{x} + \vec{\varepsilon} \cdot \vec{\sigma}(x^0 + \vec{\sigma} \cdot \vec{x}) + (x^0 + \vec{\sigma} \cdot \vec{x}) \vec{\varepsilon}^* \cdot \vec{\sigma}$$

**Rappel :**

$$\begin{aligned} \vec{a} \cdot \vec{\sigma} \vec{b} \cdot \vec{\sigma} &= a_i \sigma_j b_j \sigma_j = a_i b_j (\delta_{ij} + i \varepsilon_{ijk} \sigma_k) \\ &= a_i b_i + i \varepsilon_{ijk} a_i b_j \sigma_k \\ &= \vec{a} \cdot \vec{b} + i(\vec{a} \wedge \vec{b}) \cdot \vec{\sigma} \end{aligned}$$

donc ici  $x'^0 + \vec{\sigma} \cdot \vec{x}' = x^0 + \vec{\sigma} \cdot \vec{x} + x^0(\vec{\varepsilon} + \vec{\varepsilon}^*) \cdot \vec{\sigma} + (\vec{\varepsilon} + \vec{\varepsilon}^*) \cdot \vec{x} + i((\vec{\varepsilon} - \vec{\varepsilon}^*) \wedge \vec{x}) \cdot \vec{\sigma}$

$$\vec{\varepsilon} = \frac{i}{2}(-\vec{\alpha} - i\vec{\beta}) \quad \left| \begin{array}{l} \vec{\varepsilon} + \vec{\varepsilon}^* = \vec{\beta} \\ \vec{\varepsilon} - \vec{\varepsilon}^* = -i\vec{\alpha} \end{array} \right.$$

$$\text{donc } \begin{cases} x'^0 = x^0 + \vec{x} \cdot \vec{\beta} = x^0 - \beta^i x_i \\ \vec{x}' = \vec{x} + \vec{\beta} x^0 + \vec{\alpha} \wedge \vec{x} \text{ donc } x'^i = x^i + \beta^i x^0 - \varepsilon_{ijk} \alpha^k x^j \end{cases}$$

Par comparaison avec  $x'^\mu = x^\mu + \omega^{\mu\nu} x_\nu$  il vient

$$\begin{cases} \omega^{0i} = -\beta^i \\ \omega^{ij} = +\varepsilon_{ijk} \alpha^k \end{cases} \quad \text{identique à L-2S (heureusement !)}$$

$$A = 1 - i\omega^{0i} J_{0i} - \frac{i}{2} \omega^{ij} J_{ij}$$

donc par comparaison avec (1) on a

$$J_{0i} = -i \frac{\sigma_i}{2} \quad J_{ij} = \varepsilon_{ijk} \frac{\sigma_k}{2} \quad (\leftarrow \varepsilon_{i'j'k} \sigma_k = \varepsilon_{i'j'k} \varepsilon_{ijk} J_{ij} = J_{i'j'} - J_{j'i'} = 2J_{i'j'})$$

$$\text{d'où } \left| \begin{array}{l} K_i = -i \frac{\sigma_i}{2} \\ J_i = \frac{\sigma_i}{2} \end{array} \right. \quad \text{puisque } \left| \begin{array}{l} K_i = J_{0i} \\ J_i = \frac{1}{2} \varepsilon_{ijk} J_{jk} = \frac{1}{2} \varepsilon_{ijk} \varepsilon_{jkk'} \frac{\sigma_{k'}}{2} \end{array} \right.$$

$$\text{ou encore } \begin{cases} M_i = \frac{\sigma_i}{2} \\ N_i = 0 \end{cases} \quad \text{représentation } \left( \frac{1}{2}, 0 \right)$$

A ces générateurs infinitésimaux est associée la matrice

$$A = e^{-i\alpha_i \frac{\sigma_i}{2}} e^{\beta_i \frac{\sigma_i}{2}} \in SL(2, \mathbb{C}) \quad A = e^{-i\theta \vec{n} \cdot \vec{J}} e^{i\phi \vec{m} \cdot \vec{k}}$$

$$\text{Posant } \begin{cases} \vec{\alpha} = \theta \vec{n} \\ \vec{\beta} = \phi \vec{m} \end{cases} \quad \text{avec } \vec{n}^2 = \vec{m}^2 = 1,$$

la transformation  $(\vec{\alpha}, \vec{\beta} = \vec{0})$  définit une rotation (**active**) autour de l'axe  $\vec{n}$  d'angle  $\theta$  :

$$\begin{cases} x'_0 = x_0 \\ \vec{x}' = \vec{x} \cos \theta + \vec{n} \wedge \vec{x} \sin \theta + (1 - \cos \theta) \vec{n} (\vec{n} \cdot \vec{x}) \quad (\text{cf L-19}) \end{cases}$$

la transformation  $(\vec{\alpha} = \vec{0}, \vec{\beta})$  définit une transformation spéciale (**active**) de Lorentz le long de

$$\text{l'axe } \vec{m} : \begin{cases} x'_0 = x_0 \cosh \phi + \vec{m} \cdot \vec{x} \sinh \phi \\ \vec{x}' = \vec{x} + x_0 \vec{m} \sinh \phi + \vec{m} (\vec{m} \cdot \vec{x}) (\cosh \phi - 1) \end{cases}$$

La représentation  $(0, \frac{1}{2})$  est définie par  $M_i = 0, N_i = \frac{\sigma_i}{2}$  soit  $K_i = \frac{i\sigma_i}{2}, J_i = \frac{\sigma_i}{2}$  et la matrice de  $SL(2, \mathbb{C})$  associée s'écrit  $B = e^{-i\alpha_i \frac{\sigma_i}{2}} e^{-\beta_i \frac{\sigma_i}{2}}$ . On vérifie bien que  $B = A^{+ -1}$ . De plus  $J_{0i} = \frac{i\sigma_i}{2}$  et  $J_{ij} = \varepsilon_{ijk} \frac{\sigma_k}{2}$ . On dispose ainsi de deux représentations du groupe de Lorentz.

### Théorème :

Ces deux représentations son inéquivalentes.

*Preuve :*

Supposons les représentations équivalentes, alors  $\forall A, \exists S$  t.q.  $(A^+)^{-1} = SAS^{-1} \Rightarrow \text{Trace } A = \text{Trace } (A^+)^{-1}$ . Cette condition est violée pour certaines transformations de  $SL(2, \mathbb{C})$  (elles ne doivent bien sûr pas appartenir à  $SU(2)$ ).

$$\begin{aligned} \text{Ex. : } A &= \begin{pmatrix} 2i & 0 \\ 0 & -\frac{i}{2} \end{pmatrix} \quad \text{Trace } A = \frac{3}{2}i \\ A^+ &= \begin{pmatrix} -2i & 0 \\ 0 & \frac{i}{2} \end{pmatrix} \quad A^{+ -1} = \begin{pmatrix} \frac{i}{2} & 0 \\ 0 & -2i \end{pmatrix} \quad \text{Trace } A^{+ -1} = -\frac{3}{2}i \end{aligned}$$

### 18.3.2 Représentations de $SL(2, \mathbb{C})$

Par morphisme, on construira naturellement l'algèbre de Lie de  $sl(2, \mathbb{C})$  dans la représentation  $D$  sous la forme :

$$A = \mathbb{1} - \frac{i}{2} \omega_{\alpha\beta} J_{4 \times 4}^{\alpha\beta} \rightarrow D(A(\omega)) = D(\Lambda) = \mathbb{1} - \frac{i}{2} \omega_{\alpha\beta} J^{\alpha\beta}$$

On en déduit que

$J^{\mu\nu}$  est un opérateur tensoriel antisymétrique à deux indices :

$$D^{-1}(\Lambda) J^{\mu\nu} D(\Lambda) = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma J^{\rho\sigma}$$

Pour une transformation infinitésimale, cette relation se ramène à l'algèbre de Lie  $D(\Lambda) = \exp -\frac{1}{2} i \varepsilon_{\gamma\delta} J^{\gamma\delta}$

$$D^{-1}(\Lambda)^{\alpha\beta'} (J^{\mu\nu})_{\beta'\alpha'} D(\Lambda)^{\alpha'}{}_\beta = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma (J^{\rho\sigma})^\alpha{}_\beta$$

$$D(\Lambda)^{\alpha'}{}_\beta = \delta^{\alpha'}{}_\beta - \frac{1}{2} i \varepsilon_{\gamma\delta} (J^{\gamma\delta})^{\alpha'}{}_\beta$$

$$\begin{aligned}
D(\Lambda)^{\alpha\beta'} &= g^{\alpha\beta'} - \frac{i}{2}\varepsilon_{\gamma\delta} (J^{\gamma\delta})^{\alpha\beta'} \\
\text{donc } D^{-1}(\Lambda)^{\alpha\beta'} &= g^{\alpha\beta'} + \frac{i}{2}\varepsilon_{\gamma\delta} (J^{\gamma\delta})^{\alpha\beta'} \\
\text{Ainsi } \left( g^{\beta'\alpha} + \frac{i}{2}\varepsilon_{\gamma\delta} (J^{\gamma\delta})^{\alpha\beta'} \right) (J^{\mu\nu})_{\beta'\alpha'} &= \left( g_{\beta}^{\alpha'} - \frac{i}{2}\varepsilon_{\gamma\delta} (J^{\gamma\delta})_{\beta}^{\alpha'} \right) \\
&= (\delta_{\rho}^{\mu} + \varepsilon_{\rho}^{\mu}) (\delta_{\sigma}^{\nu} + \varepsilon_{\sigma}^{\nu}) (J^{\rho\sigma})_{\beta}^{\alpha}
\end{aligned}$$

qui s'écrit encore :

$$\begin{aligned}
&+ \frac{i}{2}\varepsilon_{\gamma\delta} (J^{\gamma\delta} J^{\mu\nu} - J^{\mu\nu} J^{\gamma\delta})_{\beta}^{\alpha} = \varepsilon_{\rho}^{\mu} J^{\rho\nu} + \varepsilon_{\sigma}^{\nu} J^{\mu\sigma} \\
&= \frac{i}{2}\varepsilon_{\gamma\delta} \left[ \underbrace{-ig^{\gamma\mu} J^{\delta\nu} + ig^{\delta\nu} J^{\mu\gamma} + ig^{\delta\mu} J^{\gamma\nu} - ig^{\gamma\nu} J^{\mu\delta}} \right]_{\beta}^{\alpha}
\end{aligned}$$

on a fait apparaître la partie antisymétrique du tenseur afin de pouvoir identifier terme à terme

$$\text{donc } [J^{\gamma\delta}, J^{\mu\nu}]_{\beta}^{\alpha} = i (g^{\nu\delta} J^{\mu\gamma} - g^{\nu\gamma} J^{\mu\delta} + g^{\mu\delta} J^{\gamma\nu} - g^{\mu\gamma} J^{\delta\nu})_{\beta}^{\alpha} \quad \mathbf{c.q.f.d.}$$

*Rem.* : la relation de transformation du tenseur  $J^{\mu\nu}$  comme un opérateur tensoriel antisymétrique de rang 2 généralise le fait que les générateurs des rotations et des boosts sont des opérateurs vectoriels (tenseurs de rang 1) sur le groupe des rotations.

## 18.4 Spineurs quadridimensionnels

### 18.4.1 Spineurs contravariants et covariants

Par définition, c'est un objet à deux composantes complexes  $\xi^{\alpha}$  ( $\alpha = 1, 2$ ) qui se transforme de la façon suivante :

$$\xi \rightarrow A\xi \quad \text{représentation notée } D^{[1/2,0]}$$

Nous venons de voir que la représentation construite à l'aide de  $B = (A^+)^{-1}$  est inéquivalente à celle construite à l'aide de  $A$ . Comme  $C^t \bar{A}^{-1} C^{-1} = \bar{A}$ , les représentations construites à l'aide de  $A^{+-1}$  et  $\bar{A}$  sont équivalentes. Introduisons donc un deuxième espace de représentation, constitué aussi de vecteurs avec deux composantes complexes  $\eta^{\dot{\alpha}}$  ( $\dot{\alpha} = 1, 2$ ), appelés spineurs pointés. Par définition la loi de transformation correspondant à l'élément  $A$  de  $SL(2, \mathbb{C})$  est :

$$\eta \rightarrow \bar{A}\eta \quad D^{[0, \frac{1}{2}]}$$

Donc  $\eta^{\dot{\alpha}}$  se transforme comme  $\bar{\xi}^{\alpha}$ .

La manipulation des indices non pointés et pointés se fait comme dans le cas tridimensionnel, par le tenseur  $C_{\alpha\beta}$  :

$$\xi_{\alpha} = C_{\alpha\beta} \xi^{\beta} \quad \eta_{\dot{\alpha}} = C_{\dot{\alpha}\dot{\beta}} \eta^{\dot{\beta}}$$

$$\text{donc} \left| \begin{array}{l} \xi_1 = \xi^2 \\ \xi_2 = -\xi^1 \end{array} \right| \quad \left| \begin{array}{l} \eta_{\dot{1}} = \eta^{\dot{2}} \\ \eta_{\dot{2}} = -\eta^{\dot{1}} \end{array} \right|$$

Examinons la façon dont se transforme  $\eta_{\dot{\alpha}}$  :

$$\eta' = \bar{A}\eta \quad \text{i.e.} \quad \eta'^{\dot{\alpha}} = \bar{A}^{\dot{\alpha}}_{\dot{\beta}} \eta^{\dot{\beta}}$$

$$C_{\dot{\alpha}'\dot{\alpha}} \eta'^{\dot{\alpha}} = C_{\dot{\alpha}'\dot{\alpha}} \bar{A}^{\dot{\alpha}}_{\dot{\beta}} C_{\dot{\beta}\dot{\beta}'}^{-1} C_{\dot{\beta}''\dot{\beta}'} \eta^{\dot{\beta}'}$$

$$\text{soit} \quad \eta'_{\dot{\alpha}'} = A^{+1}{}^{\dot{\alpha}'}_{\dot{\beta}''} \eta_{\dot{\beta}''}$$

De même  $\xi_{\beta}$  se transforme à l'aide de  ${}^t A^{-1}$

### 18.4.2 Produit scalaire invariant

$\langle \xi, \xi' \rangle = {}^t \xi C \xi' = \xi^{\alpha} \xi'_{\alpha} = \xi^1 \xi'^2 - \xi^2 \xi'^1$  est invariant sous  $SL(2, \mathbb{C})$  puisque  $\det A = 1$  (ou encore :  ${}^t A C A = C$  car  $C A C^{-1} = {}^t A^{-1} A \in SL(2, \mathbb{C})$  donc  ${}^t A C A C^{-1} = 1$  **c.q.f.d.**)

$$\langle \eta, \eta' \rangle = {}^t \eta C \eta' = \eta^{\dot{\alpha}} \eta'_{\dot{\alpha}} = \eta^{\dot{1}} \eta'^{\dot{2}} - \eta^{\dot{2}} \eta'^{\dot{1}}$$

Dans le cas tridimensionnel la combinaison positive  $\rho = \xi^1 \bar{\xi}^1 + \xi^2 \bar{\xi}^2 = {}^t \xi_{\alpha} \bar{\xi}^{\alpha}$  est un invariant qui peut s'interpréter, si  $\xi$  est une fonction d'onde, comme une densité de probabilité de présence.

Dans le cas quadridimensionnel, ce n'est pas un invariant, car  $A$  n'est pas unitaire (sauf si c'est une rotation d'espace). Pour  $A$  quelconque,  $\xi^{\alpha} \rightarrow A \xi^{\alpha}$  et  $\bar{\xi}^{\alpha} \rightarrow \bar{A} \bar{\xi}^{\alpha}$  mais en général  ${}^t A \cdot \bar{A} \neq 1$  sauf si  $A \in SU(2)$ . Ceci n'est pas gênant pour l'interprétation de  $\rho$  car en théorie relativiste, la densité de probabilité ne doit pas être un scalaire mais plutôt la composante temporelle d'un quadrivecteur.

### 18.4.3 Spineurs d'ordre supérieur

Ils se définissent par produits tensoriels des représentations fondamentales constituées des spineurs et des spineurs pointés (i.e. des spineurs d'ordre 1).

$$\begin{aligned} \text{Ex. spineurs d'ordre 2 : } \quad & \xi^{\alpha\beta} \sim \xi^{\alpha} \theta^{\beta} \\ & \phi^{\alpha\dot{\beta}} \sim \xi^{\alpha} \eta^{\dot{\beta}} \\ & \xi^{\dot{\alpha}\dot{\beta}} \sim \eta^{\dot{\alpha}} \omega^{\dot{\beta}} \end{aligned}$$

*Rem.* : les spineurs mixtes d'ordre 2  $\phi^{\alpha\dot{\beta}}$  correspondent aux quadrivecteurs  $x^{\mu}$ .

- L'ordre d'un spineur est indiqué par un couple de nombres  $(k, l)$  représentant les nombres d'indices pointés et non pointés.

- Les indices pointés et non pointés ne peuvent se mélanger sous  $L_{+}^{\uparrow}$ , donc il n'est pas nécessaire de spécifier l'ordre des indices.

- Covariance d'une égalité spinorielle : l'ordre doit être le même dans les deux membres de l'égalité (en tenant compte de la conjugaison complexe).

$$\text{ex. : } \eta^{\alpha\dot{\beta}} = \bar{\xi}^{\dot{\alpha}\beta} \text{ est covariante}$$

- la contraction de spineurs par le tenseur  $C_{\alpha\beta}$  n'a de sens que sur des indices de même espèce (la contraction sur des indices de nature différente n'est pas covariante) :



$$\begin{aligned}\xi^\beta C_{\beta\alpha} \eta^\alpha &\rightarrow \xi'^{\beta'} C_{\beta'\alpha'} \eta'^{\alpha'} = A^\beta_{\beta'} \xi^{\beta'} C_{\beta\alpha} A^\alpha_{\alpha'} \eta'^{\alpha'} \\ A^\beta_{\beta'} C_{\beta\alpha} A^\alpha_{\alpha'} &= ({}^t A C A)^{\beta'}_{\alpha'} \\ \text{or } {}^t A C A &= C \underbrace{C^{-1} {}^t A C}_{=I} A = C A^{-1} A = C\end{aligned}$$

En revanche

$$\begin{aligned}\xi^\beta C_{\beta\dot{\alpha}} \eta^{\dot{\alpha}} &\rightarrow \xi'^{\beta'} C_{\beta'\dot{\alpha}'} \eta'^{\dot{\alpha}'} = A^\beta_{\beta'} \xi^{\beta'} C_{\beta\dot{\alpha}} \bar{A}^{\dot{\alpha}}_{\dot{\alpha}'} \eta'^{\dot{\alpha}'} \\ &= A^\beta_{\beta'} C_{\beta\dot{\alpha}} \bar{A}^{\dot{\alpha}}_{\dot{\alpha}'} \\ &= ({}^t A C \bar{A})^{\beta'}_{\dot{\alpha}'} = C A^{-1} \bar{A} \neq C\end{aligned}$$

On obtient une succession de tenseurs de rang moindre en contractant par paires les indices d'un tenseur.

#### 18.4.4 Représentation irréductibles de $SL(2, \mathbb{C})$

Le résultat démontré pour les tenseurs 3d subsiste : les représentations irréductibles d'ordre supérieur de  $SL(2, \mathbb{C})$  sont données par les spineurs  $\xi^{\alpha_1 \alpha_2 \dots \alpha_k} \beta_1 \beta_2 \dots \beta_l$  séparément symétriques sur les  $k$  indices  $\alpha$  et  $l$  indices  $\beta$ .

Un spineur complètement symétrique de rang  $(k, l)$  possède  $(k+1)(l+1)$  composantes indépendantes, ce qui constitue la dimension de la représentation irréductible réalisée par ce tenseur spinoriel.

Représentations irréductibles  $D^{[j_1, j_2]}$  de dimension finie de  $SL(2, \mathbb{C})$  : elles sont caractérisées par deux nombres  $j_1$  et  $j_2$  entiers ou demi-entiers, sont de dimension  $(2j_1+1)(2j_2+1)$ , et données sur une base de polynôme homogène de degré  $2j_1+2j_2$

$$\begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} \rightarrow A \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} \quad \begin{pmatrix} \xi^3 \\ \xi^4 \end{pmatrix} \rightarrow \bar{A} \begin{pmatrix} \xi^3 \\ \xi^4 \end{pmatrix}$$

$$|j_1 j_2 m_1 m_2 \rangle \rightarrow \sum_{m'_1 m'_2} D^{[j_1, j_2]}_{m_1, m'_1, m_2, m'_2}(A) |j_1 j_2 m'_1 m'_2 \rangle$$

avec  $m_k = -j_k, -j_k+1, \dots, j_k \quad k=1, 2$

$$|j_1 j_2 m_1 m_2 \rangle = \frac{(\xi^1)^{j_1+m_1} (\xi^2)^{j_1-m_1} (\xi^3)^{j_2+m_2} (\xi^4)^{j_2-m_2}}{\sqrt{(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!}}$$

donc  $D^{[j_1, j_2]}(A) = D^{j_1, 0}(A) \otimes D^{[0, j_2]}(A)$

$$D^{[0, j]}(A) = D^{[j, 0]}(\bar{A})$$

$D^{[0, j]}(A)$  et  $D^{[j, 0]}(A)$  sont inéquivalentes, l'équivalence a lieu si et seulement si  $A \in SU(2)$ .

$$D^{[0, 0]}(A) = 1$$

$$D^{[\frac{1}{2}, 0]}(A) = A$$

$$D^{[0, \frac{1}{2}]}(A) = \bar{A}$$

$$D^{[\frac{1}{2}, \frac{1}{2}]}(A) = A \otimes \bar{A}$$

multispineurs :  $a^{\alpha_1 \alpha_2 \dots \alpha_i \dot{\beta}_1 \dot{\beta}_2 \dots \dot{\beta}_j} \in \mathbb{C}^{2 \otimes i} \otimes \mathbb{C}^{2 \otimes j}$   
 sur lesquels  $\left(D^{[\frac{1}{2}, 0]}\right)^{\otimes i} \otimes \left(D^{[0, \frac{1}{2}]}\right)^{\otimes j}$  agit par

$$\left(D^{[\frac{1}{2}, 0]}\right)^{\otimes i} \otimes \left(D^{[0, \frac{1}{2}]}\right)^{\otimes j} : a \rightarrow A^{\alpha_1}_{\gamma_1} \dots A^{\alpha_i}_{\gamma_i} \bar{A}^{\dot{\beta}_1}_{\delta_1} \dots \bar{A}^{\dot{\beta}_j}_{\delta_j} a^{\gamma_1 - \gamma_i \delta_1 \dots \delta_j}$$

pour  $i + j > 1$ , ces représentations sont réductibles, les représentations réductibles s'obtenant par symétrisation séparé par rapport aux indices pointés et non pointés. Le lien entre multispineurs symétriques de rang  $(2i, 2j)$  et  $D^{[i, j]}$  est immédiat en utilisant la preuve faite pour  $SU(2)$  (cf  $\Phi_{jm} \dots$ ).

Le produit scalaire  $d^2 \varphi d^2 \varphi^* e^{-\varphi^+ \cdot \varphi}$  n'étant plus invariant, ces représentations sont en général **non unitaires**, ce qui n'est pas étonnant puisque  $SL(2, \mathbb{C})$  est non compact.

### 18.4.5 Application : lien entre spineurs mixtes et quadrivecteurs

On sait que  $\vec{W} = {}^t \xi \vec{\sigma} \xi$  et  $\vec{V} = {}^t \bar{\eta} \vec{\sigma} \eta$  sont des vecteurs pour les rotations (voir cours sur les spineurs 3d).  $\left(\begin{array}{l} \xi \text{ a pour composantes } \xi^\alpha \\ \eta \text{ a pour composantes } \eta_{\dot{\alpha}} \end{array}\right)$

$\bar{\xi}^\alpha \xi^\alpha$  et  $\bar{\eta}_{\dot{\alpha}} \eta_{\dot{\alpha}}$  étant des scalaires sous les rotations (mais non invariants sous les boosts, voir plus haut), vérifions que  $W^\mu = \bar{\xi}^\alpha \sigma^\mu \xi^\alpha$  et  $V^\mu = \bar{\eta}_{\dot{\alpha}} \sigma^\mu \eta_{\dot{\alpha}}$  se transforment comme les composantes contravariantes de deux quadrivecteurs

$$W^\mu = \begin{pmatrix} \bar{\xi}^\alpha \xi^\alpha \\ \bar{\xi}^\alpha \vec{\sigma} \xi^\alpha \end{pmatrix} \quad V_\mu = \begin{pmatrix} \bar{\eta}_{\dot{\alpha}} \eta_{\dot{\alpha}} \\ \bar{\eta}_{\dot{\alpha}} \vec{\sigma} \eta_{\dot{\alpha}} \end{pmatrix} \quad W'^\mu = \begin{pmatrix} \bar{\xi}'^\alpha \xi'^\alpha \\ \bar{\xi}'^\alpha \vec{\sigma} \xi'^\alpha \end{pmatrix} = \begin{pmatrix} \bar{\xi}^\alpha A^+ A \xi^\alpha \\ \bar{\xi}^\alpha A^+ \vec{\sigma} A \xi^\alpha \end{pmatrix}$$

Considérons un boost infinitésimal :  $A = e^{d\vec{\beta} \cdot \frac{\vec{\sigma}}{2}}$   $A^+ = e^{d\vec{\beta} \cdot \frac{\vec{\sigma}}{2}}$   $A^+ A = e^{d\vec{\beta} \cdot \vec{\sigma}}$

$$\text{donc } \bar{\xi}^\alpha A^+ A \xi^\alpha = \bar{\xi}^\alpha \xi^\alpha + d\vec{\beta} \cdot \left(\bar{\xi}^\alpha \vec{\sigma} \xi^\alpha\right)$$

$$\bar{\xi}^\alpha A^+ \vec{\sigma} A \xi^\alpha = \bar{\xi}^\alpha \left(1 + d\vec{\beta} \cdot \frac{\vec{\sigma}}{2}\right) \vec{\sigma} \left(1 + d\vec{\beta} \cdot \frac{\vec{\sigma}}{2}\right) \xi^\alpha$$

supposons par exemple que  $d\vec{\beta} = d\beta \vec{u}_x$

$$\begin{aligned} \bar{\xi}^\alpha \left(1 + d\vec{\beta} \cdot \frac{\sigma_1}{2}\right) \vec{\sigma} \left(1 + d\vec{\beta} \cdot \frac{\sigma_1}{2}\right) \xi^\alpha &= \bar{\xi}^\alpha \vec{\sigma} \xi^\alpha + \bar{\xi}^\alpha \underbrace{\left(\sigma_1 \vec{\sigma} + \vec{\sigma} \sigma_1\right)}_{2\delta_{1\dot{2}}} \xi^\alpha \frac{d\beta}{2} \\ &= \bar{\xi}^\alpha \vec{\sigma} \xi^\alpha + d\beta \bar{\xi}^\alpha \xi^\alpha \end{aligned}$$

d'où  $W'^\mu = \begin{pmatrix} 1 & d\beta \\ d\beta & 1 \\ & & 1 \\ & & & 1 \end{pmatrix} W^\mu$  ce qui est bien la loi de transformation des composantes

contravariantes d'un 4-vecteur sous un boost actif

de même :  $V'_\mu = \begin{pmatrix} \bar{\eta}_{\dot{\alpha}} A^{-1} A^{+^{-1}} \eta_{\dot{\alpha}} \\ \bar{\eta}_{\dot{\alpha}} A^{-1} \vec{\sigma} A^{+^{-1}} \eta_{\dot{\alpha}} \end{pmatrix}$

$$A^{+^{-1}} = e^{-d\vec{\beta} \cdot \frac{\vec{\sigma}}{2}}$$

$$A^{-1} = e^{-d\vec{\beta} \cdot \frac{\vec{\sigma}}{2}}$$

donc  $V^\mu = \begin{pmatrix} 1 & -d\beta & & \\ -d\beta & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} V^\mu$  transformation des composantes covariantes d'un 4-vecteur

sous un boost actif

Comme  $\bar{\xi}^\alpha \sim \eta^{\dot{\alpha}}$ ,  $W^\mu$  peut encore s'écrire

$$W^\mu = \eta^{\dot{\alpha}} \underset{\sim}{\sigma}^\mu \xi^\alpha$$

$$\text{soit } \begin{cases} W^0 = \eta^{\dot{\alpha}} \xi^\alpha = \eta^{\dot{1}} \xi^1 + \eta^{\dot{2}} \xi^2 \\ W^1 = {}^t \begin{pmatrix} \eta^{\dot{1}} \\ \eta^{\dot{2}} \end{pmatrix} \sigma_1 \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} = \begin{pmatrix} \eta^{\dot{1}} \eta^{\dot{2}} \end{pmatrix} \begin{pmatrix} \xi^2 \\ \xi^1 \end{pmatrix} = \xi^2 \eta^{\dot{1}} + \xi^1 \eta^{\dot{2}} \\ W^2 = {}^t \begin{pmatrix} \eta^{\dot{1}} \\ \eta^{\dot{2}} \end{pmatrix} \sigma_2 \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix} = i \left( \xi^1 \eta^{\dot{2}} - \xi^2 \eta^{\dot{1}} \right) \\ W^3 = \xi^1 \eta^{\dot{1}} - \xi^2 \eta^{\dot{2}} \end{cases}$$

Ainsi au spineur mixte de rang 2  $\varphi^{\alpha\dot{\alpha}}$  correspond le quadrivecteur

$$W^\mu : \begin{cases} W^0 = \frac{1}{\sqrt{2}} \left( \varphi^{1\dot{1}} + \varphi^{2\dot{2}} \right) \\ W^1 = \frac{1}{\sqrt{2}} \left( \varphi^{1\dot{2}} + \varphi^{2\dot{1}} \right) \\ W^2 = \frac{i}{\sqrt{2}} \left( \varphi^{1\dot{2}} - \varphi^{2\dot{1}} \right) \\ W^3 = \frac{1}{\sqrt{2}} \left( \varphi^{1\dot{1}} - \varphi^{2\dot{2}} \right) \end{cases} \quad (\text{normalisation pour que } W^2 = \varphi_{\alpha\dot{\beta}} \varphi^{\alpha\dot{\beta}})$$

$$\Leftrightarrow : \begin{cases} \varphi^{1\dot{1}} = \varphi_{2\dot{2}} = \frac{1}{\sqrt{2}}(a^0 + a^3) \\ \varphi^{2\dot{2}} = \varphi_{1\dot{1}} = \frac{1}{\sqrt{2}}(a^0 - a^3) \\ \varphi^{1\dot{2}} = -\varphi_{2\dot{1}} = \frac{1}{\sqrt{2}}(a^1 - ia^2) \\ \varphi^{2\dot{1}} = -\varphi_{1\dot{2}} = \frac{1}{\sqrt{2}}(a^1 + ia^2) \end{cases}$$

soit  $\varphi^{\alpha\dot{\beta}} = \frac{1}{\sqrt{2}} a^\mu \sigma_\mu$  de même  $\varphi_{\dot{\alpha}\beta} = \frac{1}{\sqrt{2}} a^\mu \underset{\sim}{\sigma}^\mu$ .

## 18.5 Inversion d'espace et bispineurs

**3d** : l'inversion d'espace commute avec les rotations

**4d** : l'inversion d'espace, notée  $I_s$ , ne commute pas avec les transformations de Lorentz pures  $\rightarrow$  cette opération permet de relier les représentations complexes conjuguées l'une de l'autre, et en particulier de transformer les spineurs non pointés en spineurs pointés. Elle permet l'introduction des bispineurs, nécessaires à la théorie des particules massives de spin demi-entier, tels l'électron ou les nucléons.

**3d** :  $\xi \rightarrow I_s \xi = P \xi$  où  $P$  est un coefficient constant, en effet  $I_s$  est diagonal pour un spineur 3d puisque  $I_s$  commute avec les rotations, d'où le résultat d'après le Lemme de Schur.

$P$  est défini avec une certaine liberté : deux inversions successives permettent de revenir à l'état initial. Mais ce retour peut être considéré soit comme une rotation d'angle  $\theta$ , soit comme une rotation d'angle  $2n$ .

Pour un spineur, qui change de signe par rotation de  $2n$ , ces deux possibilités ne sont pas équivalentes.

$$\begin{aligned} \text{deux possibilités : } P^2 = 1 &\Rightarrow P = \pm 1 \\ P^2 = -1 &\Rightarrow P = \pm i \end{aligned}$$

Nous choisissons la seconde possibilité.

*Rem.* : la parité ne peut avoir de caractère absolu, puisque les spineurs changent de signe dans une rotation de  $2n$ , qui peut toujours être faite en même temps que l'inversion.

En revanche la parité du scalaire  $\psi^\alpha\psi_\alpha$  est absolue, tous les spineurs changeant **simultanément** de signe dans une rotation de  $2n$ , ce qui n'affecte pas la parité de ce scalaire.

**4d** : notons  $L_{\vec{V}}$  la transformation de Lorentz pure de vitesse  $\vec{V}$

$$\text{Alors } I_s L_{\vec{V}} = L_{-\vec{V}} I_s$$

$$\text{Donc } [I_s, L_{\vec{V}}] \neq 0 \text{ (pour } \vec{V} \neq \vec{0}\text{)}$$

Ainsi les composantes d'un spineur quadridimensionnel ne peuvent se transformer en elles-mêmes (sinon  $I_s$  serait multiple de l'identité, puisque  $I_s$  commute avec les rotations, et donc  $I_s$  commuterait avec les transformations de Lorentz pure!). La seule solution est que les composantes de  $\xi^\alpha$  se transforment dans d'autres quantités :

$$\xi^\alpha \rightarrow i\eta_{\dot{\alpha}} \quad \eta_{\dot{\alpha}} \rightarrow i\xi^\alpha$$

qui vérifie bien  $P^2 = -1$ .

On a donc

$$\xi_\alpha \rightarrow -i\eta_{\dot{\alpha}} \quad \eta^{\dot{\alpha}} \rightarrow -i\xi_\alpha$$

*Rem.* : Comme l'inversion d'espace ne change pas la valeur du spin, on doit avoir  $\xi^\alpha \rightarrow i\eta_{\dot{\beta}}$  avec  $\beta = \alpha$ .

**En effet** : sous une rotation les 4-spineurs se comportent comme des 3-spineurs. Or ces derniers vérifient  $\bar{\psi}_\alpha \sim \psi^\alpha$ .

Comme  $\eta^{\dot{\alpha}} \sim \bar{\xi}^{\dot{\alpha}}$ ,  $\eta_{\dot{\alpha}} \underset{SL(2,\mathbb{C})}{\sim} \bar{\xi}_\alpha \underset{SL(2,\mathbb{C})}{\sim} \xi^\alpha$ , donc si  $\xi^1$  correspond à la projection du spin  $1/2$  et  $\xi^2$  à la projection du spin  $-1/2$ , il en est de même pour  $\eta_1$  et  $\eta_2$ .

Pour combiner de façon cohérente inversion et groupe de Lorentz restreint, on doit donc considérer des espaces de représentation contenant simultanément les spineurs et les spineurs pointés. Au minimum, il faut considérer les couples  $(\xi^\alpha, \eta_{\dot{\alpha}})$ , appelés bispineurs d'ordre 1. Partant de cette représentation minimale, on peut construire par produits tensoriels les représentations irréductibles du groupe de Lorentz complet.

Produit scalaire à partir de  $(\xi^\alpha, \eta_{\dot{\alpha}})$  et  $(\theta^\alpha, H_{\dot{\alpha}})$  :

- soit  $\xi^\alpha\theta_\alpha + \eta_{\dot{\alpha}}H^{\dot{\alpha}}$  ce qui donne un scalaire de Lorentz ( $\rightarrow \eta_{\dot{\alpha}}H^{\dot{\alpha}} + \xi^\alpha\theta_\alpha$  sous  $I_s$ )

$$\text{rem. : } \xi^\alpha\theta_\alpha + \eta_{\dot{\alpha}}H^{\dot{\alpha}} = \xi^\alpha\theta_\alpha - \eta^{\dot{\alpha}}H_{\dot{\alpha}}$$

- soit  $\xi^\alpha\theta_\alpha - \eta_{\dot{\alpha}}H^{\dot{\alpha}}$  ce qui donne un pseudo-scalaire de Lorentz

$$\text{rem. : } \xi^\alpha\theta_\alpha - \eta_{\dot{\alpha}}H^{\dot{\alpha}} = \xi^\alpha\theta_\alpha + \eta^{\dot{\alpha}}H_{\dot{\alpha}}$$

*Ex.* : spineurs mixtes d'ordre 2

$$\text{soit } \varphi^{\alpha\dot{\beta}} \sim \xi^\alpha H^{\dot{\beta}} + \theta^\alpha \eta_{\dot{\beta}} \xrightarrow{I_s} \varphi_{\dot{\alpha}\beta}$$

correspond à un vrai vecteur  $a^\mu$ , i.e. qui se transforme selon  $(a^0, \vec{a}) \xrightarrow{I_s} (a^0, -\vec{a})$

soit  $\varphi^{\alpha\beta} \sim \xi^\alpha H^\beta - \theta^\alpha \eta_\beta \xrightarrow{I_s} -\varphi_{\dot{\alpha}\dot{\beta}}$

correspond à un 4 pseudo-vecteur :  $(a^0, \vec{a}) \xrightarrow{I_s} (-a^0, \vec{a})$

**Preuve :** il suffit d'examiner les formules d'inversion p. SL-17

*Rem.* : l'opération de parité s'écrit

$$\begin{pmatrix} \xi^1 \\ \xi^2 \\ \eta_1 \\ \eta_2 \end{pmatrix}_s = i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \eta_1 \\ \eta_2 \end{pmatrix}$$

l'opération de parité doit conserver le spin. On pourrait donc a priori mélanger les spineurs de la façon suivante par parité :

$$\gamma_P = i \begin{pmatrix} \xi^1 & \eta_1 & \xi^2 & \eta_2 \\ A & & 0 & \\ 0 & & B & \end{pmatrix} \quad \gamma_P^2 = -\mathbb{1}_{4 \times 4} = - \begin{pmatrix} A^2 & 0 \\ 0 & B^2 \end{pmatrix}$$

donc  $A^2 = B^2 = \mathbb{1}_{2 \times 2}$ .

$A$  et  $B$  ont donc pour valeurs propres 1 et  $-1$  (car  $(\frac{A}{B} - 1)(\frac{A}{B} + 1) = 0$ ).

Comme  $A$  et  $B$  agissent sur des espaces qui sont en somme directs, on peut les diagonaliser simultanément : or  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  a pour valeurs propres 1 et  $-1$ .

Don on peut encore écrire  $P\gamma_P P_i^{-1} = i \begin{pmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & \\ & & 1 & 0 \end{pmatrix}$  **c.q.f.d.**



# Chapitre 19

## Le groupe de Poincaré

### 19.1 Structure

## 19.2 Représentations



# Chapitre 20

## L'équation de Klein-Gordon

### 20.1 Construction

## 20.2 Solutions

# Chapitre 21

## L'équation de Dirac

### 21.1 L'équation de Dirac dans la représentation spinorielle

Nous cherchons une équation d'onde agissant sur un champ spinoriel, invariante sous l'action du groupe de Lorentz et l'opération de parité.

Nous avons vu qu'il existe deux représentations irréductibles inéquivalentes de dimension minimale du groupe de Lorentz restreint :

$$\xi^\alpha \rightarrow A\xi^\alpha \quad \text{et} \quad \eta^{\dot{\alpha}} \rightarrow \bar{A}\eta^{\dot{\alpha}} \quad \text{ou encore} \quad \eta_{\dot{\alpha}} \rightarrow A^{+^{-1}}\eta$$

Les bispineurs  $(\xi^\alpha, \eta_{\dot{\alpha}})$  permettent d'obtenir une représentation de dimension minimale du groupe de Lorentz complet, de spin  $\frac{1}{2}$ . Soit donc  $\psi = (\xi^\alpha, \eta_{\dot{\alpha}})$ . Notons  $\psi^\alpha$  ( $\alpha = 1, 2, 3, 4$ ) les quatres composantes **indépendantes** de  $\psi$ . A l'opérateur quadri-impulsion  $P_\mu = i\hbar \frac{\partial}{\partial x^\mu}$  correspond les spineurs opératoriels

$$\tilde{P} = P^\mu \sigma_\mu = P^0 + \vec{\sigma} \cdot \vec{P}$$

(cf. cours sur Lorentz : SL-17 (on change de normalisation (facteur  $\frac{1}{\sqrt{2}}$ )) ou SL-2).

$\tilde{P}$  est une matrice  $2 \times 2$  dont les éléments  $P^{\alpha\dot{\beta}}$  dépendent linéairement des  $P_\mu$ . Nous disposons également d'une constante, la masse  $m$ . Enfin

$$P_{\dot{\beta}\alpha} = P^\mu \sigma_{\dot{\beta}\alpha} = P^0 - \vec{\sigma} \cdot \vec{P}$$

qui a pour éléments de matrice  $P_{\dot{\beta}\alpha}$ .

La covariance relativiste impose donc la forme de l'équation d'onde de degré minimal qu'il est possible d'écrire pour les bispineurs

$$\begin{cases} P^{\alpha\dot{\beta}}\eta_{\dot{\beta}} = m\xi^\alpha \\ P_{\dot{\beta}\alpha}\xi^\alpha = m\eta_{\dot{\beta}} \end{cases}$$

Multiplions la deuxième équation par  $P^{\gamma\dot{\beta}}$ . Comme

$$\begin{aligned} \left(P^0 + \vec{P} \cdot \vec{\sigma}\right)^{\gamma\dot{\beta}} \left(P^0 - \vec{P} \cdot \vec{\sigma}\right)_{\dot{\beta}\alpha} &= \left(P^{02} - (\vec{P} \cdot \vec{\sigma})(\vec{P} \cdot \vec{\sigma})\right)_{\alpha}^{\gamma} \\ &= \left(P^{02} - \vec{P}^2\right)_{\alpha}^{\gamma} = P^2 \delta_{\alpha}^{\gamma} \\ &\quad \uparrow 2 \times 2 \end{aligned}$$

(et de même  $P_{\dot{\gamma}\alpha}P^{\alpha\dot{\beta}} = \delta_{\dot{\gamma}}^{\dot{\beta}}$ ) on en déduit :  $P^2\delta^{\gamma}_{\alpha}\xi^{\alpha} = m^2\xi^{\gamma}$  soit  $\|P^2\xi^{\gamma} = m^2\xi^{\gamma}$ .

De même  $P^2\eta_{\dot{\gamma}} = m^2\eta_{\dot{\gamma}}$ . Ainsi  $\left\| \begin{array}{l} (\square + m^2)\xi^{\alpha} = 0 \\ (\square + m^2)\eta_{\dot{\beta}} = 0 \end{array} \right.$

Les deux spineurs vérifient séparément l'équation de Klein-Gordon, ce qui montre que  $m$  est bien la masse de la particule.

C'est la présence de la masse dans l'équation d'onde qui impose de considérer simultanément les deux spineurs  $\xi^{\alpha}$  et  $\eta^{\dot{\alpha}}$ .

En revanche pour une particule sans masse, on peut écrire l'équation

$$P^{\alpha\dot{\beta}}\eta_{\dot{\beta}} = 0 \quad : \quad \text{spineur de Weyl}$$

**Parité :**

$$\begin{array}{ll} \xi^{\alpha} \rightarrow i\eta_{\dot{\alpha}} & \eta_{\dot{\alpha}} \rightarrow i\xi^{\alpha} \\ \xi_{\alpha} \rightarrow -i\eta^{\dot{\alpha}} & \eta^{\dot{\alpha}} \rightarrow -i\xi_{\alpha} \end{array}$$

Donc  $P^{\alpha\dot{\beta}}\eta_{\dot{\beta}} = m\xi^{\alpha}$  devient  $i(-i)P_{\dot{\alpha}\beta}i\xi^{\beta} = im\eta_{\dot{\alpha}}$  soit  $P_{\dot{\alpha}\beta}\xi^{\beta} = m\eta_{\dot{\alpha}}$   
et  $P_{\dot{\beta}\alpha}\xi^{\alpha} = m\eta_{\dot{\beta}}$  devient  $i(-i)P^{\beta\dot{\alpha}}i\eta_{\dot{\alpha}} = im\xi^{\beta}$  soit  $P^{\beta\dot{\alpha}}\eta_{\dot{\alpha}} = m\xi^{\beta}$

La parité échange donc les deux équations.

Pour  $m \neq 0$  on a introduit une équation invariante par parité.

*Rem.* : on pourrait arbitrairement introduire une phase  $e^{i\varphi}$  de sorte que

$$\left\{ \begin{array}{l} P^{\alpha\dot{\beta}}\eta_{\dot{\beta}} = e^{i\varphi}m\xi^{\alpha} \\ P_{\dot{\beta}\alpha}\xi^{\alpha} = e^{-i\varphi}m\eta_{\dot{\beta}} \end{array} \right.$$

le choix de phases opposées permet de toujours avoir  $\left\{ \begin{array}{l} (\square + m^2)\xi^{\alpha} = 0 \\ (\square + m^2)\eta_{\dot{\beta}} = 0 \end{array} \right.$  et donc que  $m$  est bien la masse.

Lorsque la particule est au repos, la distinction entre spineurs  $(\frac{1}{2}, 0)$  et  $(0, \frac{1}{2})$  n'a plus de sens car ce sont les boosts qui permettent de distinguer ces deux types de spineurs. Il est donc légitime de poser  $\xi^{\alpha}(m, \vec{p} = 0) = \eta_{\dot{\alpha}}(m, \vec{p} = 0)$ , ce qui revient à fixer la phase à zéro puisque

$$\text{pour } \vec{p} = \vec{0}, \left| \begin{array}{l} m\eta_{\dot{\alpha}} = me^{i\varphi}\xi^{\alpha} \\ m\xi^{\alpha} = me^{-i\varphi}\eta_{\dot{\alpha}} \end{array} \right.$$

$$\left\{ \begin{array}{l} (P_0 + \vec{P} \cdot \vec{\sigma})\eta = m\xi \\ (P_0 - \vec{P} \cdot \vec{\sigma})\xi = m\eta \end{array} \right. \quad (21.1)$$

s'écrit, en notation matricielle :

$$\begin{pmatrix} 0 & P_0 + \vec{P} \cdot \vec{\sigma} \\ P_0 - \vec{P} \cdot \vec{\sigma} & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = m \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (21.2)$$

Posons

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} \quad (21.3)$$

L'équation matricielle précédente se met alors sous la forme usuelle de l'équation de Dirac

$$(\mathcal{P} - m)\psi = 0 \quad (21.4)$$

avec

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (21.5)$$

et

$$\mathcal{P} = \gamma^\mu p_\mu = \gamma \cdot p = \gamma^0 p_0 - \vec{\gamma} \cdot \vec{p} = i\gamma^0 \frac{\partial}{\partial t} + i\vec{\gamma} \cdot \vec{\nabla} \quad (21.6)$$

car

$$\vec{P} = -i \frac{\partial}{\partial x^i} = -i\vec{\nabla} \quad (21.7)$$

(rappel :  $\partial_\mu = \frac{\partial}{\partial x^\mu} = \left( \frac{\partial}{\partial x^0}, \vec{\nabla} \right)$  et  $\partial^\mu = \frac{\partial}{\partial x_\mu} = \left( \frac{\partial}{\partial x^0}, -\vec{\nabla} \right)$ ).

Les matrices  $\gamma^\mu = (\gamma^0, \vec{\gamma})$  sont appelées matrices gamma ou matrices de Dirac.

## 21.2 Matrices $\gamma$

Nous avons obtenu l'équation de Dirac à partir des représentations spinorielles de  $SL(2, \mathbb{C})$ . Elle explicite naturellement l'invariance relativiste. Cependant dans les applications un changement de représentation peut être utile : on peut effectuer des transformations linéaires unitaires sur les quatre composantes du bispineur  $\psi : \psi \rightarrow \psi' = U\psi$  avec  $U$  une matrice  $4 \times 4$  unitaire. (On se limite aux transformations unitaires pour garder l'interprétation physique de  $\psi^+(x)\psi(x)$  comme étant une densité de probabilité de présence : voir plus loin).

Partant de  $(\gamma \cdot p - m)\psi = 0$ , l'équation transformée s'obtient par :

$$(\gamma \cdot p - m)U^+\psi' = 0 \text{ soit } \underbrace{(U\gamma U^+ \cdot p - m)}_{\gamma'}\psi' = 0$$

d'où  $(\gamma' \cdot p - m)\psi' = 0$  avec  $\gamma' = U\gamma U^+$

Les matrices  $\gamma$  obtenues dans la représentation spinorielle satisfont :

$$\| \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \mathbf{1}$$

en effet :

$$\gamma^0 \gamma^0 = 1 \quad (21.8)$$

$$\gamma^i \gamma^j = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} \begin{pmatrix} 0 & -\sigma_j \\ \sigma_j & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_i \sigma_j & 0 \\ 0 & -\sigma_i \sigma_j \end{pmatrix} \quad (21.9)$$

$$\gamma^i \gamma^j + \gamma^j \gamma^i = \begin{pmatrix} -(\sigma_i \sigma_j + \sigma_j \sigma_i) & 0 \\ 0 & -(\sigma_i \sigma_j + \sigma_j \sigma_i) \end{pmatrix} \quad (21.10)$$

or

$$\sigma_i \sigma_j = \delta_{ij} + i\varepsilon_{ijk} \sigma_k \quad (21.11)$$

donc

$$\gamma^i \gamma^j + \gamma^j \gamma^i = \begin{pmatrix} -2\delta_{ij} & 0 \\ 0 & -2\delta_{ij} \end{pmatrix} = -2\delta_{ij} \mathbf{1} \quad (21.12)$$

$$\gamma^0 \gamma^i = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \quad (21.13)$$

$$\gamma^i \gamma^0 = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \quad (21.14)$$

d'où

$$\gamma^0 \gamma^i + \gamma^i \gamma^0 = 0 \quad (21.15)$$

**c.q.f.d.**

En particulier  $\gamma^{\mu^2} = g^{\mu\mu} \mathbf{1}$ . Enfin  $\gamma^{0\alpha} = \gamma^0$  et  $\gamma^{i\alpha} = -\gamma^i$

$$\gamma^{0+} = \gamma^0 \quad , \quad \gamma^{i+} = -\gamma^i \quad \text{et} \quad [\gamma^\mu, \gamma^\nu] = 2g^{\mu\nu} \mathbf{1} \quad (21.16)$$

sont indépendantes de la représentation unitaire choisie.

On peut démontrer (cf. Greiner p. 104, extrait de Good. 1955) que toutes les représentations de l'algèbre de Dirac  $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \mathbf{1}$  (où les  $\gamma^\mu$  sont des  $4 \times 4$  matrices) qui satisfont  $\gamma^{0\alpha} = \gamma^0$  et  $\gamma^{i\alpha} = -\gamma^i$  sont unitairement équivalentes.

Autres propriétés :

$$\gamma^{\mu+} = \gamma^0 \gamma^\mu \gamma^+ = \gamma_\mu. \quad (21.17)$$

Cette propriété est également invariante par changement de représentation unitaire :

$$\begin{aligned} \gamma^{\mu+} &= (U \gamma^\mu U^+)^+ = U \gamma^{\mu+} U = U \gamma^0 \gamma^\mu \gamma^0 U^+ = \underbrace{U \gamma^0 U^+}_{\gamma^0} U \gamma^\mu U^+ \underbrace{U \gamma^0 U^+}_{\gamma^0} \\ &= \gamma^0 \gamma^{\mu+} \gamma^0 \quad \text{c.q.f.d.} \end{aligned}$$

On introduit également la matrice  $\gamma^s$  :

$$\gamma^s = \gamma_s = i \gamma^0 \gamma^1 \gamma^2 \gamma^3 = \frac{i}{4!} \varepsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \quad (21.18)$$

(par permutation cyclique sur 4 indices  $\rightarrow$  signe  $-$  (car  $\gamma^{3^2} = -1$ )

$$\begin{aligned} \gamma_s^2 &= -\gamma^0 \gamma^1 \gamma^2 \gamma^3 \gamma^0 \gamma^1 \gamma^2 \gamma^3 = \gamma^0 \gamma^1 \gamma^2 \gamma^3 \gamma^3 \gamma^0 \gamma^1 \gamma^2 \\ &= -\gamma^0 \gamma^1 \gamma^2 \gamma^0 \gamma^1 \gamma^2 \\ &= -\gamma^0 \gamma^1 \gamma^2 \gamma^2 \gamma^0 \gamma^1 \\ &= +\gamma^0 \gamma^1 \gamma^0 \gamma^1 = -\gamma^0 \gamma^1 \gamma^1 \gamma^0 = +\gamma^{0^2} = 1 \\ &\quad \uparrow \gamma^{2^2} = -1 \end{aligned}$$

Donc  $\|\gamma_s^2 = \mathbf{1}$ . Enfin  $\|\{\gamma_s, \gamma^\mu\} = 0$ .

En effet,  $\gamma_s = i \varepsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma$

$\uparrow$  sans sommation sur les indices

$$\begin{aligned} \gamma_s \gamma^{\mu'} &= i \varepsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma^{\mu'} = -i \varepsilon_{\mu\nu\rho\sigma} \gamma^{\mu'} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \\ &= -\gamma^{\mu'} \gamma_s \end{aligned}$$

car  $\mu'$  est égal à l'un des 4 indices  $\mu, \nu, \rho$  ou  $\sigma$

D'où la propriété :  $\gamma_s \gamma^{\mu'} + \gamma^{\mu'} \gamma_s = 0$

On définit également  $\| \sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu]$

expressions dans la représentation spinorielle (appelée aussi chirale) :

$$\begin{aligned} \gamma_s &= i\gamma^0\gamma^1\gamma^2\gamma^3 = i\gamma^0\gamma^1 \begin{pmatrix} -\sigma_2\sigma_3 & 0 \\ 0 & -\sigma_2\sigma_3 \end{pmatrix} \\ &= i \begin{pmatrix} \sigma_1 & 0 \\ 0 & -\sigma_1 \end{pmatrix} \begin{pmatrix} -i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & -\sigma_1^2 \end{pmatrix} = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} = \gamma_s \end{aligned}$$

$$\left\| \begin{aligned} \sigma^{0i} &= \frac{i}{2} [\gamma^0, \gamma^i] = \frac{i}{2} \begin{pmatrix} 2\sigma_i & 0 \\ 0 & -2\sigma_i \end{pmatrix} = i \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \\ \sigma^{ij} &= \frac{i}{2} \begin{pmatrix} -2i\varepsilon_{ijk}\sigma_k & 0 \\ 0 & -2i\varepsilon_{ijk}\sigma_k \end{pmatrix} = \varepsilon_{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \end{aligned} \right. \quad (21.19)$$

car  $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$

$\gamma_\mu\gamma^\nu + \gamma^\nu\gamma_\mu = 2g_\mu^\nu\mathbf{1}$  donc  $\gamma_\mu\gamma^\mu = g_\mu^\mu\mathbf{1} = 4\mathbf{1}$   $\| \gamma_\mu\gamma^\mu = 4\mathbf{1}$

$$\left\| \begin{aligned} Tr \gamma^\mu &= 0 \\ Tr \mathbf{1} &= 4 \\ Tr \gamma_s &= 0 \end{aligned} \right.$$

Enfin, en posant  $\not{a} = a_\mu\gamma^\mu$ ,

$$\not{a}^2 = a_\mu\gamma^\mu a_\nu\gamma^\nu = \frac{1}{2}a_\mu a_\nu (\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) = a_\mu a_\nu g^{\mu\nu}\mathbf{1} = a^2\mathbf{1}$$

$$\| \not{a}^2 = a^2\mathbf{1}$$

$$\begin{aligned} \not{a}\not{b} + \not{b}\not{a} &= a_\mu\gamma^\mu b_\nu\gamma^\nu + b_\mu\gamma^\mu a_\nu\gamma^\nu = a_\mu b_\nu (\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) \\ &= 2a_\mu b_\nu g^{\mu\nu}\mathbf{1} = 2a \cdot b\mathbf{1} \end{aligned}$$

$$\| \not{a}\not{b} + \not{b}\not{a} = 2ab\mathbf{1}$$

$$\star \| Trace \not{a}\not{b} = \frac{1}{2}Trace (\not{a}\not{b} + \not{b}\not{a}) = \frac{1}{2}Trace 2a\mathbf{1} = 4ab$$

$\star \|$  La trace d'un nombre impair de matrices  $\gamma$  est nulle :

$$\begin{aligned} Trace \phi_1 \cdots \phi_n &= Trace \phi_1 \cdots \phi_n \gamma_s^2 = Trace \gamma_s \phi_1 \cdots \phi_n \gamma_s \\ &= (-1)^n Trace \phi_1 \cdots \phi_n \underbrace{\gamma_s^2}_{1} \end{aligned}$$

donc pour  $n$  impair,  $Trace \phi_1 \cdots \phi_n = 0$

$$\star \text{ Cas pair : } \left\| \begin{aligned} Trace \phi_1 \cdots \phi_n &= a_1 a_2 Trace \phi_3 \cdots \phi_n \\ &\quad - a_1 a_3 Trace \phi_2 \phi_4 \cdots \phi_n \\ &\quad + \cdots + a_1 a_n Trace \phi_2 \cdots \phi_{n-1} \end{aligned} \right.$$

En effet

$$\begin{aligned}
\text{Trace } \phi_1 \phi_2 \cdots \phi_n &= -\text{Trace } \phi_2 \phi_1 \cdots \phi_n + 2g_{\mu\nu} a_1^\mu a_2^\nu \text{Trace } \phi_3 \cdots \phi_n \\
&= -\text{Trace } \phi_2 \phi_3 \phi_1 \cdots \phi_n - 2g_{\mu\nu} a_1^\mu a_3^\nu \text{Trace } \phi_2 \phi_4 \cdots \phi_n \\
&\quad + 2g_{\mu\nu} a_1^\mu a_2^\nu \text{Trace } \phi_3 \cdots \phi_n \\
&= \cdots = 2a_1 a_2 \text{Trace } \phi_3 \cdots \phi_1 - 2a_1 a_3 \text{Trace } \phi_2 \phi_4 \cdots \phi_n \\
&\quad + \cdots + 2a_1 a_n \text{Trace } \phi_2 \cdots \phi_{n-1} - \underbrace{\text{Trace } \phi_2 \phi_3 \cdots \phi_n \phi_1}_{= \text{Trace } \phi_1 \phi_2 \cdots \phi_n} \\
&= \text{Trace } \phi_1 \phi_2 \cdots \phi_n
\end{aligned}$$

d'où le résultat

★ en particulier  $\|\text{Trace } \gamma_s \gamma^\mu = 0$  (nombre impair de matrices  $\gamma$ )

$\text{Trace } \gamma^\mu \gamma^\nu \gamma^s = i \text{Trace } \gamma^\mu \gamma^\nu \gamma^0 \gamma^1 \gamma^2 \gamma^3$  par permutation circulaire, ceci se réécrit au signe près  $i \varepsilon_{\mu\nu\rho\sigma} \underbrace{\text{Tr } \gamma^\rho \gamma^\sigma}_{4g^{\rho\sigma} \rightarrow 0}$  (car  $\mu, \nu = 0, 1, 2$  ou  $3$ )

donc  $\|\text{Trace } \gamma^\mu \gamma^\nu \gamma^s = 0$

$$\begin{aligned}
★ \text{Tr } (\gamma^s \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma) &= \frac{i}{4!} \text{Tr } \varepsilon_{\mu'\nu'\rho'\sigma'} \gamma^{\mu'} \gamma^{\nu'} \gamma^{\rho'} \gamma^{\sigma'} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \\
&= i \varepsilon_{\mu\nu\rho\sigma} \text{Tr } \gamma^\sigma \gamma^\rho \gamma^\nu \underbrace{\gamma^\mu \gamma^\mu}_{g^{\mu\mu} \mathbf{1}} \gamma^\nu \gamma^\rho \gamma^\sigma \text{ (sans sommation)} \\
&= -i \varepsilon_{\mu\nu\rho\sigma} \text{Tr } \mathbf{1} \\
&= -4i \varepsilon_{\mu\nu\rho\sigma} = 4i \varepsilon^{\mu\nu\rho\sigma}
\end{aligned}$$

$$\|\text{Tr } \gamma^s \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma = 4i \varepsilon^{\mu\nu\rho\sigma}$$

$$★ \|\text{Tr } \sigma_{\mu\nu} = \text{Tr } \frac{i}{2} [\gamma^\mu, \gamma^\nu] = 0$$

## 21.3 Propriétés de l'équation de Dirac

### 21.3.1 Equation conjuguée

Effectuons une conjugaison complexe sur l'équation de Dirac

$$P_\mu = i\partial_\mu \quad \text{donc} \quad p_\mu^* = -p_\mu$$

$$(\not{p} - m)\psi = 0 \quad \text{i.e.} \quad (p_0\gamma^0 + p_i\gamma^i - m)\psi = 0$$

$$\text{donc } (p_0^*\gamma^{0*} + p_i^*\gamma^{i*} - m)\psi^* = 0$$

$$\text{or } \left| \begin{array}{l} \gamma^{0+} = \gamma^0 \\ \gamma^{i+} = -\gamma^i \end{array} \right. \quad \text{donc } \left| \begin{array}{l} \gamma^{0*} = {}^t\gamma^0 \\ \gamma^{i*} = -{}^t\gamma^i \end{array} \right.$$

$$\text{d'où } ({}^t\gamma^0 p_0^\mu - {}^t\gamma^i p_i^* - m)\psi^* = 0$$

$$\text{ou encore } {}^t\psi^* (-\gamma^0 \overleftarrow{p}_0 + \gamma^i \overleftarrow{p}_i - m) = 0$$

$$\text{les opérateurs différentiels agissant à gauche : } \overleftarrow{p}_\mu = -i\partial_\mu \quad p_\mu = \vec{p}_\mu = i\partial_\mu$$

c'est contraire à l'utilisation habituelle d'un opérateur dans son action duale mais c'est la convention usuelle!).



En multipliant à droite par  $\gamma_0$ , on a donc

$$\psi^+ \gamma^0 \left( -\gamma^0 \overleftarrow{p}_0 + \underbrace{\gamma^0 \gamma^i \gamma^0}_{-\gamma^i} \overleftarrow{p}_i - im \right) = 0$$

$$-\overline{\psi}(\overleftarrow{p} + m) = 0$$

$$\text{soit } \boxed{\overline{\psi}(\overleftarrow{p} + m) = 0} \quad \Leftrightarrow \quad \boxed{\overline{\psi}(\overrightarrow{p} + m) = 0}$$

avec  $\overline{\psi} = \psi^+ \gamma^0$ , spineur conjugué de Dirac

### 21.3.2 Le courant

$$\begin{aligned} (\not{p} - m)\psi = 0 & \quad \text{donc} & \quad \overline{\psi}(\not{p} - m)\psi = 0 \\ \overline{\psi}(\overleftarrow{p} + m) = 0 & & \quad \overline{\psi}(\overleftarrow{p} + m)\psi = 0 \end{aligned}$$

En ajoutant les deux équations on a donc

$$\begin{aligned} \overline{\psi}(\not{p}\psi) + (\overline{\psi}(\overleftarrow{p}))\psi &= 0 \\ &= \overline{\psi}\gamma^\mu(p_\mu\psi) + (p_\mu\overline{\psi})\gamma^\mu\psi \end{aligned}$$

donc  $p_\mu(\overline{\psi}\gamma^\mu\psi) = 0$  avec  $p_\mu = i\partial_\mu$   
ce qui prouve que le courant  $\|j^\mu = \overline{\psi}\gamma^\mu\psi$  est conservé.

densité de charge :  $j^0 = \overline{\psi}\gamma^0\psi = \psi^+\psi \geq 0$

On peut donc l'interpréter comme une densité de probabilité.

$$\vec{j} = \psi^+ \vec{\alpha} \psi \quad \text{avec} \quad \vec{\alpha} = \gamma^0 \vec{\gamma}$$

$$\frac{\partial p}{\partial \gamma} + \text{dim } \vec{j} = 0$$

### 21.3.3 L'Hamiltonien

L'équation de Dirac étant linéaire dans la dérivation par rapport au temps, elle peut être mise sous forme hamiltonienne :

$$(\not{p} - m) = 0 \quad \text{s'écrit} \quad \left( ik\gamma_0 \frac{\partial}{\partial t} + ikc\vec{\gamma} \cdot \vec{\nabla} - mc^2 \right) \psi = 0$$

par multiplication à gauche par  $\gamma_0$  on obtient

$$\| ik \frac{\partial}{\partial t} \psi = H \psi$$

avec

$$\begin{aligned} H &= c\vec{\alpha} \cdot (-ik\vec{\nabla} + \beta mc) \\ &= c\vec{\alpha} \cdot \vec{p} + \beta mc^2 \end{aligned}$$

$$\alpha^i = \gamma^0 \gamma^i = \begin{pmatrix} -\sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \quad \beta = \gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (21.20)$$

### 21.3.4 La représentation standard (ou représentation de Dirac)

La représentation de la fonction d'onde en tant que bispineur  $(\xi, \eta)$  est la représentation spinorielle, qui est la plus naturelle du point de vue mathématique. Cependant, pour discuter la limite non relativiste de l'équation de Dirac, il est plus simple d'utiliser la représentation standard. Historiquement c'est la représentation qui a d'abord été considérée par Dirac.

On l'obtient à partir de la représentation spinorielle par la transformation suivante :

$$\begin{pmatrix} \phi \\ \chi \end{pmatrix} = U \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (21.21)$$

avec

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (21.22)$$

$$\begin{cases} \phi = \frac{1}{\sqrt{2}}(\xi + \eta) \\ \chi = \frac{1}{\sqrt{2}}(\xi - \eta) \end{cases} \quad (21.23)$$

$$\gamma_{standard} = U \gamma_{spinorielle} U^+ \quad (21.24)$$

$$\gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \quad \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (21.25)$$

Les matrices intervenant dans la forme hamiltonienne sont

$$\alpha^i = \gamma^0 \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad \beta = \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (21.26)$$

**Parité :**

$$\begin{cases} \phi \rightarrow i\phi \\ \chi \rightarrow -i\chi \end{cases} \quad (21.27)$$

$$\gamma^s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (21.28)$$

L'équation de Dirac s'écrit donc maintenant :

$$\begin{cases} p_0 \phi - \vec{p} \cdot \vec{\sigma} \chi = m \phi \\ p_0 \chi - \vec{p} \cdot \vec{\sigma} \phi = -m \chi \end{cases} \quad (21.29)$$

ou encore :

$$\begin{cases} i \frac{\partial}{\partial t} \phi = -i \vec{\sigma} \cdot \frac{\partial}{\partial \vec{x}} \chi + m \phi \\ i \frac{\partial}{\partial t} \chi = -i \vec{\sigma} \cdot \frac{\partial}{\partial \vec{x}} \phi - m \chi \end{cases} \quad (21.30)$$

L'énergie est positive, donc dans le référentiel au repos,  $\chi = 0$ . C'est bien entendu en accord avec la remarque p. D-3 : au repos  $\xi = \eta$  donc  $\chi = 0$ . La description de la particule se fait donc par un spineur tridimensionnel  $\phi$ . Dans l'approximation non relativiste, la norme de  $k$  reste petite devant celle de  $\phi$  (voir plus loin). On appelle pour cette raison  $\chi$  les petites composantes du spineur quadridimensionnel, et  $\phi$  les grandes composantes.

### 21.3.5 Le couplage avec le champ électromagnétique

Il suffit de remplacer  $\partial_\mu$  par la dérivée covariante  $D_\mu = \partial_\mu + ieA_\mu$  soit encore  $p_\mu \rightarrow n_\mu = p_\mu - eA_\mu$   $e$  : charge de la particule ( $e = -|e|$  pour l'électron) forme hamiltonienne en multipliant l'équation de Dirac à gauche par  $\gamma_0$  :

$$[i\hbar\gamma_0\pi_0\psi - c\vec{\gamma} \cdot \vec{n} - mc^2]\psi = 0 \quad \text{avec} \quad \vec{n} = \vec{p} - e\vec{A}$$

soit

$$i\hbar\frac{\partial}{\partial t}\psi = (c\vec{\alpha} \cdot \vec{\pi} + \beta mc^2)\psi + ceA_0\psi \quad A_0 = \frac{V}{c} \quad (21.31)$$

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\psi &= (c\vec{\alpha} \cdot \vec{p} + \beta mc^2)\psi - ce(\vec{\alpha} \cdot \vec{A} - A_0)\psi \\ &= (H_0 + H_{int})\psi \end{aligned}$$

*Remarque* : nous avons effectué ici la substitution minimale  $p^\mu \rightarrow p^\mu - eA^\mu$ . On peut imaginer des couplages non minimaux (voir TD).

### 21.3.6 L'invariance relativiste

Il s'agit de vérifier que pour une transformation du groupe de Poincaré complet :  $x' = \Lambda x + a$ , l'équation de Dirac est invariante de forme. Autrement dit, si  $\psi(x)$  est solution de l'équation  $(i\gamma^\mu\partial_\mu - m)\psi(x) = 0$ , alors  $\psi'(x') = S(\Lambda)\psi(x)$  est solution de l'équation

$$(i\gamma^\mu\partial'_\mu - m)\psi'(x') = 0$$

**Inversion d'espace :**

$$I_s : x = (x^0, \vec{x}) = \bar{x} = (x^0, -\vec{x})$$

$\psi(x) \rightarrow i\gamma^0\psi(\bar{x})$  d'après la forme matricielle obtenue pour la parité (voir p. SL-21) :

$$\begin{pmatrix} \xi^1 \\ \xi^2 \\ \eta_1 \\ \eta_2 \end{pmatrix}_s = i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \eta_1 \\ \eta_2 \end{pmatrix} \quad (21.32)$$

donc  $\psi_s(\bar{x}) = i\gamma^0\psi(x)$ .

Montrons que

$$\begin{aligned} &\left(i\gamma^0\frac{\partial}{\partial x^0} + i\gamma^k\frac{\partial}{\partial x^k} - m\right)\psi_s(\bar{x}) = 0 \\ \Leftrightarrow &\left(i\gamma^0\frac{\partial}{\partial x^0} - i\gamma^k\frac{\partial}{\partial x^k} - m\right)i\gamma_0\psi(x) = 0 \\ \Leftrightarrow &i\gamma^0\left(i\gamma^0\frac{\partial}{\partial x^0} + i\gamma^k\frac{\partial}{\partial x^k} - m\right)\psi(x) = 0 \end{aligned}$$

qui est l'équation de Dirac (car  $\gamma^0\gamma^k\gamma^0 = -\gamma^k$ ) **c.q.f.d.**

Bien entendu cette invariance ne doit pas nous surprendre puisque l'équation de Dirac a été construite de manière explicitement covariante.

l'invariance par le sous-groupe des translations est immédiate

sous-groupe de Lorentz restreint :

Le champ  $\psi$  (bispineur) se transforme sous  $D^{[\frac{1}{2},0]} \oplus D^{[0,\frac{1}{2}]}$ , i.e. sous une transformation de Lorentz

$$\begin{aligned}\Lambda(A) : x &\rightarrow x' = \Lambda(x) , \\ \psi(x) &\rightarrow \psi'(x') = S(\Lambda(A))\psi(x)\end{aligned}$$

$$\begin{aligned}\xi^\alpha &\rightarrow A\xi^\alpha \\ \eta_{\dot{\alpha}} &\rightarrow (A^+)^{-1}\eta_{\dot{\alpha}}\end{aligned}$$

D'où

$$S(\Lambda(A)) = \begin{pmatrix} 1 & 0 \\ 0 & (A^+)^{-1} \end{pmatrix} \quad (21.33)$$

L'invariance de forme impose donc :

$$(i\gamma^\mu \partial'_\mu - m) \psi'(x') = 0 \quad \Leftrightarrow (i\gamma^\mu \partial_\mu - m) \psi(x) = 0$$

or

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial x'^\nu} \cdot \frac{dx'^\nu}{dx^\mu} = \Lambda^\nu{}_\mu \frac{\partial}{\partial x'^\nu}$$

donc  $(i\gamma^\mu \Lambda^\nu{}_\mu \partial'_\nu - m) S^{-1}(\Lambda) \psi'(x') = 0$

$$(iS(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu \partial'_\nu - m) \psi'(x') = 0$$

l'invariance de forme conduit donc à  $S(\Lambda)\gamma^\mu S^{-1}(\Lambda)\Lambda^\nu{}_\mu = \gamma^\nu$

$$\text{soit } \boxed{S^{-1}(\Lambda)\gamma^\nu S(\Lambda) = \Lambda^\nu{}_\mu \gamma^\mu}$$

Ainsi  $\gamma^\nu$  forme un opérateur vectoriel sous le groupe de Lorentz.

**Propriété :**

$$S^+ = \begin{pmatrix} A^+ & 0 \\ 0 & A^{-1} \end{pmatrix} \quad (21.34)$$

donc

$$S\gamma^0 S^+ = \begin{pmatrix} A & 0 \\ 0 & A^{+^{-1}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} A^+ & 0 \\ 0 & A^{-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \gamma^0 \quad (21.35)$$

de même pour  $S^+\gamma^0 S$

donc  $\boxed{S\gamma^0 S^+ = \gamma^0}$  et  $\boxed{S^+\gamma^0 S = \gamma^0}$  ( $\Lambda \in \mathcal{L}_+^\uparrow$ );  $S\gamma^0 S^+ = -\gamma^0$  et  $S^+\gamma^0 S = -\gamma^0$  ( $\Lambda \in \mathcal{L}_-^\uparrow$ )  
 $\uparrow i^2 = -1$

En utilisant les résultats obtenues dans le cours sur les représentations du groupe de Lorentz, il est aisé d'écrire les générateurs  $M_{\mu\nu}$  :

$$S(\Lambda)\psi(\Lambda^{-1}x) = \left(1 - \frac{i}{2}\omega^{\mu\nu}M_{\mu\nu}\right)$$

qui correspondent au moment angulaire total :

$$M_{\mu\nu} = \underbrace{i(x_\mu \partial_\nu - x_\nu \partial_\mu)} + J_{\mu\nu}$$

voir p. L-26

$J_{\mu\nu}$  correspond au moment angulaire intrinsèque : ce sont les générateurs infinitésimaux de la transformation  $S(\Lambda)$

$$S(\Lambda) = \left( 1 - \frac{i}{2} \omega^{\mu\nu} J_{\mu\nu} \right)$$

à la représentation  $D^{[\frac{1}{2}, 0]}$  correspond  $J_{0i} = -i\frac{\sigma_i}{2}$      $J_{ij} = \varepsilon_{ijk} \frac{\sigma_k}{2}$

à la représentation  $D^{[0, \frac{1}{2}]}$  correspond  $J_{0i} = i\frac{\sigma_i}{2}$      $J_{ij} = \varepsilon_{ijk} \frac{\sigma_k}{2}$

donc sur les bispineurs,

$$J_{0i} = \begin{pmatrix} -i\frac{\sigma_i}{2} & 0 \\ 0 & i\frac{\sigma_i}{2} \end{pmatrix} = \frac{1}{2}\sigma_{0i} \quad J_{ij} = \frac{\varepsilon_{ijk}}{2} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} = \frac{1}{2}\sigma_{ij} \quad (21.36)$$

Ainsi

$$J_{\mu\nu} = \frac{1}{2}\sigma_{\mu\nu} = \frac{i}{4}[\gamma_\mu, \gamma_\nu] \quad (21.37)$$

**Ainsi :**

1) à la rotation  $R$  **active** d'angle  $\theta$  autour de l'axe  $\vec{n}$

$$\begin{cases} x'_0 = x_0 \\ \vec{x}' = \vec{x} \cos \theta + \vec{n} \wedge \vec{x} \sin \theta + (1 - \cos \theta) \vec{n} (\vec{n} \cdot \vec{x}) \end{cases}$$

correspond la transformation

$$\begin{aligned} S(R) &= \exp -\frac{i}{2} \omega^{ij} J_{ij} \\ &= \exp -\frac{i}{2} \omega^{ij} \frac{1}{2} \varepsilon_{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix} \end{aligned}$$

or  $\omega^{ij} = \varepsilon_{ijk} n^k \theta$  donc

$$\begin{aligned} S(R) &= \begin{pmatrix} \exp -\frac{i}{2} \theta \vec{n} \cdot \vec{\sigma} & 0 \\ 0 & \exp -\frac{i}{2} \theta \vec{n} \cdot \vec{\sigma} \end{pmatrix} \\ &= \begin{pmatrix} \cos -\frac{\theta}{2} - i \vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} & 0 \\ 0 & \cos -\frac{\theta}{2} - i \vec{\sigma} \cdot \vec{n} \sin \frac{\theta}{2} \end{pmatrix} \end{aligned}$$

$$S(R) = \cos \frac{\theta}{2} \mathbb{1}_{4 \times 4} - i \sin \frac{\theta}{2} \vec{n} \cdot \vec{\Sigma} \quad (21.38)$$

avec

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} = \gamma^s \gamma^0 \vec{\gamma} \quad (21.39)$$

↑ dans les représentations spinorielle et standard

2) au boost  $\Lambda$  **actif** de rapidité  $\phi$  selon l'axe  $\vec{n}$

$$\begin{cases} x'_0 = x_0 \operatorname{ch} \phi + \vec{n} \cdot \vec{x} \operatorname{sh} \phi \\ \vec{x}' = \vec{x} + x_0 \vec{n} \operatorname{sh} \phi + \vec{n}(\vec{n} \cdot \vec{x})(\operatorname{ch} \phi - 1) \end{cases} \quad (21.40)$$

correspond la transformation

$$\begin{aligned} S(\Lambda) &= \exp -i\omega^{0i} J_{0i} \\ &= \exp i\phi n^i \left( -\frac{i}{2} \right) \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \\ &= \begin{pmatrix} \exp \frac{\phi}{2} \vec{n} \cdot \vec{\sigma} & 0 \\ 0 & \exp -\frac{\phi}{2} \vec{n} \cdot \vec{\sigma} \end{pmatrix} \\ S(\Lambda) &= \begin{pmatrix} \operatorname{ch} \frac{\phi}{2} + \operatorname{sh} \frac{\phi}{2} \vec{n} \cdot \vec{\sigma} & 0 \\ 0 & \operatorname{ch} \frac{\phi}{2} - \operatorname{sh} \frac{\phi}{2} \vec{n} \cdot \vec{\sigma} \end{pmatrix} \end{aligned} \quad (21.41)$$

soit

$$S(\Lambda) = \operatorname{ch} \frac{\phi}{2} \mathbf{1}_{4 \times 4} + \vec{\alpha} \cdot \vec{n} \operatorname{sh} \frac{\phi}{2} \quad (21.42)$$

avec

$$\vec{\alpha} = \gamma_0 \vec{\gamma} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -\vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix} \quad (21.43)$$

dans la représentation spinorielle

### 21.3.7 Les constantes du mouvement

Nous venons de voir qu'à une rotation active d'angle  $\theta$  autour de  $\vec{n}$  correspond la loi de transformation sur un bispineur

$$\begin{aligned} \psi(x) \rightarrow \psi'(x) &= S(R(\theta, \vec{n})) \psi(R^{-1}x) \\ &= \left( 1 - \frac{i}{2} \omega^{\mu\nu} M_{\mu\nu} \right) \psi(x) \quad \text{pour une transformation infinitésimale} \end{aligned}$$

$$\text{avec } \begin{cases} \omega_{0i} = 0 \\ \omega^{ij} = \theta \varepsilon_{ijk} n^k \end{cases} \quad (\text{cf. L-25})$$

$$\text{et } M_{ij} = \underbrace{i(x_i \partial_j - x_j \partial_i)}_{x_i P_j - x_j P_i} + J_{jk}$$

$$\begin{aligned} \omega^{ij} (x_i P_j - x_j P_i) &= \frac{1}{2} \varepsilon_{ijk} (x_i P_j - x_j P_i) n^k \theta = \varepsilon_{ijk} x_i P_j n^k \theta \\ &= \underbrace{\varepsilon_{ijk} x^i P^j}_{L^k} n^k \theta = \vec{L} \cdot \vec{n} \theta \end{aligned}$$

Donc sur un bispineur,

$$\|\psi'(x) = \left[ 1 - i\theta\vec{n} \cdot \left( \vec{L} + \frac{\vec{\Sigma}}{2} \right) \right] \psi(x)$$

Le moment cinétique total est donc  $\vec{J} = \vec{L} + \frac{\hbar}{2}\vec{\Sigma}$ , ce qui suggère d'interpréter  $\vec{L}$  comme un moment angulaire orbital et  $\vec{S} = \frac{\hbar}{2}\vec{\Sigma}$  comme un terme de spin.

Les relations de commutation  $[S^i, S^j] = i\hbar\varepsilon^{ijk}S^k$  se vérifient par calcul direct, puisque

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \quad (21.44)$$

qui traduit le fait que les bispineurs forment une représentation  $D^{[\frac{1}{2},0]} + D^{[0,\frac{1}{2}]}$  de  $SL(2, \mathbb{C})$ , où  $D^{[\frac{1}{2},0]} \simeq D^{[0,\frac{1}{2}]}$  dans le cas des rotations.

$\vec{S}^2 = \frac{3\hbar^2}{4}\mathbb{1} = \hbar s(s+1)\mathbb{1}$  avec  $s = \frac{1}{2}$  : le spineur de Dirac décrit bien une particule de spin  $\frac{1}{2}$  !

$\vec{J}$  est une constante de mouvement pour l'hamiltonien  $H_0 = \vec{\alpha} \cdot \vec{p} + \beta m$

$$\left[ H_0, \vec{L} \right]_k = \left[ \vec{\alpha} \cdot \vec{p}, \vec{L} \right]_k = [\alpha_i p_i, \varepsilon_{i'j'k} x_{i'} p_{j'}] = -i\alpha_i \varepsilon_{ijk} p_j = -i(\vec{\alpha} \wedge \vec{p})_k$$

$$\left[ H_0, \vec{\Sigma} \right] = \left[ \vec{\alpha} \cdot \vec{p}, \vec{\Sigma} \right] + m \left[ \beta, \vec{\Sigma} \right]$$

En remarquant que dans la représentation standard,

$$\vec{\alpha} = \sigma^1 \otimes \vec{\sigma} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad (21.45)$$

$$\beta = \gamma_0 = \sigma^3 \otimes I = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad (21.46)$$

$$\vec{\Sigma} = I \otimes \vec{\sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}, \quad (21.47)$$

on obtient

$$[\beta, \vec{\Sigma}] = \sigma^3 \otimes \vec{\sigma} - \sigma^2 \otimes \vec{\sigma} = 0 \quad (21.48)$$

$$\begin{aligned} \left[ \vec{\alpha} \cdot \vec{p}, \vec{\Sigma} \right] &= [\sigma_1 \otimes \vec{\sigma} \cdot \vec{p}, 1 \otimes \vec{\sigma}] \\ &= \sigma_1 \otimes (\vec{\sigma} \cdot \vec{p}) \vec{\sigma} - \sigma_1 \otimes \vec{\sigma} (\vec{\sigma} \cdot \vec{p}) = \sigma_1 \otimes [\sigma_i p_i, \sigma_j] \\ &= \sigma_1 \otimes 2i\varepsilon_{ijk} \sigma_k p_i \\ &= 2i\sigma_1 \otimes \sigma_k \varepsilon_{ijk} p_i \\ &= 2i\varepsilon_{ijk} p_i \alpha_k = 2i\vec{\alpha} \wedge \vec{p} \end{aligned}$$

D'où

$$[H_0, \vec{J}] = 0. \quad (21.49)$$

Ainsi  $\vec{J}$  est une constante du mouvement, mais  $\vec{L}$  et  $\vec{S}$  ne le sont pas séparément.

De plus pour pouvoir identifier  $\vec{L} = \vec{r} \wedge \vec{p}$  à un opérateur de moment cinétique orbital il faudrait que  $\vec{r}$  puisse être identifié à un opérateur de position. Or  $\frac{d\vec{r}}{dt} = i[H, \vec{r}]$  donne  $\frac{d\vec{r}}{dt} = c\vec{\alpha}$  dont les valeurs propres sont  $\pm c$ !

De plus  $\left[\frac{dx_i}{dt}, \frac{dx_j}{dt}\right] \neq 0$ . Donc  $\frac{dx_i}{dt}$  ne peut pas être identifié à l'opérateur de vitesse  $\frac{\vec{p}}{m}$ .

Pour définir un bon opérateur de position, il faut passer dans la représentation de Foldy-Wouthuysen (voir TD).

**Conclusion :** un spineur de Dirac se transforme sous une rotation comme une particule de spin  $\frac{1}{2}$  mais il n'est pas légitime d'identifier  $\frac{\hbar}{2}\vec{\Sigma}$  à un opérateur de spin.

*Rem. :* comme  $[H_0, \vec{\Sigma}] = 2i\vec{\alpha} \wedge \vec{p}$ , il est immédiat que  $\frac{\hbar}{2} \frac{\vec{\Sigma} \cdot \vec{p}}{p}$  commute avec  $H$ . C'est l'opérateur d'hélicité (voir plus loin).

### Rotation des fermions sous $\theta = 2\pi$

Pour une rotation d'angle  $2n$  autour de l'axe  $Oz$  on obtient  $\exp -i\theta \frac{\vec{\sigma} \cdot \vec{n}}{2} = \exp -in\sigma_z = -1$  : la fonction d'onde change de signe sous une telle rotation. Pour un électron isolé un tel changement de signe ne peut pas être mis en évidence.

On peut en revanche mettre en évidence cet effet dans une expérience d'interférences.

Figure

Le principe de l'expérience consiste à séparer un faisceau de neutrons en deux parties. Le faisceau 2 traverse un champ magnétique  $\vec{B}$ . Au niveau classique et au niveau quantique, il y a précession du spin dans le champ. On peut ajuster le temps de passage pour que  $\theta = 2\pi$ , puisque  $\theta = g \frac{eBt}{2m}$  est l'angle dont a tourné le spin dans le champ  $\vec{B}$ .

### 21.3.8 Lois de transformations de quantités bilinéaires sous $\mathbf{L}_+^\uparrow$

$$\psi(x) \rightarrow \psi'(x') = S(\Lambda)\psi(x)$$

or  $\bar{\psi} = \psi^+ \gamma^0$  donc  $\bar{\psi}(x) \rightarrow \bar{\psi}'(x') \rightarrow \psi^+(x) S^+(\Lambda) \gamma^0$   
comme  $S^* \gamma^0 S = \gamma^0$  (cf. p. D-15),

$$\bar{\psi}'(x') = \psi^+(x) \gamma^0 S^{-1} = \bar{\psi} S^{-1}$$

donc :

1)  $\bar{\psi}' \psi' = \bar{\psi} S^{-1} S \psi = \bar{\psi} \psi$  est un scalaire



Ceci n'est pas un résultat nouveau : nous avons vu lors de l'étude des représentations spinorielles du groupe de Lorentz que  $\xi^\alpha \xi'_\alpha + \eta_{\dot{\alpha}} \eta'^{\dot{\alpha}}$  est un scalaire sous le groupe de Lorentz (cf. SL-11).

Comme

$$\bar{\xi}^\alpha \sim \eta^{\dot{\alpha}} \quad (21.50)$$

et

$$\xi_\alpha \sim \bar{\eta}_{\dot{\alpha}}, \quad (21.51)$$

ce scalaire s'écrit encore  $\bar{\psi}\psi$  :

$$\psi = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \eta^{\dot{1}} \\ \eta^{\dot{2}} \end{pmatrix} \quad (21.52)$$

$$\begin{aligned} \bar{\psi}\psi = \psi^\dagger \gamma^0 \psi &= \begin{pmatrix} \bar{\xi}^1 \\ \bar{\xi}^2 \\ \bar{\eta}^{\dot{1}} \\ \bar{\eta}^{\dot{2}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \xi^1 \\ \xi^2 \\ \eta^{\dot{1}} \\ \eta^{\dot{2}} \end{pmatrix} \\ &= \bar{\xi}^1 \eta_{\dot{1}} + \bar{\xi}^2 \eta_{\dot{2}} + \bar{\eta}^{\dot{1}} \xi^1 + \bar{\eta}^{\dot{2}} \xi^2 \quad \mathbf{c.q.f.d.} \end{aligned}$$

$$2) \bar{\psi}' \sigma^{\mu\nu} \psi' = \bar{\psi} S^{-1} \sigma^{\mu\nu} S \psi = \Lambda^\mu_\rho \Lambda^\nu_\sigma \bar{\psi} \sigma^{\rho\sigma} \psi$$

puisque  $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$  est un opérateur tensoriel de rang 2 :

$$\begin{aligned} \bar{\psi} S^{-1} \sigma^{\mu\nu} S \psi &= \bar{\psi} S^{-1} \frac{i}{2} [\gamma^\mu, \gamma^\nu] S \psi = \frac{i}{2} [\bar{\psi} S^{-1} \gamma^\mu S S^{-1} \gamma^\nu S \psi - \bar{\psi} S^{-1} \gamma^\nu S S^{-1} \gamma^\mu S \psi] \\ &= \frac{i}{2} \Lambda^\mu_\rho \Lambda^\nu_\sigma (\bar{\psi} \gamma^\rho \gamma^\sigma \psi - \bar{\psi} \gamma^\sigma \gamma^\rho \psi) \\ &= \Lambda^\mu_\rho \Lambda^\nu_\sigma \bar{\psi} \sigma^{\rho\sigma} \psi \end{aligned}$$

3)  $\bar{\psi}' \gamma^\mu \psi'$  se transforme comme un **vecteur**

$$4) \bar{\psi}' \gamma^s \psi' = \bar{\psi} S^{-1} \gamma^s S \psi$$

$$\gamma^s = i\gamma^0 \gamma^1 \gamma^2 \gamma^3 = \frac{i}{4!} \varepsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma$$

$$\begin{aligned} S^{-1} \gamma^s S &= \frac{i}{4!} \Lambda^\mu_\alpha \Lambda^\nu_\beta \Lambda^\rho_\gamma \Lambda^\sigma_\delta \gamma^\alpha \gamma^\beta \gamma^\gamma \gamma^\delta \varepsilon_{\mu\nu\rho\sigma} \\ &= \frac{i}{4!} \det \Lambda \varepsilon_{\alpha\beta\gamma\delta} \gamma^\alpha \gamma^\beta \gamma^\gamma \gamma^\delta = \det \Lambda \gamma^s \end{aligned}$$

donc  $\bar{\psi}' \gamma^s \psi' = (\det \Lambda) \bar{\psi} \gamma^s \psi$  : pseudoscalaire

Ceci est à mettre en parallèle avec le résultat obtenu dans l'étude des représentations bispinorielles du groupe de Lorentz : on a vu que  $\xi^\alpha \xi'_\alpha - \eta_{\dot{\alpha}} \eta'^{\dot{\alpha}}$  est un pseudo scalaire, et en représentation spinorielle

$$\gamma^s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (21.53)$$

$$\begin{aligned} 5) \bar{\psi}' \gamma^\mu \gamma^s \psi' &= \bar{\psi} S^{-1} \gamma^\mu S S^{-1} \gamma^s S \psi \\ &= \Lambda^\mu_\rho \bar{\psi} \gamma^\rho \gamma^5 \det \Lambda \quad \text{pseudovecteur} \end{aligned}$$

De même 3) et 5) peuvent s'interpréter facilement dans le contexte des représentations du groupe de Lorentz.

### Limite non relativiste de l'équation de Dirac

L'énergie d'une particule relativiste de masse  $m$  comprend son énergie au repos  $mc^2$ . On l'élimine en posant

$$\psi = \bar{\Psi} e^{imc^2 t / \hbar}$$

qui vérifie l'équation :

$$\left( i\hbar \frac{\partial}{\partial t} + mc^2 \right) \psi = \left[ c\vec{\alpha} \cdot \left( -i\hbar \vec{\nabla} - e\vec{A} \right) + \beta mc^2 + eV \right] \psi$$

Dans la représentation standard

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}, \quad (21.54)$$

cette équation s'écrit encore :

$$\left( i\hbar \frac{\partial}{\partial t} - eV + mc^2 \right) \begin{pmatrix} \phi \\ \chi \end{pmatrix} = c(\vec{p} - e\vec{A}) \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} + mc^2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} \quad (21.55)$$

soit :

$$\begin{cases} (i\hbar \frac{\partial}{\partial t} - eV) \phi = c\vec{\sigma} \cdot (\vec{p} - e\vec{A}) \chi \\ (i\hbar \frac{\partial}{\partial t} - eV + 2mc^2) \chi = c\vec{\sigma} (\vec{p} - e\vec{A}) \phi \end{cases}$$

Dans l'approximation non relativiste, toutes les énergies sont petites devant l'énergie au repos relativiste.

Donc dans la deuxième équation,  $\chi \simeq \frac{1}{2mc} \vec{\sigma} \cdot (\vec{p} - e\vec{A}) \phi$  :

En insérant cette expression dans la première équation, on trouve donc :

$$\left( i\hbar \frac{\partial}{\partial t} - eV \right) \phi = \frac{1}{2m} \left( \vec{\sigma} \cdot (\vec{p} - e\vec{A}) \right)^2 \phi$$

or  $\vec{\sigma} \cdot \vec{a} \vec{\sigma} \cdot \vec{b} = \vec{a} \cdot \vec{b} + i\vec{\sigma} \cdot \vec{a} \wedge \vec{b}$

donc  $(\vec{\sigma} \cdot (\vec{p} - e\vec{A}))^2 = (\vec{p} - e\vec{A})^2 + i\vec{\sigma} (\vec{p} - e\vec{A}) \wedge (\vec{p} - e\vec{A})$ .

Comme  $\vec{p}$  et  $\vec{A}$  ne commutent pas, le dernier terme n'est pas nul. Il vaut

$$-e\hbar \vec{\sigma} (\vec{\nabla} \wedge \vec{A} + \vec{A} \wedge \vec{\nabla}) = -e\hbar \vec{\sigma} \cdot r\vec{t}\vec{A} = -e\hbar \vec{B} \cdot \vec{B}$$

D'où l'équation de Pauli

$$i\hbar \frac{\partial}{\partial t} \phi = \left[ \frac{1}{2m} (\vec{p} - e\vec{A})^2 + eV - \frac{e\hbar}{2m} \vec{\sigma} \cdot \vec{B} \right] \phi$$

La différence par rapport à l'équation de Schrödinger vient donc du dernier terme qui est un terme de couplage entre le champ magnétique et un moment magnétique  $\mu$ .

Considérons un champ magnétique dans la jauge  $A_0 = 0$ , avec  $\vec{A} \neq \vec{0}$ . Dans le cas particulier d'un champ magnétique constant,  $\vec{A} = \frac{1}{2}\vec{B} \wedge \vec{r}$  est solution

$$\frac{1}{2m}\vec{n}^2 = \frac{1}{2m}(\vec{p} - e\vec{A})^2 = \frac{1}{2m}\left(\vec{p}^2 - e(\vec{p}\vec{A} + \vec{A}\vec{p})\right) + O\left(e^2\vec{A}^2\right)$$

En négligeant les termes en  $(e\vec{A})^2$  (approximation de champ faible),

$$\begin{aligned}\vec{p} \cdot \vec{A} - \vec{A} \cdot \vec{p} &= [\vec{p}, \vec{A}] + 2\vec{A} \cdot \vec{p} = \frac{\hbar}{i}\partial^k A^k + (\vec{B} \wedge \vec{r}) \cdot \vec{p} \\ &= \frac{\hbar}{2i}\underbrace{\partial^k \varepsilon_{ijk} B_i r_j}_{\varepsilon_{ijk} B_i \delta_j^k} = (\vec{B}, \vec{r}, \vec{p}) = \vec{L} \cdot \vec{B} \\ \varepsilon_{ijk} B_i \delta_j^k &= 0\end{aligned}$$

$$\text{Donc } i\hbar\frac{\partial}{\partial t}\phi = \left[ \frac{\vec{p}^2}{2m} - \frac{e}{2m}\vec{L} \cdot \vec{B} - \frac{e\hbar}{2m}\vec{\sigma} \cdot \vec{B} \right] \phi$$

Interprétation du terme  $-\frac{e\vec{L} \cdot \vec{B}}{2m}$  :

Considérons le couplage d'une boucle de courant avec un champ magnétique :

$$i = \frac{|e|}{T} = \frac{|e|v}{2nr}$$

figure

Le moment magnétique vaut donc :  $\mu = iS = nr^2\frac{e|v|}{2nr} = \frac{|e|}{2m}mrv = \frac{|e|\hbar}{2m}L$ , et  $\vec{\mu} = \frac{e}{2m}\vec{L}$ .

L'hamiltonien d'interaction :

$$H = -\vec{\mu} \cdot \vec{B} = -\frac{e}{2m}\vec{L} \cdot \vec{B} = -\mu_B \vec{L} \cdot \vec{B} \quad (\text{avec } \mu_B = \frac{e\hbar}{2m} : \text{magnéton de Bohr})$$

décrit donc le couplage entre le champ magnétique extérieur et le moment magnétique de l'atome engendré par la rotation de l'électron sur son orbite.

Interprétation du terme  $-\frac{e\hbar}{2m}\vec{\sigma} \cdot \vec{B} = -\mu_B \hbar \vec{\sigma} \cdot \vec{B}$ .

En introduisant l'opérateur de spin  $\vec{S} = \hbar\frac{\vec{\sigma}}{2}$ , le terme de couplage s'écrit  $-\vec{\mu} \cdot \vec{B}$  avec  $\vec{\mu} = g\mu_B\vec{S}$ .  $g$  est le rapport gyromagnétique. On constate que la théorie de Dirac prédit  $g = 2$ .

Les corrections radiatives calculées dans le cadre de l'électrodynamique quantique (ainsi que la prise en compte de corrections faibles et hadroniques) sont connues avec une précision relative de  $10^{-7}$ , qui est également la précision expérimentale sur la mesure de  $g$ . Ces deux résultats sont en accord avec cette précision fantastique !

*Rem.* : dans le cas de particules composées de type méson ( $q\bar{q}$ ) ou baryon ( $qqq$ ) le rapport gyromagnétique n'est plus donné par la théorie de Dirac

$$g \text{ (neutron)} = -3,83$$

$$g \text{ (proton)} = 5,59$$

On peut traiter phénoménologiquement un tel couplage en introduisant un couplage non minimal au champ électromagnétique :

$$\left[ i\not{\partial} - e\not{A} - m + \frac{\mathcal{S}g}{L} \frac{e}{4m} \sigma^{\mu\nu} F_{\mu\nu} \right] \psi = 0$$

avec  $g = 2 + \mathcal{S}g$  (voir TD).

## 21.4 Les solutions ondes planes

Nous cherchons des ondes planes solution de l'équation de Dirac. Une solution générale peut alors s'écrire comme superposition d'ondes planes. On cherche donc  $\psi(x) = e^{-ipx}\psi(p)$  solution de

$$-i\gamma^\mu \partial_\mu \psi + mc\psi = 0, \quad \text{soit } (\not{p} - m)\psi(p) = 0 \quad (k=1)$$

Ce système d'équations linéaires homogènes possède une solution si et seulement si  $\det(\not{p} - mc) = (p^2 - m^2c^2)^2 = 0$ .

Donc  $p^2 = p_0^2 - \vec{p}^2 = m^2c^2$  d'où le spectre  $p_0 = \pm\sqrt{\vec{p}^2 + m^2c^2}$ .

Le spectre d'énergie

$$E = p_0c \in [-\infty, -mc^2[ \cup ]mc^2, \infty[ \quad (21.56)$$

contient donc à la fois des solutions d'énergie positive et négative.

Dans la représentation standard, on obtient donc :

$$\begin{pmatrix} p^0 - m & -\vec{\sigma} \cdot \vec{p} \\ \vec{\sigma} \cdot \vec{p} & -p^0 - m \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = 0 \quad (c=1) \quad (21.57)$$

soit

$$(1) \begin{cases} (p^0 - m)\phi = \vec{\sigma} \cdot \vec{p}\chi \\ \vec{\sigma} \cdot \vec{p}\phi = (p^0 + m)\chi \end{cases} \quad (21.58)$$

d'où

$$\chi = \frac{\vec{\sigma} \cdot \vec{p}}{p^0 + m} \phi \quad (21.59)$$

$$\psi(p) = \begin{pmatrix} \phi \\ \frac{\vec{\sigma} \cdot \vec{p}}{p^0 + m} \phi \end{pmatrix} \quad \phi \text{ étant un spineur quelconque} \quad (21.60)$$

Nous avons montré plus haut que  $\bar{\psi}(x)\psi(x)$  est un scalaire de Lorentz. Il va nous servir à normaliser les ondes planes.

$$\begin{aligned} \bar{\psi}(x)\psi(x) &= \bar{\psi}(p)\psi(p) = \phi^\dagger \phi \left[ 1 - \frac{\vec{p}^2}{(p^0 + m)^2} \right] \\ 1 - \frac{\vec{p}^2}{(p^0 + m)^2} &= \frac{(p^0 + m)^2 - (p^0^2 - m^2)}{(p^0 + m)^2} = \frac{p^0 + m - (p^0 - m)}{p^0 + m} = \frac{2m}{p^0 + m} \end{aligned}$$

★ Pour les solutions d'énergie positive  $p^0 = \sqrt{\vec{p}^2 + m^2}$  on peut donc choisir pour normalisation  $\bar{\psi}_+(p)\psi_+(p) = +1$ ,

↑ + signifie énergie positive soit

$$\psi_+(p) = \sqrt{\frac{p^0 + m}{2m}} \begin{pmatrix} \phi \\ \frac{\vec{\sigma} \cdot \vec{p}}{p^0 + m} \phi \end{pmatrix} \quad \text{avec } \phi^\dagger \phi = 1 \quad p^0 = +\sqrt{\vec{p}^2 + m^2} \quad (21.61)$$

★ Pour les solutions d'énergie négative, comme  $p_0 < -m$ , on choisit pour normalisation  $\bar{\psi}_-(p)\psi_-(p) = -1$ . En utilisant (1), on peut écrire la solution en terme du spineur  $\chi$  :

$$\phi = \frac{\vec{\sigma} \cdot \vec{p}}{p^0 - m} \chi = \frac{-\vec{\sigma} \cdot \vec{p}}{-p^0 + m} \chi$$

d'où, puisque  $\bar{\psi}(p)\psi(p) = \left( \frac{\vec{p}^2}{(p^0 - m)^2} - 1 \right) \chi^\dagger \chi = \underbrace{\frac{2m}{p^0 - m}}_{< 0 \text{ pour } p^0 < 0} \chi^\dagger \chi$

$$\text{soit } \|\psi_-(p) = \sqrt{\frac{-p^0 + m}{2m}} \begin{pmatrix} \frac{-\vec{\sigma} \cdot \vec{p}}{-p^0 + m} \chi \\ \chi \end{pmatrix} \quad \text{avec } \chi^\dagger \chi = 1 \quad p^0 = -\sqrt{\vec{p}^2 + m^2} \quad (21.62)$$

Par convention on définit les notations suivantes :

$$\left\{ \begin{array}{l} u(p) = \psi_+(p_0, \vec{p}) = \sqrt{\frac{p^0 + m}{2m}} \begin{pmatrix} \varphi \\ \frac{\vec{\sigma} \cdot \vec{p}}{p^0 + m} \varphi \end{pmatrix} \quad \psi_+(x) = e^{-ipx} u(p) \\ \hspace{15em} \text{solutions d'énergie positive} \\ v(p) = \psi_-(-p_0, -\vec{p}) = \sqrt{\frac{p^0 - m}{2m}} \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{p^0 + m} \chi \\ \chi \end{pmatrix} \quad \psi_-(x) = e^{+ipx} v(p) \\ \hspace{15em} \text{solutions d'énergie négative} \\ \text{avec maintenant } p^0 = +\sqrt{\vec{p}^2 + m^2} \text{ dans les } \mathbf{deux \ formules} \\ \text{normalisations invariantes de Lorentz : } \bar{u}(p)u(p) = -\bar{v}(p)v(p) = 1 \end{array} \right. \quad (21.63)$$

autre méthode pour construire les solutions ondes planes : partons des solutions au repos définies par

$$-i\gamma^0 \partial_0 \psi + mc\psi = 0 \quad \text{où } \gamma_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (21.64)$$

donc

$$\psi_+(x, 1) = e^{-imc^2 t} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = e^{-imc^2 t} u(0, 1) \quad (21.65)$$

$$\psi_+(x, -1) = e^{-imc^2 t} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = e^{-imc^2 t} u(0, -1) \quad (21.66)$$

$$\psi_-(x, 1) = e^{imc^2t} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = e^{imc^2t} v(0, 1) \quad (21.67)$$

$$\psi_-(x, -1) = e^{imc^2t} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = e^{imc^2t} v(0, -1) \quad (21.68)$$

Effectuons à présent un boost qui amène le vecteur  $(mc, \vec{0})$  sur le vecteur  $(p_0, \vec{p})$  :

$$\begin{cases} x'^0 = x^0 \operatorname{ch} \phi + \vec{n} \cdot \vec{x} \operatorname{sh} \phi \\ \vec{x}' = \vec{x} - x_0 \vec{n} \operatorname{sh} \phi + \vec{n} (\vec{n} \cdot \vec{x}) (\operatorname{ch} \phi - 1) \end{cases}$$

$$\text{avec} \quad \operatorname{ch} \phi = \frac{p_0}{mc} \quad \vec{n} = \frac{\vec{p}}{|\vec{p}|}$$

$$\psi'(x) = S(\Lambda) \psi(\Lambda^{-1}x)$$

$$(\Lambda^{-1}x)_0 = x_0 \operatorname{ch} \phi - \vec{n} \cdot \vec{x} \operatorname{sh} \phi = \frac{x_0 p_0 - \vec{x} \cdot \vec{p}}{mc} \quad \text{donc } mc^2t \text{ devient } p^\mu x_\mu$$

$$\begin{aligned} \psi'_+(x, \sigma) &= \left( \operatorname{ch} \frac{\phi}{2} + \vec{\alpha} \cdot \vec{n} \operatorname{sh} \frac{\phi}{2} \right) u(0, \sigma) e^{-ipx} \\ &= \left( \operatorname{ch} \frac{\phi}{2} + \vec{\alpha} \cdot \frac{\vec{p}}{|\vec{p}|} \operatorname{sh} \frac{\phi}{2} \right) u(0, \sigma) e^{-ipx} \end{aligned}$$

or

$$\begin{cases} \operatorname{ch}^2 \frac{\phi}{2} = \frac{1 + \operatorname{ch} \phi}{2} & \operatorname{ch} \frac{\phi}{2} = \sqrt{\frac{1}{2} + \frac{p_0}{2mc}} = \sqrt{\frac{mc + p_0}{2mc}} \\ \operatorname{sh}^2 \frac{\phi}{2} = \frac{\operatorname{ch} \phi - 1}{2} & \operatorname{sh} \frac{\phi}{2} = \sqrt{\frac{p_0}{2mc} - \frac{1}{2}} = \sqrt{\frac{p_0 - mc}{2mc}} \end{cases}$$

$$|\vec{p}| = \sqrt{p_0^2 - m^2 c^2} = \sqrt{(p_0 - mc)(p_0 + mc)}$$

donc

$$\psi'_+(x, \sigma) = \frac{p_0 + mc + \vec{\alpha} \cdot \vec{p}}{\sqrt{(2mc)(mc + p_0)}} u(0, \sigma) e^{-ipx} = u(p, \sigma) e^{-ipx}$$

soit

$$\left\| \psi'_+(x, \sigma) = \frac{\not{p} + m}{\sqrt{2mc(mc + p_0)}} u(0, \sigma) e^{-ipx} = u(p, \sigma) e^{-ipx} \right.$$

car  $\not{p} = p_0 \gamma_0 - \vec{p} \cdot \vec{\gamma} = p_0 \gamma_0 - \gamma_0 \vec{p} \cdot \vec{\alpha} = p_0 \gamma_0 + \vec{p} \cdot \vec{\alpha} \gamma_0$  car  $\{\vec{\alpha}, \gamma_0\} = 0$

or  $\gamma_0 u(0, \sigma) = u(0, \sigma)$  donc  $(\not{p} + mc)u(O, \sigma) = (p_0 + mc + \vec{\alpha} \cdot \vec{p})u(0, \sigma)$

De même

$$\begin{aligned}\psi'_-(x, \sigma) &= \left( ch \frac{\phi}{2} + \vec{\alpha} \frac{\vec{p}}{|\vec{p}|} sh \frac{\phi}{2} \right) v(0, \sigma) e^{ipx} \\ &= \frac{p_0 + mc + \vec{\alpha} \cdot \vec{p}}{\sqrt{2mc(mc + p_0)}} v(0, \sigma) e^{ipx} = v(p, \sigma)\end{aligned}$$

$$\gamma_0 v(0, \sigma) = -v(0, \sigma)$$

donc

$$\left\| \psi'_-(x, \sigma) = \frac{-\not{p} + m}{\sqrt{2mc(mc + p_0)}} v(0, \sigma) e^{ipx} = v(p, \sigma) \right.$$

qui satisfait  $(\not{p} + m)v(p, \sigma) = 0$

$$\bar{u}(p, \sigma) = \frac{u^+(0, \sigma)(\not{p}^+ + m)\gamma^0}{\sqrt{2mc(mc + p_0)}} = \frac{u^+(0, \sigma)\gamma^0(\not{p} + m)\gamma^0\gamma^0}{\sqrt{2mc(mc + p_0)}}$$

$$\left\| \bar{u}(p, \sigma) = \frac{\bar{u}(0, \sigma)(\not{p}^+ + m)}{\sqrt{2mc(mc + p_0)}} \right.$$

et de même

$$\left\| \bar{v}(p, \sigma) = \bar{v}(0, \sigma) \frac{(-\not{p} + m)}{\sqrt{2mc(mc + p_0)}} \right.$$

qui satisfait  $\bar{v}(p, \sigma)(\not{p} + m) = 0$

### Relations d'orthogonalité

$$\text{Dans le référentiel au repos } \bar{u}(0, \sigma)u(0, \sigma') = \delta_{\sigma\sigma'}$$

$$\bar{v}(0, \sigma)v(0, \sigma') = -\delta_{\sigma\sigma'}$$

$$\bar{u}(0, \sigma)v(0, \sigma') = 0 \text{ et } \bar{v}(0, \sigma)u(0, \sigma') = 0$$

Ces scalaires sont invariants de Lorentz, donc

$$\left\| \begin{aligned}\bar{u}(p, \sigma)u(p, \sigma') &= -\bar{v}(p, \sigma)v(p, \sigma') = \delta_{\sigma\sigma'} \\ \bar{u}(0, \sigma)v(0, \sigma') &= \bar{v}(0, \sigma)u(0, \sigma') = 0\end{aligned} \right.$$

### Opérateurs de projection

D'après ce que nous venons de montrer, l'opérateur

$$\Lambda_+(p) = \sum_{\sigma=1, -1} u(p, \sigma)\bar{u}(p, \sigma)$$

projette sur les solutions d'énergie positive :

$$\Lambda_+(p)u(p, \sigma) = u(p, \sigma)$$

$$\Lambda_+(p)v(p, \sigma) = 0$$

Calculons cet opérateur

$$\Lambda_+(p) = \frac{1}{2m(m+p_0)}(\not{p} + m)\frac{1+\gamma^0}{2}(\not{p} + m) \quad (c=1)$$

$$\begin{aligned} (\not{p} + m)\gamma^0(\not{p} + m) &= (\gamma^0 p^0 - \vec{\gamma} \cdot \vec{p} + m)\gamma^0(\not{p} + m) \\ &= (p^0 + \gamma^0 \vec{\gamma} \cdot \vec{p} + m\gamma^0)\gamma^0(\not{p} + m) \\ &= \gamma^0(2\gamma^0 p^0 - \gamma^0 p^0 + \vec{\gamma} \cdot \vec{p} + m)\gamma^0(\not{p} + m) \\ &= \gamma^0(2\gamma^0 p^0 - \not{p} + m)(\not{p} + m) \\ &= 2p^0(\not{p} + m) \quad \text{car } p^2 = m^2 \end{aligned}$$

Donc  $\Lambda_+(p) = \frac{1}{2m(m+p_0)}[p^0(\not{p} + m) + m(\not{p} + m)]$  car  $p^2 = m^2$

$$\left\| \Lambda_+(p) = \frac{\not{p} + m}{2m} \right.$$

De même

$$\begin{aligned} \Lambda_-(p) &= - \sum_{\sigma=1,-1} v(p, \sigma)\bar{v}(p, \sigma) \\ &= \frac{1}{2m(m+p_0)}(\not{p} - m)\frac{1-\gamma^0}{2}(\not{p} - m) \\ &= \frac{1}{2m(m+p_0)}[-p^0(\not{p} - m) - m(\not{p} - m)] \end{aligned}$$

$$\left\| \Lambda_-(p) = \frac{-\not{p} + m}{2m} \right.$$

Ces opérateurs vérifient  $\Lambda_{\pm}^2(p) = \Lambda_{\pm}(p)$  (projecteurs)

$$\text{tr } \Lambda_{\pm}(p) = 2$$

$$\Lambda_+(p) + \Lambda_-(p) = \mathbb{1}$$

Calculons la densité de probabilité  $e = j^0(p) = \bar{\psi}(p)\gamma^0\psi(p)$  pour ces ondes planes

$$\begin{aligned} \bar{\psi}_+(x, \sigma)\gamma^0\psi_+(x, \sigma') &= \bar{u}(p, \sigma)\gamma^0u(p, \sigma') \\ &= \bar{u}(p, \sigma)\frac{\{\not{p}, \gamma^0\}}{2m}u(p, \sigma') \end{aligned}$$

$$\text{car } \begin{cases} (\not{p} - m)u(p, \sigma') = 0 \\ \bar{u}(p, \sigma)(\not{p} - m) = 0 \end{cases}$$

donc  $j^0(p) = \bar{u}(p, \sigma)\frac{p^0}{m}u(p, \sigma') = \frac{p^0}{m}\delta_{\sigma\sigma'}$

pour les solutions d'énergie négative :

$$j^0(p) = \bar{v}(p, \sigma)\gamma^0v(p, \sigma') = -\bar{v}(p, \sigma)\frac{\{\not{p}, \gamma^0\}}{2m}v(p, \sigma')$$



$$\text{car} \quad \begin{cases} (\not{p} + m)v(p, \sigma') = 0 \\ \bar{v}(p, \sigma)(\not{p} + m) = 0 \end{cases}$$

$$j^0(p) = \frac{p^0}{m} \delta_{\sigma\sigma'}$$

On vérifie bien que la densité de probabilité est positive à la fois pour les solutions d'énergie positive et négative.

De plus  $j^0(p)d^3x$  est bien invariant de Lorentz :  $j^0(p)$  augmente d'un facteur  $\frac{E}{m}$  par rapport au repos mais  $d^3x$  diminue du même facteur.

Terminons l'étude des solutions onde plane en faisant la remarque suivante : deux solutions d'énergie opposées mais de tri-impulsion identique sont orthogonales :

$$\begin{aligned} \psi_+(x, \sigma) &= e^{-i(p^0 x^0 - \vec{p} \cdot \vec{x})} u(p, \sigma) \\ \psi_-(x, \sigma') &= e^{i(p^0 x^0 + \vec{p} \cdot \vec{x})} v(\tilde{p}, \sigma') \quad \text{avec } \tilde{p} = (p^0 - \vec{p}) \end{aligned}$$

$$\begin{aligned} \psi_-(x, \sigma') \psi_+(x, \sigma) &= e^{2-ip^0 x^0} \bar{v}(\tilde{p}, \sigma') \gamma^0 u(p, \sigma) \\ &= \frac{1}{2} e^{2-ip^0 x^0} \bar{v}(\tilde{p}, \sigma') \left( -\frac{\not{p}}{m} \gamma^0 + \gamma^0 \frac{\not{p}}{m} \right) u(p, \sigma) \end{aligned}$$

or  $\tilde{p} \gamma^0 = \gamma^0 \not{p}$  d'où le résultat :  $\bar{\psi}_-(x, \sigma') \psi_+(x, \sigma) = 0$

## 21.5 Spin et hélicité

### 21.5.1 quadrivecteur de Pauli Lubanski

Nous avons vu lors de l'étude des représentations du groupe de Poincaré que les générateurs du petit groupe  $L_p$  sont représentés dans  $\mathcal{H}_p$  par les composantes de l'opérateur de Pauli Lubanski

$$W_\mu = \frac{1}{2} \varepsilon_{\mu\nu\rho\sigma} J^{\nu\rho} p^\sigma$$

Dans la représentation  $(\frac{1}{2}, 0) + (0, \frac{1}{2})$ , nous avons vu que  $J^{\mu\nu} = \frac{1}{2} [\sigma^\mu, \sigma^\nu] = \frac{i}{4} [\gamma^\mu, \gamma^\nu]$ .

**Rappel :** comme  $W \cdot p = 0$  on peut décomposer  $W$  sur une tétrade  $(\frac{p}{m}, n_i(p))$  notée  $[p]$  :

$$p \cdot n_i = 0$$

$$n_i^2 = -1$$

$$W(\mu) = m \sum_{i=1}^3 S_i(p) n_\mu^i(p)$$

Les  $S_i = -\frac{1}{m} W \cdot n_i(p)$  vérifient  $[S_i(p), S_j(p)] = i\varepsilon_{ijk} S_k(p)$ . Dans la représentation de Dirac,

$$\begin{aligned} S_i(p) &= -\frac{i}{2m} \varepsilon_{\mu\nu\rho\sigma} [\gamma^\nu, \gamma^\rho] p^\sigma n_i^\mu \\ &= +\frac{i}{4m} \varepsilon_{\mu\nu\rho\sigma} \gamma^\nu \gamma^\rho p^\sigma n_i^\mu \end{aligned}$$

or  $\varepsilon_{\mu\nu\rho\sigma}\gamma^\rho\gamma^\sigma = i\gamma^s[\gamma_\mu, \gamma_\nu]$

donc  $S_i(p) = -\frac{\gamma^s}{4m}[\not{p}, \not{h}_i]$

comme  $\not{p}\not{h} + \not{h}\not{p} = 2n \cdot p = 0$ ,  $S_i(p) = -\frac{\gamma^s}{2m}\not{p}\not{h}_i = \frac{\gamma^s}{2m}\not{h}_i\not{p}$ .

La restriction de  $S_i(p)$  aux solutions d'énergie positive (i.e.  $\not{p}n = mn$ ) s'écrit donc

$$\boxed{S_i^+(p) = \frac{1}{2}\gamma^s\not{h}_i(p)}$$

Dans le cas des solutions d'énergie négative, qui satisfont  $\not{p}v = -mv$ ,

$$\boxed{S_i^-(p) = -\frac{1}{2}\gamma^s\not{h}_i(p)}$$

Ainsi la construction d'états de spin donné est équivalente à la recherche des états propres de  $\vec{S}^2 = \frac{3}{4}\mathbf{1}$  et  $S_3^\pm(p) = \pm\frac{1}{2}\gamma_s\not{h}_3(p)$  où le quadrivecteur  $n_3^\nu(p)$  se déduit du vecteur  $(0, \vec{S})$  par la transformation de Lorentz qui amène  $(mc, \vec{0})$  sur le vecteur  $p^\mu$ . Nous allons voir qu'il existe une tétrade particulière appelée tétrade d'hélicité, définie par  $n_3(p) = \left(\frac{|\vec{p}|}{m}, \frac{p_0}{m} \frac{\vec{p}}{|\vec{p}|}\right)$  telle que  $S_3(p) = \frac{1}{2} \frac{\vec{S} \cdot \vec{p}}{|\vec{p}|} =$  opérateur d'hélicité.

En mécanique quantique relativiste, on définit donc le spin d'un système en le rapportant à un système d'axes liés au vecteur  $p^\mu$ . On ne peut pas, comme en mécanique non relativiste définir des états  $|\vec{p}, \sigma\rangle = |\vec{p}\rangle |\sigma\rangle$ . Ceci est directement lié au fait que le groupe de Poincaré est le produit semi-direct du sous-groupe des translations d'espace temps et du groupe de Lorentz.

### 21.5.2 Spin et hélicité

Dans un référentiel où la particule est au repos,  $\hat{p} = (m, \vec{0})$ . Dans ce repère, prenons pour axe de quantification le vecteur  $\hat{n}_3 = \hat{s} = (0, \vec{s})$ , avec  $\vec{s}^2 = 1$ . On a bien  $\hat{p} \cdot \hat{s} = 0$ .

$$S_3^+(p) = \frac{1}{2}\gamma^s \cdot \hat{p} = \frac{1}{2} \begin{pmatrix} \vec{\sigma} \cdot \vec{s} & 0 \\ 0 & -\vec{\sigma} \cdot \vec{s} \end{pmatrix}$$

On doit donc chercher un spineur à deux composantes  $\varphi(\sigma)$  qui diagonalise  $\vec{\sigma} \cdot \vec{s}$  :  $\frac{1}{2}\vec{\sigma} \cdot \vec{s}\varphi(\sigma) = \frac{1}{2}\sigma\varphi(\sigma)$ . Il suffit pour cela de construire les éléments  $\pm U$  de  $SU(2)$  qui induisent la rotation  $R$  amenant  $\vec{h}$  sur  $\vec{n}$ . Les 3-spineurs  $U|\frac{1}{2}m\rangle$  sont alors états propres de  $J_{\vec{s}} = UJ_{\vec{k}}U^{-1}$  de valeur propre  $m$ . (Voir cours sur les opérateurs tensoriels irréductibles).

figure

$$\vec{n} \begin{cases} -\sin \varphi \\ \cos \varphi \\ 0 \end{cases}$$

$$\begin{aligned} U &= \cos \frac{\theta}{2} - i \sin \frac{\theta}{2} \vec{\sigma} \cdot \vec{n} \\ &= \cos \frac{\theta}{2} - i \sin \frac{\theta}{2} (-\sin \varphi \sigma_1 + \cos \varphi \sigma_2) \\ &= \begin{pmatrix} \cos \frac{\theta}{2} & i \sin \frac{\theta}{2} \sin \varphi & -\sin \frac{\theta}{2} \cos \varphi \\ i \sin \frac{\theta}{2} \sin \varphi + \sin \frac{\theta}{2} \cos \varphi & & \cos \frac{\theta}{2} \end{pmatrix} \\ &= \begin{pmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} e^{-i\varphi} \\ \sin \frac{\theta}{2} e^{i\varphi} & \cos \frac{\theta}{2} \end{pmatrix} \end{aligned}$$

$$\text{d'où } \varphi(1) = U \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\varphi} \end{pmatrix} \quad \varphi(-1) = U \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -\sin \frac{\theta}{2} e^{-i\varphi} \\ \cos \frac{\theta}{2} \end{pmatrix}$$

Donc le spineurs de Dirac  $u(\hat{p}, \sigma) = \begin{pmatrix} \varphi(\sigma) \\ 0 \end{pmatrix}$  vérifie

$$S_3^+(\hat{p})u(\hat{p}, \sigma) = \frac{1}{2}\sigma u(\hat{p}, \sigma)$$

Pour construire des états d'impulsion quelconque  $p$  il faut maintenant effectuer un boost  $\Lambda$  qui amène  $\hat{p} = (m, \vec{0})$  sur  $p = (p_0, \vec{p})$ .

$$\Lambda : \quad x' = \Lambda x \quad \text{avec} \quad \begin{cases} x'_0 = x^0 \operatorname{ch} \phi + \frac{\vec{p} \cdot \vec{x}}{|\vec{p}|} \operatorname{sh} \phi \\ \vec{x}' = \vec{x} + x_0 \frac{\vec{p}}{|\vec{p}|} \operatorname{sh} \phi + \frac{\vec{p}(\vec{p} \cdot \vec{x})}{\vec{p}^2} (\operatorname{ch} \phi - 1) \end{cases}$$

avec  $\operatorname{ch} \phi = \frac{p^0}{m}$  et  $\operatorname{sh} \phi = \frac{|\vec{p}|}{m}$

$$\text{En effet tout boost s'écrit} \begin{cases} x'_0 = x^0 \operatorname{ch} \phi + (\vec{n} \cdot \vec{x}) \operatorname{sh} \phi \\ \vec{x}' = \vec{x} - \vec{n}(\vec{x} \cdot \vec{n}) + (x^0 \operatorname{sh} \phi + (\vec{n} \cdot \vec{x}) \operatorname{ch} \phi) \vec{n} \end{cases}$$

$$\text{la contrainte étant ici} \begin{cases} p^0 = m \operatorname{ch} \phi \\ \vec{p} = m \operatorname{sh} \phi \vec{n} \end{cases} \quad (\hat{p} \rightarrow p)$$

$$\text{donc} \begin{cases} \operatorname{ch} \phi = \frac{p^0}{m} \\ \vec{n} = \frac{\vec{p}}{m \operatorname{sh} \phi} = \frac{\vec{n}}{|\vec{p}|} \end{cases} \quad \mathbf{c.q.f.d.}$$

$\Lambda(p)$  amène  $\hat{S} = (0, \vec{S})$  sur le vecteur  $n^\mu = (n^0, \vec{n})$  :

$$\begin{cases} n^0 = \frac{\vec{p} \cdot \vec{S}}{|\vec{p}|} \operatorname{sh} \phi = \frac{\vec{p} \cdot \vec{S}}{m} \\ \vec{n} = \vec{S} + \frac{\vec{p}(\vec{p} \cdot \vec{S})}{\vec{p}^2} (\operatorname{ch} \phi - 1) = \vec{S} + \frac{\vec{p}(\vec{p} \cdot \vec{S})}{m(m + p_0)} \end{cases}$$

$n$  vérifie  $n^\mu p_\mu = \xi^\mu \hat{p}_\mu = 0$ .

On dispose ainsi d'une caractérisation relativiste des états de spin d'impulsion quelconque :

$$\sqrt{\frac{p_0 + m}{2m}} \begin{pmatrix} \varphi(\sigma) \\ \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + m} \varphi(\sigma) \end{pmatrix} = u(p, \sigma)$$

satisfait  $\frac{1}{2} \gamma_s \not{p} u(p, \sigma) = \frac{\sigma}{2} u(p, \sigma)$

Pour les solutions d'énergie négative,

$$v(p, \sigma) = \sqrt{\frac{p_0 + m}{2m}} \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + m} \chi(\sigma) \\ \chi(\sigma) \end{pmatrix} \quad \text{vérifie} \quad -\frac{1}{2} \gamma_s \not{p} v(p, \sigma) = \frac{1}{2} \sigma v(p, \sigma)$$

$$\text{si } \vec{\sigma} \cdot \vec{s} \chi(\sigma) = \sigma \chi(\sigma)$$

### Hélicité

Choisissons pour axe de quantification le vecteur  $\vec{s} = \frac{\vec{p}}{|\vec{p}|}$

$$\begin{aligned} \text{alors} \quad n^0 &= \frac{\vec{p}^2}{m|\vec{p}|} = \frac{|\vec{p}|}{m} \\ \vec{n} &= \frac{\vec{p}}{|\vec{p}|} + \frac{|\vec{p}|\vec{p}}{m(m+p^0)} = \frac{\vec{p}}{|\vec{p}|} \frac{m^2 + mp_0 + p_0^2 - m^2}{m(m+p_0)} \\ \text{soit} \quad \vec{n} &= \frac{\vec{p}p_0}{|\vec{p}|m} \end{aligned}$$

Par construction  $n^\mu = (n^0, \vec{n}) = \left( \frac{|\vec{p}|}{m}, \frac{\vec{p}p_0}{|\vec{p}|m} \right)$  vérifie  $p \cdot n = 0$  et  $n^2 = -1$ . On obtient alors

$$\gamma^5 \not{p} = \frac{m}{|\vec{p}|} \begin{pmatrix} \vec{\sigma} \cdot \vec{p} & 0 \\ 0 & \vec{\sigma} \cdot \vec{p} \end{pmatrix}$$

d'où

$$S_3(p) = \frac{\gamma^s \not{p}}{2m} = \frac{1}{2|\vec{p}|} \begin{pmatrix} \vec{\sigma} \cdot \vec{p} & 0 \\ 0 & \vec{\sigma} \cdot \vec{p} \end{pmatrix} = \frac{\vec{\Sigma} \cdot \vec{p}}{2|\vec{p}|}$$

qui est l'opérateur hélicité introduit plus haut.

Comme  $[H, S_3(p)] = 0$ , on peut donc rechercher des spineurs qui diagonalisent  $H$  et qui ont une hélicité donnée.

**Limite de masse nulle** :  $m \rightarrow 0$ ,  $p^0 \rightarrow |\vec{p}|$  donc

$$n \rightarrow \left( \frac{|\vec{p}|}{m}, \frac{\vec{p}}{m} \right) = \frac{1}{m} (p^0, \vec{p}) = \frac{p}{m}$$

donc

$$\gamma^5 \not{p} = \gamma^5 \frac{\not{p}^2}{m} = \gamma^5 m \quad \text{soit} \quad \boxed{S_3(p) = \frac{1}{2} \gamma^5}$$

exemple : neutrino  $\frac{1}{2} \gamma^s \psi_\nu(p) = -\frac{1}{2} \psi_\nu(p)$

Dans la théorie de l'interaction faible les couplages sont de la forme "V-A" :

$$\bar{\psi}_c \gamma_\mu (1 - \gamma^5) \psi_\nu$$

(courant vectoriel :  $\bar{\psi}_c \gamma_\mu \psi_\nu$  axial :  $\bar{\psi}_c \gamma_\mu \gamma^5 \psi_\nu$ )

donc seule la partie gauche du champ du neutrino est effectivement couplée aux autres champs.

Dans la représentation chirale,  $\gamma_s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . Donc :

$$\begin{aligned} \gamma_s \begin{pmatrix} \varphi \\ 0 \end{pmatrix} &= + \begin{pmatrix} \varphi \\ 0 \end{pmatrix} && \text{état d'hélicité positive} \\ \gamma_s \begin{pmatrix} 0 \\ \chi \end{pmatrix} &= - \begin{pmatrix} 0 \\ \chi \end{pmatrix} && \text{état d'hélicité négative} \end{aligned}$$

$\varphi$  et  $\chi$  sont des spineurs de Weyl qui vérifient  $\begin{cases} \frac{\partial \chi}{\partial t} = \vec{\sigma} \cdot \vec{\nabla} \chi \\ \frac{\partial \varphi}{\partial t} = -\vec{\sigma} \cdot \vec{\nabla} \varphi \end{cases}$

## 21.6 La conjugaison de charge

La théorie développée par Dirac en 1928 est celle d'un électron de masse  $m$ , de spin  $\frac{1}{2}$ , de charge  $e = -|e|$  et obéissant à la statistique de Fermi-Dirac (à l'époque ceci signifiait satisfaire le principe d'exclusion de Pauli, la notion d'algèbre Grassmannienne n'étant pas encore élaborée). Nous avons vu que l'équation que Dirac écrivit prédit, dans le cas d'une particule libre, un spectre d'énergie donné par  $\pm E_p$ , i.e.  $[-\infty, -mc^2] \cup [mc^2, +\infty[$ . Il reste de même nature si l'on considère l'équation de Dirac avec un potentiel s'annulant à l'infini.

Si on couple cette équation avec la théorie quantique du rayonnement, en considérant un électron atomique, l'électron dans un état excité rayonne en émettant un photon ou plusieurs photons et passe dans un état d'énergie inférieure, jusqu'à finalement atteindre son état fondamental qui est supposé d'énergie finie. Etat donné le spectre que nous avons obtenu, si l'état fondamental est d'énergie finie, comme ce n'est pas l'état de plus basse énergie (l'énergie peut descendre jusqu'à  $-\infty$ ), il existe toujours une certaine probabilité de transition vers un état d'énergie inférieure : l'état fondamental d'un atome est instable !

En 1930 Dirac a proposé, pour résoudre cette contradiction avec l'expérience (un atome est stable), que dans les conditions usuelles tous les états d'énergie négative sont remplis. Le principe d'exclusion de Pauli empêche alors un électron d'énergie positive de tomber dans un état d'énergie négative. L'ensemble des électrons d'énergie négative constitue la mer de Dirac. C'est le vide de la théorie. Il est constitué d'un gaz de Fermi de densité infinie. Les propriétés physiques observables (énergie, charge, impulsion, spin) sont celles du système complet moins celle vide.

Supposons qu'un photon d'énergie  $\hbar\omega > 2mc^2$  soit absorbé par un électron de la mer de Dirac. Il devient alors un électron d'énergie positive, et un **trou** est créé dans la mer de Dirac. L'interprétation de ce trou, compatible avec les lois de conservation de l'énergie, de la charge, de l'impulsion et du moment cinétique, est de considérer qu'il décrit une particule d'énergie positive  $p^0$ , de charge  $-e = |e|$ , de masse  $m$ , d'impulsion  $\vec{p}$  et de spin  $\vec{\sigma}$  (l'énergie de l'électron d'énergie négative est  $-p^0$ , son impulsion  $-\vec{p}$ , son spin  $-\vec{\sigma}$ , sa charge  $e = -|e|$ ). Dirac pensait que ces trous étaient des protons, mais on peut alors montrer qu'un atome d'hydrogène se désintégrerait en deux photons en environ  $10^{-10}$  s !

En 1932 Anderson découvrit le positron, auquel s'identifie le trou d'énergie négative. C'est l'antiparticule associée à l'électron.

La réaction que nous venons d'examiner

$$\begin{array}{ccc} e_{E<0}^- & +\gamma \rightarrow & e_{E>0}^- \\ \text{trou} & \text{photon} & \text{électron} \end{array}$$

s'interprète donc comme  $\gamma \rightarrow e_{E>0}^- + e_{E>0}^+$ . Cette réaction peut avoir lieu dans le champ de Coulomb d'un noyau.

Ainsi, suivant l'interprétation des trous, la théorie de Dirac, élaborée initialement comme étant la théorie d'une particule, l'électron, conduit naturellement à une théorie des électrons et des positrons, et pour sa cohérence nécessite une analyse multiparticulaire.

Le rôle disymétrique joué dans l'équation de Dirac par les électrons et les positrons peut se corriger par la conjugaison de charge que nous allons maintenant étudier. La structure multiparticulaire trouve sa cohérence dans le cadre de la théorie quantique des champs (voir cours sur QED).

Pour que les  $e^-$  et les  $e^+$  jouent un rôle symétrique, il faudrait que ces derniers obéissent à l'équation de Dirac, mais avec une charge  $-e$ . Cherchons donc une transformation  $\psi \rightarrow \psi_c$  telle que

$$\begin{aligned} (i\rlap{-}\partial - e\lambda - m)\psi &= 0 \\ (i\rlap{-}\partial + e\lambda - m)\psi_c &= 0 \end{aligned}$$

Cette transformation doit être locale, et involutive à une phase près (puisque son carré doit redonner à une phase près la fonction d'onde de départ).

$$\bar{\psi} \left( -i\overleftarrow{\partial} - e\lambda - m \right) = 0$$

$$\text{soit } -i\partial_\mu \psi^+ \gamma^0 \gamma^\mu - e\psi^+ \gamma^0 \gamma^\mu A_\mu - \psi^+ \gamma^0 m = 0.$$

Transposons cette équation :

$$\left[ {}^t\gamma^\mu (-i\partial_\mu - eA_\mu) - m \right] \underbrace{{}^t\bar{\psi}}_{\gamma^0 \psi^*} = 0 \quad ({}^t\gamma^0 = \gamma^0)$$

Supposons qu'il existe  $C$  tel que  $C {}^t\gamma^\mu C^{-1} = -\gamma^\mu$ . L'équation précédente se réécrit :

$$\left( {}^t\gamma^\mu C^{-1} (-i\partial_\mu - eA_\mu) - C^{-1} m \right) C {}^t\bar{\psi} = 0$$

$$\text{soit } C^{-1} \underbrace{[C {}^t\gamma^\mu C^{-1} (-i\partial_\mu - eA_\mu) - m]}_{-\gamma^\mu} C {}^t\bar{\psi} = 0$$

donc  $(i\rlap{-}\partial + e\lambda - m)\psi_c = 0$  en posant  $\boxed{\psi_c = C {}^t\bar{\psi}}$ .

Dans la représentation standard, on peut vérifier que

$$C = i\gamma^2 \gamma^0 = \begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (21.69)$$

est solution.

$C$  vérifie (voir cours sur les spineurs tridimensionnels) :  $-C = C^{-1} = {}^t C = C^+$  et  $\widehat{C} = i\sigma_2$  vérifie  $\widehat{C}\sigma_i^*\widehat{C}^{-1} = -\sigma_i$

$$C = i\gamma^2\gamma^0 \quad \text{donc} \quad \|\psi_c = i\gamma^2\gamma^0 {}^t(\psi^+\gamma_0) = i\gamma^2\psi^*$$

Examinons comment une solution d'énergie positive  $\psi_+ = u(p)e^{-ipx}$  se transforme sous la conjugaison de charge :

$$\psi_c = u_c(p)e^{ipx}$$

avec

$$\begin{aligned} u_c(p) &= i \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix} u(p)^* \\ &= \begin{pmatrix} 0 & \widehat{c} \\ -\widehat{c} & 0 \end{pmatrix} u(p)^* = \begin{pmatrix} 0 & \widehat{c} \\ -\widehat{c} & 0 \end{pmatrix} \begin{pmatrix} \varphi^* \\ \frac{(\vec{\sigma}\cdot\vec{p})^*\varphi^*}{p^0+m} \end{pmatrix} \sqrt{\frac{p^0+m}{2m}} \\ &= \begin{pmatrix} \widehat{c}(\vec{\sigma}\cdot\vec{p})^*\varphi^* \\ -\widehat{c}\varphi^* \end{pmatrix} \sqrt{\frac{p^0+m}{2m}} = \begin{pmatrix} -\widehat{c}(\vec{\sigma}\cdot\vec{p})^*\widehat{c}^{-1}\chi^* \\ \chi \end{pmatrix} \sqrt{\frac{p^0+m}{2m}} \end{aligned}$$

où l'on a posé  $\chi = -C\varphi^*$

$$\text{d'où } u_c(p) = \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{p^0+m} \\ \chi \end{pmatrix} \sqrt{\frac{p^0+m}{2m}} = v(p)$$

Comment se transforme le spin ?

$$\begin{cases} \frac{1}{2}\gamma_s \not{h} u(p, \sigma) = \frac{\sigma}{2} u(p, \sigma) & \text{avec } \vec{\sigma} \cdot \vec{s}\varphi = \sigma\varphi \\ -\frac{1}{2}\gamma_s \not{h} v(p, \sigma') = \frac{\sigma'}{2} v(p, \sigma') & \text{que vaut } \sigma' \text{ lorsque } v(p) \text{ est obtenu par} \\ & \text{conjugaison de charge à partir de } u(p) ? \end{cases}$$

$\chi = -c\varphi^*$ . De  $\vec{\sigma} \cdot \vec{s}\varphi = \sigma\varphi$  on tire  $-\vec{\sigma}^* \cdot \vec{s} - \varphi^* = -\sigma\varphi^*$  soit

$$-C^{-1} \underbrace{C\vec{\sigma}^* \cdot \vec{s} C^{-1}}_{-\vec{\sigma} \cdot \vec{s}} C\varphi^* = -\sigma C^{-1} C\varphi^*$$

d'où  $-\vec{\sigma} \cdot \vec{s}\chi = \sigma\chi$  soit  $\vec{\sigma} \cdot \vec{s}\chi = -\sigma\chi$

Le spin de l'état obtenu par conjugaison de charge est donc renversé :  $\|u_c(p, \sigma) = v(p, -\sigma)$

## 21.7 Lagrangien et propagateur

La densité lagrangienne correspondant à l'équation de Dirac est

$$\mathcal{L}_D = \frac{i}{2} (\bar{\psi}\gamma^\mu\partial_\mu\psi - (\partial_\mu\bar{\psi}\gamma^\mu\psi) - m\bar{\psi}\psi)$$

Les équations du mouvement donnent en effet :

$$\frac{\delta\mathcal{L}}{\delta\bar{\psi}} - \partial_\mu \frac{\delta\mathcal{L}}{\delta\partial_\mu\bar{\psi}} = 0 \quad \text{soit} \quad \frac{i}{2}\gamma^\mu\partial_\mu\psi + \frac{i}{2}\gamma^\mu\partial_\mu\psi - m\psi = 0$$

donc  $\gamma^\mu i\partial_\mu \psi - m\psi = 0$  de même la variation par rapport à  $\psi$  donne

$$-\frac{i}{2}\partial_\mu \bar{\psi}\gamma^\mu - \frac{i}{2}\partial_\mu \bar{\psi}\gamma^\mu - m\bar{\psi} = 0$$

donc  $-i\gamma^\mu i\partial_\mu \bar{\psi} - m\bar{\psi} = 0$

on obtient donc bien l'équation de Dirac  $(\not{p} - m)\psi = 0$  et sa conjuguée  $\bar{\psi}(\overleftarrow{\not{p}} + m) = 0$ .

Comme la densité lagrangienne est linéaire en  $\psi$  ou  $\bar{\psi}$ , l'action n'a ni minimum ni maximum le lagrangien s'annule lorsque  $\psi$  et  $\bar{\psi}$  sont solution des équations du mouvement. En particulier, le signe devant le lagrangien est arbitraire. Si les champs s'annulent à l'infini, on peut ajouter la différentielle totale  $\frac{i}{2}\partial_\mu(\bar{\psi}\gamma^\mu\psi)$  à la densité lagrangienne, et donc réécrire l'action :

$$S = \int d^4x (\bar{\psi}(i\not{\partial}\psi) - m\bar{\psi}\psi)$$

En présence d'un champ électromagnétique extérieur, le couplage minimal transforme cette action en

$$S = \int d^4x (\bar{\psi}(i\not{\partial} - e\not{A})\psi - m\bar{\psi}\psi) = \int d^4x \mathcal{L}_D + \int d^4x \mathcal{L}_{int}$$

où  $\mathcal{L}_{int} = -j^\mu A_\mu$  et  $j^\mu = e\bar{\psi}\gamma^\mu\psi$ .

Au niveau classique (comme au niveau quantique), les règles seront donc les suivantes, en accord avec l'étude de la conjugaison de charge :

$$\begin{array}{ll} \text{particule entrante : } u & \text{particule sortante : } \bar{u} \\ \text{antiparticule entrante : } \bar{v} & \text{antiparticule sortante : } v \end{array}$$

et en tenant compte du couplage  $e\bar{\psi}\gamma^\mu\psi$ .



Sixième partie

Les symétries en action



# Chapitre 22

## Théories de Yang-Mills

Le principe qui guide la construction des théories de Yang-Mills est le suivant : en rendant locales les symétries globales, on peut engendrer la dynamique du système. Ce principe sera le fil conducteur de ce chapitre<sup>1</sup>, et des suivants, puisqu'il fournit de l'information tant au niveau classique qu'au niveau quantique.

En QED,  $A_0$  produit des états de norme négative. Grâce à l'identité de Ward, qui découle de l'invariance de jauge locale, on montre que ces états de norme négative sont exactement compensés par les modes longitudinaux. Il ne reste alors que deux degrés de liberté physique, correspondant aux états de polarisation transverse.

La même propriété est valable pour les théories de jauge non abéliennes. Les identités de Ward associées sont alors plus complexes, mais la conclusion sera la même. L'existence de cette symétrie permet également de montrer que la théorie est renormalisable, i.e. que les contre-terms sont invariants de jauge (et qu'en particulier le photon et le gluon sont bien des particules de masse nulle, à tous les ordres en perturbation).

### 22.1 Théorie de jauge abélienne

Nous allons effectuer un petit retour sur QED, en montrant que c'est la seule théorie, en partant du lagrangien de Dirac libre, qui puisse être obtenue si l'on impose l'invariance de jauge locale et la renormalisabilité, ainsi bien entendu que l'invariance de Lorentz.

Considérons donc le lagrangien de l'électron libre

$$\mathcal{L}_0 = \bar{\psi}(x) (i\gamma^\mu \partial_\mu - m) \psi(x)$$

Il est invariant sous la transformation globale U(1) suivante :

$$\begin{aligned} \psi(x) &\rightarrow \psi'(x) = e^{-i\alpha} \psi(x) & (\alpha \text{ réel}) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}'(x) = e^{i\alpha} \bar{\psi}(x). \end{aligned}$$

---

1. Ce chapitre utilise des notions élémentaires de théorie des groupes et algèbres de Lie, rappelées brièvement. Elle ne sont pas indispensables pour la suite du cours et peuvent être sautées en première lecture

Construisons une théorie qui soit invariante de jauge locale, i.e. laissée invariante par

$$\begin{cases} \psi(x) \rightarrow \psi'(x) = e^{-i\alpha(x)}\bar{\psi}(x) \\ \bar{\psi}(x) \rightarrow \bar{\psi}'(x) = e^{i\alpha(x)}\psi(x). \end{cases}$$

Alors

$$\begin{aligned} \bar{\psi}(x)\partial_\mu\psi(x) \rightarrow \bar{\psi}'(x)\partial_\mu\psi'(x) &= \bar{\psi}(x)e^{i\alpha(x)}\partial_\mu(e^{-i\alpha(x)}\psi(x)) \\ &= \bar{\psi}(x)\partial_\mu\psi(x) - \underbrace{i\bar{\psi}(x)\partial_\mu\alpha(x)\psi(x)}_{\text{terme non compensé!}} \end{aligned}$$

Introduisons une dérivée covariante  $D_\mu$ , construite à partir de  $\partial_\mu$ , de manière à compenser le terme en trop :  $D_\mu = \partial_\mu + ?$

Le terme à ajouter doit porter un indice de Lorentz : c'est donc un champ vectoriel (champ car il doit dépendre de  $x$  puisque  $\partial_\mu\alpha(x)$  dépend de  $x$ ). On le choisit réel, et on pose donc

$$\boxed{D_\mu = \partial_\mu + ieA_\mu.} \quad (22.1)$$

$e$  est ici un paramètre arbitraire (ce sera la charge électrique). Si

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x)$$

sous une transformation de jauge, alors

$$\begin{aligned} D_\mu\psi(x) \rightarrow [D_\mu\psi(x)]' &= [(\partial_\mu + ieA_\mu)\psi(x)]' = [\partial_\mu + ieA_\mu + i\partial_\mu\alpha(x)]e^{-i\alpha(x)}\psi(x) \\ &= e^{-i\alpha(x)}(\partial_\mu + ieA_\mu)\psi(x) = e^{-i\alpha(x)}D_\mu\psi(x) \end{aligned}$$

Donc

$$\boxed{D_\mu\psi(x) \rightarrow [D_\mu\psi(x)]' = e^{-i\alpha(x)}D_\mu\psi(x),} \quad (22.2)$$

i.e.  $D_\mu\psi(x)$  possède la même loi de transformation que  $\psi(x)$

A présent le terme  $\bar{\psi}(x)D_\mu\psi(x)$  est invariant de jauge

Le lagrangien  $\mathcal{L}'_0 = \bar{\psi}i\gamma^\mu(\partial_\mu + ieA_\mu)\psi - m\bar{\psi}\psi$  est donc invariant de jauge.

Ce champ de jauge n'est pas dynamique. En effet, le lagrangien ne possède pas de terme dérivatif en  $\partial_0 A^\mu$ , donc le moment conjugué de  $A^\mu$  est nul. Nous devons donc modifier le lagrangien.

Résultat préliminaire :

$$\boxed{F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \text{ est invariant de jauge.}} \quad (22.3)$$

*Preuves :*

- directement : le résultat est immédiat d'après la loi de transformation de  $A_\nu$
- $(D_\mu D_\nu - D_\nu D_\mu)\psi = ieF_{\mu\nu}\psi$

Comme  $D_\nu\psi \rightarrow [D_\nu\psi]' = e^{-i\alpha(x)}D_\nu\psi$  (i.e.  $D_\nu\psi$  se transforme comme  $\psi$  comme nous l'avons vu en (22.2)), on en déduit que  $(D_\mu D_\nu)\psi = D_\mu(D_\nu\psi)$  se transforme comme  $D_\mu\psi$ , donc

$$[D_\mu D_\nu\psi]' = e^{-i\alpha(x)}D_\mu D_\nu\psi,$$

d'où

$$[(D_\mu D_\nu - D_\nu D_\mu)\psi]' = e^{-i\alpha(x)}(D_\mu D_\nu - D_\nu D_\mu)\psi(x)$$

qui s'écrit encore

$$[F_{\mu\nu}\psi]' = F'_{\mu\nu}\psi' = e^{-i\alpha(x)}F_{\mu\nu}\psi = (F_{\mu\nu}\psi)e^{-i\alpha(x)}$$

d'où

$$F'_{\mu\nu} = F_{\mu\nu},$$

Le terme le plus simple qui soit de dimension inférieure ou égale à 4 (critère de renormalisabilité qui sera justifié dans les chapitres suivants), scalaire de Lorentz et invariant de jauge est

$$\mathcal{L}_A = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (\text{à une divergence totale près})$$

Le facteur 1/4 est conventionnel. Il est là pour assurer que les équations du mouvement obtenues sont bien les équations de Maxwell correctement normalisées. Le lagrangien de QED s'écrit donc finalement :

$$\boxed{\mathcal{L}_{QED} = \bar{\psi} i \gamma^\mu (\partial_\mu + ieA_\mu) \psi - m \bar{\psi} \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.} \quad (22.4)$$

*Remarques :*

- le photon est de masse nulle car un terme en  $A_\mu A^\mu$  n'est pas invariant de jauge
- le couplage du photon à la matière est complètement déterminé par l'exigence d'avoir une théorie de jauge locale (ce couplage apparaît lors du remplacement  $\partial_\mu \rightarrow D_\mu$ ). Un couplage non minimal du type  $\bar{\psi}\sigma_{\mu\nu}\psi F^{\mu\nu}$  n'est pas possible car il briserait la renormalisabilité (seuls les termes de dimension  $\leq 4$  sont permis). Un tel terme de couplage est cependant intéressant pour décrire, à basse énergie, un couplage phénoménologique au proton ou au neutron (pour décrire le moment magnétique qui est différent de celui de la théorie de Dirac corrigée par QED par exemple). On parle alors de théorie effective, au sens où, si l'on introduit une coupure UV à la main, il devient licite de considérer un lagrangien non renormalisable.
- rien n'oblige à ce que la charge électrique élémentaire soit la même pour toutes les particules chargées. Ceci est dû au fait que le groupe de jauge est U(1). Comme les constantes de structure du groupe sont nulles, il n'y a pas de relation entre les différentes charges des particules de matière couplées au champ de jauge.
- le lagrangien ne comporte pas de terme d'interaction du boson de jauge avec lui-même : le photon ne possède pas de charge U(1). Donc sans matière, QED est une théorie libre.

Les deux premières remarques restent valables pour les théories de jauge non abélienne. En revanche, comme une théorie de jauge abélienne est non linéaire (à cause de la structure du groupe), les deux dernières remarques sont susceptibles d'être modifiées.

Le principe que vous venons de mettre en oeuvre pour l'électromagnétisme a été généralisé en 1954 par Yang et Mills pour le groupe  $SU(2)$  d'isospin. Nous allons examiner ici plus généralement le cas d'un groupe de jauge  $G$  dont l'algèbre de Lie est la somme directe de sous-algèbres de Lie simples compactes et de sous algèbres  $U(1)$ . Commençons par donner quelques éléments utiles de théorie des groupes et de leur représentations.

## 22.2 Vade-mecum de théorie des groupes

### 22.2.1 Algèbre de Lie

Pour un groupe de Lie donné, l'algèbre de Lie s'obtient par différentiation au voisinage de l'identité, suivant toutes les directions indépendantes. On note alors  $n$  la dimension du groupe, qui est le nombre de générateurs de sa  $\mathbb{R}$ -algèbre de Lie. Par exemple  $n = N^2 - 1$  pour  $SU(N)$ . L'algèbre de Lie est simplement l'ensemble des combinaisons linéaires de ces générateurs (sur  $\mathbb{R}$  si on parle de la  $\mathbb{R}$ -algèbre de Lie, sur  $\mathbb{C}$  si on parle de la  $\mathbb{C}$ -algèbre de Lie), muni d'un crochet de Lie (c'est le commutateur dans le cas des algèbres de matrices). Les éléments de l'algèbre de Lie vérifient l'identité de Jacobi

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0, \quad (22.5)$$

$X, Y, Z$  étant 3 combinaisons linéaires arbitraires des générateurs.

Nous noterons conventionnellement par  $T^a$  ( $a = 1, \dots, n$ ) les générateurs de l'algèbre de Lie. Ces générateurs vérifient la relation

$$[T_a, T_b] = i f_{ab}^c T_c, \quad (22.6)$$

où  $f_{ab}^c$  sont les constantes de structure du groupe. Dans le cas qui nous intéressera des groupes de Lie compacts semi-simples (voir ci-dessous), la covariance des indices sur le groupe ne joue aucun rôle et on pourra écrire  $f_{abc} = f_{ab}^c$ , et donc

$$\boxed{[T_a, T_b] = i f_{abc} T_c.} \quad (22.7)$$

### 22.2.2 Représentations

#### Représentation d'un groupe

Soit  $V$  un espace vectoriel sur  $\mathbb{R}$  ou  $\mathbb{C}$ , de dimension  $d$ . Une représentation linéaire d'un groupe  $G$  dans  $V$  est un homomorphisme  $D$  de  $G$  dans le groupe  $GL(d)$  des opérateurs linéaires et inversibles de  $V$ . On a donc

$$\boxed{\forall g_1, g_2 \in G, D(g_1 g_2) = D(g_1) D(g_2).}$$

En particulier  $D(e) = \mathbb{1}$  et  $D(g^{-1}) = D(g)^{-1}$ ,  
où  $e$  et  $\mathbb{1}$  sont les éléments neutres de  $G$  et de  $GL(d)$ .

$V$  est appelé espace de représentation, et  $d = \dim V$  est la dimension de la représentation.  $D(g)$  s'écrit donc comme une matrice  $d \times d$ . Il ne faut surtout pas confondre cette dimension avec celle du groupe. De façon générale, un groupe donné possède plusieurs représentations, de dimensions différentes. On ne confondra pas cette dimension avec celle de la représentation du groupe ou de son algèbre de Lie (les dimensions sont égales dans le cas de la représentation adjointe).

### Représentation d'une algèbre de Lie $\mathfrak{g}$

Soit  $V$  un espace vectoriel sur  $\mathbb{R}$  ou  $\mathbb{C}$ . On dit que  $D$  est une représentation linéaire de  $\mathfrak{g}$  dans l'espace vectoriel  $V$  si

$$\begin{aligned} [D(A), D(B)] &= D([A, B]) \\ D(\alpha A + \beta B) &= \alpha D(A) + \beta D(B) \end{aligned}$$

#### Théorème :

Une représentation  $D_G$  d'un groupe  $G$  induit par différentiation une représentation  $D_A$  de son algèbre de Lie. Cela signifie que l'on passe d'une représentation du groupe à une représentation de son algèbre de Lie exactement comme on passe du groupe à son algèbre de Lie.

### Représentations unitaires

Si  $\forall g \in G$ , les opérateurs  $D(g)$  sont unitaires, la représentation est dite unitaire.

### Représentations équivalentes

Deux représentations  $D$  et  $D'$  sont dites équivalentes si et seulement si il existe un opérateur  $S$  inversible t.q.

$$\forall g \in G, D'(g) = S^{-1}D(g)S.$$

$S$  est appelé opérateur d'entrelacement.

### Quelques représentations utiles

#### Représentation fondamentale :

C'est la représentation la plus simple (en dehors de la représentation "triviale" pour laquelle  $T(g) = 1$ ). Elle consiste simplement à représenter chaque élément du groupe par lui-même. Ainsi un champ de matière  $\psi$  appartenant à l'espace de représentation de la représentation fondamentale de  $G$ , se transforme suivant

$$\psi \xrightarrow{G} \psi'_i = U_{ij}\psi_j \quad U \in G \quad (U : \text{matrice } N \times N \text{ pour } SU(N))$$

soit localement :

$$\psi \xrightarrow{G} \psi' = \psi - i\alpha^a T^a \psi$$

(ou  $\psi' = \psi + i\alpha^a T^a \psi$  si l'on préfère cette paramétrisation) qui doit se comprendre comme

$$\psi'_i = \psi_i - i\alpha^a T_{ij}^a \psi_j \quad (22.8)$$

où  $\alpha$  possède  $n$  composantes : l'indice  $a$  est une étiquette pour le  $a$ -ième générateur tandis que  $i$  et  $j$  sont des indices matriciels pour la représentation de ce générateur agissant sur les vecteurs  $\psi_j$  à  $N$  composantes.

La relation de transformation (22.8) fait intervenir la matrice  $U$  avec ses coefficients  $U_{ij}$ . Nous n'avons pas utilisé la notation covariante pour la représentation fondamentale pour simplifier la discussion. C'est néanmoins utile (et indispensable !) dans le cas où les représentations fondamentales complexes conjuguées ne sont pas équivalentes. Ainsi, dans le cas de  $SU(N)$ , dès que  $N > 2$ , les représentations conjuguées ne sont pas équivalentes. Par exemple,  $3$  et  $\bar{3}$  ne sont pas équivalentes sous  $SU(3)$ . On écrira alors  $\psi^i$  pour une représentation (conventionnellement, c'est celle correspondant aux quarks) et  $\psi_i (= \psi^{*i})$  pour sa complexe conjuguée (c'est celle des antiquarks) avec

$$\begin{array}{ll} \psi'^i = U^i_j \psi^j & U^i_j \equiv (U)_{ij} \\ \psi'_i = U_i^j \psi_j & U_i^j \equiv (\bar{U})_{ij} \end{array} \quad \boxed{\begin{array}{l} \text{matriciellement :} \\ \psi' = U\psi \\ \psi^{*'} = \bar{U}\psi^* \end{array}}$$

Comme pour les fermions on utilise  $\bar{\Psi} = \Psi^\dagger \gamma^0$  plutôt que  $\Psi^*$ , il est utile d'écrire la loi de transformation de  $\bar{\Psi}$  sous le groupe. Partant de

$$\psi^{*'} = \bar{U}\psi^*$$

on déduit que

$$\psi^{\dagger'} = {}^t \Psi^* {}^t \bar{U} = {}^t \Psi^* U^\dagger$$

et donc, puisque  $U$  n'agit que sur les indices du groupe (et non sur les indices spinoriels),

$$\bar{\Psi}' = \bar{\Psi} U^\dagger.$$

Ainsi, en résumé, pour des fermions (cas des quarks)

$$\boxed{\begin{array}{l} \psi' = U\psi \\ \bar{\Psi}' = \bar{\Psi} U^\dagger \end{array}} \quad (22.9)$$

*Représentation adjointe :*

Une représentation particulière joue un rôle important dans la construction de Yang-Mills : il s'agit de la représentation adjointe. Elle est définie de la façon suivante.



★ sur le groupe :

Sur le groupe  $G$ , l'application  $g \rightarrow ad\ g$  définie par  $ad\ g(h) = ghg^{-1}$  est une représentation de  $G$  sur lui-même.

*Preuve :*

$$\begin{aligned} \forall g_1, g_2 \in G, \forall h \in G, ad(g_1 g_2)(h) &= (g_1 g_2) h (g_1 g_2)^{-1} = g_1 (g_2 h g_2^{-1}) g_1^{-1} \\ &= ad_{g_1}(ad_{g_2}(h)) \end{aligned}$$

ce qui prouve que l'on a bien construit un morphisme.

★ sur l'algèbre :

A cette représentation sur le groupe correspond la représentation  $X \rightarrow adX$  définie par  $adX(Y) = [X, Y]$ .

Ceci signifie que les générateurs jouent deux rôles simultanément : celui de générateur et celui de vecteur de base sur lesquels agissent les générateurs. *La dimension de la représentation adjointe est donc égale à celle du groupe.*

*Preuve :*

Il suffit de différentier  $\exp(-itT_a)T_b \exp(itT_a)$  en  $t = 0$  pour en déduire l'expression de la représentation adjointe sur l'algèbre de Lie. En utilisant le théorème qui assure qu'une représentation d'un groupe induit par différentiation une représentation de son algèbre de Lie, ceci achève la preuve. Il est cependant instructif de le vérifier directement de deux façons différentes :

◇ de manière intrinsèque, il faut montrer que  $[adX_1, adX_2] = ad[X_1, X_2]$ .

$$\begin{aligned} \forall Y \in g, [adX_1, adX_2](Y) &= adX_1 adX_2(Y) - adX_2 adX_1(Y) \\ &= [X_1, [X_2, Y]] - [X_2, [X_1, Y]] = [X_1, [X_2, Y]] + [X_2, [Y, X_1]] = -[Y, [X_1, X_2]] \end{aligned}$$

soit encore  $[adX_1, adX_2](Y) = [[X_1, X_2], Y]$ , ce qui prouve le résultat.

◇ matriciellement, en utilisant les constantes de structure (22.7) on peut construire cette représentation par les  $n$  matrices  $T_a^A$  ( $A$  comme adjointe) de coefficients

$$(T_a^A)_{cb} = i f_{abc}. \quad (22.10)$$

Cette définition est bien cohérente puisque

$$T_a(T_b) = [T_a, T_b] = i f_{abc} T_c \quad (22.11)$$

et donc par identification du coefficient  $c$  de l'image de  $T_b$  par  $T_a$  on en tire bien (22.10). Ceci se visualise sous la forme

$$\left( \begin{array}{c} b \\ \parallel \\ \parallel \\ \parallel \\ \parallel \end{array} \right)_{n \times n} = D(T_a)$$

$\nwarrow$   
 $\sum_c [D(T_a)]_{cb} T_c = \text{image du vecteur } T_b$

Sous cette forme, il suffit donc de vérifier que  $[D(T_a), D(T_b)]_{\alpha\beta} = D([T_a, T_b])_{\alpha\beta}$  pour ce convaincre que l'on a bien une représentation, ce qui se vérifie grâce à l'identité de Jacobi.

*Remarque :*

La dimension de la représentation adjointe est égale à la dimension du groupe. En revanche les générateurs  $T_a$  peuvent être dans une représentation complètement arbitraire. La seule chose qui compte finalement est que leur nombre soit la dimension du groupe (qu'il ne faut surtout pas confondre avec la dimension  $d$  de la représentation des générateurs  $T_a$  (i.e. les  $T_a$  sont des matrices  $d \times d$ )). On obtient ainsi différentes formes isomorphes de la représentation adjointe.

*Exemple :*  $SU(2)$

*représentation fondamentale :* matrices  $\frac{\vec{\sigma}}{2}$  (3 matrices  $2 \times 2$ )

*représentation adjointe :* c'est la représentation vectorielle que l'on peut aussi fabriquer à partir des rotations infinitésimales  $(J_k)_{ij} = -i \varepsilon_{kij} = -i \varepsilon_{ijk}$  (3 matrices  $3 \times 3$ ).

---

En tant que matrice,  $T_a^A$  agit comme une matrice  $n \times n$  ( $n =$  dimension du groupe). En revanche si l'on décide de décrire  $T_a^A$  comme un opérateur construit à partir d'une représentation donnée (en général la représentation fondamentale) de dimension  $d$ , alors

$$T_a^A(T_b) = [T_a, T_b] \quad T_a : \text{pour } a \text{ fixé, c'est une matrice } d \times d$$

$\nwarrow$   
 $n$  indices possibles

Notations usuelles dans la littérature :  $ad \Leftrightarrow T^A \Leftrightarrow T$  pour l'adjointe  
 $T^f \Leftrightarrow t$  pour la fondamentale

Dans la suite nous utiliserons  $T$  pour l'adjointe et  $t$  pour la(les) fondamentale(s). Il est assez usuel d'utiliser les indices  $i, j, k$  pour les coefficients des générateurs de la (des) représentation(s) fondamentale(s), et  $a, b, c$  pour les coefficients des générateurs de la représentation adjointe.

### 22.2.3 Compacité, simplicité et semi-simplicité

#### Compacité

On dit qu'un groupe est compact si le domaine dans lequel vivent les  $n$  paramètres du groupe est compact. Ainsi le groupe des rotations est compact, mais le groupe de Lorentz ne l'est pas, à cause des boosts qui sont codés par des rapidités<sup>2</sup> variant de  $-\infty$  à  $+\infty$ . Notons dans ce dernier cas que l'on pourrait avoir la (fausse) impression qu'un boost pouvant être codé par une vitesse variant de  $-c$  à  $+c$ , cette notion de compacité est toute relative. Il n'en est rien, car c'est bien le codage en terme des rapidités qui permet d'obtenir une structure de groupe de Lie.

#### Simplicité

La notion d'algèbre simple (et semi-simple) est importante en mathématique d'un point de vue de classification. Pour nous, elle joue un rôle important puisqu'elle assurera la positivité de l'énergie associée au lagrangien ainsi construit (voir commentaire page 331).

Pour la définir, on introduit la notion d'idéal : c'est un sous-espace  $\mathcal{I}$  de l'algèbre  $g$  stable par multiplication (au sens du crochet de Lie). Cela signifie que  $[\mathcal{I}, g] \subset \mathcal{I}$ , c'est-à-dire que le commutateur de n'importe quel élément de  $\mathcal{I}$  avec un élément quelconque de  $g$  appartient à  $\mathcal{I}$ . Ceci étant dit, un idéal est abélien si tous les éléments de  $\mathcal{I}$  commutent entre eux.

Nous pouvons maintenant définir la notion d'algèbre de Lie simple : c'est une algèbre qui n'a pas d'autre idéal que  $\{0\}$ . Une algèbre est semi-simple si elle n'a pas d'autre idéal abélien que  $\{0\}$  (ce qui n'interdit pas qu'elle possède d'autre idéal, d'où l'appellation semi-simple). Exemple : l'algèbre de Lie du groupe  $SO(4)$ , que l'on note  $so(4)$ , peut s'écrire sous la forme  $so(4) = so(3) \oplus so(3)$  (ceci se vérifie en partant des générateurs de  $so(4)$  et en les combinant astucieusement. Ceci permet de montrer que  $so(4)$  n'est pas simple, bien que semi-simple (on ne peut isoler de générateur abélien dans  $so(4)$ ).

On montre (Cartan) que toute algèbre de Lie semi-simple est somme directe d'algèbres de Lie simple. Les algèbres de Lie simple sont complètement classifiées (classification due à Cartan).

Dans la composante connexe de l'identité, on peut écrire, le groupe étant compact,  $U(\alpha) = e^{-i\alpha^a T^a}$ . Ceci n'est possible pour un groupe de Lie arbitraire (i.e. qui n'est pas forcément compact) que dans un voisinage de l'identité, alors que la compacité permet d'avoir accès à toute la composante connexe de l'identité, qui est par définition l'ensemble des éléments du groupe que l'on peut obtenir à partir de l'identité en faisant varier continûment les paramètres.

*Pour un groupe compact, on démontre que toute représentation irréductible est de dimension finie et équivalente à une représentation unitaire.* Cette hypothèse étant supposée ici vérifiée, on considèrera donc dans toute la suite que les matrices  $U$  sont unitaires :  $U U^\dagger = U^\dagger U = 1$  et donc  $U^{-1} = U^\dagger$ .

---

2. Rappelons que la rapidité  $\phi$  est reliée aux paramètres d'un boost de Lorentz par les relations  $\text{ch}\phi = \gamma$  et  $\text{sh}\phi = \gamma\beta$ .

## 22.3 Théories de jauge non abéliennes

### 22.3.1 Principe de jauge et couplage matière-champ de jauge

Examinons le terme cinétique  $\bar{\psi}(x)\partial_\mu\psi(x)$ . En utilisant (22.9), on obtient

$$\begin{aligned} \bar{\psi}(x)\partial_\mu\psi(x) \xrightarrow{G} \bar{\psi}'(x)\partial_\mu\psi'(x) &= \bar{\psi}(x)\partial_\mu\psi(x) \\ &+ \underbrace{\bar{\psi}(x)U^{-1}(\alpha) [\partial_\mu U(\alpha)] \psi(x)}_{\text{terme non compensé}}. \end{aligned}$$

Nous allons suivre la même procédure que pour QED :

- on ajoute à  $\partial_\mu$  un champ de jauge
- comme  $U \underset{|\alpha| \ll 1}{\sim} 1 - i\alpha^a T^a$ ,  $\partial_\mu U \sim -i\partial_\mu \alpha^a(x) T^a$

$\alpha^a$  porte un indice de la représentation adjointe ( $N^2 - 1$  indices pour  $SU(N)$ ), donc le champ de jauge doit obligatoirement être choisi dans la représentation **adjointe** (indépendamment du choix de la représentation  $T^a$ , en principe la fondamentale). On pose donc

$$\boxed{D_\mu = \partial_\mu - igT^a A_\mu^a}, \quad (22.12)$$

que l'on peut comparer à  $D_\mu = \partial_\mu + ieA_\mu$  en QED : dans ce cas trivial, le générateur de la représentation adjointe (comme de la représentation fondamentale) est simplement le nombre 1 (on peut difficilement faire plus simple comme matrice  $1 \times 1$  hermitienne!!). Noter que les signes devant les parties faisant intervenir les champs de jauge sont conventionnels.

Nous allons maintenant exiger la loi de transformation

$$\boxed{D_\mu\psi \rightarrow (D_\mu\psi)' = U(\alpha)D_\mu\psi} \quad (22.13)$$

analogue de  $(D_\mu\psi)' = e^{-i\alpha(x)}D_\mu\psi$  pour QED. Alors

$$(\partial_\mu - igT^a A_\mu^a)(U(\alpha)\psi) = U(\alpha)(\partial_\mu - igT^a A_\mu^a)\psi$$

soit

$$[\partial_\mu U(\alpha) - igT^a A_\mu^a U(\alpha)]\psi = -igU(\alpha)T^a A_\mu^a \psi$$

d'où finalement

$$\boxed{T^a A_\mu^a = U(\alpha)T^a A_\mu^a U^{-1}(\alpha) - \frac{i}{g}[\partial_\mu U(\alpha)]U^{-1}(\alpha)}. \quad (22.14)$$

Dans le cas d'une transformation de jauge globale,  $\alpha^a = \text{constante}$ , et cette relation se simplifie en

$$T^a A_\mu^a = U(\alpha)T^a A_\mu^a U^{-1}(\alpha), \quad (22.15)$$

qui n'est autre que la représentation adjointe du groupe  $G$  sur son algèbre de Lie.

Si l'on développe (22.14) autour de l'identité, on obtient, en partant de  $U(\alpha) \sim 1 - i\alpha^a T^a$ ,

$$T^a A_\mu^a = T^a A_\mu^a - i\alpha^b A_\mu^c [T^b, T^c] - \frac{1}{g}T^a \partial_\mu \alpha^a.$$

Or  $[T^b, T^c] = if^{bca}T^a$ . Donc puisque  $f^{abc}$  est complètement antisymétrique, on en déduit que

$$\boxed{A'^a_\mu = A^a_\mu + f^{abc}\alpha^b A^c_\mu - \frac{1}{g}\partial_\mu\alpha^a.} \quad (22.16)$$

Si l'on considère à nouveau le cas particulier d'une transformation de jauge globale  $\alpha^a =$  constante, on obtient

$$A'^a_\mu = A^a_\mu + f^{abc}\alpha^b A^c_\mu$$

qui est bien la loi de transformation de la représentation adjointe de l'algèbre de Lie de  $G$ . C'est bien sûr la version locale de (22.15).

*Remarque :*

On comparera utilement cette écriture avec la définition habituelle d'une représentation adjointe d'une algèbre de Lie : ici on a écrit la représentation sur les coordonnées (les  $A^a_\mu$ ) plutôt que sur les vecteurs de base de l'algèbre de Lie (les  $T^a$ ). C'est bien sûr équivalent (on dit **dual** en mathématiques).

Pour s'en convaincre il suffit d'écrire

$$\begin{aligned} A'^a_\mu &= \underbrace{\left(e^{-i\alpha_c T^c_A}\right)}_{ab} A^b_\mu \\ &= D^A(U) \text{ dans la représentation adjointe} \end{aligned}$$

La relation (22.14) s'écrit alors, pour une transformation de jauge globale,

$$\begin{array}{ccc} \underbrace{T^a (\delta_{ab} - i\alpha_c (T^A_c)_{ab})}_{\text{développement de } D^A(U)} A^b_\mu & = & \underbrace{(1 - i\alpha^b T^b)}_{\text{développement de } D^F(U)} T^a A^a_\mu (1 + i\alpha^b T^b) \\ \text{au voisinage de } U = \mathbb{1} & & \text{au voisinage de } U = \mathbb{1} \end{array}$$

qui donne, en égalant les coefficients de  $A^a_\mu$ ,

$$\underbrace{T^{a'} (T^A_c)_{a'a}}_{\text{image du vecteur de base } T^a \text{ par } T^A_c} = [T^c, T^a] \quad (22.17)$$

qui est bien la représentation adjointe connue. Noter (ce n'est pas une erreur d'écriture) la position des indices : l'indice courant est à **gauche** de la matrice  $T^A_c$  (comme dans tout changement de base).

En écrivant  $[T^c, T^a] = if_{caa'}T^{a'}$  on a bien sûr  $(T^A_c)_{a'a} = -if_{ca'a}$ . Au risque de se répéter, (22.17) peut être vue de deux façons : comme une identité matricielle (membre de gauche), avec une matrice agissant sur l'espace constitué des générateurs de la représentation considérée, ou comme une identité opératorielle (membre de droite), avec cette fois un calcul de commutateurs qui fait donc intervenir des produits de matrices avec des indices dans la représentation considérée :

•  $[T^c, T^a]$  est une matrice dans la représentation des champs de matière (en général c'est la représentation fondamentale, mais ce n'est pas obligatoire) :

$[T^c, T^a]_{ij} = T_{ik}^c T_{kj}^a - T_{ik}^a T_{kj}^c$  où  $i, j$  sont des indices de la représentation des champs de matière ( $i, j = 1 \cdots N$  pour une représentation fondamentale de  $SU(N)$ ).

•  $(T_c^A)_{a'a}$  est un élément de matrice de la représentation adjointe ( $a, a' = 1 \cdots n$  où  $n$  est la dimension du groupe).

Notons que la relation (22.16) peut encore s'écrire

$$\boxed{A_\mu^a = A_\mu^a - \frac{1}{g} (D_\mu \alpha)^a} \quad (22.18)$$

où  $D_\mu$  agit ici dans la *représentation adjointe*, que l'on notera aussi  $(D_A)_\mu$  (la lettre  $A$  signifie adjointe).

*Preuves :*

• c'est immédiat en écrivant l'action de  $(D_A)_\mu$ , qui est une matrice  $n \times n$  ( $n$  étant la dimension du groupe), sur  $\alpha$  (vecteur à  $n$  composantes) :

$$(D_\mu \alpha)^a = \partial_\mu \alpha^a - ig(T_A^c)_{ab} A_\mu^c \alpha_b$$

avec  $(T_A^c)_{ab} = -if_{cab}$ .

• le même résultat peut bien sûr s'obtenir directement en utilisant la définition de  $(D_A)_\mu$ ,

$$(D_A)_\mu^{ab} = \partial_\mu \delta^{ab} - ig(T_A^c)_{ab} A_\mu^c = \partial_\mu \delta^{ab} - gf_{abc} A_\mu^c,$$

ce dernier point de vue faisant ressortir le fait que  $D_\mu$  dépend de la représentation sur laquelle elle agit.

*Notation :*

on écrira

$$\boxed{A_\mu = A_\mu^a T^a}, \quad (22.19)$$

qui est donc un champ appartenant à la représentation adjointe de l'algèbre de Lie de  $G$ . Comme  $A_\mu$  ne porte plus explicitement d'indice de la représentation du groupe, il ne faut pas se laisser abuser par la notation! (ainsi dans le plan euclidien habituel,  $\vec{V} = x\vec{e}_x + y\vec{e}_y$  ne porte pas non plus d'indice, mais on y est plus habitué!).

### 22.3.2 Dynamique du champ de jauge

Il nous reste maintenant à déterminer la dynamique du champ de jauge, c'est-à-dire à déterminer l'équivalent de  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  pour QED. Comme nous l'avons fait pour QED, considérons donc  $(D_\mu D_\nu - D_\nu D_\mu)\psi$  :

$$\begin{aligned} D_\mu D_\nu - D_\nu D_\mu &= (\partial_\mu - igT^a A_\mu^a) (\partial_\nu - igT^b A_\nu^b) - (\mu \leftrightarrow \nu) \\ &= -igT^a \partial_\mu A_\nu^a - igT^a A_\mu^a \partial_\nu - g^2 T^a T^b A_\mu^a A_\nu^b \\ &\quad + igT^a \partial_\nu A_\mu^a + igT^a A_\nu^a \partial_\mu + g^2 T^b T^a A_\mu^a A_\nu^b \\ &= -igT^c (\partial_\mu A_\nu^c - \partial_\nu A_\mu^c) - g^2 if_{abc} A_\mu^a A_\nu^b T^c \end{aligned}$$

qui est donc *un objet se transformant suivant la représentation adjointe*, puisqu'il a la même structure (par rapport au groupe de jauge) qu'un champ tel que  $A^\mu$ . Nous fixerons les coordonnées de cet objet dans la base  $T^a$  en posant  $D_\mu D_\nu - D_\nu D_\mu = -igT^a F_{\mu\nu}^a$ .

Alors

$$\boxed{F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f_{abc} A_\mu^b A_\nu^c} \quad (22.20)$$

On notera la présence d'un terme supplémentaire *non linéaire* par rapport au cas abélien.

Comme  $D_\mu \psi$  et  $\psi$  se transforment de la même façon sous le groupe de jauge, on en déduit que

$$[(D_\mu D_\nu - D_\nu D_\mu) \psi]' = U(\alpha) (D_\mu D_\nu - D_\nu D_\mu) \psi.$$

Or

$$[(D_\mu D_\nu - D_\nu D_\mu) \psi]' = -igT^a F_{\mu\nu}^{\prime a} \psi' = -igT^a F_{\mu\nu}^{\prime a} U(\alpha) \psi,$$

d'où par comparaison

$$\boxed{T^a F_{\mu\nu}^{\prime a} = U(\alpha) T^a F_{\mu\nu}^a U(\alpha)^{-1}} \quad (22.21)$$

qui traduit bien le fait attendu que  $F_{\mu\nu}^a$  se transforme suivant la représentation adjointe. Si l'on développe autour de  $\alpha = 0$ , on aura bien sûr  $F_{\mu\nu}^{\prime a} = F_{\mu\nu}^a + f_{abc} \alpha_b F_{\mu\nu}^c$ , y compris si la transformation de jauge est locale (pas de terme en  $\partial_\mu \alpha^a$  comme pour  $A_\mu^{\prime a}$ ).

Nous pouvons à présent écrire la seule quantité scalaire, invariante de jauge et de dimension 4 disponible pour compléter le lagrangien (à une constante multiplicative près) :

$$\boxed{\text{Tr}(T^a F_{\mu\nu}^a)(T^b F_{\mu\nu}^b)} \text{ est invariant de jauge.} \quad (22.22)$$

*Preuves :*

- directement :

$$\begin{aligned} T^a F_{\mu\nu}^a T^b F_{\mu\nu}^b &\xrightarrow{G} U(\alpha) T^a F_{\mu\nu}^a U(\alpha)^{-1} U(\alpha) T^b F_{\mu\nu}^b U(\alpha)^{-1} \\ &= U(\alpha) T^a F_{\mu\nu}^a T^b F_{\mu\nu}^b U(\alpha)^{-1}. \end{aligned}$$

Par invariance de la trace par permutation cyclique, on en déduit le résultat.

- mathématiquement,  $\text{Tr} T^a T^b$  est une métrique invariante, proportionnelle à la métrique de Cartan-Killing  $\text{Tr} T_a^A T_b^A$ . Donc si l'on examine la loi de transformation de (22.22) de manière duale par rapport au groupe de jauge (i.e. on interprète la transformation de jauge comme un changement de repère affectant les vecteurs de base  $T^a$  de la représentation (point de vue passif) plutôt qu'en terme d'un changement des coordonnées  $F_{\mu\nu}^a$  (point de vue actif), le résultat est immédiat.

---

Dans toute théorie physiquement raisonnable, la positivité de l'énergie doit être satisfaite (non pas au sens strict, le signe absolu de l'énergie n'ayant pas de sens puisque l'énergie est toujours définie à une constante près) : l'*adjonction* d'un mode physique supplémentaire à un état donné doit *augmenter* l'énergie. C'est aussi la raison pour laquelle, dans une théorie scalaire par



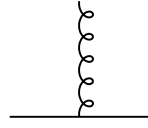
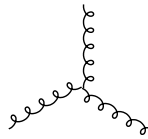
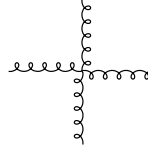
exemple, le terme cinétique est de la forme  $\frac{1}{2} Z \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2$ , où  $Z$  est *strictement positif* (on choisit en général  $Z = 1$  pour la théorie non renormalisée). C'est la raison pour laquelle on exige que  $Tr T^a T^b$  soit une métrique *définie positive*. On démontre alors que ceci est équivalent à choisir pour groupe de jauge un groupe de Lie dont l'algèbre de Lie soit la somme directe de sous-algèbres de Lie simples compactes et de sous algèbres  $U(1)$ .

En outre, ce résultat permet de montrer que  $Tr T^a T^b$  est proportionnelle à  $\delta^{ab}$ .

Nous pouvons maintenant écrire le lagrangien complet de Yang-Mills

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi$$

Développons ce lagrangien pour comprendre sa structure. Pour simplifier, on suppose que l'algèbre de Lie du groupe  $G$  est simple, ce qui évite d'introduire plusieurs types de charge. Ceci s'applique donc directement au cas  $G = SU(N)$  :

$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi$	terme cinétique (quark)	
$-\frac{1}{4} (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) (\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu})$	terme cinétique (gluon)	
$+g A_\mu^a \bar{\psi} \gamma^\mu t^a \psi$	couplage quark-gluon	
$-g f_{abc} \partial_\mu A_\nu^a A_\nu^b A^{c\mu}$	couplage à 3 gluons	
$-\frac{1}{4} g^2 f_{abc} f_{ade} A_\mu^b A_\nu^c A^{d\mu} A^{e\nu}$	couplage à 4 gluons	

(22.23)

*Remarques :*

- le terme cinétique des gluons n'a pas de sens pour le moment, car la composante  $A^0$  ne se propage pas. Nous étudierons la solution de ce problème, lié à l'invariance de jauge, dans le chapitre suivant.

- les bosons de jauge sont de masse nulle
- les interactions ont été fixées complètement par le fait d'imposer l'invariance de jauge locale.

- on retrouve le **même** couplage  $g$  pour la matière et pour les interactions entre bosons de jauge : lorsque le groupe de jauge est simple (comme pour QCD), il n'y a aucune liberté. La contrainte vient de la non-linéarité des relations de commutation de l'algèbre de Lie : ayant fixé une échelle  $g$  pour définir  $D_\mu$ , cette échelle réapparaît dans l'expression de  $F_{\mu\nu}^a$  (lors du



calcul de  $D_\mu D_\nu - D_\nu D_\mu$ , on a explicitement utilisé les constantes de structure), et donc dans le lagrangien. Si le groupe est simple, ce choix implicite de la forme de la transformation de jauge pour les bosons  $A_\mu$ , valable pour tout le groupe, interdit un autre choix pour d'autres particules de matière.

Dans le cas du modèle standard,  $G = SU(2) \times U(1) \times SU(3)$ , qui n'est pas simple. Les couplages sont indépendants suivant chaque sous-groupe.

- même sans matière, la théorie de Yang-Mills n'est pas libre. Par exemple, des états dits "boules de glue" (glueballs en anglais) pourraient exister en QCD, comme états liés purement gluoniques.

- le terme  $-\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu}$  correspond à l'action la plus générale pour la partie de Yang-Mills à condition d'imposer la conservation de la parité  $P$  ou du produit  $CP$ . Sinon, on peut également inclure dans le lagrangien le terme  $-\frac{1}{2}\theta_{ab}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}^a F_{\rho\sigma}^a$ . On peut montrer que ce terme est une dérivée totale, de sorte qu'il ne contribue ni aux équations du mouvement ni aux règles de Feynman. En revanche, en tant que terme topologique (son unique effet vient des termes de bord dans l'intégration de cette dérivée totale, et peut dépendre de la topologie à l'infini), il peut avoir des effets quantiques. QCD est une théorie qui conserve  $P$  et  $CP$  dans les limites expérimentales actuelles, ce qui fournit des contraintes sur l'ordre de grandeur de ce terme supplémentaire, en particulier par la mesure du moment dipolaire du neutron (qui est plus petit que  $10^{-25}e$  cm). Une possibilité de violation de  $P$  et de  $CP$  par la présence d'un tel terme reste cependant possible. Elle peut être décrite à l'aide d'un degré de liberté supplémentaire scalaire, appelé *axion*, qui est l'un des candidats possibles pour la matière noire. Ce mécanisme donnant lieu à l'apparition d'une violation de  $P$  et de  $CP$  en QCD (qui ne se manifesterait qu'à grande énergie) est basé sur un mécanisme de brisure spontanée de symétrie  $U(1)$  : à basse énergie, le champ axionique  $\phi$  est dans un minimum de potentiel qui respecte les symétries  $P$  et de  $CP$ . Un terme supplémentaire provient d'un terme dit d'anomalie (on parle d'anomalie lorsqu'une symétrie vérifiée au niveau classique ne l'est plus au niveau quantique). A grande énergie (de l'ordre d'une échelle de brisure  $M$  très grande par rapport aux énergie actuelles), cette symétrie est spontanément brisée, et combinée au terme d'anomalie, elle fait apparaître une phase dans la matrice de masse (ainsi que des termes de couplage entre axion et hadrons), de façon semblable au mécanisme donnant lieu à des termes de violation de  $CP$  dans le secteur electro-faible à cause de l'existence d'une troisième génération qui rend impossible de choisir la matrice de Cabibbo-Kobayashi-Maskawa purement réelle (pour le secteur électrofaible, un deuxième mécanisme existe, basé sur l'existence de plusieurs doublets de Higgs). Notons que dans le cas du secteur électro-faible, la violation de  $P$  n'est pas liée à l'échelle d'énergie, contrairement à celle de  $CP$ . En revanche, dans le mécanisme envisagé ici pour QCD, une telle violation de  $P$  et de  $CP$  ne peut apparaître qu'à très grande énergie.



# Appendices



# Annexe A

## Abréviations

$\mathbb{S}^n$  : sphère unité de dimension  $n$ , dans l'espace euclidien de dimension  $n + 1$  :

$$\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} / x^2 = 1\}.$$

$E^n$  : boule unité de l'espace euclidien  $\mathbb{R}^n$  de dimension  $n$ .

$$E^n = \{x \in \mathbb{R}^n / x^2 \leq 1\}.$$

$\mathbb{S}^{n-1}$  est donc le bord de  $E^n$ .



# Annexe B

## Quelques éléments de topologie

On donne ici quelques bases de la topologie. Cet appendice est constitué pour l'essentiel d'une revue des définitions et concepts élémentaires, ainsi que des résultats principaux. Seules quelques preuves sont données. On consultera des ouvrages ou cours [10] pour plus de détails.

### B.1 Les concepts de base de la topologie

#### B.1.1 Topologie générale

##### Ouverts, fermés, voisinages

**Définition 2.1 :** *Ouverts*

Soit  $E$  un ensemble. On peut définir sur  $E$  une topologie  $T$  qui est l'ensemble des ouverts de  $E$ . Une topologie  $T$  est constituée de la façon suivante :

- ◇ l'ensemble  $E$  et l'ensemble vide  $\emptyset$  font partie de  $T$
- ◇  $T$  est stable par intersection finie :  $\forall U, V \in T, U \cap V \in T$
- ◇  $T$  est stable par réunion (finie ou infinie) :  
Soit  $I$  un ensemble d'indices. Alors si  $\forall i \in I, U_i \in T, \bigcup_{i \in I} U_i \in T$ .

**Définition 2.2 :** *Base d'ouverts*

Une base d'ouverts d'une topologie est un ensemble d'ouverts tel que tout ouvert de la topologie soit une réunion d'éléments de cet ensemble.

Un espace à base dénombrable est constitué d'une base dénombrable.

**Définition 2.3 :** *Point isolé*

Un point  $x$  d'un espace topologique  $(E, T)$  est isolé si le singleton  $\{x\}$  est un ensemble ouvert.

**Définition 2.4 :** *Espace discret*

Un espace topologique discret est un espace pour lequel tout point est isolé.

*Exemple :*

L'ensemble  $\mathbb{N}$  est discret.

On notera que tout espace  $E$  peut être muni d'une topologie  $T$  discrète. Il suffit pour cela de définir  $T$  comme l'ensemble des parties de  $E$ .

A l'opposé, la topologie grossière est constituée de  $E$  et de  $\emptyset$ .

**Définition 2.5 :** *Fermé*

Une partie de l'espace topologique  $(E, T)$  est dite fermée lorsque son complémentaire dans  $E$  est un ouvert.

**Définition 2.6 :** *Intérieur d'une partie*

L'intérieur  $\overset{\circ}{A}$  d'une partie  $S$  de  $(E, T)$  est l'union de tous les ouverts inclus dans  $S$ . On vérifie facilement que c'est le plus grand (au sens de l'inclusion) ouvert inclus dans  $S$ .

**Définition 2.7 :** *Voisinage*

Si  $E$  est un espace topologique et  $x$  un point de  $E$ , un voisinage  $V$  de  $x$  est un sous-ensemble de  $E$  qui contient un ouvert  $U$  contenant lui-même  $x$ .

De façon équivalente, on dit que  $x$  est dans l'intérieur de  $V$ .

A noter qu'il est possible de définir un espace topologique à partir de la donnée de voisinages plutôt que de la donnée des ouverts.

On note  $\mathcal{V}(x)$  l'ensemble des voisinages de  $x$ .

## Adhérence

**Définition 2.8 :** *Point adhérent*

Soit  $A$  une partie d'un espace topologique  $E$ . Un point  $x$  adhérent à  $A$  est un point tel que tout voisinage de  $x$  rencontre  $A$ . De façon équivalente, tout ouvert contenant  $x$  rencontre  $A$ .

**Définition 2.9 :** *Adhérence*

Soit  $A$  une partie d'un espace topologique  $E$ . L'adhérence de  $A$ , notée  $\bar{A}$ , est l'ensemble des points adhérents à  $A$ .

On a de façon évidente  $A \subset \bar{A}$ .

**Définition 2.10 :** *Adhérence (bis)*

De façon équivalente, l'adhérence  $\bar{A}$  d'une partie  $A$  d'un espace topologique  $E$  est le plus petit fermé contenant  $A$ .



Ceci permet d'en déduire la proposition suivante :

**Proposition 2.11 :**

Une partie  $A$  d'un espace topologique  $E$  est fermée si et seulement si  $\bar{A} = A$ .

**Définition 2.12 :** *Partie dense*

Soit  $E$  un espace topologique. Une partie  $A$  de  $E$  est dense dans  $E$  si  $\bar{A} = E$ .

*Exemples :*

- ◇  $\mathbb{Q}$  est dense dans  $\mathbb{R}$
- ◇ L'ensemble des nombres irrationnels est dense dans  $\mathbb{R}$
- ◇ Sur un intervalle fermé  $[a, b]$ , l'espace des fonctions polynomiales  $\mathbb{C}[X]$  est dense dans l'espace des fonctions continues à valeurs complexes définies sur l'intervalle  $[a, b]$  (pour la norme  $\| \cdot \|_\infty$  définie par  $\|f\|_\infty = \sup\{|f(x)| / x \in [a, b]\}$ )
- ◇  $GL_n(\mathbb{R})$  est dense dans  $M_n(\mathbb{R})$ .

**Espace topologique séparé**

Un espace séparé, encore appelé espace de Hausdorff, est un espace topologique dans lequel deux points distincts quelconques admettent toujours des voisinages disjoints.

C'est par exemple le cas des espaces métriques. En effet, deux points  $x$  et  $y$  quelconques étant donnés, en considérant la distance  $d(x, y)$  entre ces deux points, il suffit de considérer les boules de rayon  $d(x, y)/3$  centrées sur chacun de ces deux points pour conclure.

**B.1.2 Topologie des espaces métriques**

Les concepts précédents prennent tout leur sens dans le cas des espaces métriques, qui nous est familier puisque l'espace euclidien à trois dimensions en fournit un exemple.

Rappelons qu'un espace métrique  $(E, d)$  est un espace  $E$  muni d'une distance  $d$ , qui vérifie les axiomes suivants :

**Définition 2.13 :** *Distance*

$$d : E^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto d(x, y)$$

est une distance si et seulement si

- ◇  $d(x, y) = d(y, x)$
- ◇  $d(x, y) = 0 \Leftrightarrow x = y$

$$\diamond d(x, z) \leq d(x, y) + d(y, z)$$

Un exemple immédiat est fourni par une norme  $N$  sur un espace vectoriel réel, qui induit de manière naturelle une distance  $d(x, y) = N(x - y)$ .

Dans le cas général, une distance  $d$  étant choisie, on peut alors construire explicitement une topologie sur  $E$  de la façon suivante (qui dépend *a priori* de la distance  $d$  choisie).

**Définition 2.14 :** *Ouvert d'un espace métrique*

Un sous-ensemble  $U$  d'un espace métrique  $(E, d)$  est appelé ouvert si,

$$\forall x \in U, \exists \epsilon > 0 / \forall y \in E, d(x, y) < \epsilon \Rightarrow y \in U.$$

De façon équivalente, les ouverts d'un espace métrique sont définis à partir de la notion de boule ouverte :

**Définition 2.15 :** *Boule ouverte*

Une boule ouverte de centre  $x$  et de rayon  $r > 0$  est l'ensemble de points situés à une distance de  $x$  strictement inférieure à  $r$  :

$$B(x, r) = \{y \in E / d(x, y) < r\}.$$

Si  $S$  est un sous-ensemble de  $E$ , le point  $x$  est *intérieur* à  $S$  s'il existe une boule ouverte centrée en  $x$  entièrement contenue dans  $S$ .

On peut alors définir les ouverts  $U$  de  $E$  comme les sous-ensembles de  $E$  tels que tout point de  $x$  soit intérieur à  $U$ .

On notera qu'une boule ouverte est bien un ouvert !

*Exemples :*

- ◇ Dans  $\mathbb{R}$ , les boules ouvertes sont les intervalles ouverts, et les ouverts de  $\mathbb{R}$  sont les réunions dénombrables de boules ouvertes.
- ◇ Dans le plan  $\mathbb{R}^2$ , les boules ouvertes sont les disques ouverts, et les ouverts de  $\mathbb{R}^2$  sont les réunions dénombrables de ces disques ouverts.

Dans un espace métrique, une base d'ouvert est fournie par un ensemble de boules ouvertes. De façon analogue, on définit les

**Définition 2.16 :** *Boule fermée*

Une boule fermée de centre  $x$  et de rayon  $r > 0$  est l'ensemble de points situés à une distance de  $x$  inférieure ou égale à  $r$  :

$$B(x, r) = \{y \in E / d(x, y) \leq r\}.$$

### B.1.3 Topologie des espaces vectoriels normés

**Définition 2.17 :** *Espace vectoriel normé*

Soit  $E$  un  $\mathbb{K}$ -espace vectoriel. On dit qu'il est normé s'il est muni d'une norme  $N : E \rightarrow \mathbb{R}^+$  qui doit vérifier les hypothèses suivantes :

- ◇ séparation :  $\forall x \in E, N(x) = 0 \Rightarrow x = 0$
- ◇ homogénéité :  $\forall \lambda \in \mathbb{K}, \forall x \in E, N(\lambda x) = |\lambda| N(x)$
- ◇  $\forall (x, y) \in E^2, N(x + y) \leq N(x) + N(y)$ .

On notera de façon générale  $\|x\| = N(x)$ .

*Exemples d'espaces vectoriels normés :*

- ◇ les corps  $\mathbb{R}$  et  $\mathbb{C}$ , munis de  $|\cdot|$
- ◇ tout espace vectoriel euclidien, *i.e.* un  $\mathbb{R}$ -espace vectoriel de dimension finie muni d'un produit scalaire  $\langle \cdot, \cdot \rangle$ , avec la norme  $N(x) = \sqrt{\langle x, x \rangle}$
- ◇  $\mathbb{R}^n$  avec l'une des normes suivantes :

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

ou

$$\|x\|_\infty = \max(|x_1|, \dots, |x_n|).$$

A noter que  $\|x\|_p \xrightarrow{p \rightarrow \infty} \|x\|_\infty$ .

- ◇  $\forall a, b \in \mathbb{R}$ , l'espace vectoriel  $E = \mathcal{C}^\infty([a, b])$  des fonctions de classe  $C^\infty$  sur le compact  $[a, b]$  (voir plus loin), muni de l'une des normes

$$\|f\|_\infty = \sup_{x \in [a, b]} |f(x)|$$

ou

$$\|f\|_p = \left( \int_a^b |f(x)|^p \right)^{\frac{1}{p}}.$$

A noter que  $\|f\|_p \xrightarrow{p \rightarrow \infty} \|f\|_\infty$ .

**Définition 2.18 :**

Sur un espace vectoriel  $E$ , deux normes sont équivalentes si et seulement si l'une des deux propriétés équivalentes suivantes est vérifiées :

- ◇ les topologies induites par ces normes sont identiques.
- ◇  $\exists \alpha, \beta \in \mathbb{R}_+^* / \forall x \in E, \alpha N_1(x) \leq N_2(x) \leq \beta N_1(x)$

L'équivalence entre ces deux définitions repose sur le fait que les boules ouvertes des normes  $N_1$  et  $N_2$  peuvent s'inclure l'une dans l'autre.

**Théorème 2.19 :**

La distance  $d$  associée à la norme  $N$  munit  $E$  d'une structure d'espace métrique. L'ensemble des résultats valables pour les espaces métriques sont donc applicable ici.

## B.2 Limite, continuité

### B.2.1 Limite

La notion de limite dans un espace topologique est une généralisation naturelle de la notion de limite pour les fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$  :

**Définition 2.20 :** *Limite*

Soient  $E$  et  $F$  deux espaces topologiques. Soit  $A$  un partie de  $E$ ,  $x_0 \in \bar{A}$ . et  $\ell \in F$ . Soit  $f$  une application de  $A$  dans  $F$ .

Alors  $f(x) \xrightarrow{x \rightarrow x_0} \ell$  si et seulement si

$$\forall V \in \mathcal{V}(\ell), \exists W \in \mathcal{V}(x_0) / \forall y \in W \cap A, f(y) \in V.$$

Dans le cas usuel où  $F$  est séparé, on démontre que la limite, si elle existe, est unique.

*Remarque :*

La définition ci-dessus de la limite diffère de celle historiquement introduite par Weierstrass. Dans la définition historique (que l'on rencontre dans de nombreux cours de mathématiques), on suppose que  $x_0 \notin A$ . Avec cette définition, une fonction peut posséder une limite en  $x_0$  qui diffère de sa valeur en  $x_0$ . Une telle fonction sera alors discontinue en  $x_0$ . Suivant la définition donnée ici, la limite au sens de Weierstrass est appelée limite épointée.

Le concept précédent de limite s'étend immédiatement aux suites.

### B.2.2 Continuité

**Définition 2.21 :** *Continuité en un point*

Soient  $E$  et  $F$  deux espaces topologiques. Soit  $f$  une application de  $E$  dans  $F$ , et  $x_0$  un point de  $E$ . La fonction  $f$  est continue au point  $x_0$  si  $f(x_0)$  est une limite de  $f$  en  $x_0$ .

**Proposition 2.22 :**

Si  $F$  est séparé, alors  $f$  est continue si et seulement si  $f$  a une limite en  $x_0$ .

*Remarque :*

Même remarque que ci-dessus concernant la définition de la limite.

**Proposition 2.23 :**

Une fonction  $f$  est continue en  $x_0$  si et seulement si l'image réciproque de tout voisinage  $W$  de  $f(x_0)$  est un voisinage de  $x_0$ , *i.e.*

$$\forall W \in \mathcal{V}(f(x_0)), f^{-1}(W) \in \mathcal{V}(x_0).$$

**Définition 2.24 :** *Continuité globale*

Soient  $E$  et  $F$  deux espaces topologiques. Une application  $f$  de  $E$  dans  $F$ , est continue si et seulement si  $f$  est continue en tout point de  $E$ .

**Théorème 2.25 :**

Soient  $E$  et  $F$  deux espaces topologiques. Une application  $f$  de  $E$  dans  $F$ , est continue si et seulement si l'une des conditions suivantes est satisfaite :

- ◇ L'image réciproque de tout ouvert de  $F$  est un ouvert de  $E$ .
- ◇ L'image réciproque de tout fermé de  $F$  est un fermé de  $E$ .

**Définitions 2.26 :** *Homéomorphisme*

- ◇ Une application  $f : E \rightarrow F$  est un homéomorphisme (on dit aussi que  $f$  est bicontinue) si et seulement si  $f$  est continue, bijective, et  $f^{-1}$  est continue.
- ◇ Deux espace topologique  $E$  et  $F$  sont homéomorphes si et seulement si il existe un homéomorphisme de  $E$  sur  $F$ .

*Exemples :*

- ◇  $f(x) = 1/x$  est un homéomorphisme de  $]0, 1[$  sur  $]1, +\infty[$ .
- ◇  $f(x) = \arctan(x)$  est un homéomorphisme de  $] -\pi/2, \pi/2[$  sur  $\mathbb{R}$ . A noter que ce dernier homéomorphisme est en, pratique très utile dans les applications numériques (dans le sens réciproque) pour représenter  $\mathbb{R}$  à l'aide d'un intervalle borné.

## B.3 Espace métrique complet

### B.3.1 Suites de Cauchy

**Définition 2.27 :** *Suite de Cauchy*

Une suite  $(x_n)_{n \in \mathbb{N}}$  dans un espace métrique  $(E, d)$  est une suite de Cauchy si

$$\forall \epsilon > 0, \exists N \in \mathbb{N} / \forall p, q \geq N, d(x_p, x_q) < \epsilon.$$

**Propriété 2.28 :**

Dans un espace métrique, toute suite convergente est de Cauchy.

**Propriété 2.29 :**

Dans un espace métrique, toute suite de Cauchy est bornée.

**Propriété 2.30 :**

Une suite de Cauchy possède au plus une valeur d'adhérence. Si elle possède une valeur d'adhérence, alors elle converge.

### B.3.2 Espace complet

**Définition 2.31 :** *Espace métrique complet*

Un espace métrique  $(E, d)$  est complet si toute suite de Cauchy de  $E$  a une limite dans  $E$ .

*Exemple :*

◇  $\mathbb{R}$  et  $\mathbb{C}$  sont complets.

Donnons la preuve pour  $\mathbb{R}$  :

- La preuve utilise le lemme des pics :  
Toute suite à valeurs dans un ensemble totalement ordonné possède une sous-suite monotone
- Considérons une suite de Cauchy dans  $\mathbb{R}$ . Elle est donc bornée. Or de toute suite réelle bornée, on peut extraire une sous-suite convergente.  
En effet, grâce au lemme de pics, de toute suite réelle  $x$  on peut extraire une sous-suite monotone  $y$ . Si  $x$  est bornée, alors  $y$  aussi. Or d'après le théorème de la limite monotone,  $y$  est alors convergente.
- On conclut en utilisant le fait que puisque  $y$  est convergente, elle possède donc une valeur d'adhérence, et donc la suite  $x$  elle-même possède une valeur d'adhérence, et converge donc.

- ◇ l'intervalle ouvert  $]0, 1[$  n'est pas complet, car la suite  $x_n = 1/n$  est de Cauchy, mais converge vers  $0 \notin ]0, 1[$ .
- ◇ l'intervalle fermé  $[0, 1]$  est complet.
- ◇  $\mathbb{Q}$  n'est pas complet, car par exemple  $\sqrt{2}$  est la limite d'une suite de Cauchy de nombre rationnels. Il suffit pour cela de considérer la suite  $x_n$  définie par  $x_0 = 3/2$  et  $x_{n+1} = x_n/2 + 1/x_n$ . La suite  $x_n^2$  est une suite de Cauchy, et de limite 2.

## B.4 Compacité

### B.4.1 Compacité en topologie générale

#### Définition 2.32 : Compacité

Soit  $E$  un espace topologique séparé. Une partie  $A$  de  $E$  est dite compacte si de tout recouvrement ouvert de  $A$ , on peut extraire un recouvrement fini (axiome de Borel-Lebesgue).

*Rappel :*

Un recouvrement d'un ensemble  $E$  est une famille  $(X_i)_{i \in I}$  d'ensembles dont l'union contient  $E$ .

La notion de valeur d'adhérence est un concept plus faible que la notion de limite, qui permet en particulier de formuler la compacité pour un espace métrique.

Une valeur d'adhérence d'une suite d'une espace topologique est un point près duquel une infinité de termes de la suite s'accroissent. Plus précisément :

#### Définition 2.33 : Valeur d'adhérence

Soit  $E$  un espace topologique,  $y \in E$  et  $(u_n)_{n \in \mathbb{N}}$  une suite dans  $E$ . Le point  $y$  est une valeur d'adhérence de la suite  $(u_n)$  si, pour tout voisinage  $V$  de  $y$ , il existe une infinité d'indices  $n$  tels que  $u_n \in V$ .

#### Théorème 2.34 :

Soit  $E$  un espace compact. Alors

- ◇ Toute suite  $(u_n)_{n \in \mathbb{N}}$  à valeur dans  $E$  admet une valeur d'adhérence
- ◇ si  $(u_n)_{n \in \mathbb{N}}$  admet une seule valeur d'adhérence, alors elle converge dans  $E$ .

On en déduit alors immédiatement le corollaire suivant :

#### Corollaire 2.35 :

Tout espace métrique compact est complet.

*Preuve :*

La preuve est immédiate en utilisant le fait que toute suite de Cauchy qui admet une valeur d'adhérence converge.

*Remarque :*

On notera que la réciproque est fautive en général. Ainsi  $\mathbb{R}$  est complet mais n'est pas compact.

**Théorème 2.36 :** *de Bolzano-Weierstrass*

Soit  $E$  un espace métrique.  $E$  est compact si et seulement si toute suite de points de  $E$  admet une valeur d'adhérence dans  $E$  (i.e. admet une sous-suite qui converge dans  $E$ ).

## B.4.2 Compacité locale

**Définition 2.37 :** *Espace localement compact*

Soit  $E$  un espace topologique séparé. Cet espace est localement compact s'il admet des voisinages compacts pour tous ses points.

Ce concept est très utile. En effet, de très nombreux résultats valables pour les espaces compacts peuvent être étendus aux espaces localement compacts.

*Exemples :*

- ◇  $\mathbb{R}$ , et plus généralement  $\mathbb{R}^n$  est localement compact.
- ◇ Les espaces homéomorphes à  $\mathbb{R}^n$  sont localement compacts.

*Exemples :*

- $\mathbb{C}^n$
- les intervalles du type  $]a, b[$  (avec  $a < b \in \mathbb{R}$ )
- le disque ouvert  $\{z \in \mathbb{C} / |z| < 1\}$  du plan complexe
- les variétés topologiques
- les espace discrets.

**Théorème 2.38 :**

Soient  $E$  un espace topologique séparé et  $A$  une partie de  $E$ .

- ◇ Si  $A$  est compacte, alors  $A$  est fermée dans  $E$ .
- ◇ Si  $E$  est compact et  $A$  est fermée dans  $E$ , alors  $A$  est compacte.
- ◇ Si  $E$  est un espace métrique et si  $A$  est compacte, alors  $A$  est bornée.

**Corollaire 2.39 :**

Soit  $E$  un espace topologique séparé.



- ◇ Une réunion finie de compact est compacte.
- ◇ Une intersection quelconque de compacts est compacte.

*Remarque :*

Un espace topologique fini  $E$  est discret si et seulement s'il est séparé. Alors  $E$  est compact.

### Compacité et continuité

#### Théorème 2.40 :

L'image d'un compact par une application continue à valeur dans un espace séparé est un compact.

Ce résultat est très utile dans la détermination de maxima de fonctions à valeurs réelles.

### B.4.3 Quelques résultats concernant les espaces vectoriels normés

#### Théorème 2.41 : (*Riesz*)

Soit  $E$  un  $\mathbb{R}$ - ou  $\mathbb{C}$ -espace vectoriel. Si la boule unité fermée est compacte, alors  $E$  est de dimension finie.

#### Définition 2.42 : *Espace de Banach*

Un espace vectoriel normé complet porte le nom d'espace de Banach.

#### Théorème 2.43 :

Soit  $E$  un  $\mathbb{R}$ - ou  $\mathbb{C}$ -espace vectoriel de dimensions finie  $n$  (en particulier sur  $\mathbb{R}^n$  ou  $\mathbb{C}^n$ )

- ◇ toutes les normes sont équivalentes
- ◇  $E$  est complet
- ◇ les parties compactes de  $E$  sont les fermés bornés (théorème de Borel-Lebesgue).
- ◇ en particulier, sur  $\mathbb{R}$ , les intervalles fermés bornés (*i.e.* du type  $[a, b]$  avec  $a, b \in \mathbb{R}$ ) sont appelés segments de  $\mathbb{R}$ . Ce sont les seuls intervalles réels compacts. En particulier un singleton est compact.

#### Théorème 2.44 : *du point fixe de Brouwer*

Toute application continue de la boule unité de  $\mathbb{R}^n$  dans elle-même admet au moins un point fixe.

### B.4.4 Compactifié d'Alexandroff

Considérons  $\mathbb{C}$ . Par projection stéréographique (voir p. 38), il est possible de construire un homéomorphisme entre  $\mathbb{C}$  et la sphère  $\mathbb{S}^2$  privé d'un point. Cet exemple est particulièrement éclairant, car si l'on ajoute un point (le pôle  $N$  qui sert à définir la projection stéréographique) à  $\mathbb{S}^2$ , on obtient bien sûr toute la sphère  $\mathbb{S}^2$ , ce qui montre que  $\mathbb{C}$  auquel on ajoute un point (dit « point à l'infini ») devient compact (par homéomorphisme, voir ci-dessous). Cette construction est plus générale, comme nous allons le voir maintenant.

**Définition 2.45 :** *Compactifié d'Alexandroff*

Soit  $E$  un espace localement compact. Soit  $\omega \notin E$ . Considérons l'ensemble  $\tilde{E} = E \cup \{\omega\}$ , que l'on munit de la topologie constituée des ouverts suivants :

- ◇ les ouverts de  $E$
- ◇ les complémentaires de compacts de  $E$  dans  $\tilde{E}$ , *i.e.* les sous-ensembles de la forme  $\{\omega\} \cup K^c$  où  $K^c$  est le complémentaire dans  $E$  d'un compact de  $E$ .

**Théorème 2.46 :**

Soient  $E$  et  $F$  des espaces localement compacts. Alors l'espace  $\tilde{E}$  est compact.

*Preuve :*

Considérons un recouvrement de  $\tilde{E}$  par des ouverts  $\{U_i\}_{i \in I}$ . Il existe donc au moins un indice  $i_0 \in I$  tel que  $\omega \in U_{i_0}$ . Alors  $K = \tilde{E} \setminus U_{i_0}$  est un compact de  $E$  par définition des ouverts de  $\tilde{E}$  (puisque  $U_{i_0}$  n'est pas un ouvert de  $E$ ). D'autre part, comme  $K \subset \tilde{E}$ , les  $U_i \cap K$ ,  $i \in I$ , forment un recouvrement de  $K$ . Par compacité de  $K$  on peut extraire de ce recouvrement un recouvrement fini  $U_{i_1} \cap K, \dots, U_{i_N} \cap K$ . L'ensemble des  $U_{i_0}, U_{i_1}, \dots, U_{i_N}$  constitue alors un recouvrement de  $\tilde{E}$ , ce qui prouve finalement la compacité de  $\tilde{E}$ .

---

## B.5 Connexité

**Définition 2.47 :** *Connexité*

Un espace topologique  $E$  est connexe si on ne peut l'écrire sous la forme  $E = U_1 \cup U_2$  avec  $U_1, U_2$  ouverts et  $U_1 \cap U_2 = \emptyset$ .

Une définition équivalente est fournie par :

**Définition 2.48 :** *Connexité (bis)*

Un espace topologique  $E$  est connexe si les seules parties de  $E$  qui soient à la fois ouvertes et fermées sont  $E$  et  $\emptyset$ .

**Définition 2.49 :** *Connexité par arc*

Un espace topologique  $X$  est connexe par arc si  $\forall x, y \in X$ , il existe une application continue  $f : [0, 1] \rightarrow X$  t.q.  $f(0) = x$  et  $f(1) = y$ .

**Définition 2.50 :** *Connexité locale par arc*

Un espace  $X$  localement connexe par arcs est un espace topologique qui possède une base d'ouverts dont chacun est connexe par arcs.

Tout espace connexe par arcs est connexe, mais la réciproque est fausse. Mais en pratique :

- ◇ Sur  $\mathbb{R}$ , les parties connexes, *i.e.* les intervalles de  $\mathbb{R}$  sont connexes par arcs.
- ◇ Tout espace vectoriel normé ou plus généralement tout espace vectoriel topologique (sur  $\mathbb{R}$ ) est localement connexe par arcs. C'est par exemple le cas de  $\mathbb{R}^n$  et  $\mathbb{C}^n$ .
- ◇ Tout espace connexe qui est localement connexe par arcs est connexe par arcs.

Donc en pratique, pour le physicien théoricien, connexité = connexité par arc !



# Annexe C

## Petite introduction aux variétés

### C.1 Variété topologique

**Définition 3.1 :** *Variété topologique*

Soit  $E$  un espace topologique séparé ayant une base dénombrable. L'espace  $E$  est une variété topologique si pour tout  $x$  de  $E$ , il existe un voisinage ouvert  $U$  de  $x$ , un entier naturel  $n$  et un ouvert  $V$  de  $\mathbb{R}^n$  tels que  $U$  et  $V$  soient homéomorphes.

Ceci correspond simplement à un espace qui ressemble localement à  $\mathbb{R}^n$ .

*Remarques :*

- ◇ La condition d'espace séparé permet d'éviter d'éventuels points doubles (penser par exemple à  $\infty$ , qui n'est pas une variété topologique)
- ◇ En toute généralité, on peut remplacer  $\mathbb{R}^n$  par un espace de Banach, ce qui permet de considérer les variétés de dimension infinie.

**Théorème 3.2 :**

L'entier  $n$  tel que  $E$  est localement homéomorphe à un ouvert de  $\mathbb{R}^n$  est unique. on l'appelle la dimension de la variété topologique  $E$ .

**Définition 3.3 :** *Carte locale*

Soit  $E$  est une variété topologique. Par définition,

$$\forall x \in E, \exists U_x \in \mathcal{V}(x) \text{ et un homéomorphisme } \varphi_x : U_x \longrightarrow \varphi_x(U_x) \subset \mathbb{R}^n .$$

Le couple  $(U_x, \varphi_x)$  est appelée carte locale de  $E$ .

*Exemple :*

La projection stéréographique de la sphère  $\mathbb{S}^2$ , privée d'un point (appelé pôle  $N$ ), sur un plan quelconque, par exemple le plan équatorial, voir fig. 3.14 (tout plan hormis celui passant par le

pôle convient), est une carte locale de  $\mathbb{S}^2$  sur  $\mathbb{R}^2$ , en prenant pour origine du plan le centre de la sphère.

**Définition 3.4 :** *Atlas*

Un atlas est une famille de cartes  $(U_i, \varphi_i)$  qui constitue un recouvrement de  $E$ .

*Exemple :*

Les deux projections stéréographiques de la sphère  $\mathbb{S}^2$ , privée respectivement du pôle  $N$  et du point diamétralement opposé  $S$ , sur  $\mathbb{R}^2$  fournissent un atlas de  $\mathbb{S}^2$ .

## C.2 Variété différentielle

**Définition 3.5 :** *Atlas de classe  $\mathcal{C}^k$*

Un atlas est de classe  $\mathcal{C}^k$  si pour tout couple  $i, j$  tels que  $U_i \cap U_j \neq \emptyset$ , l'application de changement de cartes (d'un ouvert de  $\mathbb{R}^n$  vers un ouvert de  $\mathbb{R}^n$ , d'où le nom)

$$\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \rightarrow \varphi_i(U_i \cap U_j)$$

est un difféomorphisme de classe  $\mathcal{C}^k$  (i.e.  $\varphi_i \circ \varphi_j^{-1}$  et sa réciproque sont de classe  $\mathcal{C}^k$ ).

**Définition 3.6 :** *Atlas compatibles*

Deux atlas de classe  $\mathcal{C}^k$  sont dits compatibles lorsque leur réunion constitue encore un atlas de classe  $\mathcal{C}^k$ .

Cette relation de compatibilité est une relation d'équivalence de façon quasi immédiate (la transitivité demande un peu de soin).

De façon intuitive, plusieurs atlas pouvant décrire la même variété (penser à différents atlas de cartes représentant la surface du globe terrestre), considérons les classes d'équivalence par rapport à cette relation de compatibilité.

**Définition 3.7 :** *Variété différentielle*

Considérons la classe d'équivalence d'un atlas donné, obtenue en combinant tous les atlas compatibles avec cet atlas (et donc entre eux), formant ainsi un *atlas maximal* (au sens de l'inclusion), qui est unique. La donnée d'un tel atlas maximal définit une variété différentielle, de classe  $\mathcal{C}^k$ .

**Définition 3.8 :** *Variété lisse*

Une variété lisse est une variété différentielle de classe  $\mathcal{C}^\infty$ .

**Définition 3.9 :** *Variété analytique et variété complexe*

On obtient une variété analytique en exigeant que les changements de carte soient analytiques.

On obtient une variété complexe en remplaçant  $\mathbb{R}$  par  $\mathbb{C}$  et en exigeant que les changements de carte soient biholomorphes.

*Exemples :*

- ◇ les ouverts de l'espace euclidien  $\mathbb{R}^n$  sont des variétés différentielles de dimension  $n$
- ◇ la sphère  $\mathbb{S}^n$  est une variété différentielle compacte de dimension  $n$
- ◇ le tore de dimension  $n$  est une variété différentielle compacte de dimension  $n$ . En particulier le tore  $\mathbb{T}^2$  est une variété différentielle de dimension 2
- ◇ l'espace projectif  $P^n(\mathbb{R})$  est une variété différentielle compacte de dimension  $n$ .
- ◇ l'espace projectif  $P^n(\mathbb{C})$  est une variété complexe compacte de dimension  $n$ .
- ◇ une surface de Riemann est une variété complexe de dimension 1.

## C.3 Immersion, plongement, submersion

Il est souvent utile de pouvoir visualiser une variété en la « plongeant » dans une variété de dimension supérieure. Ainsi visualise-t-on la sphère  $\mathbb{S}^2$  dans l'espace  $\mathbb{R}^3$ .

### C.3.1 Immersion

**Définition 3.10 :** *Immersion*

Soient  $V$  et  $W$  deux variétés et  $f$  une application différentiable de  $V$  dans  $W$ . On dit que  $f$  est une immersion si  $f$  est une injection, *i.e.* si  $\forall x \in V$ ,  $\text{rang}(Tf(x)) = \dim(V)$ .

### C.3.2 Plongement

**Définition 3.11 :** *Plongement*

Soient  $V$  et  $W$  deux variétés et  $f$  une application différentiable de  $V$  dans  $W$ . On dit que  $f$  est un plongement si  $f$  est une immersion et si  $f$  est un homéomorphisme de  $V$  sur  $f(V)$ .

*Exemple :*

La théorie des nœuds concerne l'étude du plongement du cercle  $\mathbb{S}^1$  dans  $\mathbb{R}^3$ .

### C.3.3 Submersion

**Définition 3.12 :** *Submersion*

Soient  $V$  et  $W$  deux variétés et  $f$  une application différentiable de  $V$  dans  $W$ . On dit que  $f$  est une submersion si  $f$  est surjective, *i.e.* si  $\forall x \in V$ ,  $\text{rang}(Tf(x)) = \dim(W)$ .

*Exemple :*

Les projections  $\pi : \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n \subset \mathbb{R}^{m+n}$ .

## C.4 Connexion affine

idée : on relie les espaces tangents (vus comme des espaces affines) en chaque point de la variété, par une transformation (une connexion affine) qui généralise la translation permettant de passer du plan affine tangent en un point de  $\mathbb{R}^n$  au plan tangent en un autre point de  $\mathbb{R}^n$ .

Pour les variétés munies d'une métrique riemannienne (ou variétés riemanniennes), une connexion est naturelle : la connexion de Levi-Civita.

Invariants d'une connexion affine :

- ◇ courbure
- ◇ torsion

## C.5 Application exponentielle



# Références



# Table des figures

1.1	Les pyramides de Gizeh, 2500 ans avant J-C. Crédits : « All Gizah Pyramids » par Ricardo Liberato — All	
1.2	Les cinq solides de Platon, seuls polyèdres réguliers convexes. De gauche à droite, le tétraèdre, l'oc	
1.3	Mouvement d'une pierre. A gauche : dans le sens du déroulement temporel réel du mouvement. A	
1.4	Deux exemplaires de coquilles polies de Nautilus pompilius, symétriques l'une de l'autre dans un r	
1.5	Un disque proto-planétaire autour de l'étoile HL Tauri, située à 450 années-lumière de la Terre. Cr	
3.1	Angles aux sommets d'un polygone. . . . .	28
3.2	Les étapes de la réduction du graphe d'un parallélépipède. . . . .	30
3.3	Diagramme non planaire d'ordre minimal. . . . .	32
3.4	Le diagramme non planaire d'ordre minimal, où chaque arête est remplacée par une double ligne a	
3.5	Le diagramme non planaire d'ordre minimal, dessiné sur un tore à un trou. . . . .	33
3.6	Un exemple de polyèdre à un trou. . . . .	33
3.7	Diagramme non planaire d'ordre minimal (où chaque arête est remplacée par une double ligne affr	
3.8	Le domaine fondamental du tore $\mathbb{T}^2$ . . . . .	34
3.9	Recollement de deux tores à un trou. . . . .	35
3.10	Le bonnet croisé. . . . .	35
3.11	Deux choix possibles de domaine fondamental pour le bonnet croisé : un disque (à gauche) ou un	
3.12	Le bonnet croisé. . . . .	36
3.13	La projection stéréographique de $P^1(\mathbb{R})$ sur $\mathbb{S}^1$ . . . . .	37
3.14	La projection stéréographique de $P^1(\mathbb{C})$ sur $\mathbb{S}^2$ . . . . .	38
3.15	L'espace projectif $P(E)$ vu comme l'union d'un hyperplan affine $\mathcal{H}$ et de l'espace projectif $P(H)$ c	
3.16	Les quatre polyèdres réguliers non convexes de Kepler-Poinsot. . . . .	41
3.17	Les polytopes en dimensions $d$ pour $0 \geq d \geq 3$ . . . . .	42
3.18	Les $n$ -simplexes pour $0 \geq n \geq 3$ . . . . .	43
4.1	Suite exacte courte. . . . .	60
6.1	réseau parallélogramme . . . . .	90
6.2	Les quatre axes de réflexion dans le cas $D_4$ . . . . .	92
6.3	<b>axes des réflexions</b> . . . . .	92
6.4	. . . . .	94
6.5	Réseau rectangulaire “ $p$ ” . . . . .	95
6.6	Réseau orthorhombique “ $C$ ” (on dessinera aussi $\begin{matrix} \nearrow t_2 \\ \searrow t_1 \end{matrix}$ ) . . . . .	95
9.1	Variation de la trajectoire. . . . .	125

12.1	Visualisation symbolique des variétés correspondant aux groupes $O(3, 1)$ et $O(4)$ . A gauche, hyper	
14.1	Rotation d'angle $\omega$ autour d'un axe orienté suivant le vecteur $\vec{n}$ . . . . .	199
14.2	A gauche : lacet $c_1$ joignant en ligne droite les deux points diamétralement opposés $P$ et $P'$ . A dro	
14.3	La succession de deux lacets $c_1$ , homotope à un lacet constant. . . . .	205
14.4	Angles d'Euler. . . . .	217
14.5	Rotation de l'orbitale $p$ d'un atome. . . . .	223

# Liste des tableaux

3.1	Les différents polyèdres connexes réguliers et leurs paramètres. . . . .	31
3.2	Caractéristique d'Euler-Descartes de quelques objets élémentaires. . . . .	31
3.3	Les différents simplexes et le nombre de leurs <i>i</i> faces. . . . .	44



# Bibliographie





# Bibliographie

- [1] P. Depovere, *La fabuleuse histoire des bâtisseurs de la chimie moderne*. De Boeck Supérieur, 2008.
- [2] M. Berger, *Géométrie. T. 1*. Nathan Paris, 1990.
- [3] M. Audin, *Géométrie (L3M1)*. EDP sciences, 2012.
- [4] W. Bertram, “Groupes de transformations, cours de master 1.”  
<http://www.iecn.u-nancy.fr/~bertram/>, 2013. Accessible le : 3-3-2015.
- [5] C. Renard, “Préparation à l’Agrégation, ENS de Cachan.”  
<http://www.normalesup.org/~crenard/geometrie-projective.pdf>, 2013. Accessible le : 3-3-2015.
- [6] C. Renard, “Préparation à l’Agrégation, ENS de Cachan.”  
<http://www.normalesup.org/~crenard/geometrie-affine.pdf>, 2013. Accessible le : 3-3-2015.
- [7] L. Landau and L. Lifschitz, *Physique théorique, Tome 1 : Mécanique*. 1971.
- [8] S. Weinberg, *The Quantum Theory of Fields : Volume 1, Foundations*. Cambridge University Press, 1995.
- [9] A. O. Barut and R. Raczka, *Theory of group representations and applications*, vol. 2. World Scientific, 1986.
- [10] E. Aubry, “Rappels de topologie pour la licence.”  
<http://math.unice.fr/~eaubry/Enseignement/L3/rappelstopo.pdf>, 2014. Accessible le : 16-3-2015.