

# Variétés et groupes de Lie

Florent Novembre

15 juin 2016

# TIPE

# Table des matières

<b>Introduction</b>	<b>3</b>
<b>Partie I : Groupes</b>	<b>5</b>
<b>1 Groupes topologiques</b>	<b>5</b>
1.1 Définitions et exemples . . . . .	5
1.2 Produit cartésien de groupes . . . . .	7
1.3 Topologie des groupes matriciels . . . . .	8
Ouverts, fermés, bornés . . . . .	8
Densité . . . . .	9
Connexité . . . . .	10
Générateurs du groupe linéaire . . . . .	12
1.4 Morphismes de groupes et homéomorphismes . . . . .	13
<b>2 Groupes quotients</b>	<b>14</b>
2.1 Relations d'équivalence et ensembles quotients . . . . .	14
2.2 Quotient d'un groupe par un sous-groupe . . . . .	15
2.3 Structure de l'ensemble quotient d'un groupe : . . . . .	16
2.4 Théorèmes d'isomorphisme . . . . .	16
2.5 Exemples de groupes quotients . . . . .	18
2.6 Groupes simples * . . . . .	19
2.7 Centre d'un groupe, groupes projectifs . . . . .	21
2.8 Groupe dérivé . . . . .	23
2.9 Groupe de Heisenberg * . . . . .	24
<b>3 Actions de groupe</b>	<b>28</b>
3.1 Définitions et exemples . . . . .	28
3.2 Espaces homogènes, action des groupes matriciels . . . . .	31
3.3 Connexité et actions de groupes . . . . .	32
3.4 Représentations linéaires des groupes . . . . .	33
<b>Partie II : Géométrie différentielle</b>	<b>35</b>
<b>4 Sous variétés et variétés</b>	<b>35</b>
4.1 Sous-variétés . . . . .	35
4.2 Groupes de matrices comme sous-variétés . . . . .	44
4.3 Variété . . . . .	45
4.4 Calcul différentiel sur une variété . . . . .	54
Applications différentiables et difféomorphismes . . . . .	54
Fonctions réelles sur une variété . . . . .	55
Courbes paramétrées sur une variété . . . . .	56
Espaces tangents . . . . .	56
Différentielle d'une application . . . . .	59
Immersion, submersions, plongements sur des variétés . . . . .	60
4.5 Théorème de plongement de Whitney dans le cas compact* . . . . .	61
<b>5 Fibré tangent</b>	<b>63</b>
5.1 Fibré tangent . . . . .	63
5.2 Champs de vecteurs . . . . .	66
Sections . . . . .	66
Vecteurs tangents et dérivations . . . . .	68
Champs de vecteurs . . . . .	69
5.3 Fibrés trivialisables . . . . .	70
5.4 Théorème de la boule chevelue* . . . . .	75

<b>Partie III : Groupes de Lie</b>	<b>82</b>
<b>6 Groupes et Algèbre de Lie</b>	<b>82</b>
6.1 Groupes de Lie et premiers exemples . . . . .	82
6.2 Algèbre de Lie . . . . .	82
6.3 Algèbre de Lie d'un groupe de Lie . . . . .	84
6.4 Champs de vecteurs invariants à gauche . . . . .	86
6.5 Flot . . . . .	89
<b>7 Exponentielle</b>	<b>92</b>
7.1 Définition à partir du flot . . . . .	92
7.2 Exponentielle matricielle : Définition, Convergence, Propriétés . . . . .	94
7.3 Théorème de Cartan* . . . . .	98
7.4 Formule de Baker-Campbell-Hausdorff * . . . . .	104
Lemme 1 . . . . .	104
Lemme 2 . . . . .	104
Formule intégrale . . . . .	105
Formule de Dynkin . . . . .	106
7.5 3ème théorème de Lie* . . . . .	107
Problème . . . . .	107
Démonstration . . . . .	108
<b>8 Groupes matriciels</b>	<b>108</b>
8.1 Algèbre de Lie des principaux groupes de matrice . . . . .	108
8.2 Surjectivité de l'exponentielle matricielle . . . . .	111
Dunford-Jordan . . . . .	112
Matrices diagonales semblables . . . . .	113
Surjectivité de $exp : D_n(\mathbb{C}) \rightarrow D'_n(\mathbb{C})$ . . . . .	113
Surjectivité de $exp : M_n(\mathbb{C}) \rightarrow GL_n(\mathbb{C})$ . . . . .	113
Cas réel et applications . . . . .	115
Surjectivité de $exp : AHerm_n(\mathbb{C}) \rightarrow U_n(\mathbb{C})$ . . . . .	115
Surjectivité de $exp : ASym_n(\mathbb{C}) \rightarrow O_n(\mathbb{R})$ . . . . .	116
Surjectivité de $exp : ASym_n(\mathbb{R}) \rightarrow SO_n(\mathbb{R})$ . . . . .	117
Surjectivité de $exp : \mathfrak{su}_n(\mathbb{C}) \rightarrow SU_n(\mathbb{C})$ et $exp : \mathfrak{sl}_n(\mathbb{C}) \rightarrow SL_n(\mathbb{C})$ . . . . .	117
8.3 Homomorphisme entre $SU_2(\mathbb{C})$ et $SO_3(\mathbb{R})$ * . . . . .	117
8.4 Caractérisation de $SL_2(\mathbb{R})$ * . . . . .	120
<b>Annexes et conclusion</b>	<b>124</b>
<b>9 Annexe I : Groupes de matrices</b>	<b>124</b>
9.1 Groupes élémentaires et représentation matricielle . . . . .	124
9.2 Groupe Linéaire . . . . .	125
9.3 Groupe Orthogonal . . . . .	125
9.4 Groupe Unitaire . . . . .	129
9.5 Groupes Spécial Linéaire, Spécial Orthogonal et Spécial Unitaire . . . . .	130
9.6 Matrices triangulaires supérieures inversibles . . . . .	130
<b>10 Annexe II : Topologie</b>	<b>131</b>
10.1 Espaces topologiques . . . . .	131
10.2 Fonctions sur des espaces topologiques . . . . .	132
10.3 Espaces compacts . . . . .	133
Motivation . . . . .	133
Espaces séparés . . . . .	134
Espaces compacts . . . . .	134
Espaces localement compacts . . . . .	136

Applications propres . . . . .	138
10.4 Partitions de l'unité . . . . .	139
10.5 Topologie dans les espaces métriques . . . . .	141
Espaces de Banach . . . . .	141
Suites extraites . . . . .	142
Compacité dans les espaces métriques . . . . .	142
Ouverts, Fermés, bornés . . . . .	144
Espaces normés, Normes . . . . .	146
10.6 Connexité . . . . .	149
Espaces topologiques connexes . . . . .	149
Application continue sur un connexe . . . . .	150
Quelques critères de connexité . . . . .	151
<b>11 Conclusion :</b>	<b>153</b>
11.1 Bibliographie . . . . .	154
11.2 Index des résultats intéressants prouvés et expliqués dans ce TIPE : . . . . .	154
11.3 Index de thèmes liés à ce TIPE . . . . .	156

# Introduction

Le point central de ce TIPE est la notion de groupe. Les groupes topologiques ou encore les groupes de Lie sont des groupes différentiables (et infinis). Ils sont intéressants surtout car ils interviennent en géométrie, et permettent de classer les symétries, qu'on retrouve partout dans la nature.

Le but de ce TIPE est de démontrer différents théorèmes sur les groupes de Lie, en particulier la formule de Baker-Campbell-Hausdorff et le théorème de Cartan. Je m'intéresse plus aux aspects locaux qu'aux aspects globaux des questions (les outils développés avec les algèbres de Lie permettent cette étude). Ces deux points de vue ne coïncident pas, mais sont néanmoins très liés.

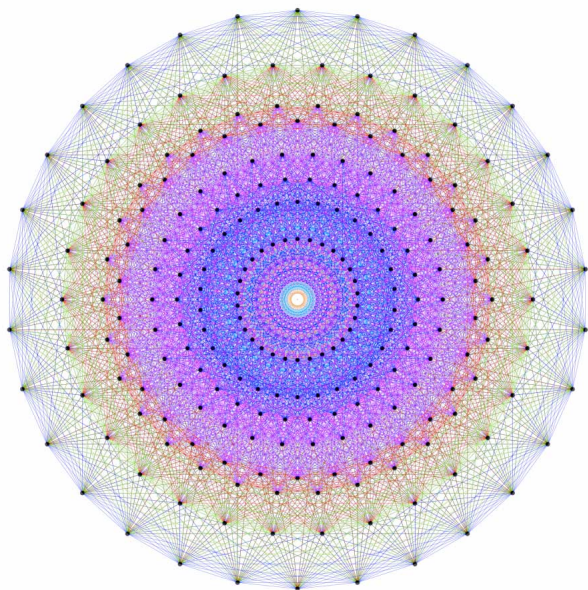
Dans un premier temps, on introduit les notions de groupes topologiques, et on montre différentes propriétés topologiques des groupes matriciels. On définit le quotient d'un groupe par un sous-groupe, dans le but d'avoir un nouveau moyen (en plus du produit cartésien, du produit semi-direct,...) de produire des groupes. Puis on parle des actions de groupes dans le but de parvenir à la théorie des représentations, indispensable lorsqu'on étudie les groupes de Lie.

Dans une seconde partie, on introduit, de manière suffisamment complète pour la suite, des notions de géométrie différentielle : les sous-variétés et les variétés, puis la relation entre les deux notions donnée par le théorème du plongement de Whitney, qu'on démontre dans le cas compact. Cette partie est nécessaire, car un groupe de Lie est une variété (et un groupe, comme l'indique son nom). Enfin, on termine cette partie avec des bases sur les fibrés tangents et les champs de vecteurs. C'est certainement le plus intéressant de cette partie, car vraiment géométrique et visuel, et aussi parce qu'elle permet de démontrer le fameux théorème de la boule chevelue concernant les sphères de dimension paires.

Enfin, la troisième et dernière partie est consacrée aux groupes de Lie. On donne un certain nombre d'exemples en utilisant des groupes tels que  $U(1)$ ,  $U(2)$ ,  $SU(2)$ ,  $SO(3)$  et en s'appuyant sur les deux grandes parties précédentes. Dans le chapitre sur les algèbres de Lie, on introduit l'algèbre de Lie d'un groupe de Lie. Dans le chapitre exponentielle, on prouve les deux résultats qui constituaient les buts de ce TIPE (théorème de Cartan et formule de Baker-Campbell-Hausdorff). Enfin, dans le dernier chapitre, on donne de nombreux exemples intéressants lorsqu'on applique les résultats obtenus en général aux groupes de matrices.

Ainsi, les groupes de Lie réunissent de façon élégante l'Algèbre, l'Analyse, la Géométrie Différentielle et la Topologie.

Exemple de symétries : représentation du groupe de Lie  $E_8$  :



**Notations utilisées :** La plupart du temps :

$G$  désigne un groupe ou un groupe de Lie.

$e$  désigne le neutre du groupe.

$H$  désigne un sous groupe de  $G$

$I$  ou  $I_n$  désigne la matrice identité de taille  $n$ .

$O$  ou  $O_n$  désigne la matrice nulle de taille  $n$ .

$\mathfrak{g}$  désigne l'algèbre de Lie du groupe  $G$ .

$x, y$  désigne les éléments d'un groupe.

$exp$  est l'application exponentielle.

$e^a$  désigne la fonction exponentielle de  $\mathbb{R}$  dans  $\mathbb{R}$  ou de  $\mathbb{C}$  dans  $\mathbb{C}$  appliquée à  $a$

$g$  désigne un élément de  $G$ .

$G/H$  désigne le quotient de 2 groupes.

$A, B$  désignent des matrices.

$N$  désigne une matrice nilpotente.

$\|\cdot\|$  désigne une norme.

$\mathbb{S}^n$  désigne la sphère dans  $\mathbb{R}^{n+1}$  (de dimension  $n$  en tant que variété), c'est à dire : l'ensemble

$$\left\{ (x_1, \dots, x_n, x_{n+1}), \sum_{i=1}^{n+1} x_i^2 = 1 \right\}.$$

$\varphi$  désigne une carte.

$\alpha, \beta, \gamma$  désignent des courbes paramétrés qui dépendent le plus souvent de la variable  $t$ .

$\lambda, \mu$  désignent des scalaires.

Il y a certainement d'autres notations utilisées mais elles sont expliquées. Celles-ci sont les principales, faites pour vous y référer si vous avez un trou de mémoire.

Avant de commencer, j'ai choisi de placer une étoile bleu \* après un théorème ou un résultat pour indiquer qu'il est important (il y en a 5 au total), et une étoile verte après un exemple particulièrement élégant (il y en a 4).

# PARTIE I : Groupes

Le but de cette première partie est d'étudier de manière très générale certains groupes topologiques (dont les groupes matriciels). Pour cela, on utilise des notions topologiques et algébriques

## 1 Groupes topologiques

Le but de ce chapitre est de définir la notion de groupes topologiques et d'étudier les propriétés topologiques des groupes matriciels en particulier, car ce sont tous des groupes topologiques de dimension finie (donc plus faciles à étudier qu'un groupe de dimension infinie).

### 1.1 Définitions et exemples

**1.1.1 Définition:** Un *groupe*  $G$  est un ensemble muni d'une opération binaire  $G \times G \rightarrow G$ ,  $(a, b) \mapsto a \cdot b$  tel que :

1. l'opération  $\cdot$  est associative, c'est à dire :  $(a \cdot b) \cdot c = a \cdot (b \cdot c)$  pour tous  $a, b, c \in G$ ;
2. il existe  $e \in G$  tel que pour tout  $a \in G$  on a  $a \cdot e = e \cdot a = a$ ;
3. pour tout  $a \in G$ , il existe  $b \in G$  tel que  $a \cdot b = b \cdot a = e$ . On dit que  $b$  est l'inverse de  $a$  et on le note  $a^{-1}$ .

Le groupe est dit *abélien* si la loi  $\cdot$  est commutative, c'est à dire  $a \cdot b = b \cdot a$  pour tous  $a, b \in G$ .

**Notations** En général :

1. On note  $+$  la loi d'un groupe abélien, avec  $e = 0$  (le neutre) et d'inverse  $a^{-1} = -a$ .
2. On note  $\cdot$  la loi d'un groupe quelconque, avec  $e = 1$  et d'inverse  $a^{-1}$ .

**1.1.2 Exemple:** *Le  $\star$  est la loi du groupe (l'opération binaire).*

À titre d'exemple de lois quelconques (donc pas de groupes)  $\mathbb{R}^2 \rightarrow \mathbb{R}$  :

$$a \star b = ab^2, a \star b = (a \times b), a \star b = a + 2b$$

À titre d'exemple de lois de groupes :  
 $a \star b = a + b$  dans  $\mathbb{R}$  ou  $\mathbb{Z}$ .

**1.1.3 Exemple:** *Groupes :*

1.  $(\mathbb{R}, +)$  et  $(\mathbb{R}^*, \cdot)$  sont des groupes abéliens.
2.  $G = \mathbb{R} \setminus \{-1\}$  avec  $a \star b = a + b + ab (= b \star a)$  est un groupe abélien. En effet,

$$(a \star b) \star c = (a + b + ab) \star c = (a + b + ab + c) + (a + b + ab)c = a + b + c + ab + bc + ac + abc = a \star (b \star c)$$

(on le voit directement à cause de la symétrie qu'il y a dans la relation). La loi est donc associative. Cherchons le neutre :

$$a \star e = a \forall a \in G \Leftrightarrow a + e + ae = a \Leftrightarrow e(a + 1) = 0 \Leftrightarrow e = 0$$

Enfin,

$$a \star b = 0 \Leftrightarrow a + b + ab = 0 \Leftrightarrow b = -\frac{a}{a+1}$$

Tant que  $a \neq -1$  (c'est à dire,  $a \in G$ ), l'inverse existe, et donc  $G$  est un groupe.

3. Si  $G = \mathbb{R}$ ,  $a \star b = a + 2b$  ne vérifie pas l'associativité, aucun ensemble muni de cette loi ne peut être un groupe.

**1.1.4 Définition:** On dit que  $H$  est un *sous-groupe* de  $G$  si  $H$  est un ensemble non vide et qu'il est stable par produit

Un moyen équivalent de le définir est :

1. Le neutre de  $G$  est dans  $H$ .
2. Pour tout  $x, y$  dans  $H$ , on a  $x \cdot y^{-1} \in H$ .

**1.1.5 Définition:** Un *groupe topologique* est un groupe muni d'une topologie compatible avec la structure de groupes, c'est à dire un groupe  $(G, \star)$  pour lequel les applications :

$$\begin{aligned} G \times G &\longrightarrow G, (x, y) \longmapsto x \star y; \\ G &\longrightarrow G, x \longmapsto x^{-1} \end{aligned}$$

sont continues.

**1.1.6 Exemple:** L'ensemble des *difféomorphismes* de  $\mathbb{R}$  (noté  $Diff(\mathbb{R})$ ) est un *groupe topologique*. De plus c'est un *groupe de dimension infinie*.

**Preuve:** Soit  $f, g, h$  des difféomorphismes de  $\mathbb{R}$ . Comme on est dans  $\mathbb{R}$ , une condition équivalente (pour qu'un fonction  $\varphi$  soit un difféomorphisme) est  $\varphi'(x) \neq 0, \forall x \in \mathbb{R}$ . Ainsi :

1.  $f \circ g$  est un difféomorphisme car  $(f \circ g)'(x) = f'(g(x)) \cdot g'(x) \neq 0, \forall x$  car  $g'(x) \neq 0 \forall x \in \mathbb{R}$  et  $f'(g(x)) \neq 0$  car  $f'(y) \neq 0, \forall y \in \mathbb{R}$ .
2.  $f \circ (g \circ h)(x) = f(g(h(x))) = (f \circ g) \circ h(x), \forall x \in \mathbb{R} \Leftrightarrow f \circ (g \circ h) = (f \circ g) \circ h$
3.  $f \circ e = f \Leftrightarrow f(e(x)) = f(x), \forall x, \forall f \in Diff(\mathbb{R})$ . En particulier, si  $f(x) = x, f(e(x)) = e(x) = f(x) = x, \forall x \in \mathbb{R}$ . Réciproquement, si  $e(x) = x$ , alors  $f(x) = f(x), \forall x$ . Le neutre est donc  $e = Id$  (l'application identité).
4.  $f \circ g = Id \Leftrightarrow g = f^{-1}$ , qui existe car  $f$  est un difféomorphisme, donc bijective.

□

**1.1.7 Définition:** Soit  $V$  un espace vectoriel de dimension  $n$ . On peut écrire  $v = \sum_{i=1}^n a_i e_i$ , avec  $a_i \in \mathbb{R}$  et  $(e_1, \dots, e_n)$  une base de  $V$ . Une fonction  $f : V \longrightarrow \mathbb{R}$  est *polynomiale* s'il existe  $P \in \mathbb{R}[X_1, \dots, X_n]$  tel que  $f(v) = P(a_1, a_2, \dots, a_n)$ .

**1.1.8 Définition:** On peut alors étendre la définition si la fonction  $f$  est définie de  $V$  dans  $W$  (2 espaces vectoriels), en disant qu'il existe  $P_1, \dots, P_q \in \mathbb{R}[X_1, \dots, X_n]$  ( $q$  est alors la dimension de  $W$ ) tels que  $f(v) = \sum_{i=1}^q P_i(a_1, \dots, a_n) w_i$ ,  $(w_1, \dots, w_q)$  étant une base de  $W$ .

On peut maintenant démontrer la proposition suivante :

**1.1.9 Proposition:** *Tous les groupes matriciels  $G$  (multiplicatifs ou additifs, et pour des matrices carrées) dont les éléments sont dans les corps  $\mathbb{R}$  ou  $\mathbb{C}$  sont des groupes topologiques.*

**Preuve:** On sait que  $M_n(\mathbb{R})$  est un  $\mathbb{R}$ -espace vectoriel de dimension  $n^2$ , et  $M_n(\mathbb{C})$  est un  $\mathbb{C}$ -espace vectoriel de dimension  $n^2$ , et comme  $\mathbb{C} \cong \mathbb{R}^2$ , c'est un  $\mathbb{R}$ -espace vectoriel de dimension  $2n^2$ .

Soit :

$$f : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \longrightarrow M_n(\mathbb{K}), (A, B) \longmapsto A + B$$

et

$$g : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \longrightarrow M_n(\mathbb{K}), (A, B) \longmapsto A \cdot B$$

Ces fonctions sont polynomiales de par la définition de l'addition et de la multiplication matricielle, donc continues.

De même, l'inverse d'une matrice (pour l'addition ou pour la multiplication) est une fonction vectorielle de  $G$  dans  $G$ . On considère :

$$f : G \longrightarrow G, A \longmapsto -A$$



et

$$g : G \longrightarrow G, A \longmapsto A^{-1} = \frac{{}^t \text{com}(A)}{\det(A)}$$

Il est évident que  $f$  est polynomiale donc continue.  $g$  n'est pas polynomiale, mais rationnelle (elle l'est en revenant à la définition de la comatrice et du déterminant).  $g$  est donc bien définie car, comme on est dans une groupe matriciel, le déterminant est toujours non nul, auquel cas il n'y aurait pas d'inverse, ce qui est impossible dans un groupe. Elle est continue car une fonction rationnelle est continue tant que son dénominateur est non nul.  $\square$

**1.1.10 Exemple:** *Le cercle  $\mathbb{S}^1$  est un groupe topologique (et même, un groupe matriciel).*

**Preuve:** Soit  $A \in U(1)$ . On a alors  $A = (a)$ , avec  $a \in \mathbb{C}$ . Donc  $AA^* = (a)(\bar{a}) = (|a|^2) = 1$ . On a donc  $|a| = 1$ , ce qui signifie qu'on peut écrire  $a$  sous la forme  $a = e^{it}$ ,  $t \in [0; 2\pi[$ , c'est à dire  $a \in \mathbb{S}^1$ . On peut donc représenter  $\mathbb{S}^1$  sous forme matricielle. Le produit de matrices devient  $e^{it} \cdot e^{is} = e^{i(t+s)}$ , et il est clairement continue.  $\square$

**1.1.11 Exemple:** *La sphère  $\mathbb{S}^3$  est un groupe topologique (et même, un groupe matriciel).*

**Preuve:** On va montrer que  $SU(2)$  (le groupe matriciel) est homéomorphe à  $\mathbb{S}^3$  Soit  $B \in SU(2) : B =$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad BB^* = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = \begin{pmatrix} |a|^2 + |b|^2 & a\bar{c} + b\bar{d} \\ \bar{c}a + \bar{d}b & |c|^2 + |d|^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

De plus,  $\det(B) = ad - bc = 1$ . On a donc 4 équations (2 sont équivalentes, celles sur les coefficients non diagonaux) sur 4 coefficients dans  $\mathbb{C}$ . Soit (1) l'équation :  $|c|^2 + |d|^2 = 1$  et (2) l'équation :  $a\bar{c} + b\bar{d} = 0 \Rightarrow a\bar{c} = -b\bar{d}$  On utilise la dernière équation, sur le déterminant, et on multiplie par  $\bar{c}$ .

$$\bar{c} = ad\bar{c} - bc\bar{c} = -b|d|^2 - b|c|^2 (1) = -b(|d|^2 + |c|^2) = -b(2)$$

Donc  $\bar{c} = -b \Leftrightarrow b = -\bar{c}$ ; Un raisonnement analogue permet de montrer, que  $d = \bar{a} \Leftrightarrow a = \bar{d}$ . Finalement, le groupe  $SU(2)$  représente l'ensemble des matrices de la forme :  $= \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$

Avec  $|a|^2 + |b|^2 = 1$ . Donc, avec la partie réelle et imaginaire,  $a = x + iy$ ,  $b = z + it$ . avec  $x^2 + y^2 + z^2 + t^2 = 1$ . On peut donc représenter  $\mathbb{S}^3$  sous comme groupe matriciel. De plus,  $SU(2)$  est isomorphe à  $\mathbb{S}^3$ , l'isomorphisme est :

$$f : U(2) \longrightarrow \mathbb{R}^4, A \longmapsto (x, y, z, t), \quad \text{avec } x^2 + y^2 + z^2 + t^2 = 1.$$

Enfin, il reste à vérifier que c'est un groupe de matrice :

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} c & d \\ -\bar{c} & \bar{d} \end{pmatrix} = \begin{pmatrix} ac - \bar{b}d & bc + \bar{a}d \\ -bc - \bar{a}d & -bd + \bar{a}c \end{pmatrix}$$

On reconnait clairement la même forme qu'une matrice de  $SU(2)$  :  $ac - \bar{b}d = -bd + \bar{a}c, bc + \bar{a}d = -bc - \bar{a}d$  et le déterminant (moins facile à voir) est :

$$\begin{aligned} (ac - \bar{b}d)(-bd + \bar{a}c) - (bc + \bar{a}d)(-bc - \bar{a}d) &= -abcd - \bar{a}bcd + |ac|^2 + |bd|^2 + abcd + \bar{a}bcd + |ad|^2 + |bc|^2 \\ &= |a|^2 (|c|^2 + |d|^2) + |b|^2 (|c|^2 + |d|^2) = |a|^2 + |b|^2 = 1 \end{aligned}$$

$\square$

## 1.2 Produit cartésien de groupes

**1.2.1 Définition:** Le produit cartésien de 2 groupes  $G$  et  $H$  est l'ensemble  $\{(g, h), g \in G, h \in H\}$ . Soit  $p = \dim G$  et  $q = \dim H$ .

La proposition suivante est très utile, car elle assure que si l'on peut représenter plusieurs groupes comme groupes matriciels, le produit cartésien de ces groupes sera également un groupe matriciel.

**1.2.2 Proposition:** *Le produit cartésien de 2 groupes matriciels est un groupe matriciel.*

**Preuve:** Soit  $G$  et  $H$  2 groupes matriciels. On note  $G \times H$  ce produit. Soit  $A, A' \in G$  et  $B, B' \in H$ . Soit

$$C = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} A' & 0 \\ 0 & B' \end{pmatrix} = \begin{pmatrix} AA' & 0 \\ 0 & BB' \end{pmatrix} \in G \times H$$

, de par le produit de matrices par blocs. L'associativité et l'existence du neutre sont évidents. Enfin :  $A^{-1}$

$$\text{et } B^{-1} \text{ existent et sont dans } H \text{ et } G \text{ respectivement. } \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} = \begin{pmatrix} AA^{-1} & 0 \\ 0 & BB^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} I_p & 0 \\ 0 & I_q \end{pmatrix} = I_{p+q}$$

Donc, tout produit cartésien de 2 groupes matriciels est également un groupe matriciel. Par récurrence, on généralise cela pour le produit cartésien de 3, 4, ...,  $n$  groupes. Par exemple :  $\mathbb{R}^3 \times SO_5(\mathbb{R}) \times Sym_4(\mathbb{R})$  est un groupe matriciel.  $\square$

### 1.3 Topologie des groupes matriciels

Avant de lire cette section, vous pouvez, vous référer à l'annexe II sur la topologie, où des rappels nécessaires sont introduits. En particulier, si les notions d'ouvert, fermés, bornés et connexes vous sont déjà familières, vous pouvez les sauter. Cependant, je conseille à tout le monde de lire au moins la partie **Espaces compacts**, pas pour cette partie, mais avant de commencer le chapitre 2.

Pour la suite, on utilisera deux normes matricielles principalement :

$$\| \| A \| \| = \| \| A \| \|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n | a_{ij} | \quad (1).$$

$$\| A \| = \| A \|_{\infty} = \max_{1 \leq i, j \leq n} | a_{ij} | \quad (2).$$

C'est toujours l'une des 2 qui est utilisée, en général, la seconde, quelquefois la deuxième. Dans tous les cas, les 2 barres sont utilisées pour désigner (2) et les 3 barres pour (1).

Attention cependant, il y a des fois où je me sers de  $\| \cdot \|$  pour la norme euclidienne d'un vecteur.

#### Ouverts, fermés, bornés

Soit :  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$

**1.3.1 Proposition:** *La fonction*

$$\det : M_n(\mathbb{K}) \longrightarrow \mathbb{K}; \det : M \longmapsto \det(M)$$

*est continue.*

**Preuve:** Pour tout  $M = (m_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{K})$ , le déterminant de  $M$

$$\det(M) = \sum_{\sigma \in S_n} \text{signe}(\sigma) \prod_{i=1}^n m_{i, \sigma(i)}. \text{ s'exprime comme un polynôme dans les coefficients de } M. \quad \square$$

**1.3.2 Proposition:** *La fonction  $f$  :*

$$f : M_n(\mathbb{K}) \longrightarrow M_n(\mathbb{K}); f : M \longmapsto M^t M$$

*est continue.*

**Preuve:** Soit  $f(M) = C = M^t M$ .  $c_{i,j} = \sum_{k=1}^n m_{i,k} m_{j,k}$  est une fonction polynomiale à  $2n$  variables. Donc la fonction  $f$  est une fonctions vectorielle de  $\mathbb{K}^{n^2}$  dans  $\mathbb{K}^{n^2}$ , dont les composantes  $c_{ij}$  sont des polynômes la fonction  $f$  est donc continue.  $\square$

**1.3.3 Proposition:** *La fonction  $g$  :*

$$g : M_n(\mathbb{C}) \longrightarrow M_n(\mathbb{C}); g : M \longmapsto M^*M$$

*est continue.*

**Preuve:** Analogue à la précédente. □

**1.3.4 Proposition:**  *$GL_n(\mathbb{K})$  est ouvert, non borné.*

**Preuve:** Pour montrer que  $GL_n(\mathbb{K})$  est ouvert, on utilise la définition de fonction continue (la pré-image d'un ensemble ouvert par une fonction continue est un ouvert).

$GL_n(\mathbb{K}) = \{A \in M_n(\mathbb{K}), \det(A) \neq 0\} = \det^{-1}(\mathbb{K}^*)$ . Or,  $\mathbb{K}^*$  est un ouvert, qu'on ait  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ . Donc  $GL_n(\mathbb{K})$  est ouvert. Il est facile de voir qu'il est non borné : pour tout  $\alpha$  dans  $\mathbb{K}$   $\alpha I_n \in M_n(\mathbb{K})$ , en particulier, si  $\alpha \rightarrow \infty$ ,  $\|\alpha I_n\| = \alpha \rightarrow \infty$ . □

**1.3.5 Proposition:**  *$O_n(\mathbb{R})$  et  $U_n(\mathbb{C})$  sont fermés et bornés.*

**Preuve:** On reprend les fonctions  $f$  et  $g$  définies ci-dessus : pour montrer que ces 2 ensembles sont fermés, on va utiliser la définition de fonction continue (la pré-image d'un ensemble fermé par une fonction continue est un fermé). On rappelle que le sigleton  $\{I_n\}$  est fermé dans  $M_n(\mathbb{K})$ .  $O_n(\mathbb{R}) = f^{-1}(I_n)$  et  $U_n(\mathbb{C}) = g^{-1}(I_n)$ . De plus, on rappelle que les vecteurs colonnes  $v$  d'une matrice unitaire ou orthogonale vérifient :  $\|v\| = 1$ , on en déduit que si  $A$  est orthogonale ou unitaire,  $\|A\| \leq 1$ , donc ces 2 ensembles sont bornés. □

**1.3.6 Proposition:**  *$SO_n(\mathbb{R})$  et  $SU_n(\mathbb{C})$  sont fermés et bornés.*

**Preuve:**  $SO_n(\mathbb{R}) \subset O_n(\mathbb{R})$  et  $SU_n(\mathbb{C}) \subset U_n(\mathbb{C})$ . Donc ces ensembles sont bornés. De plus,  $SO_n(\mathbb{R}) = \{A \in O_n(\mathbb{R}), \det(A) = 1\}$  et  $SU_n(\mathbb{C}) = \{A \in U_n(\mathbb{C}), \det(A) = 1\}$ , ce sont donc des ensembles fermés. □

**1.3.7 Proposition:**  *$SL_n(\mathbb{K})$  est fermé et non borné.*

**Preuve:**  $SL_n(\mathbb{K}) = \det^{-1}(1)$  donc  $SL_n(\mathbb{K})$  est fermé. Pour montrer qu'il n'est pas borné, on va montrer qu'il existe une suite d'éléments de  $SL_n(\mathbb{K})$  qui est divergente (de norme qui tend vers  $+\infty$ ). Soit  $A_p =$

$$\begin{pmatrix} p & \cdots & 0 \\ 0 & \frac{1}{p} & 0 \\ 0 & \cdots & I_{n-2} \end{pmatrix}$$

$A_p \in SL_n(\mathbb{K})$  et :  $\|A_p\| = p \Rightarrow \lim_{p \rightarrow \infty} A_p = \lim_{p \rightarrow \infty} p = \infty$ . Donc  $SL_n(\mathbb{K})$  n'est pas borné. □

### Densité

**1.3.8 Definition:** Une partie  $X$  de  $E$  est dite *dense* dans  $E$  si  $Adh(X) = E$ .

**1.3.9 Proposition:**  *$GL_n(\mathbb{K})$  est dense dans  $M_n(\mathbb{K})$*

**Preuve:** Soit  $A \in M_n(\mathbb{K})$ . Si  $A \in GL_n(\mathbb{K})$ , il n'y a rien à faire (la suite constante égale à  $A$  est inversible et converge vers  $A$ ). Sinon on pose :  $A_p = A + \frac{1}{p}I_n$ .  $\exists n$  tel que  $\forall p > n$ ,  $A_p \in GL_n(\mathbb{K})$ . En effet, dans le cas contraire, il existerait une infinité de valeurs de  $p$  telles que  $\det(A_p) = 0 \Leftrightarrow \ker(A - (-\frac{1}{p}I_n)) \neq 0$ , autrement dit,  $A$  posséderait une infinité de valeurs propres (de forme  $-\frac{1}{p}$ ) ce qui est impossible. Par l'absurde, on déduit le résultat annoncé. Enfin :  $\lim_{p \rightarrow \infty} A_p = A$ . On vient de prouver que pour toute matrice  $A$  de  $M_n(\mathbb{K})$ , il existe une suite de  $GL_n(\mathbb{K})$  (en considérant la suite  $(A_p)_{p > n}$ ) telle que cette suite converge vers  $A$ , autrement dit, l'adhérence de  $GL_n(\mathbb{K})$  est  $M_n(\mathbb{K})$ . □

Soit  $D_n(\mathbb{C})$  l'ensemble des matrices diagonalisables de  $M_n(\mathbb{C})$ .

**1.3.10 Proposition:**  $D_n(\mathbb{C})$  est dense dans  $M_n(\mathbb{C})$ .

**Preuve:** Toute matrice est trigonalisable dans  $M_n(\mathbb{C})$ . De plus : les valeurs propres d'une matrice triangulaire sont les coefficients sur la diagonale. Enfin : un théorème stipule que si une matrice a des valeurs propres toutes distinctes, alors elle est diagonalisable. Ces 3 rappels en tête : Soit  $A \in M_n(\mathbb{C})$ .  $\exists P \in GL_n(\mathbb{C})$  telle que  $A = PTP^{-1}$  avec  $T$  triangulaire supérieure : soit  $A$  est diagonalisable, de sorte qu'on peut prendre  $T = D$  avec  $D$  diagonale. Sinon, soit  $(\lambda_k)_{1 \leq k \leq n}$  l'ensemble des valeurs propres de  $A$ . Puisque  $A$  n'est pas diagonalisable, au moins 2 valeurs propres sont égales. On utilise la notation  $(\lambda_{i,k})$  où  $k$  permet de compter le nombre de valeurs propres distinctes et  $i$  permet de numéroter les valeurs propres égales. Autrement dit :  $\lambda_{i,k} = \lambda_{j,k} \forall i, j \in \mathbb{N}$ . On pose  $\alpha_{i,k,n} = \lambda_{i,k} + \frac{i}{n}$ . ( $\alpha_{i,k,n}$  est une suite). On a :  $\lim_{n \rightarrow \infty} \alpha_{i,k,n} = \lambda_{i,k}$ . De plus,  $\alpha_{i,k,n} \neq \alpha_{j,k,n} \forall i \neq j$ . Enfin,  $\exists n_0$  tel que  $\alpha_{i,k,n} \neq \alpha_{j,p,n} \forall p \neq k \forall n > n_0$  : en effet,  $p \neq k \Rightarrow \alpha_{i,k,n}$  et  $\alpha_{j,p,n}$  n'ont pas la même limite, donc si on suppose :  $\lambda_{i,k} < \lambda_{j,p}$ , pour  $n$  assez grand, on aura :  $\alpha_{i,k,n} < \lambda_{j,p} < \alpha_{j,p,n}$  (par la définition de  $\alpha_{i,k,n}$  et  $\alpha_{j,p,n}$ ). Pour  $n$  suffisamment grand, on obtient l'inégalité  $\forall k \neq p$ . Pour finir, soit  $B_n = PT'_n P^{-1}$  où  $T'_n$  est égale à la matrice  $T$  dont les coefficients diagonaux  $\lambda_{i,k}$  sont remplacés par les  $\alpha_{i,k,n}$ .  $\forall n \in \mathbb{N}$ , les valeurs propres de  $T'_n$  sont distinctes, donc  $T'_n$  est diagonalisable, donc  $B_n$  est diagonalisable.  $\lim_{n \rightarrow \infty} B_n = A$ . L'adhérence de  $D_n(\mathbb{C})$  est donc  $M_n(\mathbb{C})$ .  $\square$

### Connexité

Avant de commencer, je vais donner la définition de connexité, et je renvoie à l'annexe II, partie connexité, pour plus de détails :

**1.3.11 Definition:** On dira que l'espace  $(X, O)$  est *connexe* s'il vérifie l'une des conditions équivalentes suivantes.

1. Si  $X$  est réunion de deux ouverts disjoints alors l'un de ces deux ouverts est vide et l'autre est égale à  $X$ .
2. Si  $X$  est réunion de deux fermés disjoints alors l'un de ces deux fermés est vide et l'autre est égale à  $X$ ;
3. Si on considère  $\{0, 1\}$  muni de la topologie discrète et  $f : X \rightarrow \{0, 1\}$  une application continue, alors  $f$  est constante sur  $X$ .
4. Les seuls ensembles à la fois ouverts et fermés de  $X$  sont  $X$  lui même et l'ensemble vide.

**1.3.12 Proposition:**  $GL_n(\mathbb{R})$  n'est pas connexe.

**Preuve:**  $GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}), \det(A) > 0\} \cup \{A \in M_n(\mathbb{R}), \det(A) < 0\}$  c'est donc l'intersection de 2 ouverts disjoints, par conséquent, il n'est pas connexe.  $\square$

**1.3.13 Proposition:**  $O_n(\mathbb{R})$  n'est pas connexe.

**Preuve:**  $O_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}), \det(A) = 1\} \cup \{A \in M_n(\mathbb{R}), \det(A) = -1\} = O_n^+(\mathbb{R}) \cup O_n^-(\mathbb{R})$  c'est donc l'intersection de 2 fermés disjoints, par conséquent, il n'est pas connexe.  $\square$

Maintenant, je vais démontrer la connexité des certains ensembles, et **je passe à chaque fois par la connexité par arcs**. Je vais donc définir ce qu'est la connexité par arcs, et je renvoie à l'annexe II sur la connexité pour plus de détails, avec des preuves.

**1.3.14 Definition:** Soient  $x$  et  $y$  deux éléments de  $X$ . On appelle *chemin d'extrémités  $x$  et  $y$*  (ou chemin joignant  $x$  et  $y$ ) de  $X$  toute application continue  $c : [0, 1] \rightarrow X$  telle que  $c(0) = x$  et  $c(1) = y$ .

**1.3.15 Definition:** On dira que  $(X, O)$  est *connexe par arcs* si tout couple d'éléments de  $X$  peut être joint par un chemin.

**1.3.16 Definition:** Le *chemin inverse* d'un chemin  $\gamma$  dans un ensemble  $E$  est le chemin  $\bar{\gamma}(t) = \gamma(1-t)$  de sorte que  $\bar{\gamma}(0) = \gamma(1)$  et  $\bar{\gamma}(1) = \gamma(0)$ .

**1.3.17 Lemme:** Soit  $G$  un groupe. Si pour tout  $x \in G$ , il existe un chemin  $\gamma : [0, 1] \rightarrow G$  continu tel que  $\gamma(0) = x$  et  $\gamma(1) = e$ ,  $e$  étant le neutre du groupe, alors  $G$  est connexe par arcs.

**Preuve:** Soit  $x, y \in G$ .  $\exists \gamma_1 : [0, 1] \rightarrow G, \gamma_2 : [0, 1] \rightarrow G$  continus tels que :  $\gamma_1(0) = x, \gamma_2(0) = y, \gamma_1(1) = \gamma_2(1) = e$ . On pose

$$\gamma_3(t) = \gamma_1(2t), 0 \leq t \leq \frac{1}{2}$$

$$\gamma_3(t) = \bar{\gamma}_2(2t - 1), \frac{1}{2} \leq t \leq 1$$

Alors,  $\gamma_3$  est le *chemin composé* de  $\gamma_1$  et  $\bar{\gamma}_2$ . Il est facile de voir qu'il est continu, est bien dans  $G$  ( $\gamma_1(1) = \bar{\gamma}_2(0) = e$ ) et part de  $x$  pour arriver à  $y$ . Ainsi, dans un groupe, ce critère est équivalent à la connexité par arcs (l'autre sens étant évident).  $\square$

**1.3.18 Proposition:**  $GL_n(\mathbb{C})$  est connexe.

**Preuve:** Soit  $A, B \in GL_n(\mathbb{C})$ ,  $P(z) = \det((1 - z)A + zB)$ . Ce polynôme n'a qu'un nombre fini de racines (polynôme de degré  $n$ ). Soit  $\gamma$  un chemin complexe ne passant pas par ces racines. La fonction  $f(t) = (1 - \gamma(t))A + \gamma(t)B$  est un chemin  $[0; 1] \rightarrow GL_n(\mathbb{C})$  tel que  $f(0) = A$  et  $f(1) = B$  (car  $\gamma(0) = 0$  et  $\gamma(1) = 1$ ). Donc  $GL_n(\mathbb{C})$  est connexe par arcs.  $\square$

**1.3.19 Proposition:**  $SO_n(\mathbb{R})$  est connexe.

**Preuve 1.2.14 :** On rappelle que toute matrice orthogonale est "diagonalisable" (donc semblable) dans une certaine base orthogonale et la diagonale est composée de rotations (blocs de taille 2) et de coefficients diagonaux (1 ou -1). Mais, comme on considère  $SO_n(\mathbb{R})$ , il y a un nombre pair de (-1). On a :  $A \in SO_n(\mathbb{R}) \Rightarrow \exists P \in O_n(\mathbb{R})$  telle que  $A = PDP^{-1}$ ,  $D$  étant la matrice diagonale par blocs. Pour cela, on pose :  $\gamma(t)_i : [0; 1] \rightarrow R_\theta$  (matrices de rotations de taille 2) =  $\begin{pmatrix} \cos(\theta_i(1-t)) & \sin(\theta_i(1-t)) \\ -\sin(\theta_i(1-t)) & \cos(\theta_i(1-t)) \end{pmatrix}$

Ainsi, avec cette transformation, on a un chemin de  $[0; 1]$  qui change n'importe quelle matrice de rotation  $R(\theta_1)$  en une matrice identité  $I_2$  composée de 1 sur la diagonale. Comme il y a un nombre pair de (-1), en les regroupant par 2 on les considère comme une rotation de  $\pi$ , et les (1) comme une rotation nulle. Selon la taille :  $n$  pair : on peut regrouper les (1) et (-1) par 2;  $n$  impair : il reste un (1) tout seul pour toute matrice. De cette manière, la matrice  $D$  est transformée en  $I_n$ . Soit donc  $\Gamma$  la "composée" l'ensemble des chemins  $\gamma_i$  qui permettent cette transformation. Autrement dit, une matrice de  $SO_n(\mathbb{R})$  est composée de rotations et éventuellement d'un coefficient 1 en plus. En remplaçant chaque rotation  $R(\theta_i)$  par la matrice  $\gamma_i$ , on obtient un chemin  $\Gamma(t)$  tel que  $\Gamma(0) = D$  et  $\Gamma(1) = I_n$ . On a :  $\alpha(t) = PD(\Gamma(t))P^{-1}$ , qui vérifie alors  $\alpha(0) = A$  et  $\alpha(1) = I_n$ .

**1.3.20 Proposition:**  $U_n(\mathbb{C})$  est connexe.

**Preuve:** Soit  $A$  unitaire, donc diagonalisable dans une base unitaire avec des valeurs propres de module 1. On a :  $A \in U_n(\mathbb{C}) \Rightarrow \exists P \in U_n(\mathbb{C})$  telle que  $A = PDP^{-1}$ ,  $D$  étant diagonale. On utilise  $\gamma(t)_i = e^{i\theta_i(1-t)}$  pour passer d'un coefficient complexe  $e^{i\theta_i}$  à 1. On note alors, de la même manière que précédemment  $\Gamma$  la "composée" l'ensemble des chemins  $\gamma_i$ , de sorte que  $\Gamma(0) = D$  et  $\Gamma(1) = I_n$ . On a :  $\alpha(t) = PD(\Gamma(t))P^{-1}$ , qui vérifie alors  $\alpha(0) = A$  et  $\alpha(1) = I_n$ .  $\square$

**1.3.21 Proposition:**  $SU_n(\mathbb{C})$  est connexe.

**Preuve:** Assez similaire à celle de  $U_n(\mathbb{C})$ , je ne vais donc pas la faire.  $\square$

### Générateurs du groupe linéaire

Voici tout d'abord une définition très spécifique, mais qui nous servira à plusieurs reprises :

**1.3.22 Définition:** Soit  $G$  un groupe. Pour une partie  $S$  de  $G$ , il existe un sous-groupe de  $G$  minimal pour l'inclusion parmi les sous-groupes contenant  $S$  à savoir, l'intersection de tous les sous-groupes contenant  $S$ . On l'appelle *sous-groupe engendré par  $S$* , et on le note  $\langle S \rangle$ . On peut décrire  $\langle S \rangle$  de manière explicite :

$$\langle S \rangle = \{x_1 x_2 \dots x_n, n \in \mathbb{N}, x_i \in S\}$$

**1.3.23 Définition:** On dit que  $S$  est une *partie génératrice* du groupe  $G$ , ou que  $G$  est engendré par  $S$ , lorsque le sous-groupe engendré par  $S$  est  $G : G = \langle S \rangle$

**1.3.24 Théorème:** Soit  $n \geq 2$  et  $\mathbb{K}$  un corps commutatif. On appelle *matrice de transvection* toute matrice de la forme  $T_{ij}(\lambda) = I_n + \lambda E_{ij}$  où  $i \neq j$  et  $\lambda \in \mathbb{K}$ . On appelle *matrice de dilatation* toute matrice diagonale  $D_i(\alpha) = I_n + (\alpha - 1)E_{ii}$  avec  $\alpha \in \mathbb{K}^*$ . Alors l'ensemble des matrices de transvection engendrent le groupe  $SL_n(\mathbb{K})$  et l'ensemble des matrices de transvection et de dilatation engendrent le groupe  $GL_n(\mathbb{K})$ .

**Preuve:** Les matrices de transvections sont toutes de déterminant 1. Le groupe qu'elles engendrent est donc inclus dans  $SL_n(\mathbb{K})$ . La multiplication à gauche (respectivement : à droite) revient à effectuer l'opération élémentaire :  $L_i \leftarrow L_i + \lambda L_j$  (respectivement :  $C_j \leftarrow C_j + \lambda C_i$ ), c'est très simple à vérifier. On note que la multiplication à gauche par la matrice  $T_{ij}(1)T_{ji}(-1)T_{ij}(1)$  permet de permuter 2 lignes au signe près. En effet, soit  $A$  une matrice et  $A = (L_k)_{1 \leq k \leq n}$ ,  $L_k$  étant la  $k$ -ième ligne. Cela revient à faire l'opération :  $L'_i = L_i + L_j$ , puis  $L'_j = L_j - L'_i = -L_i$  et  $L''_i = L'_i + L'_j = L_j$ . L'apparition de ce signe  $-$  est normale car permuter 2 lignes change le signe du déterminant. Maintenant, soit  $A \in GL_n(\mathbb{K})$ . En appliquant l'algorithme du pivot de Gauss, on va transformer  $A$  en une matrice de dilatation, mais en utilisant que des transvections. Comme  $A$  est inversible, sa première colonne est non nulle. Si  $a_{i1} \neq 0$  avec  $i \geq 2$ , l'opération  $L_1 \leftarrow L_1 - (\frac{a_{i1}}{a_{11}})L_i$  permet de mettre un coefficient 1 à la place de  $a_{11}$ . Si tous les coefficients  $a_{i1}$  sont nuls  $\forall i \geq 2$ , on effectue l'échange des lignes  $L_1 \leftarrow L_2$  et  $L_2 \leftarrow -L_1$  pour se ramener au cas précédent. En utilisant le coefficient (1,1) comme pivot, une succession d'opérations sur les lignes puis sur les colonnes permet d'annuler tous les autres coefficients de la première ligne et de la première colonne. Autrement dit, il existe des matrices de transvections  $M_1 \dots M_p$  et  $N_1, \dots, N_q$  telles que :

$$M_p \dots M_1 A N_1, \dots, N_q = \begin{pmatrix} 1 & 0 \\ 0 & A_1 \end{pmatrix}$$

où  $A_1 \in GL_{n-1}(\mathbb{K})$ . On recommence le même algorithme sur la matrice  $A_1$  et ainsi de suite. On aboutit à la fin de cet algorithme à une matrice diagonale  $D_n(\det(A)) = \text{diag}(1, 1, \dots, \alpha)$ ; où le scalaire  $\alpha$  est tout simplement le déterminant de  $A$ . On vient donc de montrer que pour toute matrice inversible  $A$ , il existe des matrices de transvections  $U_1 \dots U_r$  et  $V_1, \dots, V_s$  telles que :  $A = U_r \dots U_1 D_n(\det(A)) V_1, \dots, V_s$ . Ainsi : si  $\det(A) = 1$  les matrices de transvection engendrent  $SL_n(\mathbb{K})$ . Sinon,  $\det(A) \neq 0$  et l'ensemble des matrices de transvections et une unique matrice de dilatation permettent d'engendrer  $GL_n(\mathbb{K})$ .  $\square$

**1.3.25 Théorème:**  $GL_n(\mathbb{R})$  est engendré par l'ensemble des matrices inversible diagonalisables.

**Preuve:** Les matrices de dilatation sont diagonales. Montrons que les matrices de transvections sont diagonalisables : on écrit  $T_{ij}(\lambda)$  avec  $i \neq j$  et  $\lambda \in \mathbb{K}$ . Soit  $D = \text{diag}(1, 2, \dots, n) : T_{ij}(\lambda) = (D^{-1}D)T_{ij}(\lambda) = D^{-1}(DT_{ij}(\lambda))$  avec  $DT_{ij}(\lambda)$  triangulaire, de même diagonale que  $D$  : elle est donc diagonalisable.  $\square$

**1.3.26 Proposition:**  $SL_n(\mathbb{K})$  est connexe, de même que  $GL_n^+(\mathbb{R}) = \{A \in M_n(\mathbb{R}), \det(A) > 0\}$  et  $GL_n^-(\mathbb{R}) = \{A \in M_n(\mathbb{R}), \det(A) < 0\}$ .

Montrons que ces groupes sont reliés à l'identité par un arc continu. Pour  $SL_n(\mathbb{K})$  : D'après le premier théorème, il existe une partie  $X$  contenue dans l'ensemble des couples  $(i, j) \in \{1, 2, \dots, n\}^2$  avec  $i \neq j$  et une famille  $(\lambda_C)_{C \in X}$  de  $\mathbb{K}$  telle que  $A$  soit un produit des transvections :  $T_C(\lambda_C)$ .

$$A = \prod_{C \in X} T_C(\lambda_C)$$

On pose alors :  $\phi : t \in [0; 1] \ A_t \mapsto A = \prod_{C \in X} T_C(t\lambda_C)$ , de sorte que  $\phi(0) = I_n$  et  $\phi(1) = A$ , donc  $SL_n(\mathbb{K})$  est connexe par arcs.

Pour  $GL_n^+(\mathbb{R})$  : On doit introduire au milieu de ce produit la matrice  $D_n(\det(A))$  du théorème précédent. Le chemin est alors le même, avec au milieu du produit la fonction  $f(t) = D_n(t\det(A) + (1-t))$ . On vérifie alors que si  $\det(A) > 0, t\det(A) + (1-t) > 0 : \forall t \neq 0, t\det(A) > 0$ , donc  $t\det(A) + (1-t) > 0$  car  $(1-t) \geq 0$ . Si  $t = 0, t\det(A) + (1-t) = 1 > 0$ . La continuité de la fonction est évidente. On a alors relié  $GL_n^+(\mathbb{R})$  à  $I_n$ .

Un raisonnement analogue permet alors de relier  $GL_n^-(\mathbb{R})$  à  $-I_n$ . On comprend alors un peu mieux la structure de  $GL_n(\mathbb{R})$ .

**Remarque :** Le résultat sur les générateurs du groupe linéaire est assez surprenant et très intéressant. On verra d'ailleurs un peu plus loin quelques résultats sur la signification de "générer" un groupe.

## 1.4 Morphismes de groupes et homéomorphismes

Soit 2 groupes  $(G, \cdot)$  et  $(H, \perp)$ . On note  $e_G$  et  $e_H$  leurs éléments neutres respectifs.

**1.4.1 Définition :** On dit qu'une application  $f : G \rightarrow H$  est un *homomorphisme de groupe* si :

1.  $f(e_G) = e_H$
2. Si  $x, y \in G, f(x \cdot y) = f(x) \perp f(y)$

De plus : Si  $G = H, f$  est un endomorphisme. Si  $f$  est bijective,  $f$  est un isomorphisme. S'il existe un isomorphisme entre 2 groupes  $G$  et  $H$ , on dit qu'ils sont isomorphes et on note  $G \cong H$ . Si  $f$  est un isomorphisme et un endomorphisme, on dit que  $f$  est un automorphisme.

**1.4.2 Proposition :** Soit  $f$  un homomorphisme de groupe (de  $G$  dans  $H$ , comme introduits ci-dessus). Soit  $x \in G : f(x^{-1}) = f(x)^{-1}$

**Preuve :**  $f(e_G) = f(x * x^{-1}) = f(x) \perp f(x^{-1}) = e_H$ , donc  $f(x^{-1}) = f(x)^{-1}$ . □

**1.4.3 Définition :** Soit  $f$  un homomorphisme de  $G$  dans  $H$ . On appelle noyau de  $f$  l'ensemble  $\ker f = \{x \in G, f(x) = e_H\}$ .

**1.4.4 Théorème :** Soit  $f$  un homomorphisme entre  $G$  et  $H$ . On a équivalence entre :

1.  $f$  est injective
2.  $\ker f = \{e_G\}$

**Preuve :** Si  $f$  est injective,  $f(x) = e_H = f(e_G) \Rightarrow x = e_G$ . Donc  $\ker f = \{e_G\}$ . Si  $\ker f = \{e_G\}, f(x) = f(y) \Rightarrow f(x) \perp f(y)^{-1} = e_H = f(x * y^{-1})$ . Mais comme  $\ker f$  ne contient que  $e_G, x * y^{-1} = e_G$  et donc,  $x = y$ .  $f$  est injective. □

**1.4.5 Proposition :** Soit  $f : G \rightarrow H$  un homomorphisme. Alors :

- $\ker f$  est un sous-groupe de  $G$  (1)
- $\text{Im } f$  est un sous groupe de  $H$  (2)

**Preuve :** (1) Soit  $x, y \in \ker f : f(x * y^{-1}) = f(x) \perp f(y)^{-1} = e_H$ . Donc  $x * y^{-1} \in \ker f$ .  $e_G$  est également dans  $\ker f$ . (2)  $e_H$  est dans  $\text{Im } f$ . Puis soit  $x', y'$  dans  $\text{Im } f : \exists x, y \in G$  tels que  $x' = f(x), y' = f(y)$  et donc :  $x' \perp y'^{-1} = f(x) \perp f(y)^{-1} = f(x * y^{-1})$  donc  $x' \perp y'^{-1}$  est dans  $\text{Im } f$ . □

**1.4.6 Proposition :** L'application composée de 2 homomorphismes est encore un homomorphisme.

**Preuve :** Triviale. □

**1.4.7 Définition :** Un *homoméomorphisme* de groupes topologiques est un isomorphisme de groupes qui est continu, et d'inverse continue.

## 2 Groupes quotients

Au cours de cette partie, on a vu différents moyens de produire des groupes :

- le produit cartésien
- l'introduction d'un sous-groupe

Il en existe d'autres :

- le quotient d'un groupe par un sous-groupe
- le produit semi direct de groupes

Ces différents moyens permettent de s'intéresser à la classification des groupes finis (un problème déjà très difficile), celle des groupes infinis étant probablement impossible. On verra en particulier les groupes simples, qui n'admettent pas de groupes quotients autres qu'eux même et le singleton  $\{e\}$  ( $e$  étant le neutre du groupe simple en question).

Ainsi, le but de ce chapitre est de définir un nouveau moyen de créer des groupes, et de voir quelle propriétés et théorèmes intéressants sont associés à ce nouveau moyen.

### 2.1 Relations d'équivalence et ensembles quotients

**2.1.1 Définition:** On considère une relation  $R$  sur un ensemble  $X$ , c'est à dire :  $R \subset X \times X$ . On dit que  $R$  est une *relation d'équivalence* si pour tout  $x, y, z \in X$ , on a que :

- $R$  est réflexive :  $xRx$
- $R$  est symétrique :  $xRy \Leftrightarrow yRx$
- $R$  est transitive :  $xRy$  et  $yRz \Rightarrow xRz$

**2.1.2 Définition:** Soit  $X$  un ensemble muni d'une relation d'équivalence  $R$ . Soit  $x$  un élément de  $X$ . On appelle *classe d'équivalence* de  $x$  (notée  $[x]$  suivant  $R$ ) l'ensemble  $\{y \in X, yRx\}$ . Un élément  $y$  d'une classe d'équivalence est appelé un représentant de la classe d'équivalence.

**2.1.3 Proposition:** Une classe d'équivalence n'est jamais vide.

**Preuve:** On a toujours  $xRx$ . □

**2.1.4 Proposition:** Si  $x$  et  $y$  sont des éléments d'une même classe d'équivalence, alors leur classe d'équivalence sont identiques.

**Preuve:** On a  $xRy$  et donc  $yRx$ . Soit  $z$  tel que  $xRz$  : par transitivité,  $yRz$ , donc  $[x] \subseteq [y]$ . De même, soit  $t$  tel que  $yRt$  : par transitivité  $xRt$  et donc  $[y] \subseteq [x]$  puis finalement,  $[x] = [y]$ . □

**2.1.5 Proposition:** L'ensemble des classes d'équivalences d'un ensemble  $X$  pour une relation d'équivalence donnée  $R$  définit une partition de  $X$ .

**Preuve:** D'une part, tout élément de  $X$  est élément d'une classe d'équivalence de la relation  $R$ . Au pire, cet élément constitue à lui seul une classe d'équivalence. D'autre part, si deux classes d'équivalence s'intersectaient en un ensemble non vide, alors de part la transitivité de la relation d'équivalence  $R$ , ceci impliquerait qu'elles seraient en fait égales. L'ensemble des classes d'équivalence d'une relation  $R$  sur  $X$  définit ainsi bien une partition de  $X$ . □

**2.1.6 Définition:** On appelle *ensemble quotient* de l'ensemble  $X$  pour la relation d'équivalence  $R$  l'ensemble des classes d'équivalence de la relation  $R$ . On note cet ensemble  $X/R$ . À tout élément de  $X$  on peut associer la classe d'équivalence correspondante. Cela définit une application  $-- : X \rightarrow X/R, x \rightarrow \bar{x}$ .

**2.1.7 Exemple:** Le plan projectif :



Le plan projectif réel  $P^2(\mathbb{R})$  est le quotient de  $\mathbb{R}^3$  par la relation d'équivalence être colinéaire. On sait qu'au moins une coordonnée est non nulle. Supposons que ce soit  $z$ . On pose  $\lambda = \frac{1}{z}$ , de sorte que  $(x, y, z)$  est équivalent à  $(\frac{x}{\lambda}, \frac{y}{\lambda}, 1)$ . Puis, si  $z = 0$  et  $y \neq 0$ , on ramène cet ensemble à  $(\frac{x}{\lambda}, 1, 0)$  en posant  $\lambda = \frac{1}{y}$ . Et enfin, si  $y = z = 0$ , alors  $x \neq 0$  et l'ensemble des points équivalents peut être ramené à  $(1, 0, 0)$ . On appelle :  $Aff_2 = \{x, y \in \mathbb{R}, (x, y, 1)\}$ ;  $Aff_1 = \{x \in \mathbb{R}, (x, 1, 0)\}$  et  $Aff_0 = \{(0, 0, 1)\}$  On a alors :

$$P^2(\mathbb{R}) = Aff_2 \cup Aff_1 \cup Aff_0$$

## 2.2 Quotient d'un groupe par un sous-groupe

Dans cette partie,  $(G, .)$  désigne un groupe et  $(H, .)$  désigne un sous-groupe de  $G$ . On considère aussi la relation : si  $x$  et  $y$  sont éléments de  $G$  :  $xRy \Leftrightarrow x^{-1}.y \in H$ .

**Proposition :** La relation  $R$  définie pour tout  $x, y \in G$  par

$$xRy \Leftrightarrow x^{-1}.y \in H$$

est une relation d'équivalence.

**Preuve:**  $xRx$  est évident car  $xx^{-1} = e_G \in H$

$xRy$  et  $yRz$  donne  $x^{-1}.y \in H$  et  $y^{-1}.z \in H \Rightarrow x^{-1}.y.y^{-1}.z = x^{-1}.z \in H$

Enfin  $xRy \Rightarrow x^{-1}.y \in H \Rightarrow (x^{-1}.y)^{-1} = y.x^{-1} \in H$ .  $\square$

**2.2.1 Definition:** On notera  $G/H$  l'ensemble  $G/R$  des classes d'équivalences de la relation  $R$  sur  $G$ .

**2.2.2 Proposition:** Soit  $x \in G$ . La classe d'équivalence de la relation  $xRy \Leftrightarrow x^{-1}.y \in H$  est l'ensemble  $xH = \{x.h, h \in H\}$ .

**Preuve:** Soit  $y \in G$  équivalent à  $x$  pour la relation  $R$  :  $\exists h \in H$  tel que  $x^{-1}.y = h \in H \Leftrightarrow y = x.h$  donc  $y$  est élément de  $xH$ . Réciproquement, si  $y$  est élément de  $xH$ , on a clairement  $xRy$ .  $\square$

**2.2.3 Definition:** L'ensemble  $xH$  s'appelle *classe à gauche de l'élément  $x$  de  $G$* , ou encore *coset gauche de  $x \in G$  par  $H$* .

**Remarque :** On aurait aussi pu définir notre relation  $R$  par  $xRy \Leftrightarrow y.x^{-1} \in H$ . Dans ce cas, la classe d'équivalence d'un élément  $x$  de  $G$  aurait été donné par l'ensemble  $Hx$ .

**2.2.4 Definition:** L'ensemble  $Hx$  s'appelle *classe à droite de l'élément  $x$  de  $G$* , ou encore *coset droit de  $x \in G$  par  $H$* .

Il est naturel de se demander pour quelle condition sur  $H$  on a coïncidence entre les classes à gauche et les classe à droite. Nous allons examiner la question un peu plus loin.

**2.2.5 Definition:** On dira que le sous groupe  $H$  de  $G$  est *distingué ou normal* si pour tout  $g \in G$  et pour tout  $h \in H$ , on a :  $g.h.g^{-1} \in H$ . On note  $H \triangleleft G$  le fait que  $H$  soit normal dans  $G$ .

**2.2.6 Proposition:** Soit donc  $H$  un sous groupe de groupe  $G$ . Les classes à gauche et à droites de la relation d'équivalence héritée de  $H$  coïncident si et seulement si  $H$  est normal. Autrement dit :  $\forall x \in G, xH = Hx \Leftrightarrow H \triangleleft G$ .

**Preuve:** Supposons que les classes à droite et à gauche coïncident.  $\forall g \in G, gH = Hg$ . Donc,  $\forall h \in H, \exists h' \in H$  tel que  $g.h = h'.g$ . Donc,  $\forall h \in H, \exists h' \in H$  tel que  $g.h.g^{-1} = h' \in H$ . Donc  $H \triangleleft G$ . Réciproquement, si  $H \triangleleft G$ ,  $gh = ghg^{-1}g = h'g$  pour un certain  $h'$  car  $H$  est normal.  $\square$

**2.2.7 Proposition:** Si  $G$  et  $G'$  sont deux groupes et que  $f : G \rightarrow G'$  est un homomorphisme de groupe, alors le noyau de  $f$  :  $\ker f$  est un sous groupe normal de  $G$ .

**Preuve:** Soit  $g \in G$  et  $x \in \ker f$  :  $f(g.x.g^{-1}) = f(g)f(x)f(g^{-1}) = f(g)e_G f(g^{-1}) = f(g)f(g^{-1}) = f(g.g^{-1}) = f(e_G) = e_G$ .  $\ker f \triangleleft G$ .  $\square$

## 2.3 Structure de l'ensemble quotient d'un groupe :

**Remarque :** Si  $x$  est élément de  $G$ , on notera  $\bar{x}$  la classe d'équivalence de  $x$  dans  $G/H$ .  $x$  sera alors représentant de la classe d'équivalence  $\bar{x}$ .

**2.3.1 Définition:** Nous allons définir une loi interne  $\perp$  sur  $G/H$  par : si  $x$  et  $y$  sont éléments de  $G$ , alors  $\bar{x} \perp \bar{y} = \overline{xy}$ . Quand aucune confusion n'est à craindre, on notera la loi interne de  $G/H$  de la même façon que celle de  $G$ . Cette loi est celle *induite de  $G$  sur  $G/H$* .

**Remarque :** Il faut vérifier que cette loi est bien définie, c'est à dire que si  $x, x', y, y'$  sont des éléments de  $G$  tels que  $\bar{x} = \bar{x}'$  et  $\bar{y} = \bar{y}'$ , alors  $\bar{x} \cdot \bar{y} = \bar{x}' \cdot \bar{y}'$ .

**2.3.2 Proposition:** *La loi définie précédemment est bien définie si et seulement si  $H$  est normal dans  $G$ .*

**Preuve:** En utilisant la remarque précédente :  $x.x'^{-1} \in H$  et  $y.y'^{-1} \in H \Rightarrow x.y.y'^{-1}.x'^{-1} \in H$ . Mais, par hypothèse,  $y.y'^{-1} \in H \Rightarrow \exists h \in H$  tel que  $y.y'^{-1} = h$  et  $x.x'^{-1} \in H \Rightarrow \exists h' \in H$  tel que  $x.x'^{-1} = h' \Rightarrow x'^{-1} = h'x^{-1}$  et donc  $x.hh'.x^{-1} \in H$  et comme tout élément  $h$  de  $H$  peut s'écrire sous la forme  $y.y'^{-1}$  cela implique que  $x \in G \Rightarrow xHx^{-1} \in H$ , et donc que  $H$  est normal dans  $G$ . Réciproque : si  $H$  est normal dans  $G$ , montrons que  $x.x'^{-1} \in H$  et  $y.y'^{-1} \in H \Rightarrow \bar{x} \cdot \bar{y} = \overline{xy} : x.y.y'^{-1}.x'^{-1} = x.h.x'^{-1} = (\exists h \in H \text{ tel que } y.y'^{-1} = h) = xhx^{-1}.x.x'^{-1} = xhx^{-1}h' (\exists h' \in H \text{ tel que } x.x'^{-1} = h') = h'.h'' (\exists h'' \in H \text{ tel que } x.h.x^{-1} = h'' \text{ car } H \text{ est normal dans } G) \text{ et } h'.h'' \in H \text{ (car } H \text{ est un sous groupe)}. \quad \square$

**2.3.3 Théorème:** *Supposons que  $H$  est un sous groupe normal de  $G$ . L'ensemble  $G/H$  muni de la loi interne induite de celle de  $G$  a une structure de groupe. De plus, si  $G$  est abélien, il en est de même de  $G/H$  équipé de la loi induite.*

**Preuve:** Comme  $H$  est normal dans  $G$ , la loi induite par celle de  $G$  sur  $G/H$  est bien définie. L'élément  $\bar{e_G}$  est le neutre de la loi  $\cdot$ . Dans  $G/H : \bar{e_G} \cdot \bar{x} = \overline{e_G x} = \overline{x e_G} = \bar{x}$ . De plus, tout élément de  $G/H$  possède un inverse, qui est  $\bar{x}^{-1}$  (et donc  $\bar{x}^{-1} = \overline{x^{-1}}$ ). L'associativité de la loi induite se démontre en passant au quotient celle de la loi de départ. De même, on démontrerait que la loi induite est commutative si la loi de départ l'est aussi, et donc que  $G/H$  est abélien si  $G$  l'est.  $\square$

## 2.4 Théorèmes d'isomorphisme

Avant de commencer, je vais devoir introduire le produit interne de sous-groupes. De façon analogue à la définition des cosets, on introduit, pour 2 sous-groupes  $H$  et  $K$  d'un groupe  $G$ , l'ensemble  $HK$ .

**2.4.1 Définition:**  $HK = \{h.k \mid h \in H, k \in K\}$ .

**2.4.2 Proposition:** *Si l'un des 2 groupes  $H$  ou  $K$  est normal dans  $G$ ,  $HK$  est un sous-groupe de  $G$ .*

**Preuve:** Supposons que  $H \triangleleft G$  :

Soit  $x, y \in HK$  : il existe  $h, k, h', k'$  tels que  $x = hk, y = h'k'$ . Donc  $xy = hkh'k' = hkh'k^{-1}kk'$ . Or,  $z = kh'k^{-1} = h'' \in H \Rightarrow xy = hh''kk' \in HK$ .

L'inverse est alors bien définie : si un élément  $x = hk \in HK, h \in H, k \in K, x^{-1} = k^{-1}h^{-1} = k^{-1}hkk^{-1}$ . Or,  $z = k^{-1}hk = h' \in H$ . Donc  $x^{-1} = h'k^{-1} \in HK$ .

Enfin, le neutre est clairement dans  $HK$ .  $\square$

**2.4.3 Proposition:** *Supposons que  $H \triangleleft G$ . Notons :  $\Pi : G \rightarrow G/H$  l'application qui à  $x \in G$  associe sa classe d'équivalence  $\bar{x}$  dans  $G/H$ . Alors  $\Pi$  est un homomorphisme de groupe,  $H = \ker \Pi$  et  $\Pi$  est surjectif.*

**Preuve:** Soit  $x, y \in G : \prod(x.y) = \overline{x.y} = \overline{x}.\overline{y} = \prod(x).\prod(y)$ . Cette propriété n'est que l'expression de la définition de la loi de groupe sur  $G/H$ . Pour l'égalité entre le noyau et  $H$ , il suffit de remarquer que tout élément de  $h$  est équivalent dans  $G/H$  au neutre de  $G/H$ . De plus, si  $\overline{x}$  est un élément de  $G/H$ ,  $x$  est un antécédent de cet élément par  $\prod$ . Donc  $\prod$  est bien surjectif.  $\square$

**2.4.4 Théorème:** Soit  $G, G'$  des groupes. Soit  $f : G \rightarrow G'$  un homomorphisme de groupe. Rappelons que  $\ker f$  est un sous groupe normal de  $G$ , et donc que  $G/\ker f$  a une structure de groupe pour la loi induite de celle de  $G$ . Rappelons aussi que l'on a un morphisme surjectif  $\prod : G \rightarrow G/\ker f$  qui, à tout élément de  $G$ , associe sa classe d'équivalence dans  $G/\ker f$ . Ajoutons encore que l'image d'un groupe par un morphisme est un sous groupe du groupe image. On peut alors affirmer qu'il existe un isomorphisme  $\overline{f} : G/\ker f \rightarrow \text{Im} f$  tel que  $\overline{f} \circ \prod = f$ .

**Preuve:** Soit  $H = \ker f$  : Construisons  $\overline{f}$ . Pour  $\overline{x} \in G/H$ ,  $\overline{f}(\overline{x}) = f(x)$  où  $x$  est un représentant de la classe d'équivalence  $\overline{x}$ .  $\overline{f}$  est bien définie car si  $y$  est un autre représentant de la classe d'équivalence associée à  $x$ , alors  $\overline{f}(\overline{y}) = f(y) = f(x.x^{-1}.y) = f(x).f(x^{-1}.y) = f(x)$ . Cette dernière égalité est vraie car  $x$  et  $y$  étant équivalents dans  $G/H$ ,  $x^{-1}.y \in \ker f$ .  $\overline{f}$  est bien définie et c'est bien un homomorphisme de groupe. Montrons maintenant que  $\overline{f}$  est injective :  $\overline{f}(\overline{x}) = \overline{f}(\overline{y}) \Rightarrow f(x) = f(y) \Rightarrow f(x).f(y)^{-1} = f(x.y^{-1}) = e_{G'} \Rightarrow x.y^{-1} \in \ker f$  donc  $\overline{x} = \overline{y}$ , ce qui prouve l'injectivité. Comme une application est surjective sur son image (par définition),  $\overline{f}$  est un isomorphisme de  $G/H$  dans  $\text{Im} f$ . Enfin, par définition de  $\overline{f}$ ,  $\overline{f} \circ \prod = f$ .  $\square$

**2.4.5 Théorème:** Soit  $G$  un groupe,  $H$  et  $K$  des sous groupes de  $G$ . On suppose que  $H \triangleleft G$ . Alors  $H \cap K$  est normal dans  $K$  et  $\frac{H}{H \cap K} \cong HK/K$ .

**Preuve:** - Montrons que  $H \cap K$  est normal dans  $G$  :  $H \cap K$  est un sous groupe de  $G$  (évident). Soit  $x \in H \cap K$  et  $g \in K$  :  $gxg^{-1} \in H$  car  $H$  est normal dans  $G$  et  $gxg^{-1} \in K$  car produit d'éléments de  $K$ . Donc  $H \cap K \triangleleft K$ .

Par la proposition au début de cette partie, on sait que  $HK$  est un sous-groupe de  $G$ .

$H \cap K$  étant normal dans  $K$ ,  $H/(H \cap K)$  a une structure de groupe pour la loi induite de celle de  $K$ . Montrons que ce groupe est isomorphe à  $HK/K$ . Définissons pour cela l'application :  $\theta : HK \rightarrow H/(K \cap H)$  par : si  $g = h.k$  est élément de  $HK$ ,  $\theta(g) = \overline{h}$  où  $\overline{h}$  désigne la classe d'équivalence de  $h$  dans  $H/(K \cap H)$ . Cette application est bien définie car si  $g$  est aussi représenté par le produit  $h'.k'$  de  $HK$ ,  $h.k = h'.k'$  et donc :  $k.k'^{-1} = h^{-1}.h'$ . Donc  $h^{-1}.h' \in H \cap K$ . Donc  $\overline{h^{-1}.h'} = e_{H/(K \cap H)}$  et  $\theta(g) = \overline{h} = \overline{h'}$ . Comme  $H \cap K$  est un sous groupe normal de  $H$ , alors  $\theta$  est un homomorphisme de groupe.

Calculons  $\ker \theta$ . Soit  $g = h.k \in HK$  tel que  $\theta(g) = \overline{h} = e_{H/(K \cap H)}$ .  $h$  est donc élément de  $H \cap K$ , donc de  $K$ . Donc  $g = h.k \in K$ .  $\ker \theta \subseteq K$ . Soit  $k \in K$  :  $\theta(k) = e_{H/(K \cap H)}$ . Donc  $\ker \theta = K$

$\theta$  est surjective. En effet, si  $\overline{h}$  est élément de  $H/(K \cap H)$ ,  $\theta(h) = \overline{h}$ . pour finir : on applique le premier théorème d'isomorphisme à  $\theta$  et on obtient  $\frac{HK}{K} \cong \frac{H}{H \cap K}$ .  $\square$

**2.4.6 Théorème:** Soient  $G, H$  et  $K$  des sous groupes de  $G$ . On suppose  $H \triangleleft G$  et  $K \triangleleft G$ . On suppose également :  $H \subseteq K$ . Alors  $K/H \triangleleft G/H$  et  $\frac{G/H}{K/H} \cong G/K$ .

**Preuve:** Soit  $g \in G$ . Notons  $\overline{x}$  sa classe d'équivalence dans  $G/H$  et  $\overline{\overline{x}}$  sa classe d'équivalence dans  $G/K$ . Montrons que  $K/H \triangleleft G/H$  : Soit  $\overline{g} \in G/H$  et  $\overline{k} \in K/H$ . Alors  $\overline{g}.\overline{k}.\overline{g}^{-1} = \overline{g.k.g^{-1}}$ . Mais  $K$  est normal dans  $G$ , donc  $g.k.g^{-1} = k' \in K$  et donc  $\overline{g.k.g^{-1}} = \overline{k'}$ . Soit  $\theta$  l'application qui à un élément  $\overline{k}$  de  $K/H$  associe l'élément  $\overline{\overline{k}}$  de  $G/K$ .  $\theta$  est bien définie et c'est un morphisme du groupe  $G/H$  dans le groupe  $G/K$ . De plus,  $H \subseteq K \Rightarrow \theta$  est surjective. Par application du premier théorème d'isomorphisme,  $\frac{G/H}{K/H} \cong G/K$ .  $\square$

## 2.5 Exemples de groupes quotients

### 2.5.1 Exemple: Soit $H = \mathbb{Z}/3\mathbb{Z}$ .

Tout d'abord,  $3\mathbb{Z}$  est normal dans  $\mathbb{Z}$ , donc  $H$  est bien un groupe. Soit un nombre  $a \in H$  : soit  $a \in \mathbb{Z}$ ,  $a + 3b$ ,  $b \in 3\mathbb{Z}$  est l'ensemble des classe d'équivalences pour la relation  $R$  définie pour le quotient. Donc, tout élément est équivalent à 0,1 ou 2 dans  $\mathbb{Z}/3\mathbb{Z}$ . Les éléments de forme  $3p$ ,  $p \in \mathbb{Z}$  sont équivalents à 0, les éléments de forme  $3p + 1$  sont équivalents à 1 et ceux de la forme  $3p + 2$  sont équivalents à 2. C'est un exemple relativement simple.

### 2.5.2 Exemple: Quotient $H' = \mathbb{R}/\mathbb{Z}$

Tout d'abord,  $\mathbb{Z}$  est normal dans  $\mathbb{R}$  donc  $H'$  est bien un groupe. Soit  $a \in \mathbb{R}$  et  $b \in \mathbb{Z}$ . L'ensemble  $a + b \in \mathbb{R}$  peut être ramené à  $[0; 1[$  en posant  $b = -[a]$  où  $[.]$  est la partie entière.

### 2.5.3 Exemple: Le quotient $U(2)/U(1)$ est homéomorphe à $\mathbb{S}^3$ et $U(1)$ est normal dans $U(2)$ .

**Remarque :** C'est une preuve alternative que  $\mathbb{S}^3$  est un groupe. Il est également intéressant de voir  $\mathbb{S}^3$  comme un quotient.

**Preuve:** Montrons que  $U(2)/U(1)$  est homéomorphe à  $SU(2)$  et donc à  $\mathbb{S}^3$ .

1/On va déjà commencer par caractériser le groupe  $U(2)$  (de manière plus précise) : En reprenant la caractérisation de  $SU(2)$  (je renvoie le lecteur à celle-ci) : soit :  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in U(2)$  On a les mêmes équations que pour  $SU^2$ , c'est à dire :

$$|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1; a\bar{c} = -b\bar{d} \Rightarrow \\ \exists \theta, \varphi, \gamma : |a| = \cos(\theta), |b| = \sin(\theta), |c| = \sin(\varphi), |d| = \cos(\varphi)$$

Le seule différence est l'équation sur le déterminant, qui au lieu d'être  $\det(A) = ad - bc = 1$  devient  $|\det(A)| = |ad - bc| = 1$ . Donc

$$|ad - bc| |\bar{c}| = |a\bar{c}d - bc\bar{c}| = |-b\bar{d}d - bc\bar{c}|$$

La dernière égalité vient du fait que  $a\bar{c} = -b\bar{d}$ . Mais comme  $|c|^2 + |d|^2 = 1$  et  $|\det(A)| = 1$ , on a :

$$|ad - bc| |c| = |c| = (-b)(|d|^2 + |c|^2) = |b|$$

On déduit que  $|c| = |b|$  et  $|a| = |d|$ , de manière analogue.

Les égalités

$$|a| = |d|, |c| = |b|$$

permettent alors de choisir  $\theta = \varphi$  car  $|\cos(\theta)| = |\cos(\varphi)|$ ;  $|\sin(\theta)| = |\sin(\varphi)|$ , il y a en fait 2 possibilités :  $\theta = \pm\varphi$ . On se restreint à expliciter une solution, pour  $\theta = \varphi$ .

Donc finalement on peut écrire :

$$a = \cos(\theta)e^{i\alpha_1}; b = \sin(\theta)e^{i\alpha_2}, c = \sin(\theta)e^{i\alpha_3}, d = \cos(\theta)e^{i\alpha_4}$$

avec  $\theta, \alpha_1, \alpha_2, \alpha_3, \alpha_4 \in [0, 2\pi[$ . On a enfin une dernière relation :

$$a\bar{c} + b\bar{d} = 0 = \cos(\theta)\sin(\theta)(e^{i(\alpha_1 - \alpha_3)} + e^{i(\alpha_2 - \alpha_4)})$$

En supposant  $\cos(\theta)\sin(\theta) \neq 0$ , on a

$$e^{i(\alpha_1 - \alpha_3)} = -e^{i(\alpha_2 - \alpha_4)} = e^{i(\alpha_2 - \alpha_4 + \pi)} \Rightarrow \alpha_1 + \alpha_4 = \alpha_2 + \alpha_3 + \pi$$

, qu'on appellera relation (1).

Donc  $A$  est de la forme :  $A = \begin{pmatrix} \cos(\theta)e^{i\alpha_1} & \sin(\theta)e^{i\alpha_2} \\ \sin(\theta)e^{i\alpha_3} & \cos(\theta)e^{i\alpha_4} \end{pmatrix}$  Avec en plus la relation (1).

2/ Ceci étant fait, on peut faire le quotient  $U(2)/U(1)$  : On va maintenant chercher à voir  $U(1)$  comme un sous groupe de  $U(2)$ . On peut poser cette fois-ci, comme représentation de  $U(1)$  :  $B = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} = e^{i\theta} I_2 \in U(2)$

On vérifie de façon immédiate que c'est in sous-groupe normal car pour  $A \in U(2)$ ,  $ABA^{-1} = e^{i\theta} AI_2A^{-1} = e^{i\theta} I_2 = B \in U(1)$  On va chercher les éléments équivalents :  $AB = \begin{pmatrix} \cos(\theta)e^{i\alpha_1+i\theta} & \sin(\theta)e^{i\alpha_2+i\theta} \\ \sin(\theta)e^{i\alpha_3+i\theta} & \cos(\theta)e^{i\alpha_4+i\theta} \end{pmatrix}$

Pour assimiler cela à un groupe que l'on connaît, on utilise la relation (1) et l'astuce  $\theta = -\frac{\alpha_1+\alpha_4}{2}$  De

sorte que :  $AB = \begin{pmatrix} \cos(\theta)e^{i\frac{\alpha_1-\alpha_4}{2}} & \sin(\theta)e^{i\alpha_2-i\frac{\alpha_1+\alpha_4}{2}} \\ \sin(\theta)e^{i\alpha_3-i\frac{\alpha_1+\alpha_4}{2}} & \cos(\theta)e^{i\frac{\alpha_4-\alpha_1}{2}} \end{pmatrix}$

Or, avec la relation (1),  $\alpha_3 = \alpha_1 + \alpha_4 - \alpha_2 + \pi \Rightarrow \alpha_3 - \frac{\alpha_1+\alpha_4}{2} = \alpha_1 + \alpha_4 - \alpha_2 - \frac{\alpha_1+\alpha_4}{2} + \pi = \frac{\alpha_1+\alpha_4}{2} + \pi - \alpha_3$  et donc, en posant :  $\psi_1 = \frac{\alpha_1-\alpha_4}{2}$  et  $\psi_2 = \alpha_2 - \frac{\alpha_1+\alpha_4}{2}$  on obtient :  $AB = \begin{pmatrix} \cos(\theta)e^{i\psi_1} & \sin(\theta)e^{i\psi_2} \\ -\sin(\theta)e^{-i\psi_2} & \cos(\theta)e^{i\psi_1} \end{pmatrix} \in SU(2)$

**Remarque :** Cela ne correspond pas exactement à la caractérisation de  $SU(2)$  qu'on a donné. Cependant, si on reprend la caractérisation de  $SU(2)$  faite précédemment  $|u|^2 + |v|^2 = 1 \Leftrightarrow \exists \theta$  tel que  $|u| = \cos(\theta)$ ,  $|v| = \sin(\theta)$  et donc,  $\exists \psi_1, \psi_2$  tels que  $u = \cos(\theta)e^{i\psi_1}$ ,  $v = \sin(\theta)e^{i\psi_2}$ .  $\square$

## 2.6 Groupes simples \*

**2.6.1 Définition:** On dit qu'un groupe est *simple* si les seuls sous groupes distingués de  $G$  sont  $G$  et  $\{e\}$ .

### 2.6.2 Exemple:

*Dans la suite de cette partie, on va très peu parler de la simplicité, pourtant c'est une notion intéressante. Pour l'illustrer, on va montrer le résultat suivant :  $SO_3(\mathbb{R})$  est simple.*

On a défini de façon très générale le groupe  $SO_n(\mathbb{R})$  dans l'annexe I, mais je vais parler plus particulièrement de  $SO_3(\mathbb{R})$ .

**2.6.3 Définition:** On définit  $SO_3(\mathbb{R})$  comme l'ensemble des matrices  $A \in M_3(\mathbb{R})$  telles que  $A^t A = I_3$  avec  $\det(A) = 1$ . On peut montrer que cela équivaut à l'existence d'une base orthonormée dans laquelle la matrice  $A$  peut s'écrire sous la forme :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Pour un certain  $\theta \in [0, 2\pi[$ .  $SO_3(\mathbb{R})$  est donc l'ensemble des rotations de l'espace euclidien  $\mathbb{R}^3$ .

**2.6.4 Lemme:** *L'ensemble des rotations d'angle  $\pi$  (les retournements) de  $SO_3(\mathbb{R})$  engendrent  $SO_3(\mathbb{R})$ .*

**Preuve:** On a vu qu'une matrice orthogonale était composée de rotations et de réflexions. On va maintenant monter un autre résultat : Les réflexions orthogonales engendrent  $O_n(\mathbb{R})$ .

On rappelle qu'une réflexion orthogonale est une matrice dont les valeurs propres sont 1 et  $-1$ , avec  $-1$  de multiplicité 1 (voir un cours d'algèbre linéaire pour plus de détails). En particulier, c'est une symétrie orthogonale (on rappelle que, par définition, une symétrie est une application linéaire  $s$  telle que  $s \circ s = Id$  ; je renvoie là encore à un cours d'algèbre linéaire pour plus de détails). Soit  $E = \mathbb{R}^n$  un espace vectoriel et  $\langle \cdot, \cdot \rangle$  le produit scalaire usuel, c'est à dire tel que  $\langle x, y \rangle = \sum_{i=1}^n x_i y_i = {}^t x y$ .

Soit  $u \in O_n(\mathbb{R})$  et  $F_u = \{x \in \mathbb{R}^n \mid u(x) = x\}$  l'espace des points fixes. On pose  $p_u = n - \dim F_u$ . On raisonne par récurrence sur  $p_u \in \mathbb{N}$  : Le cas  $p_u = 0$  est simple car il correspond à  $u = I_n = Id$ .

Supposons  $p_u > 0$ . Soit  $x \in F_u^\perp \setminus \{0\}$  et  $y = u(x)$ .  $x \notin F_u \Rightarrow y \neq x$ , et comme  $F_u$  et  $F_u^\perp$  sont stables par

$u, y \in F_u^\perp$ . De plus,  $\langle x - y, x + y \rangle = \|x\|^2 - \|y\|^2 = \|x\|^2 - \|u(x)\|^2 = 0$  ( $u$  est une isométrie). Donc  $x + y$  et  $x - y$  sont orthogonaux. Soit  $\tau$  la réflexion orthogonale associée au vecteur  $x - y$  (c'est à dire, telle que  $E_{-1}$  (l'espace propre associé à la valeur propre  $-1$ ) =  $Vect(x - y)$ ). On a  $\tau(x - y) = y - x$  et  $\tau(x + y) = x + y$  (car vecteur orthogonal :  $\tau(a + b) = a - b$  si  $a \in E_{-1}$  et  $b \in E_{-1}$ , cela par définition d'une réflexion orthogonale). On déduit alors, en utilisant la linéarité de  $\tau$ , que  $\tau(x + y) - \tau(x - y) = 2\tau(y) = 2x \Rightarrow \tau(y) = \tau(u(x)) = x$ . De plus,  $x - y \in F_u^\perp$  (c'est un espaces vectoriel) et donc  $\forall z \in F_u, z$  est orthogonal à  $x - y$  donc  $\tau|_{F_u} = Id_{F_u}$  (là encore, par définition d'une réflexion orthogonale). Donc  $F_u \subset F_{\tau u}$ , mais  $x \in F_{\tau u} \setminus F_u$ , donc  $p_u > p_{\tau \circ u}$ . On utilise donc notre hypothèse de récurrence sur  $\tau \circ u : \tau \circ u = \tau_1 \dots \tau_r$ , où les  $\tau_i$  sont des réflexions orthogonales et  $r \leq p_{\tau u}$ . Mais alors on a  $u = \tau \tau_1 \dots \tau_r$  ( $\tau^2 = Id$ , car c'est une symétrie) et  $r + 1 \leq p_u$ .

On va maintenant prouver le lemme : pour  $SO_3(\mathbb{R})$ , si  $u \neq I_3 = Id$ , on écrit  $u = \tau_1 \tau_2$ , qui sont 2 symétries orthogonales. Or, dans  $SO_3(\mathbb{R})$ , les symétrie orthogonales peuvent, s'écrire, dans une base adaptée, sous la forme  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$  Donc leur opposé est une rotation d'angle  $\pi$ , donc  $u = (-\tau_1)(-\tau_2)$ , qui est un

produit de 2 retournements, car les  $-\tau$  sont alors de la forme  $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ , c'est à dire : des rotations d'angles  $\pi$ . On note que ce lemme n'est pas valable si  $n < 3$ .

Pour  $n > 3$ , montrons (en bonus) qu'on a la même chose pour  $SO_n(\mathbb{R})$  : on peut écrire  $u = \tau_1 \dots \tau_{2p}$ ,  $2p \leq n$  car le déterminant d'une réflexion orthogonale est  $-1$  et celui d'un endomorphisme de  $SO_n(\mathbb{R})$  : est 1 (nécessité d'un nombre pair de réflexions orthogonales). Montrons que la composition d'une paire de réflexions orthogonales peut s'écrire comme composition de 2 retournements : soit  $\tau_1, \tau_2$  2 réflexions orthogonales par rapport aux hyperplans  $H_1$  et  $H_2$  (respectivement), c'est à dire, qui laissent fixe  $H_1$  et  $H_2$  (respectivement). Soit  $V$  un sous-espace vectoriel de  $H_1 \cap H_2$  de dimension  $n - 3$ .  $(\tau_1 \tau_2)|_V = \tau_1 \tau_2|_V = Id$ , donc  $\tau_1 \tau_2(V^\perp) \subseteq V^\perp$ , mais, d'après le cas  $n = 3$ , on peut trouver 2 retournements  $\sigma_1, \sigma_2$  tels que  $\tau_1 \tau_2|_{V^\perp} = \sigma_1 \sigma_2$ . Il reste à prolonger ces 2 retournements par l'identité sur  $V$  et on obtient le résultat voulu.  $\square$

### 2.6.5 Théorème: Le groupe $SO_3(\mathbb{R})$ des rotations de l'espace euclidien $\mathbb{R}^3$ est simple.

**Preuve:** Soit  $G$  un sous-groupe normal de  $SO_3(\mathbb{R})$ , et  $G_0$  la composante connexe de  $G$  contenant  $Id = I_3$ .

On rappelle au préalable que la composante connexe d'un élément  $x \in U$  ( $U$  est un ensemble) est le plus grand ensemble connexe contenant  $X$ .

On a :

(i)  $G_0$  est un sous groupe de  $G$ . En effet, l'application  $(g, h) \in G_0 \times G_0 \mapsto gh^{-1}$  (notons la  $\rho$ ) est continue (structure de groupe topologique des groupes matriciels) et par connexité de  $G_0 \times G_0$ , elle est d'image connexe incluse dans  $G$  et contenant  $G_0$ . Elle est même égale à  $G_0$  car l'image est connexe (il est évident que  $Im(\rho) \subseteq G_0$ ). Si il existait  $g, h$  tels que  $gh^{-1} \notin G_0$ , alors  $G_0 \subset Im(\rho)$  mais, l'inclusion étant stricte et  $Im(\rho)$  étant un ensemble connexe, cela contredirait le fait que  $G_0$  soit une composante connexe de  $G$  contenant  $Id$  (en effet, la composante connexe d'un élément est le plus grand ensemble connexe contenant c'est élément).

(ii)  $G_0$  est un sous-groupe normal de  $SO_3(\mathbb{R})$  : en effet, l'application  $(g, h) \in SO_3(\mathbb{R}) \times G_0 \mapsto ghg^{-1}$  (notons la  $\varphi$ ) est continue (structure de groupe topologique des groupes matriciels) et par connexité de  $SO_3(\mathbb{R}) \times G_0$ , elle est d'image connexe incluse dans  $G$  et contenant  $G_0$ . Un raisonnement analogue au précédent permet de montrer que l'image de  $\varphi$  est exactement  $G_0$ .

(iii) Supposons tout d'abord que  $G_0 \neq \{Id\}$ . On va alors montrer que  $G_0$  contient une rotation d'angle  $\pi$  (un retournement). Pour cela, soit  $g \in SO_3(\mathbb{R})$ . Il existe une base orthornormée dans laquelle  $g$  peut

s'écrire sous la forme : 
$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

On a alors  $tr(g) = 2\cos(\theta) + 1$ , et l'application  $\psi : g \in SO_3(\mathbb{R}) \mapsto \cos(\theta) = \frac{tr(g)-1}{2}$  est continue. On veut montrer qu'il existe  $R \in G_0$  tel que  $\psi(R) = -1 \Leftrightarrow tr(R) = -1 = 1 + 2\cos(\theta) \Leftrightarrow \theta = \pi$  (on a bien équivalence). Pour cela, on remarque que, pour une matrice de rotation  $r \in SO_3(\mathbb{R})$ ,  $\theta = \frac{\pi}{2} \Leftrightarrow tr(r)1 + 2\cos(\theta) = 1$ . Dans ce cas, comme  $r$  correspond à une rotation de  $\frac{\pi}{2}$ ,  $r^2$  correspond à une rotation de  $\pi$ . On rappelle en effet le résultat suivant : dans la même base, la composée de  $n$  rotations par rapport au même axe, d'angles  $(\theta_i)_{1 \leq i \leq n}$  est une rotation d'angle  $\sum_{i=1}^n \theta_i$ . Donc,  $\exists r$  telle que  $\psi(r) = 0 \Rightarrow \exists R = r^2$  telle que  $\psi(R) = -1$ .

Montrons qu'il existe  $s \in G_0$  tel que  $\psi(s) \leq 0$ . Par hypothèse,  $G_0$  contient une rotation  $g$  distincte de l'identité. Quite à considérer  $g^{-1}$ , on peut supposer qu'une mesure  $\theta$  de son angle appartienne à  $]0, \pi[$ . Si  $\cos(\theta) \leq 0$ , on choisit  $s = g$ . Sinon, cela signifie que  $\theta \in ]0, \frac{\pi}{2}[$ , et on cherche alors  $N \in \mathbb{N}$  tel que  $N\theta \in [\frac{\pi}{2}, \pi[$  :  $N = E(\frac{\pi}{2\theta}) + 1$  convient car

$$\frac{\pi}{2\theta} < E(\frac{\pi}{2\theta}) + 1 < \frac{\pi}{2\theta} + 1 \Rightarrow \frac{\pi}{2} < N\theta < \frac{\pi}{2} + \theta < \pi$$

On choisit alors  $s = g^N$ . L'image de  $G_0$  par  $\psi$  est un connexe de  $\mathbb{R}$  contenant  $\psi(s) \leq 0$ , et  $\psi(Id) = 1$ . Donc il existe  $r \in G_0$  tel que  $\psi(r) = 0$ . Par conséquent,  $G_0$  contient bien un retournement.

Mais  $G_0$  est normal dans  $SO_3(\mathbb{R})$ , par conséquent, il contient l'ensemble des retournements de  $SO_3(\mathbb{R})$ , qui engendrent  $SO_3(\mathbb{R})$ , donc  $G_0 = SO_3(\mathbb{R})$ , donc  $G = SO_3(\mathbb{R})$ .

(iv) Si  $G_0 = \{Id\}$  : montrons alors que  $G = \{Id\}$  : tout d'abord, on peut remarquer que l'ensemble des composantes connexes sont réduites à des sigletons. En effet, soit  $G'$  une composante connexe de  $G$ , l'application  $(g, h) \in G' \times G' \mapsto gh^{-1} \in G$  (notons la  $\Phi$  est continue, donc d'image connexe incluse dans  $G$  et contenant l'identité (on prend  $g \in G' \Rightarrow \Phi(g, g) = Id$ ), ce qui est imposé  $G' = \{g\}$ , car dans le cas contraire, s'il existe un élément  $h \neq g$ ,  $\Phi(g, h) \neq Id$  (car il y a une unique inverse dans un groupe), mais dans ce cas,  $Id \subset Im(\Phi)$  (strictement), ce qui contredit le fait que  $G_0 = \{Id\}$ .

Donc, si  $G \neq \{Id\}$ , et si  $g \in G \setminus \{Id\}$ , l'application  $h \in SO_3(\mathbb{R}) \mapsto hgh^{-1} \in G$  est constante égale à  $g$ , car continue et d'image connexe dans  $G$  (même raisonnement que précédemment). Donc,  $\forall h \in SO_3(\mathbb{R})$ ,  $hgh^{-1} = h \Leftrightarrow hg = gh$ , ce qui est absurde car cela signifie  $g \in Z(SO_3(\mathbb{R}))$  (centre de  $SO_3(\mathbb{R})$ ), qui est seulement l'identité (voir un peu en dessous).  $\square$

Remarque : pour mieux visualiser  $SO_3(\mathbb{R})$  (car a priori, le fait qu'il soit simple n'a rien d'intuitif), on peut utiliser les représentations (on en parlera un peu plus loin). Voici un lien très intéressant concernant représentations et le groupe  $SO_3(\mathbb{R})$  (et d'autres groupes) :

<http://gregegan.customer.netspace.net.au/ORTHOGONAL/03/WavesExtra.html#DC>  
(The Double Covers of SO(3) and SO(4))

## 2.7 Centre d'un groupe, groupes projectifs

**2.7.1 Définition :** On appelle *centre du groupe*  $(G, \perp)$  le sous ensemble de  $G : \{x \in G \forall y \in G, x \perp y = y \perp x\}$ . On note  $Z(G)$  ce sous ensemble.  $Z(G)$  est l'ensemble de éléments de  $G$  qui commutent avec tout les autres éléments de  $G$ .

**Remarque :** Le centre d'un groupe contient toujours l'élément neutre de ce groupe et donc toujours non vide.

**2.7.2 Proposition :** *Le centre d'un groupe est un sous groupe distingué de ce groupe.*

**Preuve :** Soit  $x \in Z(G)$  et  $g \in G : g \perp x \perp g^{-1} = g \perp g^{-1} \perp x = x \in Z(G)$ .  $\square$

**2.7.3 Definition:** Le *groupe projectif* d'un groupe matriciel  $H$  est  $H/Z(H)$  où  $Z(H)$  est le centre de  $H$ . On en déduit donc les groupes projectifs des groupes dont on a déterminé le centre ci-dessus.

**Remarque :** Lorsque je détermine les centres, je donne la forme nécessaire qu'ils doivent avoir (condition nécessaire pour qu'un élément soit dans le centre). Mais, bien évidemment, on doit également vérifier que la condition est suffisante. Or dans tous les cas (sauf Heisenberg, où c'est un peu plus compliqué, on le verra partie 1.6), il est clair qu'une matrice de forme  $\alpha I_n$  commute avec toutes les autres (c'est pourquoi je ne le vérifie pas en dessous, et à chaque fois qu'un centre est de cette forme, donc tachez de le retenir, d'autant plus que c'est évident).

**Remarque :** Toujours lorsque je détermine le centre, on retrouve l'équation matricielle  $AB = BA$ , suivie d'une équation : cette équation avec des sommes traduit à chaque fois que si 2 matrices  $C$  et  $D$  sont égales, leurs coefficients sont égaux, c'est à dire :  $c_{ij} = d_{ij} \forall i, j \in \mathbb{N}$ .

**2.7.4 Proposition:** Le centre de  $GL_n(\mathbb{K})$  est  $\alpha I_n$ ,  $\alpha \in \mathbb{K}$  et  $\alpha \neq 0$ .

**Preuve:** On utilise  $B = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$

Avec  $p + q = n$ . Cette matrice est clairement dans  $GL_n(\mathbb{K})$ . Soit  $B \in GL_n(\mathbb{K})$ .

$$AB = BA \Rightarrow \sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n b_{ik} a_{kj} \forall i, j \in \mathbb{N}$$

Comme  $B$  est diagonale, on a :  $a_{ij} b_{jj} = b_{ii} a_{ij}$ . Si  $i, j > p$  ou  $i, j < p + 1$ , on a une égalité triviale car  $b_{ii} = b_{jj} = \pm 1$ . Supposons que  $i < j$ . Alors on peut trouver  $p$  et  $q$  tels que  $i \leq p < j$ . On a alors  $a_{ij} = -a_{ij} \Rightarrow a_{ij} = 0$ . Donc  $A$  est diagonale. Montrons que les coefficients diagonaux sont égaux : on

considère la matrice  $C = \begin{pmatrix} I_{p_1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & I_{p_2} \end{pmatrix}$

Avec  $p_1 + p_2 + 2 = n$ . C'est une matrice de permutation  $2 \times 2$  qui change le signe du déterminant (car on peut considérer que la matrice  $B$  est une permutation des 2 lignes de la matrice  $I_n$ ).  $AC = CA \Rightarrow a_{p_1+1, p_1+1} = a_{p_1+2, p_1+2}$  En faisant varier  $p_1$  et  $p_2$ , on déduit que les coefficients diagonaux sont égaux. Si on les pose égaux à  $\alpha$ , on déduit la proposition.  $\square$

**2.7.5 Proposition:** Le centre de  $O_n(\mathbb{R})$  est  $\pm I_n$ .

**Preuve:** On remarque que les 2 matrices  $B$  et  $C$  utilisés dans la preuve précédente sont orthogonales. La seule chose qui change, c'est qu'on a  $\det(A) = \pm 1 = \alpha^n$  car  $A$  est orthogonale,  $\Rightarrow \alpha = \pm 1$ .  $\square$

**2.7.6 Proposition:** Le centre de  $SL_n(\mathbb{K})$  est  $\alpha I_n$ ,  $\alpha \in \mathbb{K}$  et  $\alpha^n = 1$ .

**Preuve:** On utilise les matrices de transvections  $T_{ij}, i \neq j$  avec des 1 sur la diagonale, des 0 partout ailleurs et le coefficient 1 sur la ligne  $i$  et la colonne  $j$ . On peut écrire  $T_{ij} = I_n + E_{ij}$  où  $E_{ij}$  est une matrice élémentaire avec des 0 partout et 1 sur la ligne  $i$  et la colonne  $j$ . Ces matrices sont toutes de déterminant 1.  $AT_{i,j} = T_{i,j}A \Rightarrow A(I_n + E_{ij}) = (A(E_{ij} + I_n)) \Rightarrow AE_{ij} = E_{ij}A$ . On a alors :  $\sum_{k=1}^n a_{pk} e_{ks} = \sum_{k=1}^n e_{pk} a_{ks} \forall p \neq s, \Leftrightarrow$  : Si  $s \neq j$  et  $p = i$ ,  $a_{ij} = 0 \forall s = j$ . Si  $p = i$  ou  $j$  et  $s \neq j$ ,  $a_{js} = 0 \forall p = i$ . Enfin, si  $p = i, s = j$ ,  $a_{ij} = a_{ji}$ . On déduit alors, en faisant varier  $i$  et  $j$ , que  $A$  est diagonale. Pour montrer

que les coefficients diagonaux sont égaux, on utilise :  $D = \begin{pmatrix} I_{p_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{p_2} \end{pmatrix}$  si  $n \geq 3$ . La matrice

"introduite" dans l'identité est une matrice de permutation de déterminant 1 qui permet (comme pour  $GL_n(\mathbb{K})$ ) de montrer que les coefficients diagonaux sont égaux.  $\square$



**2.7.7 Proposition:** *Le centre de  $U_n(\mathbb{C})$  est  $\alpha I_n$ ,  $\alpha \in \mathbb{C}$  et  $|\alpha| = 1$ .*

**Preuve:** Là encore, les matrices utilisées pour  $GL_n(\mathbb{K})$  sont valables. La seule condition supplémentaire est  $|\det(A)| = 1 = \alpha^n \Rightarrow |\alpha| = 1$ .  $\square$

**2.7.8 Exemple:** *de groupes projectifs  $PGL_n(\mathbb{R})$*

$PGL_n(\mathbb{R}) = GL_n(\mathbb{R}) / Z(GL_n(\mathbb{R})) = A_1 \cup A_2 \cup \dots \cup A_n$  avec  $A_1$ , l'ensemble des matrices inversibles de coefficient  $a_{11} = 1$ ;  $A_2$  l'ensemble des matrices inversibles telles que  $a_{11} = 0$  et  $a_{12} = 1$ , et de façon générale,  $A_k$  l'ensemble des matrices inversibles telles que  $a_{11} = a_{12} = \dots = a_{1,(k-1)} = 0$  et  $a_{1,k} = 1$ . On retrouve ce résultat avec un raisonnement analogue à celui du plan projectif, et en remarquant que  $A_{n+1}$  n'est pas dans  $GL_n(\mathbb{R})$ , car les matrices de cet ensemble ont une ligne nulle.

**2.7.9 Exemple:** *Après cet exemple un peu théorique, regardons quelque chose de plus simple :  $PSL_2(\mathbb{R})$*

D'après ce qu'on a fait précédemment, soit  $\alpha^2 = 1 \Rightarrow \alpha = \pm 1$ . Donc,  $PSL_2(\mathbb{R}) = SL_2(\mathbb{R}) / \{\pm I_2\}$ .

On va déjà caractériser  $SL_2(\mathbb{R})$  :

On écrit  $A \in SL_2(\mathbb{R})$ ,  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  De plus,  $\det(A) = 1 = ad - bc$ . Supposons  $a \neq 0$  : Donc  $d = \frac{1+bc}{a}$

et  $A = \begin{pmatrix} a & b \\ c & \frac{1+bc}{a} \end{pmatrix}$

On note  $B_1$  l'ensemble de ces matrices.

Si  $a = 0$ . Supposons que  $b \neq 0$ . On s'arrange (en multipliant par  $I_2$  ou  $-I_2$  pour que  $b > 0$ ) On a  $\det(A) = -bc = 1 \Rightarrow c = -\frac{1}{b}$ . Donc  $A = \begin{pmatrix} 0 & b \\ -\frac{1}{b} & d \end{pmatrix}$

On note  $B_2$  l'ensemble de ces matrices. Enfin,  $a = b = 0$  n'est pas possible car dans ce cas,  $\det(A) = 0 \neq 1$ . Donc  $SL_2(\mathbb{R}) = B_1 \cup B_2$ . Plus précisément, on peut même écrire  $SL_2(\mathbb{R}) = C_1 \cup C_2 \cup C_3 \cup C_4$  avec  $C_1 = \{A \in B_1, a_{11} > 0\}$ ,  $a_{11}$  étant le coefficient de la première ligne et de la première colonne de la matrice  $A$ . Puis  $C_2 = \{A \in B_1, a_{11} < 0\}$ ,  $C_3 = \{A \in B_2, a_{12} > 0\}$  et  $C_4 = \{A \in B_2, a_{12} < 0\}$ . On remarque que  $B_1 = C_1 \cup C_2$  et  $B_2 = C_3 \cup C_4$ .

Il ne nous reste plus qu'à faire le quotient par  $\{\pm I_2\}$  : on remarque que tout élément de  $C_2$  admet un représentant qui est dans  $C_1$ , que tout élément de  $C_4$  admet un représentant qui est dans  $C_3$ . Par conséquent,  $PSL_2(\mathbb{R}) = C_1 \cup C_3$ .

## 2.8 Groupe dérivé

**2.8.1 Definition:** On appelle *commutateur* des éléments  $g$  et  $g'$  de  $G$  l'élément de  $G$  noté  $\llbracket g, g' \rrbracket$  et donné par  $\llbracket g, g' \rrbracket = g.g'.g^{-1}.g'^{-1}$ .

**2.8.2 Definition:** (proposition) L'ensemble des commutateurs de éléments de  $G$  est un sous groupe distingué de  $G$  appelé sous groupe dérivé de  $G$  et noté  $D(G)$ .

**Preuve:** Soit  $h \in D(G) : \exists x, y \in G$  tels que  $h = x.y.x^{-1}.y^{-1}$ . Soit  $z \in G : z.h.z^{-1} = z.x.y.x^{-1}.y^{-1}.z^{-1} = (z.x.z^{-1}).(z.y.z^{-1}).(z.x^{-1}.z^{-1}).(z.x^{-1}.z^{-1})$ . Soit  $u = (z.x.z^{-1})$  et  $v = (z.y.z^{-1})$  ( $u, v \in G$ ) :  $z.h.z^{-1} = u.v.u^{-1}.v^{-1}$ , donc  $D(G) \triangleleft G$

L'intérêt du sous groupe dérivé d'un groupe réside dans le fait que  $G/D(G)$  est le plus grand quotient abélien de  $G$ . Cela signifie que si  $H$  est un sous groupe normal de  $G$  tel que  $G/H$  est abélien, alors on a une injection de  $G/H$  dans  $G/D(G)$ .  $\square$

**2.8.3 Proposition:**  *$G/D(G)$  est le plus grand quotient abélien de  $G$ .*

**Preuve:** Montrons tout d'abord que  $G/D(G)$  est abélien (On ne rappellera pas que  $G/D(G)$  a, comme  $D(G)$  est normal dans  $G$ , une structure de groupe). Soient  $\bar{g}$  et  $\bar{g}'$  des éléments de  $G/D(G)$ . Alors comme  $\llbracket g, g' \rrbracket$  est un élément de  $D(G)$ , on a  $\bar{g} = \llbracket g, g' \rrbracket = g.g'.g^{-1}.g'^{-1} = \bar{g}.g'.g^{-1}.g'^{-1}$ . Donc  $\bar{g}.g' = g'.\bar{g}$ . Comme

$g$  et  $g'$  sont quelconques dans  $G$ , on a bien montré que  $G/D(G)$  est abélien. Remarquons maintenant que si  $H$  est un sous groupe normal de  $G$  tel que  $G/H$  est abélien, alors si  $g$  et  $g'$  sont éléments de  $H$  et si  $\bar{g}$  désigne la classe d'équivalence de  $g$  dans  $G/H$ ,  $\bar{g}.\bar{g}' = \bar{g}'.\bar{g}$ . Soit encore  $\overline{g.g'.g^{-1}.g'^{-1}} = \bar{e}$ . Cette dernière égalité prouve que  $\llbracket g, g' \rrbracket$  est un élément de  $H$ . Comme  $g$  et  $g'$  sont quelconques dans  $G$ , nous venons d'établir que  $D(G) \subset H$ . Montrons l'existence d'une injection de  $G/H$  dans  $G/D(G)$ . Soit  $\theta : G/H \rightarrow G/D(G)$  telle que si  $g$  est un élément de  $G$ ,  $\theta(\bar{g}) = \bar{g}$ . On vérifie que  $\theta$  est bien définie et que c'est un homomorphisme de groupe. Si  $\theta(\bar{g}) = \bar{e}$  alors cela signifie que  $g \in D(G)$ . Mais d'après l'inclusion précédente, cela signifie aussi que  $g \in H$  et donc que  $\bar{g} = \bar{e}$ .  $\square$

## 2.9 Groupe de Heisenberg \*

On peut se demander, avec tous les groupes matriciels que nous avons vu, si tous les groupes représentés dans une matrice (c'est à dire, les groupes dont la loi de groupe peut être représentés par une matrice) sont des groupes matriciels. Nous verrons ici que ce n'est pas le cas.

**2.9.1 Lemme:** *Si  $H$  est un sous-groupe additif de  $\mathbb{R}$  non réduit à  $\{0\}$ , soit  $a = \inf\{x \in H, x > 0\}$ , soit il existe  $a$  tel que  $H = a\mathbb{Z}$ .*

**Preuve:** Soit  $H$  un sous groupe additif de  $\mathbb{R}$  non réduit à  $\{0\}$  et  $a = \inf\{x \in H, x > 0\} = \inf H^{+*}$ . Cet ensemble est minoré par 0 et  $H^{+*}$  est non vide donc il existe au moins un élément dans  $H^{+*}$ , et  $a$  est donc bien défini.

Supposons que  $a \neq 0$  et montrons par l'absurde que  $a \in H$  : Si  $a \notin H$ , alors  $\exists x, y \in H$  tels que  $a < x < y < 2a$  donc  $a - x < 0 < y - x < a$  :

Pourquoi cela? Par la caractérisation de la borne inférieure. Soit un ensemble  $A$ .  $\forall \varepsilon, \exists x \in A$  tel que  $y - \varepsilon \leq \inf(A)$ . Puisque  $a$  n'est pas dans  $H$ ,  $a < y$ . En prenant  $\varepsilon = \frac{a}{2}$ , on trouve qu'il existe au moins un élément  $x$  de sorte que  $a < y \leq a + \frac{a}{2} = \frac{3}{2}a < 2a$ . Puis on pose ensuite  $\varepsilon' = \frac{y-a}{2}$ , on obtient alors  $x - \varepsilon' \leq \inf(A)$ , soit  $a < x \leq x + \frac{y-a}{2} < y < 2a$ .

Remarque : ce raisonnement permet de montrer que si la borne inférieure  $a$  de l'ensemble  $A$  (respectivement : supérieure) n'est jamais atteinte, il doit exister une infinité d'éléments (par exemple une suite  $x_n \in A^{\mathbb{N}}$ ) de sorte que  $\lim_{n \rightarrow \infty} x_n = a$ .

On reprend le raisonnement :  $y - x \in H$  et  $y - x < a$  ce qui contredit le fait que  $a$  est le plus petit élément positif de  $H$ . Ainsi, nécessairement :  $a > 0 \Rightarrow a \in H$ .

Puis :  $a + a = 2a$  est dans  $H$ , puis par récurrence,  $H \subseteq a\mathbb{Z}$  où  $\mathbb{Z}$  est l'ensemble des entiers relatifs. Montrons qu'il y a égalité : s'il existe  $x \in \mathbb{R}$  et  $k \in \mathbb{Z}$  tel que  $ka < x < (k+1)a$ , alors  $0 < x - ka < a$ , mais si on pose  $y = x - ka$ , cela contredit le fait que  $a$  est le plus petit élément positif de  $H$ . Conclusion :  $H = a\mathbb{Z}$

Puis : si  $a = 0$  : Soit  $z$  dans  $\mathbb{R}$ . Soit  $E(x)$  la partie entière.  $\forall b, \exists x \in H$  tel que  $0 \leq x < b$ , donc  $z = E(\frac{z}{x})x + r$  avec  $0 < r < b$  car  $E(\frac{z}{x}) \leq \frac{z}{x} < E(\frac{z}{x}) + 1$  donc  $E(\frac{z}{x})x \leq z < E(\frac{z}{x})x + x$ , mais comme  $x < b$ , c'est bon. Finalement,  $\forall b \in \mathbb{R}$ , il existe un élément de  $H$  ( $nx$  est dans  $H$  car  $H$  est un sous-groupe additif) tel que  $|z - nx| < b$  et cela  $\forall z \in \mathbb{R}$ . Donc  $H$  est dense dans  $\mathbb{R}$ .  $\square$

**2.9.2 Proposition:** *Soit  $w$  un irrationnel : on pose  $G = \{nw + p, n, p \in \mathbb{Z}\}$ .  $G$  est dense dans  $\mathbb{R}$ .*

**Preuve:**  $G$  est un sous groupe additif de  $\mathbb{R}$  (facile à vérifier). Soit  $a = \inf\{x \in G, x > 0\}$  : d'après la preuve du lemme ci-dessus, si  $a = 0$ ,  $G$  est dense dans  $\mathbb{R}$ , mais si  $a \neq 0$ ,  $G = a\mathbb{Z}$ . Supposons par l'absurde que  $a \neq 0$ . On a alors  $G = a\mathbb{Z}$  : donc, pour  $n_1 > 0, n_2 > 0, p_1 > 0, p_2 \in \mathbb{Z}, \exists k, k' \in \mathbb{Z}$  tels que  $n_1 w + p_1 = ka$  et  $n_2 w + p_2 = k'a$

Prenons  $n_2 = n_1$ ,  $p_2 > p_1$  :  $n_2 w + p_2 = k' a, k' > k$ . Donc  $a = \frac{n_1 w + p_1}{k} = \frac{n_1 w + p_2}{k'}$ , ce qui donne  $wn_1(k' - k) = (p_2 - p_1) > 0$  et enfin  $w = \frac{p_2 - p_1}{n_1(k' - k)}$  : absurde!  
Par conséquent,  $a = 0$  et  $G$  est dense dans  $\mathbb{R}$ .  $\square$

On peut maintenant parler du groupe de Heisenberg

On note  $SUT_n(\mathbb{R})$  l'ensemble des matrices triangulaires supérieures de taille  $n$  où tous les coefficients sur la diagonale sont des 1. On vérifie très simplement que c'est un groupe matriciel. De plus, soit  $A \in SUT_n(\mathbb{R})$  : on peut écrire :  $A = I_n + B$  avec  $B \in T_n^+(\mathbb{R})$ , l'ensemble des matrices réelles triangulaires supérieures strictes.

**2.9.3 Proposition:** *Le centre de  $SUT_n(\mathbb{R})$  est l'ensemble :* 
$$\left\{ I_n + \alpha E_{1,n} \mid E_{1,n} = \begin{pmatrix} 0 & \cdots & 1 \\ \vdots & \ddots & 0 \\ 0 & \cdots & 0 \end{pmatrix} \right\}$$

**Preuve:** Comme d'habitude : soit  $A$  une matrice de  $SUT_n(\mathbb{R})$  commutant avec l'ensemble des matrice  $B$  appartenant à ce même groupe :

$$AB = BA \Leftrightarrow (A' + I_n)(B' + I_n) = (B' + I_n)(A' + I_n); A', B' \in T_n^+(\mathbb{R}) \Leftrightarrow A'B' = B'A'$$

Donc  $Z(SUT_n(\mathbb{R})) = I_n + Z(T_n^+(\mathbb{R}))$  (le centre est le même à addition près, cependant,  $T_n^+$  n'est pas un groupe, il est donc inexact de parler de son centre, d'où les "). Déterminons le "centre". Rappel : soit  $T = (t_{i,j})_{1 \leq i, j \leq n}$ , si  $T \in T_n^+(\mathbb{R})$ ,  $t_{i,j} = 0 \forall i > j$ . Donc, on pose  $A' = (a_{i,j})_{1 \leq i, j \leq n}$ ,  $B' = E_{i,j}, i < j$  (matrice élémentaires de  $T_n^+(\mathbb{R})$ , qui, de plus, en forment une base).

$$A'B' = B'A' \Leftrightarrow \sum_{k=1}^n a_{p,k} e_{k,l} = \sum_{k=1}^n e_{p,k} a_{k,l} \forall p, l \in \mathbb{N}$$

Mais  $e_{k,l} = 0$  si  $k \neq l$ , 1 sinon ; et de la même manière :  $e_{p,l} = 0$  si  $p \neq l$ , 1 sinon. Le système se réduit alors à :

$$\begin{aligned} 0 &= 0 \text{ si } p \neq i, l \neq j \\ a_{p,i} &= 0 \text{ si } l = j, p \neq i \\ a_{j,l} &= 0 \text{ si } p = i, l \neq j \\ a_{i,i} &= a_{j,j} \text{ si } p = i, l = j \end{aligned}$$

Mais  $A'$  est triangulaire supérieure stricte, donc  $a_{q,r} = 0 \forall q > r$ . Finalement, il s'agit de regarder comment adapter  $i, j; i < j$  pour obtenir des simplifications :

On peut choisir  $j = n, i = 1, 2, \dots, n - 1$  pour obtenir  $a_{p,i} = 0 \forall p, \forall i < n$ .

On peut ensuite choisir  $i = 1, j = 2, 3, \dots, n$  pour obtenir  $a_{j,l} = 0 \forall l, \forall j > 1$ .

On a donc  $a_{q,r} = 0 \forall r \neq n, q \neq 1$  car l'intersection des 2 systèmes ci-dessus n'est pas valable pour  $i = n, j = 1$  qui correspond alors à  $a_{1,n}$ . Ce coefficient est donc non nul. Donc le centre de  $SUT_n(\mathbb{R})$  est bien de la forme annoncée.  $\square$

On introduit  $Z_n = Z(SUT_n(\mathbb{Z}))$ , c'est à dire le centre du groupe de Heisenberg dont les éléments sont des entiers relatifs. On vérifie facilement que  $SUT_n(\mathbb{Z})$  est bien un groupe.

On vérifie que  $A = I_n + kE_{1,n}, k \in \mathbb{Z}$  est bien la description de ce centre. On définit le groupe de Heisenberg comme étant  $Heis_n = SUT_n(\mathbb{R})/Z_n$ . Vérifions que c'est un groupe : il suffit de prouver que  $Z_n$  est normal dans  $SUT_n(\mathbb{R})$ . Pour cela, soit  $T \in SUT_n(\mathbb{R}), Z \in Z_n$ .

$$TZT^{-1} = T(I_n + kE_{1,n})T^{-1} = I_n + kTE_{1,n}T^{-1} = I_n + kE_{1,n} \in Z_n$$

(la dernière égalité se vérifie par un bref calcul)

On vérifie alors facilement que ce groupe correspond à  $SUT_n(\mathbb{R})$ , mais avec le coefficient  $a_{1,n}$  de la première ligne et de la  $n$ -ième colonne. dans  $\mathbb{R}/\mathbb{Z} \cong \mathbb{S}^1$ .

Groupe de Heisenberg  $Heis_3$

$$\text{Le groupe de Heisenberg de taille 3 est : } = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

Avec  $(x, y) \in \mathbb{R}^2, z \in \mathbb{R}/\mathbb{Z}$ . On notera  $\bar{w}$  un élément  $w$  de  $\mathbb{R}$  ramené à  $\mathbb{R}/\mathbb{Z}$ . On rappelle que  $\bar{w} = w - [w]$ ,  $[.]$  étant la partie entière. On remarque que l'on a pas, en général :  $\overline{w_1 + w_2} = \bar{w}_1 + \bar{w}_2$  mais qu'on a cependant :  $\overline{\bar{w}_1 + \bar{w}_2} = \overline{w_1 + w_2}$  et  $\overline{\bar{w}_1 + w_2} = \overline{w_1 + w_2}$  (on le prouve en utilisant le rappel).

**2.9.4 Proposition:** *C'est un groupe topologique tel que (en notant  $\cdot$  la loi de groupe).*

$$\begin{aligned} (x, y, z) \cdot (x', y', z') &= (x + x', y + y', \overline{z + z' + xy'}) \\ (x, y, z)^{-1} &= (-x, -y, \overline{-z - xy}) \\ (x, y, z) \cdot (x', y', z') \cdot (x, y, z)^{-1} \cdot (x', y', z')^{-1} &= (0, 0, \overline{xy' - x'y}) \end{aligned}$$

*De plus, la loi de groupe est bien continue (l'addition de deux éléments est continue, et pour la dernière partie de la loi de groupe (troisième composante), il y a également continuité si on considère que  $\bar{1} = 0$ ).*

On remarque que la troisième ligne est en fait un calcul du commutateur de 2 éléments du groupe de Heisenberg de dimension 3, c'est à dire  $[[x, y, z], [x', y', z']] = [[w, w']]$  ( $w, w' \in Heis_3$ ).

**Preuve:** La loi de groupe est :  $(x, y, z) \cdot (x', y', z') = (x + x', y + y', z + z' + xy')$ , qu'on vérifie en faisant la multiplication de 2 matrices de ce groupe, mais  $z + z' + xy' \in \mathbb{R}/\mathbb{Z}$ . On a donc prouvé la première relation. Déterminons le neutre : soit  $e = (e_1, e_2, e_3), x = (x_1, x_2, x_3) : x \cdot e = x \forall x \in \mathbb{R}^3 = (x_1 + e_1, x_2 + e_2, x_3 + e_3 + x_1 e_2) = (x_1, x_2, x_3)$  donc  $e_1 = e_2 = e_3 = 0$ . Le neutre est  $(0, 0, 0)$ . On remarque maintenant que la seconde relation découle directement de la première et que la troisième découle de la seconde. En effet (pour la troisième),

$$\begin{aligned} (x, y, z) \cdot (x', y', z') \cdot (x, y, z)^{-1} \cdot (x', y', z')^{-1} &= (x + x', y + y', \overline{z + z' + xy'}) \cdot (x, y, z)^{-1} \cdot (x', y', z')^{-1} \\ &= (x', y', \overline{z + z' + xy' - z - xy} + (x + x') \times (-y)) \cdot (x', y', z')^{-1} = (x', y', \overline{z + z' - z + xy + xy' - xy - yx'}) \cdot (x', y', z')^{-1} = \\ &= (x', y', \overline{z' + xy' - x'y}) = (0, 0, \overline{z' + xy' - x'y - z' - x'y' + x' \times -y'}) = \\ &= (0, 0, \overline{z' + xy' - x'y - z' + x'y' - x'y'}) = (0, 0, \overline{xy' - x'y}) \end{aligned}$$

Donc le commutateur est dans  $\mathbb{R}/\mathbb{Z}$  et le couvre entièrement.  $\square$

On va maintenant montrer que le groupe  $Heis_3$  n'est pas un groupe matriciel.

**2.9.5 Théorème:** *Il n'existe aucun homomorphisme continu de groupe*

$$\varphi : Heis_3 \longrightarrow GL_n(\mathbb{C})$$

*de noyau trivial  $\ker \varphi = 1(I_3)$  (neutre du groupe).*

**Preuve:** Supposons l'existence d'un tel homomorphisme, et supposons également que  $n$  est l'entier minimal ayant cette propriété.  $\forall g \in Heis_3$ , la matrice  $\varphi(g)$  agit (ou opère) sur les vecteurs de  $\mathbb{C}^n$ . On va identifier  $Z(Heis_3)$  à l'aide du cercle  $\mathbb{S}^1$ . Le cercle a ce qu'on appelle un générateur topologique  $z_0$ , c'est un élément dont les puissances engendrent le cercle. Cette notion de générateur est très utile dans la classification des groupes de Lie compact, mais on ne développera pas cela. Explicitons un tel générateur : si  $r$  est irrationnel on sait que  $\forall s \in \mathbb{R}, \exists p, q \in \mathbb{Z}$ ,

$$|s - pr - q| \leq \epsilon$$

(pour la preuve, voir plus haut : chapitre 9.1). Cela implique que  $e^{2i\pi r}$  est un générateur topologique du cercle car ses puissances sont denses dans  $\mathbb{S}^1$ . Maintenant, soit  $\lambda$  une valeur propre pour la matrice  $\varphi(z_0)$ , et un vecteur propre  $v$ . Quitte à remplacer  $z_0$  par  $z_0^{-1} = \frac{1}{z_0}$ , on peut supposer que  $|\lambda| \geq 1$ .

Rappel : si  $\lambda$  est valeur propre non nulle d'une application linéaire  $f$  (ce qui est possible car  $f$  est inversible et que toute application linéaire admet une valeur propre au moins dans  $\mathbb{C}$ ) et que celle-ci est inversible (ce qui est le cas ici car le noyau est trivial),

$$f(x) = \lambda x \Rightarrow f^{-1}(\lambda x) = \lambda f^{-1}(x) = x \Rightarrow f^{-1}(x) = \frac{1}{\lambda}x$$

De plus, comme  $\varphi$  est un homomorphisme,

$$\varphi(z_0^{-1}) = \varphi(z_0)^{-1}$$

Supposons  $|\lambda| > 1$ , alors

$$\varphi(z_0^k)v = \varphi(z_0)^k v = \lambda^k v$$

2 moyens de voir les choses : dans ce cas  $\varphi$  admet une infinité de valeur propre, ce qui est impossible car on est en dimension finie.

Ou :

$$\|\varphi(z_0^k)\| \geq |\lambda|^k$$

(en utilisant une des propriétés de la norme). Donc,

$$\lim_{k \rightarrow \infty} \|\varphi(z_0^k)\| = \infty$$

Cela implique que  $\varphi(\mathbb{S}^1)$  est non borné. Mais  $\varphi$  est un homomorphisme continu et  $\mathbb{S}^1$  est compact, donc  $\varphi(\mathbb{S}^1)$  doit être borné.

La seule possibilité est donc  $|\lambda| = 1$ . Comme  $\varphi$  est un homomorphisme de groupe,  $\forall g \in Heis_3$ , on a :

$$\varphi(z_0)\varphi(g)v = \varphi(z_0g)v = \varphi(gz_0)v = \varphi(g)\varphi(z_0)v = \lambda\varphi(g)v$$

Donc  $\varphi(g)v$  est un autre vecteur propre pour la valeur propre  $\lambda$ . On pose :

$$V_\lambda = \{v \in \mathbb{C}^n : \exists k \leq 1, t.g(\varphi(z_0) - \lambda I_n)^k v = 0\}$$

Alors  $V_\lambda \subseteq \mathbb{C}^n$  est un espace vectoriel qui est fermé par les actions de toutes les matrices  $\varphi(g)$  avec  $g \in Heis_3$ . Soit  $k_0 \geq 1$  le plus grand nombre pour lequel il existe un vecteur  $v_0 \in V_\lambda$  qui vérifie :

$$(\varphi(z_0) - \lambda I_n)^{k_0} v_0 = 0; (\varphi(z_0) - \lambda I_n)^{k_0-1} v_0$$

(cet entier existe, car on rappelle que l'espace  $E_\lambda = (f - \lambda Id)^k$  correspond au sous-espace caractéristique associé à la valeur propre  $\lambda$ ,  $k$  étant la multiplicité de la valeur propre. Réciproquement, si cet espace est non vide, cela implique que  $\lambda$  est une valeur propre de multiplicité au moins  $k$ . Si  $k_0 > 1$ , cela signifie qu'il existe des vecteurs non nuls  $u, v \in V_\lambda$  tels que

$$\varphi(z_0)u = \lambda u + v; \varphi(z_0)v = \lambda v$$

(de par la décomposition par blocs de Jordan d'une matrice). Dans ce cas,

$$\varphi(z_0^k)u = \varphi(z_0)^k u = \lambda^k u + k\lambda^{k-1}v$$

(se montre par récurrence), et puisque  $v \neq 0$ ,

$$\|\varphi(z_0^k)\| = \|\varphi(z_0)^k\| \geq |\lambda u + kv| \rightarrow \infty$$

quand  $k \rightarrow \infty$ . Cela contredit également le fait que  $\varphi(\mathbb{S}^1)$  est borné. Donc  $k_0 = 1$  et  $V_\lambda$  est l'espace propre associé à la valeur propre  $\lambda$ .

Ce raisonnement permet de prouver la proposition suivante :

**2.9.6 Proposition:** Soit  $G$  un groupe de topologique compact et  $\rho : G \longrightarrow GL_n(\mathbb{C})$  un homomorphisme continu de groupes. Alors,  $\forall g \in G$ ,  $\rho(g)$  est diagonalisable.

En effet, ce qu'on a fait précédemment permet de dire que les seules valeurs propres possible sont de module 1, et que  $V_\lambda$  est le sous espace propre. Mais par définition de  $V_\lambda$ ,

$$(\varphi(z_0) - \lambda I_n)v = 0 \forall v \in \mathbb{C}^n \Rightarrow \varphi(z_0)v = \lambda v$$

, ce qui prouve que  $\rho(g)$  est diagonalisable et qu'il n'y a qu'une unique valeur propre possible.

Après avoir choisi une base pour  $V_\lambda$ , on obtient un homomorphisme continu  $\theta : Heis_3 \longrightarrow GL_d(\mathbb{C})$  tel que  $\theta(z_0) = \lambda I_d$ . Par continuité, tout élément de  $\mathbb{S}^1$  a la forme d'un scalaire multiplié par  $I_d$ . La minimalité de  $n$  implique que  $d = n$ , et on peut donc écrire  $\varphi(z_0) = \lambda I_n$ . Par l'équation des commutateurs de  $Heis_3$ , tout élément de  $\mathbb{S}^1$  peut s'écrire :

$$z = ghg^{-1}h^{-1} \in Heis_3; g, h \in Heis_3$$

Comme le déterminant  $det$  et  $\varphi$  sont des homomorphismes continus de groupe,

$$\begin{aligned} det(\varphi(z)) &= det(\varphi(ghg^{-1}h^{-1})) = det(\varphi(g))det(\varphi(h))det(\varphi(g^{-1}))det(\varphi(h^{-1})) \\ &= det(\varphi(g))det(\varphi(h))det(\varphi(g)^{-1})det(\varphi(h)^{-1}) = det(\varphi(g))det(\varphi(h))det(\varphi(g))^{-1}det(\varphi(h))^{-1} = 1 \end{aligned}$$

Donc il existe une fonction  $\mu : \mathbb{S}^1 \longrightarrow \mathbb{C}^\times$  telle que  $\forall z \in \mathbb{S}^1$ ,

$$\varphi_z = \mu(z)I_d; \mu(z)^d = 1$$

Mais comme  $\mathbb{S}^1$  est connexe par arcs,  $\mu(z) = 1 \forall z \in \mathbb{S}^1$ . Donc  $\forall z \in \mathbb{S}^1$  l'unique valeur propre possible est 1. Donc  $\mathbb{S}^1 \subseteq ker \varphi$ , ce qui contredit le fait que  $ker \varphi$  est trivial.

Donc le groupe de Heisenberg  $Heis_3$  ne peut être assimilé à un groupe matriciel. Un raisonnement analogue permet de prouver qu'aucun des groupes de Heisenberg  $Heis_n$  ne peut être un groupe de matrice.  $\square$

### 3 Actions de groupe

Le but de ce chapitre est de montrer l'utilité des actions de groupes pour mieux comprendre leur structure, et de parvenir à la notion de représentation d'un groupe dans un espace vectoriel (un outil plus simple que les groupes).

#### 3.1 Définitions et exemples

Dans tout le paragraphe qui vient,  $X$  désignera un ensemble et  $(G, \cdot)$  un groupe. On notera  $e$  le neutre de  $G$ .

**3.1.1 Définition:** On dira que le groupe  $G$  agit (ou opère) sur l'ensemble  $X$  si il existe une application  $\theta : G \times X \longrightarrow X$  telle que :

- Pour tout  $x$  dans  $X$ ,  $\theta(e, x) = x$
- Pour tout  $g_1, g_2 \in G$ ,  $\theta(g_1, \theta(g_2, x)) = \theta(g_1 \cdot g_2, x)$ .

On dira aussi que  $\theta$  définit une action de  $G$  sur  $X$ .

D'un point de vue équivalent, on peut montrer que  $G$  opère sur l'ensemble  $X$  si l'on dispose d'un morphisme de groupe dit *associé* à l'action,  $\phi : G \longrightarrow S_X$ , le groupe symétrique de  $X$  (c'est à dire le groupe des permutations de  $X$  (=bijections de  $X$  sur lui-même)). Un tel morphisme est appelé *représentation* du groupe  $G$ . Il est lié à l'action par  $g \cdot x = (\phi(g))(x)$ ,  $\forall g \in G, \forall x \in X$ .

**Remarque :** Afin de simplifier les notations, et quand aucune confusion n'est à craindre, on écrira, si  $g \in G$  et  $x \in X$  alors  $\theta(g, x) = g.x$ .

**3.1.2 Definition:** Soit  $\theta$  une action de  $G$  sur  $X$ . On dira que l'action est *fidèle* si  $\theta$  vérifie : pour tout  $x \in X$  et  $g.x = x$  alors  $g = e$ .

**3.1.3 Definition:** On dira que l'action  $\theta$  de  $G$  sur  $X$  est *transitive* si  $\forall x, y \in X \exists g \in G / g.x = y$ .

**3.1.4 Definition:** Soit  $x \in X$  et soit  $\theta$  une action de  $G$  sur  $X$ .

- On appelle *stabilisateur* de  $x$  et on note  $stab(x)$  le sous ensemble de  $G$  noté par  $stab(x) = \{g \in G / g.x = x\}$ .

- On appelle *orbite* de  $x$  et on note  $w(x)$  le sous ensemble de  $X$  donné par  $\{g.x / g \in G\}$

**3.1.5 Definition:** Si  $g$  est un élément de  $G$  et que  $\theta$  est une action de  $G$  sur  $X$ , on appelle *fixateur* de  $g$  et on note  $fix(g)$  ou  $X^g$  le sous ensemble de  $X$  donné par  $fix(g) = \{x \in X / g.x = x\}$ . De même si  $K$  est une partie de  $G$ , on notera  $X^K$  l'ensemble des  $x \in X$  tels que pour tout  $g \in K, g.x = x$ .

On s'intéresse maintenant à un groupe  $G$  agissant sur un ensemble  $X$  via une action  $\theta$ .

**3.1.6 Proposition:** La relation sur  $X$  définie par : si  $x, y \in X, x \sim y \Leftrightarrow y \in w(x)$  est une relation d'équivalence sur  $X$ .

**Preuve:** Comme d'habitude réflexivité, symmétrie, transitivité,... □

**Remarque :** Les actions de groupes nous permettront donc de partitionner des ensembles suivant des classes d'équivalence.

**3.1.7 Proposition:** Si  $x$  est un élément de  $X$  alors  $Stab(x)$  est un sous groupe de  $G$ .

**Preuve:** Remarquons que  $e$  est toujours un élément de  $Stab(x)$ . Remarquons aussi que si  $g \in Stab(x)$  alors  $x = e.x = (g^{-1}.g).x = (g^{-1}.(g.x)) = g^{-1}.x$ . Donc  $g^{-1}$  est un élément de  $Stab(x)$ . Soient maintenant  $g, g' \in Stab(x)$ , alors on vérifie facilement que  $g.g'^{-1} \in Stab(x)$ . □

**3.1.8 Proposition:** Si  $x$  et  $y \in X$  sont éléments d'une même orbite alors  $Stab(x)$  et  $Stab(y)$  sont des sous groupes conjugués de  $G$ .

**Preuve:** Comme  $x$  et  $y$  sont dans une même orbite, il existe  $h \in G$  tel que  $h.x = y$ . Mais alors si  $g \in Stab(x), h.g.h^{-1} \in Stab(y) : h.g.h^{-1}.y = h.g.x = h.x = y$ . Donc  $h.Stab(x).h^{-1} \subset Stab(y)$ . On montrera de même que  $h^{-1}.Stab(y).h \subset Stab(x)$ , Ce qui nous prouve que  $h.Stab(x).h^{-1} = Stab(y)$ . Et que ces deux groupes sont conjugués. □

**3.1.9 Proposition:** Soit  $x \in X$ . On a une bijection entre  $G/Stab(x)$  et  $w(x)$ .

**Preuve:** Afin de démontrer cela nous allons définir une application  $f : G/Stab(x) \rightarrow w(x)$  par : si  $\bar{g} \in G/Stab(x)$  alors  $f(\bar{g}) = g.x$ .

Montrons que  $f$  est bien définie : Si  $g$  et  $g'$  sont des représentants de  $\bar{g}$  alors il existe  $h \in Stab(x)$  tel que  $g' = g.h$ . Donc  $g'.x = g.h.x = g.x$ .  $f$  ne dépend donc pas du représentant de  $\bar{g}$  choisie et est donc bien définie.

Montrons que  $f$  est injective : Si  $g.x = g'.y$  alors  $g'^{-1}.g$  est un élément de  $Stab(x)$ . Autrement dit  $\overline{g'^{-1}.g} = \bar{e}$  et  $\bar{g} = \bar{g}'$ .  $f$  est donc injective.

$f$  est surjective : Si  $y \in w(x)$  alors par définition de  $w(x)$ , il existe  $g \in G$  tel que  $g.x = y$ .

Donc  $f(\bar{g}) = y$ . La proposition est maintenant démontré. □

Liens entre quotient de groupes et actions de groupes : on va faire le lien par des observations simples, ce qui nous permettra d'introduire la sous-partie suivante.

**3.1.10 Proposition:** Si un sous-groupe  $H$  de  $G$  n'est pas normal dans  $G$  (et même si c'est le cas), alors  $G$  agit de façon transitive sur  $G/H$ .

**Preuve:** Soit  $x \in G/H : \forall g, gx \in G/H$  car  $x = zh, z \in G, h \in G/H, g \in G$  et donc  $gx = gzh$ .  $x$  et  $gx$  ont donc le même représentant (ils appartiennent à la même classe d'équivalence). Soit  $x, y \in G/H$  et  $\bar{x} = x', \bar{y} = y'$  les représentants de la classe d'équivalence, de sorte qu'il existe  $h, h' \in H$  tels que  $x = x'h, y = y'h'$ . Soit  $g = yh'h^{-1}x'^{-1} = yx^{-1} : gx = yh'h^{-1}x'^{-1}x'h = yh' = y$ . L'action est bien transitive.  $\square$

**3.1.11 Proposition:** L'action de  $G$  sur  $G/Stab_G(x)$  est toujours transitive (et elle arrive dans  $w(x)$ ).

**Preuve:** Soit 2 éléments  $y, z$  de  $w(x) : \exists g, g' \in G/Stab_G(x)$  tels que  $y = gx, z = g'x \Rightarrow z = g'g^{-1}gx = g'g^{-1}y$ . L'action est bien transitive. Autre moyen : on vérifie que  $G/Stab_G(x)$  n'est pas normal dans  $G$ .  $\square$

**3.1.12 Proposition:** Si une action d'un groupe  $G$  sur un ensemble  $X$  est transitive, alors  $w(x) = w(y), \forall x, y \in X : l'orbite est unique$ .

**Preuve:** Par la transitivité, soit  $x, y \in X : \exists g \in G$  tel que  $y = gx \Rightarrow g'y = g'gx, \forall g' \in G$ . Donc,  $\forall z \in w(y), \exists g'$  tel que  $z = g'y = g'gx \in w(x)$ . On raisonne de la même manière avec  $w(x)$ , ce qui prouve le résultat par double inclusion.  $\square$

**3.1.13 Exemple:**  $\mathbb{S}^2$  n'est pas un groupe car  $\mathbb{S}^2$  est homéomorphe au quotient  $SU(2)/U(1) \cong \mathbb{S}^3/\mathbb{S}^1$  mais  $U(1)$  n'est pas normal dans  $SU(2)$ .

**Preuve:** On a déjà prouvé que, si  $A$  appartient à  $SU(2)$ ,  $A$  est de la forme :  $\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$  avec  $|u|^2 + |v|^2 = 1$

On cherche alors à représenter  $U(1)$  plongé dans  $SU(2)$ , autrement dit, on cherche à voir  $U(1)$  comme un sous-groupe de  $SU(2)$ . On vérifie que si on pose  $u = e^{i\theta} (\theta \in [0; 2\pi[), v = 0$ , on obtient comme représentation de  $U(1) : B = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \in SU(2)$

Il s'agit bien d'un sous groupe de  $SU(2)$  et si on multiplie cet élément par un élément de  $SU(2) : BA = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} = \begin{pmatrix} ue^{i\theta} & ve^{i\theta} \\ -\bar{v}e^{-i\theta} & \bar{u}e^{-i\theta} \end{pmatrix}$

Or,

$$\begin{aligned} |u|^2 + |v|^2 = 1 &\Rightarrow \exists \phi \\ t.q \ |u| = \cos(\phi), \ |v| = \sin(\phi) \end{aligned}$$

Donc  $\exists \alpha_1$  et  $\alpha_2$  tel que

$$u = \cos(\phi)e^{i\alpha_1}, \ v = \sin(\phi)e^{i\alpha_2}$$

Donc, en posant  $\theta = -\alpha_1$  et  $\alpha_3 = \alpha_2 - \alpha_1 : BA = C = \begin{pmatrix} \cos(\phi) & \sin(\phi)e^{i\alpha_3} \\ -\sin(\phi)e^{-i\alpha_3} & \cos(\phi) \end{pmatrix}$

Donc toute matrice de  $SU(2)$  est dans la même classe d'équivalence qu'une matrice de cette forme. En posant

$$a = \cos(\phi), \ b = \operatorname{Re}(\sin(\phi)e^{i\alpha_3}), \ c = \operatorname{Im}(\sin(\phi)e^{i\alpha_3})$$

, en sachant que  $\det(C) = 1$  (se vérifie facilement), on a

$$\det(C) = a^2 + b^2 + c^2$$

, et que  $a, b, c$  varient dans  $[-1, 1] : SU(2)/U(1)$  est donc homéomorphe à la sphère  $\mathbb{S}^2$ .

Vérifions maintenant que  $U(1)$  n'est pas un sous-groupe normal : on reprend les notations ci dessus en prenant  $\theta$  quelconque : calculons tout d'abord  $A^{-1} : \det(A) = 1 \Rightarrow$  : en utilisant la formule d'inversion d'une matrice  $2 \times 2$ .  $A^{-1} = \begin{pmatrix} \bar{u} & -v \\ \bar{v} & u \end{pmatrix}$  Soit :  $D = ABA^{-1} = \begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \begin{pmatrix} \bar{u} & -v \\ \bar{v} & u \end{pmatrix}$

$$= \begin{pmatrix} ue^{i\theta} & ve^{-i\theta} \\ -\bar{v}e^{i\theta} & \bar{u}e^{-i\theta} \end{pmatrix} \begin{pmatrix} \bar{u} & -v \\ \bar{v} & u \end{pmatrix} = \begin{pmatrix} |u|^2 e^{i\theta} + |v|^2 e^{-i\theta} & -2iuv\sin(\theta) \\ -2i\bar{u}\bar{v}\sin(\theta) & |v|^2 e^{i\theta} + |u|^2 e^{-i\theta} \end{pmatrix}$$



On vérifie très simplement que ce n'est pas un élément de  $U(1)$ . Finalement,  $SU(2)/U(1)$  n'est pas un groupe, donc  $\mathbb{S}^2$  non plus.  $\square$

On déduit donc que  $\mathbb{S}^3$  agit sur  $\mathbb{S}^3/\mathbb{S}^1 \cong \mathbb{S}^2$ .

## 3.2 Espaces homogènes, action des groupes matriciels

**3.2.1 Définition:** Un espace homogène est un ensemble muni d'une action transitive d'un groupe.

On va maintenant s'intéresser à l'action de certains de ces groupes :  $GL_n(\mathbb{K})$  et  $O_n(\mathbb{R})$  sur  $\mathbb{R}^n$  ou  $\mathbb{C}^n$ .

Le stabilisateur dans  $G$  de  $x \in E$  avec  $E$  un espace vectoriel, est  $\{g \in G \mid gx = x\}$ . En décomposant :  $x = \sum_{i=1}^n x_i e_i$  avec  $(e_1, \dots, e_n)$  base de  $E$ .  $gx = x \Rightarrow \sum_{i=1}^n gx_i e_i = \sum_{i=1}^n x_i e_i \Rightarrow ge_i = e_i$  si  $x_i \neq 0$ . C'est pourquoi, dans la suite, on ne va s'intéresser qu'au stabilisateur d'un seul vecteur, le reste se déduisant de façon plus ou moins symétrique. Ce vecteur sera  $e_n = (0, 0, \dots, 0, 1)$  en colonne.

**3.2.2 Proposition:** Le stabilisateur de  $GL_n(\mathbb{K})$  pour  $e_n$  est de la forme :

$$\left\{ A = \begin{pmatrix} B & 0 \\ v & 1 \end{pmatrix}, B \in GL_{n-1}(\mathbb{K}), v \in M_{1, n-1}(\mathbb{K}) \right\}$$

**Preuve:**  $Ae_n = \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{n,n} \end{pmatrix} = e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$  Donc  $a_{i,n} = 0 \forall i$  tel que  $1 \leq i \leq n-1$  et  $a_{n,n} = 1$ . Le stabilisateur

est de la forme annoncée. Reste à vérifier que  $B$  est inversible. On considère les vecteurs lignes de  $B$  :

$B = \begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix}$  Si  $B$  est inversible,  $l_1, l_2, \dots, l_{n-1}$  sont linéairement indépendants. De plus, il est clair que

$b_n = (v, 1)$  n'est pas combinaison linéaire de  $l_1, l_2, \dots, l_{n-1}$ , car leur dernière composante est nulle. Donc  $B$  est inversible. Réciproquement, si  $B$  n'est pas inversible,  $b_1, b_2, \dots, b_{n-1}$  sont liés, donc  $B$  n'est pas inversible. On a donc nécessairement  $B$  inversible et  $v$  quelconque. On note  $GL(n-1)$  le groupe de la forme annoncée. C'est un exemple d'espace homogène (espace sur lequel un groupe agit de façon transitive).  $\square$

**3.2.3 Proposition:** Le stabilisateur de  $O_n(\mathbb{R})$  pour  $e_n$  est de la forme :

$$\left\{ A = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}, B \in O_{n-1}(\mathbb{R}) \right\}$$

**Preuve:**  $Ae_n = \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{n,n} \end{pmatrix} = e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$  Donc  $a_{i,n} = 0 \forall i$  tel que  $1 \leq i \leq n-1$  et  $a_{n,n} = 1$ .

Donc :  $A = \begin{pmatrix} B & 0 \\ v & 1 \end{pmatrix}$  Comme  $A \in O_n(\mathbb{R})$ ,  $A^t A = I_n$  :

$$A^t A = \begin{pmatrix} B & 0 \\ v & 1 \end{pmatrix} = \begin{pmatrix} {}^t B & {}^t v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} B^t B & B^t v \\ v^t B & 1 + |v|^2 \end{pmatrix} = I_n$$

En particulier,  $1 + |v|^2 = 1 \Rightarrow v = 0$ .  $B^t B = I_{n-1} \Rightarrow B$  est orthogonale. On note  $O(n-1)$  le groupe de la forme annoncée.

C'est un autre exemple d'espace homogène.  $\square$

**3.2.4 Proposition:** Le groupe  $O_n(\mathbb{R})$  agit transitivement sur la sphère  $\mathbb{S}^{n-1}$ .

**Preuve:** On sait qu'une matrice orthogonale préserve la norme. Tout vecteur de norme 1 est également de norme 1. L'action sur la sphère est évidente. Une fois qu'on a déterminé le stabilisateur, on détermine

l'orbite de  $e_n$  : soit  $A$  orthogonale :  $Ae_n = \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{n,n} \end{pmatrix}$ , qui correspond à la dernière colonne de la matrice

$A$ . Ce vecteur est bien de norme 1. En choisissant cette colonne pour une matrice orthogonale, on vérifie que  $\forall x \in \mathbb{S}^{n-1} \exists A \in O_n(\mathbb{R})$  telle que  $x = Ae_n$ . De plus, on a une bijection de  $Stab(e_n)$  dans  $Orb(e_n)$ . L'action est donc transitive : soit  $x, y \in \mathbb{R}^n$  de norme 1 :  $\exists A, B \in O_n(\mathbb{R})$  tels que :  $x = Ae_n$  et  $y = Be_n$ . Soit  $e_n = A^{-1}x$ , et finalement,  $y = BA^{-1}x$ , avec  $BA^{-1} \in O_n(\mathbb{R})$ .  $\square$

**3.2.5 Proposition:** *Le groupe  $U_n(\mathbb{C})$  agit transitivement sur la sphère  $\mathbb{S}^{2n-1}$ .*

**Preuve:** Analogie à  $O_n(\mathbb{R})$ .  $\square$

**3.2.6 Proposition:** *Le groupe  $GL_n(\mathbb{K})$  agit de façon transitive sur  $(\mathbb{K}^n)^*$ .*

**Preuve:** Là encore, on détermine l'orbite :  $Ae_n = \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{n,n} \end{pmatrix}$  Soit  $z = Ae_n : z \neq 0$  car  $A$  inversible :  $\Rightarrow$

$e_n = A^{-1}z \neq 0$ . De plus, en choisissant  $z \neq 0$  comme dernière colonne de la matrice  $A$ , on peut compléter cette colonne (ce vecteur) en une base de  $\mathbb{K}^n$ , pour rendre  $A$  inversible. Ainsi  $\forall z \in \mathbb{S}^{n-1} \exists A \in GL_n(\mathbb{K})$  telle que  $z = Ae_n$ . De plus, on a une bijection de  $Stab(e_n)$  dans  $Orb(e_n)$ . L'action est donc transitive : soit  $x, y \in \mathbb{K}^n$  de norme 1 :  $\exists A, B \in GL_n(\mathbb{K})$  tels que :  $x = Ae_n$  et  $y = Be_n$ . Soit  $e_n = A^{-1}x$ , et finalement,  $y = BA^{-1}x$ , avec  $BA^{-1} \in GL_n(\mathbb{K})$ .  $\square$

**3.2.7 Proposition:** *Le groupe  $GL_n(\mathbb{K})$  agit de façon transitive sur lui même (c'est à dire, l'ensemble des vecteurs formant une base de  $K$ ).*

**Preuve:** De manière générale, on peut démontrer qu'un groupe opère transitivement (à droite ou bien à gauche si le groupe est non comutatif) sur lui même : soit  $G$  un groupe et soit  $g, g' \in G : gg^{-1}g' = g'$ . Donc :  $\forall g, g' \in G, \exists z = g^{-1}g'$  tels que  $g' = gz$ . Donc c'est le cas pour  $GL_n(\mathbb{K})$ .  $\square$

### 3.3 Connexité et actions de groupes

**3.3.1 Proposition:**  *$GL_n(\mathbb{C})$  est connexe.*

**Preuve:** On le fait pas récurrence :

On reprend  $GL(n-1)(\mathbb{C})$ . Un théorème de la partie action de groupe expliquait qu'il y a une bijection entre  $G/Stab_G(x)$  et  $Orb_G(x)$ . Ici, cela signifie : il existe une bijection entre  $GL_n(\mathbb{C})/GL(n-1)(\mathbb{C})$  et  $\mathbb{C}^{n*}$  (de par l'action de  $GL_n(\mathbb{C})$ ). On utilise une récurrence :

Initialisation :

soit tout d'abord  $G = GL_2(\mathbb{C})$ .  $GL_1(\mathbb{C})$  est connexe (évident)  $\Rightarrow H = GL(1)(\mathbb{C})$  est connexe, de même que  $G/H = \mathbb{C}^{2*}$ . Donc : par un critère de connexité sur le quotient de groupes (voir dernière proposition de l'annexe de topologie),  $G$  est connexe. Héritéité :

Soit  $G = GL_{n+1}(\mathbb{C})$  Supposons  $= GL_n(\mathbb{C})$  connexe, il en est alors de même pour  $H = GL(n)(\mathbb{C})$ . De plus,  $G/H = \mathbb{C}^{(n+1)*}$  est connexe donc  $G = GL_{n+1}(\mathbb{C})$  est connexe. Conclusion :

$GL_n(\mathbb{C})n$  est connexe. Pour la suite, on ne refera jamais la récurrence complète car elles se ressemblent toutes, mais seulement l'initialisation.  $\square$

**3.3.2 Proposition:**  *$SO_n(\mathbb{R})$  est connexe.*

**Preuve:** On le fait par récurrence. Comme ces démonstrations sont toutes analogues, je ne fais que l'initialison.

On reprend  $SO(n-1)(\mathbb{R})$  (qu'on définit de la même manière que pour  $O_n(\mathbb{R})$ ). Un théorème de la partie action de groupe expliquait qu'il y a une bijection entre  $G/Stab_G(x)$  et  $Orb_G(x)$ . Ici, cela signifie : il existe une bijection entre  $SO_n(\mathbb{R})/O(n-1)(\mathbb{R})$  et  $\mathbb{S}^{n-1}$  (de par l'action de  $SO_n(\mathbb{R})$ ). On utilise une récurrence :

Initialisation :

soit tout d'abord  $G = SO_2(\mathbb{R})$ .  $SO_1(\mathbb{C})$  est connexe (évident)  $\Rightarrow H = SO(1)(\mathbb{R})$  est connexe, de même que  $G/H = \mathbb{S}^1$ . (remarque : il est évident qu'un sphère est connexe). Donc : par un critère de connexité sur le quotient de groupe,  $G$  est connexe.  $\square$

**3.3.3 Proposition:**  $U_n(\mathbb{C})$  est connexe.

**Preuve:** On le fait par récurrence. Comme ces démonstrations sont toutes analogues, je ne fais que l'initialison.

On reprend  $U(n-1)(\mathbb{C})$  (qu'on définit de la même manière que pour  $O_n(\mathbb{R})$ ). Un théorème de la partie action de groupe expliquait qu'il y a une bijection entre  $G/Stab_G(x)$  et  $Orb_G(x)$ . Ici, cela signifie : il existe une bijection entre  $U_n(\mathbb{C})/U(n-1)(\mathbb{C})$  et  $\mathbb{S}^{2n-1}$  (de par l'action de  $U_n(\mathbb{C})$ ). On utilise une récurrence :

Initialisation :

soit tout d'abord  $G = U_2(\mathbb{C})$ .  $U_1(\mathbb{C})$  est connexe (évident)  $\Rightarrow H = U(1)(\mathbb{R})$  est connexe, de même que  $G/H = \mathbb{S}^1$ . Donc : par un critère de connexité sur le quotient de groupe,  $G$  est connexe.  $\square$

Remarque : des preuves analogues permettent de montrer que  $GL_n^+(\mathbb{R}), GL_n^-(\mathbb{R})$  et  $SU_n(\mathbb{C})$  sont connexes. Cependant, ce résultat de connexité est moins fort que le précédent (car il ne dit rien sur la connexité par arcs).

## 3.4 Représentations linéaires des groupes

**3.4.1 Définition:** Un groupe  $G$  est *représenté* dans un espace vectoriel  $E$  (sur un corps qui sera presque toujours  $\mathbb{R}$  ou  $\mathbb{C}$ ) si on a un homomorphisme  $D$  du groupe  $G$  dans un sous-groupe  $\mathcal{G}$  du groupe des transformations linéaires  $GL(E)$ , c'est à dire le groupe des automorphismes de  $E$  muni de la composition des applications :

$$\forall g \in G, g \mapsto D(g) \in GL(E)$$

$$\forall g, g' \in G, D(g.g') = D(g).D(g')$$

$$D(e) = I$$

$$\forall g \in G, D(g^{-1}) = (D(g))^{-1}$$

( $I$  désigne l'opérateur identité dans  $GL(E)$ ).

**3.4.2 Définition:** La représentation qui à tout  $g \in G$  associe  $I \in GL(E)$  est appelée *triviale* ou représentation identité.

**3.4.3 Définition:** La *dimension*  $p$  d'une représentation est la dimension de l'espace de représentation  $E$ .

**3.4.4 Définition:** Une représentation est *fidèle* si  $G$  est isomorphe à  $\mathcal{G}$ , c'est à dire  $\ker D = \{e\}$ , ou encore si  $D(g) = D(g') \Leftrightarrow g = g'$ . Sinon, le noyau de l'homomorphisme  $D$  est un sous-groupe normal  $H$  et la représentation est fidèle de  $G/H$  dans  $E$ .

Si  $E$  est de dimension finie  $p$ , on peut choisir une base  $e_i, i = 1, \dots, p$ , et associer à tout  $g \in G$  la matrice représentative de  $D(g)$  :

$$D(g)e_j = e_i \mathcal{D}_{ij}(g)$$

(avec la convention de sommation sur les indices répétés, voir convention d'Einstein sur le net pour les détails ;  $i$  indice de ligne et  $j$  indice de colonnes).

**3.4.5 Exemple:** Le groupe des rotations dans le plan admet une représentation de dimension 2, avec des matrices

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

Maintenant, au lieu d'entrer dans les détails de cette partie des mathématiques, on va directement passer au seul type de représentation qui nous intéresse : la représentation adjointe. Mais avant cela, quelques prérequis.

**3.4.6 Définition:** On dit que deux sous groupes  $H$  et  $K$  de  $G$  sont *conjugués* si il existe un élément  $g$  de  $G$  tel que  $g.H.g^{-1} = K$ .

**3.4.7 Proposition:** Si deux sous groupes de  $G$  sont conjugués alors ils sont en bijection. Dans le cas où leur cardinal est fini, ils ont le même cardinal.

**Preuve:** Soient  $H$  et  $K$  deux sous groupes conjugués de  $G$ . Soit donc  $g$  dans  $G$  tel que  $g.H.g^{-1} = K$ . Définissons  $\theta : H \rightarrow K$  par  $\theta(h) = g.h.g^{-1}$ . L'égalité  $g.H.g^{-1} = K$  nous permet d'être convaincu de la surjectivité de  $\theta$ . Pour l'injectivité : si  $g.h.g^{-1} = g.h'.g^{-1}$  alors en multipliant cette égalité à gauche par  $g^{-1}$  et à droite par  $g$ , on obtient  $h = h'$ .  $\square$

On note, pour tout élément  $g \in G$ ,  $aut_g : G \rightarrow G, x \mapsto aut_g(x) := gxg^{-1}$  l'automorphisme intérieur de  $G$  associé à  $g$  (c'est un automorphisme de  $G$ ). Alors l'application  $g \mapsto aut_g$  dans  $S_G$  est un morphisme de groupe (facile à voir).

**3.4.8 Définition:** L'action de groupe associée, définie par  $g \cdot x := aut_g(x) = gxg^{-1}$  est appelée *l'action par conjugaison* de  $G$  sur lui-même.

**3.4.9 Définition:** Pour tout  $x \in G$ , l'orbite de  $x$  sous cette action est appelée la *classe de conjugaison* de  $x$  et est notée

$$C_x = \{gxg^{-1} \mid g \in G\}$$

Ses éléments sont appelés les conjugués de  $x$ .

# PARTIE II : Géométrie différentielle

Cette partie concerne des notions de géométrie différentielle. Elle est composée de 2 chapitres.

## 4 Sous variétés et variétés

Le but de ce chapitre est d'introduire la notion de sous-variété, de variété, ainsi que des notions de géométrie différentielle telles que l'espace tangent, qui seront indispensables pour la suite.

### 4.1 Sous-variétés

La géométrie différentielle a pour objet l'étude locale et globale d'objets géométriques appelés variétés différentielles et leurs transformations. Ces objets interviennent dans de nombreux domaines des mathématiques, allant des systèmes dynamiques, équations aux dérivés partielles, géométrie algébrique ou arithmétique et probabilités.

Avant de définir la notion de variété différentielle, nous allons définir et étudier les sous-variétés différentielles de  $\mathbb{R}^n$  qui en sont un exemple remarquable, généralisation naturelle des courbes et surfaces régulières.

Au voisinage de chacun des points, une sous-variété ressemble à un sous-espace vectoriel de  $\mathbb{R}^n$ , si l'on décide que deux objets se ressemblent lorsqu'ils sont difféomorphes- au moins localement.

L'une des principales difficultés de la Géométrie différentielle est la diversité des points de vue possibles, qui correspondent aux différentes manières dont elles interviennent naturellement. Il est indispensable de savoir passer d'un point de vue à l'autre, et le géomètre différentiel est une sorte de jongleur, qui pour chaque problème sait prendre en compte les différents points de vue appropriés qui constituent la première étape de sa résolution.

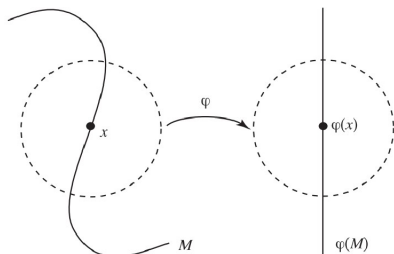
#### 1) Définition, premières propriétés.

Dans ce chapitre, on supposera que tous les difféomorphismes sont de classe  $C^\infty$ .

**4.1.1 Définition:** (A) Soit  $M$  un sous-ensemble de  $\mathbb{R}^n$ . On dit que  $M$  est une *sous-variété* de dimension  $p$  si en chaque point  $x_0$  de  $M$  il existe un voisinage  $U$  de  $x_0 \in \mathbb{R}^n$  et un difféomorphisme local  $\varphi$  de  $U$  dans un ouvert de  $\mathbb{R}^n$  tel que l'on ait

$$\varphi(U \cap M) = (\mathbb{R}^p \times \{0\}) \cap \varphi(U)$$

$\varphi$  est appelé un difféomorphisme rectifiant. Plus explicitement, cela signifie que  $\forall x \in U \cap M$ , on a  $\varphi(x) = (\varphi_1(x), \dots, \varphi_p(x), 0, \dots, 0)$ .



L'entier  $p$  est appelé dimension de  $M$  en  $x_0$ .

**Remarque :** (A) le difféomorphisme  $\varphi$  s'appelle une "carte locale" de  $M$  en  $x_0$ . On va voir que  $p$  est bien défini et ne dépend pas du choix de  $\varphi$ .

(B) Être une sous-variété est une propriété "locale", ce qui signifie qu'elle se vérifie en se restreignant au voisinage de chaque point.

L'unicité de la dimension en  $x$  et son indépendance de  $x$  pour une variété connexe résultent de la proposition suivante :

**4.1.2 Proposition:** *La dimension en un point  $x$  d'une variété est définie de manière unique, et notée  $\dim_x M$ . Si la variété  $M$  est connexe, cette dimension ne dépend pas du point choisi.*

Si la variété  $M$  n'est pas connexe, il s'agit d'une union disjointe de variétés connexes.

**Preuve:** En effet, si on avait deux difféomorphismes  $\varphi_1, \varphi_2$  tels que  $\varphi_1(U \cap M) = \varphi_1(U) \cap \mathbb{R}^p \times \{0\}$  et  $\varphi_2(U \cap M) = \varphi_2(U) \cap \mathbb{R}^q \times \{0\}$  alors  $\psi = \varphi_2 \circ \varphi_1^{-1}$  serait un difféomorphisme local défini sur un voisinage de 0 envoyant  $\mathbb{R}^p \times \{0\}$  sur  $\mathbb{R}^q \times \{0\}$ . On en déduit que  $d\psi(0)$  envoie  $\mathbb{R}^p \times \{0\}$  dans  $\mathbb{R}^q \times \{0\}$ . Or  $d\psi(0)$  est une application linéaire bijective donc injective, ce qui entraîne  $p \leq q$ . Par symétrie on voit que  $p = q$ .

Maintenant si  $\varphi : U \rightarrow V$  est une carte en  $x$ , et  $y$  est un point de  $U$ ,  $\varphi$  donne par translation une carte en  $y$  : le difféomorphisme  $\psi$  donné par  $\psi(z) = \varphi(z) - \varphi(y)$  est une carte en  $y$ . On en déduit que  $\dim_x M = \dim_y M$  et donc  $D_p = \{x \in M \mid \dim_x M = p\}$  est ouvert. Son complémentaire est réunion d'ouverts  $M - D_p = \bigcup_{j \neq p} D_j$ , il est donc ouvert, ce qui entraîne que  $D_p$  est à la fois ouvert et fermé. On en conclut que si  $D_p$  est non vide et  $M$  connexe,  $M = D_p$ . □

**4.1.3 Définition:** Soit  $U$  un ouvert de  $\mathbb{R}^n$ ,  $V$  un ouvert de  $\mathbb{R}^{n-p}$  et un entier  $p$  tel que  $1 \leq p \leq n - 1$ . On dit que  $F : U \rightarrow V$  (application  $C^1$ ) est une submersion si  $dF(x)$  est surjective  $\forall x \in U$ .

**4.1.4 Définition:** Soit un entier  $p$  tel que  $1 \leq p \leq n - 1$ ,  $U$  un ouvert de  $\mathbb{R}^p$ ,  $V$  un ouvert de  $\mathbb{R}^n$  et  $j : U \rightarrow V$ . On dit que  $j$  (application  $C^1$ ) est une immersion si  $dj(x)$  est injective  $\forall x \in U$ .

**4.1.5 Définition:** Un plongement est une immersion injective qui est un homéomorphisme sur son image.

**4.1.6 Définition:** Une application  $f$  est dite propre si l'image réciproque de tout compact par  $f$  est compacte.

Remarque : quand on parle de différentielle bijective, surjective ou injective, cela signifie que la matrice Jacobienne de la fonctions vectorielles de plusieurs variables  $F = (f_1, f_2, \dots, f_k)$  (matrice dont la ligne  $i$  est constituée des dérivés partielles en  $x_1, x_2, \dots, x_n$  de  $f_i$ ) est la représentation matricielle d'une application linéaire bijective, surjective ou injective en tous points  $(x_1, x_2, \dots, x_n)$  fixés.

Avant de prouver la proposition suivante, on va avoir besoin de ces deux théorèmes (supposés connus pour tout étudiant de L2) :

**4.1.7 Théorème:** (Fonctions implicites) Soient  $E_1, E_2, F$  des espaces de Banach  $U, V$  des ouverts respectivement dans  $E_1, E_2$ , et  $f : U \times V \rightarrow F$  une application  $C^k$  où  $k \geq 1$ . On suppose  $f(x_0, y_0) = 0$  et  $\frac{\partial}{\partial y} f(x_0, y_0)$  inversible dans  $\mathfrak{L}(E_2, F)$  d'inverse continue. Alors il existe des voisinages  $U'$  de  $x_0$  et  $V'$  de  $y_0$  et une application  $\varphi \in C^k(U', V')$  tels que

$$\forall (x, y) \in U' \times V', f(x, y) = 0 \iff y = \varphi(x)$$

**4.1.8 Théorème:** (Théorème d'inversion locale) Si  $U, V$  sont des ouverts dans les espaces de Banach  $E_1, E_2$ , et  $\varphi : U \rightarrow V$  est  $C^k, k \geq 1$ , et  $d\varphi(x_0)$  est un isomorphisme, il existe des ouverts  $U'$  et  $V'$  contenant  $x_0$  et  $y_0 = \varphi(x_0)$  et une application  $\psi \in C^k(V', U')$  telle que  $\varphi \circ \psi = \psi \circ \varphi = Id$ . L'application  $\varphi$  est donc un difféomorphisme de  $U'$  sur  $V'$ .

Utiliser la définition ci-dessus pour démontrer qu'un ensemble est une sous-variété est parfois peu commode. Il vaut mieux connaître le plus grand nombre de définitions équivalentes pour utiliser chaque fois celle qui est la mieux adaptée. En voici quelques unes :

**4.1.9 Proposition:** Soit  $M$  un sous-ensemble de  $\mathbb{R}^n$ . Alors  $M$  est une sous-variété si et seulement si pour chaque point de  $x_0$  de  $M$  :

(B) (**équation**) Il existe une submersion  $C^\infty, F : W \rightarrow \mathbb{R}^{n-p}$  (c'est à dire  $dF(x_0)$ , est surjective) telle que  $W \cap M = F^{-1}(0)$ .

**Remarque :** La condition (B) signifie que si on pose  $F(x) = (F_1(x), \dots, F_{n-p}(x))$ , les formes linéaires  $dF_1(x), \dots, dF_{n-p}(x)$  sont linéairement indépendantes.

(C) (**graphe**) Il existe un changement de coordonnées (qu'on peut même choisir linéaire),  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , et une application  $C^\infty, f : \mathbb{R}^p \rightarrow \mathbb{R}^{n-p}$  telle que

$$W \cap M = W \cap \{A(z, f(z)) \mid z \in \mathbb{R}^p\}$$

**4.1.10 Exemple:** Pour  $p = 0$  une variété est simplement un ensemble de points isolés. Pour  $p = 1$  on retrouve la notion de courbe plane régulière injective propre du plan ( $n = 2$ ) ou de l'espace ( $n = 3$ ). Pour  $p = n$  une sous-variété de dimension  $n$  de  $\mathbb{R}^n$  est un ouvert de  $\mathbb{R}^n$ .

**Preuve:** On note (DEF) la propriété de la définition d'une sous-variété.

(i) Montrons que (DEF) entraîne la propriété (B).

Soit  $\varphi : M \cap W \rightarrow \mathbb{R}^p \times \{0\}$ . On note  $(u_1, \dots, u_p, v_1, \dots, v_{n-p})$  les coordonnées sur  $\mathbb{R}^n$  et  $F : x \mapsto (v_1(\varphi(x)), \dots, v_{n-p}(\varphi(x)))$ . Alors il est clair que

$$x \in M \cap W \iff F(x) = 0, \text{ et } x \in W$$

Or  $dF(0)$  est surjective car c'est la composée de la différentielle de  $(u, v) \mapsto v$  et de  $d\varphi$ .

(ii) Montrons que  $(B) \implies (C)$ .

Soit en effet  $(u, v) \in W$ . Alors :

$$F(u, v) \in W \quad \text{et} \quad F(u, v) = 0 \iff v = f(u) \quad \text{et} \quad u \in U$$

Donc  $M$  est localement le graphe de  $f$ .

-  $(C) \implies (DEF)$

Quitte à faire un changement linéaire de variable, on peut supposer que  $M \cap W$  est le graphe de  $f : U \rightarrow V$ .

Alors  $\varphi : (u, v) \mapsto (u, v - f(u))$  envoie  $M \cap W$  sur  $\mathbb{R}^p \times \{0\} \cap (U \times V)$ .

□

Notons qu'une sous-variété n'est pas a priori fermée. Par exemple la spirale, image de  $s \mapsto (\exp(-s)\cos(s), \exp(-s)\sin(s))$  est une sous-variété, d'après le critère  $(C)$ . de même une droite privée d'un point est une sous-variété du plan. Si on veut exclure ces exemples (mais on exclut du même coup les ouverts de  $\mathbb{R}^n$ ), on peut se restreindre aux sous-variétés fermées. La plupart des sous-variétés que nous utiliserons sont fermées.

**4.1.11 Proposition:** Soit  $M$  un sous-ensemble de  $\mathbb{R}^n$  et  $W$  un voisinage de  $x_0$ . Alors  $M$  est une sous-variété si et seulement si pour chaque point de  $x_0$  de  $M$  :

(D) (**nappe paramétrée**) Il existe un plongement  $j : \mathbb{R}^p \rightarrow \mathbb{R}^n, C^\infty$  et est définie sur un voisinage  $U$  de 0, (c'est à dire,  $dj(0)$  est injective, et bicontinue) telle que  $j(0) = x_0$  et

$$j : U \rightarrow M \cap W$$

(Une application bicontinue est bijective, continue et d'inverse continue, on dit aussi que c'est un homéomorphisme)

**Preuve:** (i)  $(C) \implies (D)$

Soit  $M \cap (U \times V)$  le graphe de  $f : U \rightarrow V$ . Et soit  $j : x \mapsto (x, f(x))$  définie sur  $U$ . Alors  $j$  a sa différentielle injective. De plus l'image d'un ouvert  $U'$  contenu dans  $U$  est égale à  $(U' \times V) \cap M$  qui est donc un voisinage de  $x_0$  dans  $M$ . L'application  $j$  est donc bicontinue.

(ii)  $(D) \implies (C)$

Soit  $j(x) = (u(x), v(x))$  définie sur  $U$ , et telle que  $j(U) = M \cap W$ . On peut supposer après changement linéaire de variable, que l'image de  $dj(0)$  est  $\mathbb{R}^p \times \{0\}$ , c'est à dire que  $dv(0) = 0$  et  $du(0)$  est un isomorphisme. Le théorème d'inversion locale permet alors de trouver des ouverts  $U' \subset U$  et  $V' \subset V$  et une application  $C^\infty, \rho : U' \cap (\mathbb{R}^p \times \{0\}) \rightarrow V'$  telle que pour tout  $(x, y) \in U' \times V'$ ,

$$u(x) = y \iff x = \rho(y)$$

Alors  $f : x \mapsto v \circ \rho(x)$  a pour graphe  $\{(x, v \circ \rho(x)) \mid x \in U'\} = \{(u(x), v(x)) \mid x \in U'\}$ . Par bicontinuité de  $j$ , on peut affirmer que  $j(U') = M \cap W'$  pour un voisinage  $W'$  de  $x_0$ . Alors  $M \cap W'$  est bien le graphe de  $f$  au dessus de  $U'$ . □

**Remarque :** Attention : Si  $j : U \rightarrow \mathbb{R}^n$  est une application dont la différentielle est injective, il n'en résulte pas que  $j(U)$  soit une sous-variété. Une telle application  $j$  s'appelle une immersion. Par abus de langage un sous-ensemble de  $\mathbb{R}^n$  tel que tout point a un voisinage qui soit une image d'une application de différentielle injective est appelé sous-variété immergée. Mais ce n'est pas une sous-variété. En particulier si  $j$  n'est pas injective,  $j(u_0) = j(u_1) = x_0$ , on peut avoir la situation suivante ( $p = 1, n = 2$ ).



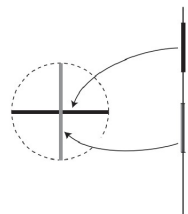


FIGURE 2.2 – Image d'une immersion qui n'est pas une sous-variété

On voit alors que  $M = j(U)$  n'est pas une sous-variété car une carte enverrait les deux vecteurs tangents à chaque branche de la courbe (qui sont linéairement indépendants) sur des vecteurs de  $\mathbb{R}$ , donc liés. Un autre cas, le plus subtil se présente si  $j$  n'est pas bicontinue, comme sur le dessin suivant.

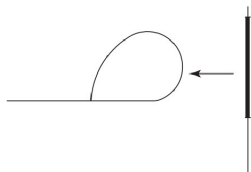
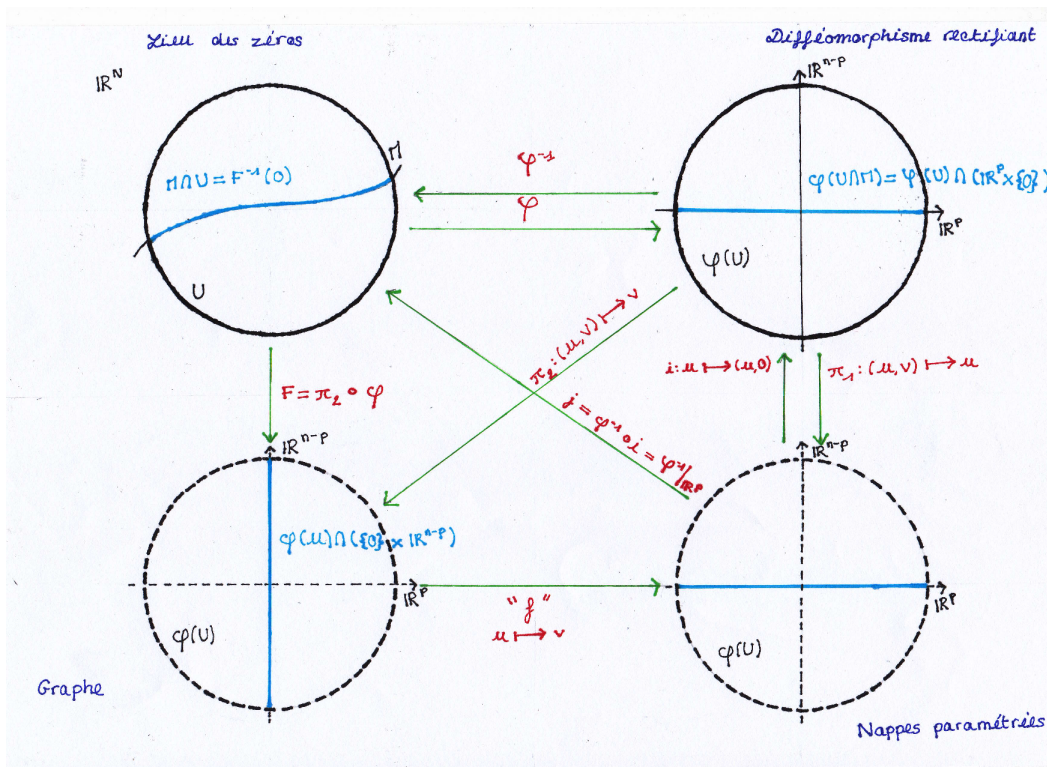


FIGURE 2.3 – Image d'une immersion injective qui n'est pas une sous-variété

On utilisera donc la caractérisation des sous-variétés comme nappes paramétrées avec une certaine prudence. Et dans ce cas il est souvent préférable d'utiliser le fait que la bijectivité est la propriété de  $j$  entraînant sa bicontinuité (on admet pour l'instant cette propriété, mais on la démontre un peu plus loin).

Avant de continuer, une petite illustration du lien entre les différentes définitions des sous-variété :



On va utiliser le théorème suivant pour montrer que toutes les immersions (respectivement les submersions) définies sur les mêmes ouverts (de  $U$  dans  $V$ ) sont toutes "équivalentes".

**4.1.12 Théorème:** (Théorème du rang constant) Soient  $U, V$  des ouverts de  $\mathbb{R}^m$  et  $\mathbb{R}^n$  et  $f : U \rightarrow V$  de classe  $C^k, k \geq 1$ , et  $d\varphi(x)$  est de rang constant  $r$  dans  $U$ . Alors il existe des ouverts  $U'$  et  $V'$  contenant  $x_0$  et  $y_0 = f(x_0)$  et des difféomorphisme locaux  $\rho$  défini sur  $U'$  et  $\sigma$  défini sur  $V'$  tels que  $\sigma \circ f \circ \rho(x_1, \dots, x_m) = (x_1, \dots, x_r, 0, 0, \dots, 0)$ .

**Preuve:** Quitte à composer avec des changements de coordonnées affines, on peut supposer que  $x_0 = 0, y_0 = 0$  et que  $df(x_0)$  est la projection de  $(x_1, \dots, x_m) \mapsto (x_1, \dots, x_r, 0, \dots, 0)$ . Soit alors  $\pi$  la projection  $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_r)$  et  $w$  l'application

$$w(x_1, \dots, x_m) = (\pi \circ f(x_1, \dots, x_m), x_{r+1}, \dots, x_m)$$

Sa différentielle en  $x_0$  est l'identité, il existe donc  $\rho$  difféomorphisme local de  $\mathbb{R}^m$  telle que  $w \circ \rho = Id$  Soit  $\pi(f(\rho(x_1, \dots, x_m))) = (x_1, \dots, x_r)$ , on s'est donc ramené à la situation où

$$f(x_1, \dots, x_m) = (x_1, \dots, x_r, g(x_1, \dots, x_m))$$

où  $g : \mathbb{R}^m \rightarrow \mathbb{R}^{n-r}$ . L'hypothèse que  $f$  est de rang  $r$  au voisinage de 0 entraîne que pour tout  $x$  dans  $U', \frac{\partial g}{\partial x_j} = 0$  pour  $j > r$  est donc que  $g(x_1, \dots, x_m) = g(x_1, \dots, 0, \dots, 0)$ . En posant  $g = (g_{r+1}, \dots, g_m)$  et composant avec le difféomorphisme

$$\sigma : (y_1, \dots, y_n) \mapsto (y_1, \dots, y_r, y_{r+1} - g_{r+1}(y_1, \dots, y_r), \dots, y_n - g_n(y_1, \dots, y_r))$$

cela conclut la démonstration. □

**Deux cas particuliers :**

**4.1.13 Corollaire:** (Forme normale et immersions) Soient  $U, V$  des ouverts de  $\mathbb{R}^m$  et  $\mathbb{R}^n$  et  $\varphi : U \rightarrow V$  de classe  $C^k, k \geq 1$ , et  $d\varphi(x_0)$  est de rang  $m$ . Alors il existe des ouverts  $U'$  et  $V'$  contenant  $x_0$  et  $y_0 = \varphi(x_0)$  et des applications  $\rho \in C^k, k(U', U')$  et  $\sigma \in C^k(V', V)$  telle que  $\sigma \circ \varphi \circ \rho(x_1, \dots, x_m) = (x_1, \dots, x_m, 0, 0, \dots, 0)$ .

**4.1.14 Corollaire:** (*Forme normale des submersions*) Soient  $U, V$  des ouverts de  $\mathbb{R}^m$  et  $\mathbb{R}^n$  et  $\varphi : U \rightarrow V$  de classe  $C^k, k \geq 1$ , et  $d\varphi(x_0)$  est de rang  $n$ . Alors il existe des ouverts  $U'$  et  $V'$  contenant  $x_0$  et  $y_0 = \varphi(x_0)$  et des applications  $\rho \in C^k, k(U', U')$  et  $\sigma \in C^k(V', V)$  telle que  $\sigma \circ \varphi \circ \rho(x_1, \dots, x_m) = (x_1, \dots, x_n)$ .

**Preuve:** des deux corollaires : Il suffit de remarquer que le rang est semi-continu inférieurement, et que si ce rang est maximal, il reste maximal dans un voisinage. On peut alors appliquer le théorème du rang constant.  $\square$

**4.1.15 Exemple:** Le cercle  $\mathbb{S}^1$  est une sous-variété de dimension 1 (on va le montrer de différentes manières).

**Preuve:** Moyen 1 : on utilise la définition initiale, (A) : pour le cercle, le difféomorphisme

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$(r \cos(\theta), r \sin(\theta)) \longmapsto (\theta, r - 1); 0 \leq \theta < 2\pi; r \geq 0$$

convient, car pour  $r = 1$ ,  $\varphi(\cos(\theta), \sin(\theta)) = (\theta, 0)$ . De plus, il est facile de vérifier qu'en chaque point  $x_0$ , il existe un ouvert  $U_{x_0}$  tel que  $\varphi$  soit un difféomorphisme.

Moyen 2 : par les nappes paramétrés, (D) : il s'agit à peu près de la même chose qu'au dessus, mais dans le sens inverse.

Soit  $j(t) = (\cos(t), \sin(t))$ . Elle vérifie  $j'(t) = (-\sin(t), \cos(t)) \neq 0, \forall t$  et  $j$  est bijective et propre pour  $0 \leq t < 2\pi$ .

Moyen 3 : par le graphe, (C) : on sait que l'équation d'un cercle est  $x^2 + y^2 = 1$ . Donc on peut exprimer  $y$  en fonction de  $x$  :  $y = f_1(x) = \sqrt{1 - x^2}$  si  $y \geq 0$  et  $y = f_2(x) = -\sqrt{1 - x^2}$  si  $y \leq 0$ . Donc  $f_1$  et  $f_2$  sont bien localement le graphe du cercle sur les zones  $y > 0$  et  $y < 0$  respectivement. Cependant, autour du point  $y = 0$ , on a un problème (théorème d'inversion locale), on doit donc, plutôt exprimer  $x$  en fonction de  $y$  :  $x = f_3(y) = \sqrt{1 - y^2}$  si  $x \geq 0$  et  $x = f_4(y) = -\sqrt{1 - y^2}$  si  $x \leq 0$ . Ceci permet de conclure.  $\square$

**4.1.16 Exemple:** La sphère  $\mathbb{S}^2$  est une sous-variété de dimension 2 (on va le montrer de différentes manières).

**Preuve:** Moyen 1 : on utilise la définition initiale, (A) : Pour la sphère, on choisit :

$$\varphi_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

$$(\cos(\theta)\cos(\varphi), \cos(\theta)\sin(\varphi), \sin(\theta)) \longmapsto (\theta, \phi, r - 1); 0 \leq \varphi < 2\pi, -\frac{\pi}{2} \leq \theta < \frac{\pi}{2}, r \geq 0$$

et

$$\varphi_2 : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

$$(r \sin(\theta)\cos(\varphi), r \sin(\theta)\sin(\varphi), r \cos(\theta)) \longmapsto (\theta, \phi, r - 1); 0 \leq \varphi < 2\pi, 0 \leq \theta < \pi, r \geq 0$$

Ces 2 difféomorphismes conviennent tant que l'on est pas dans un voisinage de  $\theta = \pm \frac{\pi}{2}$  pour  $\varphi_1$  et de  $\theta = 0$  ou  $\theta = \pi$  pour  $\varphi_2$ .

Moyen 2, par les nappes paramétrées :

$$j_1(\theta, \varphi) = (\cos(\theta)\cos(\varphi), \cos(\theta)\sin(\varphi), \sin(\theta)); 0 \leq \varphi < 2\pi, -\frac{\pi}{2} \leq \theta < \frac{\pi}{2}$$

et

$$j_2(\theta, \varphi) = (\sin(\theta)\cos(\varphi), \sin(\theta)\sin(\varphi), \cos(\theta)); 0 \leq \varphi < 2\pi, 0 \leq \theta < \pi$$

Ces 2 fonctions sont bijectives et propres, et on vérifie par le calcul que les dérivées partielles sont linéairement indépendantes, sauf aux pôles de chacune des 2 paramétrisations.

Moyen 3, par le graphe : l'équation de la sphère est  $x^2 + y^2 + z^2 = 1$ , soit  $z = f_1(x, y) = \sqrt{1 - x^2 - y^2}$  si  $z \geq 0$  et  $z = f_2(x, y) = -\sqrt{1 - x^2 - y^2}$  si  $z \leq 0$ . Là encore, ces graphes ne sont pas valables au voisinage du point  $z = 0$ , on peut donc choisir  $x = f_3(y, z) = \sqrt{1 - y^2 - z^2}$  si  $x \geq 0$  et  $x = f_4(y, z) = -\sqrt{1 - y^2 - z^2}$  si  $x \leq 0$ . Enfin, si  $x$  et  $z$  sont tous deux proches de 0, on doit définir  $y = f_5(x, z) = \sqrt{1 - x^2 - z^2}$  si  $y \geq 0$  et  $y = f_6(x, z) = -\sqrt{1 - x^2 - z^2}$  si  $y \leq 0$ . Cela permet de conclure.  $\square$

On va maintenant généraliser ces 2 résultats pour toutes les sphères, en utilisant la définition (B), de loin la plus simple à utiliser :

**4.1.17 Exemple:** La sphère  $\mathbb{S}^{n-1}$  est une sous-variété de  $\mathbb{R}^n$ .

**Preuve:** L'équation de la sphère de dimension  $n - 1$  est  $\sum_{i=1}^n x_i^2 = 1$ . Soit  $F(x_1, \dots, x_n) = (\sum_{i=1}^n x_i^2) - 1$ .  $\frac{\partial F}{\partial x_i} = 2x_i$ .  $dF(x_1, \dots, x_n) = 2(x_1, \dots, x_n) \neq 0$  car au moins l'un des  $x_i$  est non nul (car  $(0, 0, \dots, 0)$  n'est pas un point de la sphère).  $\square$

**4.1.18 Definition:** Le tore de révolution  $T^2$  est le donut, c'est à dire le graphe de la fonction

$$f(u, v) = ((r + R\cos(u))\cos(v), (r + R\cos(u))\sin(v), R\sin(u)); 0 \leq u < 2\pi; 0 \leq v < 2\pi$$

avec  $r, R > 0; r > R$ .

**4.1.19 Exemple:** Le tore

$$j : (\theta, \varphi) \mapsto ((r - \rho\cos(\theta))\cos(\varphi), (r - \rho\cos(\theta))\sin(\varphi), \rho\sin(\theta))$$

où  $\rho < r$  est une sous-variété de dimension 2.

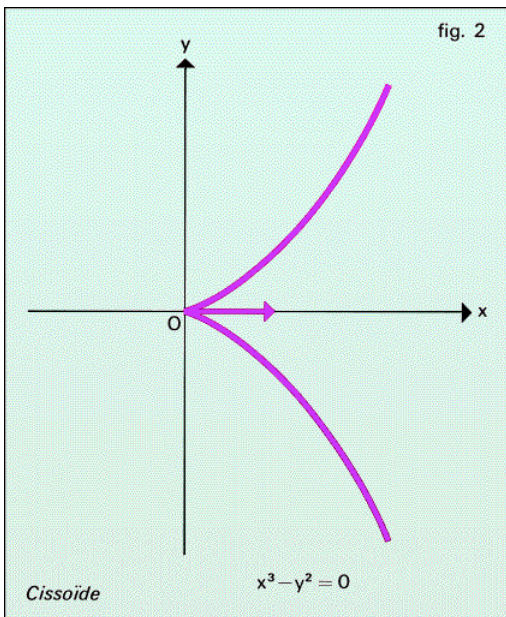
**Preuve:** Pour montrer cela, on va utiliser les nappes paramétrées : En effet, les vecteurs

$$\frac{\partial j}{\partial \theta}(\theta, \varphi) = (\rho\cos(\varphi)\sin(\theta), \rho\sin(\varphi)\sin(\theta), \rho\cos(\theta))$$

$$\frac{\partial j}{\partial \varphi}(\theta, \varphi) = (-(r - \rho\cos(\theta))\sin(\varphi), (r - \rho\cos(\theta))\cos(\varphi), 0)$$

sont linéairement indépendants. De plus  $j$  est localement bicontinue (et propre). Soit en effet  $j(\theta_0, \varphi_0) = (x_0, y_0, z_0)$ . Si  $z$  est dans un intervalle de longueur inférieure à  $1/2$ , il existe une unique détermination continue de  $\theta(z) = \arcsin(\frac{z}{\rho})$  égale à  $\theta_0$  en  $z_0$ . Comme  $\rho < r$ , on a  $r - \rho\cos(\theta(z)) > 0$  et en choisissant une détermination continue de  $\varphi(y, z) = \arcsin(\frac{y}{r - \rho\cos(\theta(z))})$  telle que  $\varphi(y_0, z_0) = \varphi_0$ , on obtient une application continue  $(x, y, z) \mapsto (\theta(z), \varphi(y, z))$  qui envoie un voisinage de  $(x_0, y_0, z_0)$  du tore sur un voisinage de  $(\theta_0, \varphi_0)$ .  $\square$

**4.1.20 Exemple:** Le support de la courbe  $\gamma(t) = (t^2, t^3)$  n'est pas une sous-variété.



**Preuve:** On utilise les coordonnées cartésiennes :  $x = t^2, y = t^3$  de sorte que  $x^3 = y^2$ . Soit  $F(x, y) = x^3 - y^2 : \nabla F(x, y) = (3x^2, 2y) = 0$  si  $(x, y) = (0, 0)$ . C'est donc ce point qui pose problème.  $\square$

**4.1.21 Exemple:** Soit  $\gamma(t) = (x_1(t), \dots, x_n(t))$  une courbe  $C^1$ . Si la dérivée de l'une des coordonnées ne s'annule jamais, le support de  $\gamma$  est une sous-variété.

**Preuve:** Si la dérivée de l'une des coordonnées ne s'annule jamais, on peut inverser cette fonction (c'est un difféomorphisme de  $\mathbb{R}$ ). Supposons que ce soit  $x_1$ . On écrit alors  $y_1 = x_1(t) \Rightarrow t = \varphi(y_1) = x_1^{-1}(y_1)$ , puis  $y_2 = x_2(x_1^{-1}(y_1)), \dots, y_n = x_n(x_1^{-1}(y_1))$  de sorte que  $\gamma$  est solution du système  $(y_2, \dots, y_n) = (x_2(x_1^{-1}(y_1)), \dots, x_n(x_1^{-1}(y_1)))$ . Donc  $\gamma$  est le lieu des zéro de la fonction  $F = (y_2 - x_2(x_1^{-1}(y_1)), \dots, y_n - x_n(x_1^{-1}(y_1)))$ , dont la différentielle est bien surjective. Cette dernière affirmation n'est pas triviale, mais se

vérifie par le calcul de la Jacobienne :  $Jac_F = \begin{pmatrix} -x'_2(x_1^{-1}(y_1)) \cdot \frac{1}{x'(x_1^{-1}(y_1))} & \cdots & \cdots & \cdots & -x'_n(x_1^{-1}(y_1)) \cdot \frac{1}{x'(x_1^{-1}(y_1))} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$

$\square$

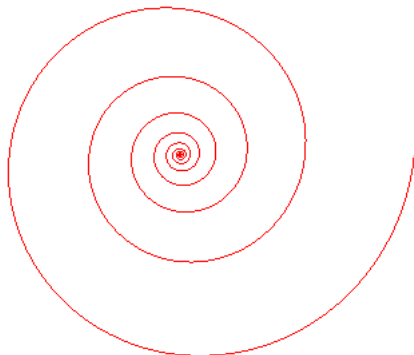
Remarque : cette condition (que l'une des coordonnées ne s'annule jamais) est plus forte que la régularité (une courbe  $\gamma(t)$  est dite régulière si  $\gamma'(t) \neq 0, \forall t$ ).

Cependant :

**4.1.22 Corollaire:** Soit  $\gamma(t) = (x_1(t), \dots, x_n(t))$  une courbe  $C^1$ . Si  $\gamma$  est régulière et que  $\gamma$  a pour ensemble de départ un ouvert de  $\mathbb{R}$ , le support de  $\gamma$  est une sous-variété.

**Preuve:** La preuve précédente permet de montrer (en se restreignant à des ouverts), que si la dérivée de l'une des coordonnées ne s'annule jamais sur un intervalle ouvert  $U$ , le support de  $\gamma$  est une sous-variété sur cet intervalle  $U$ . Or, par définition, d'une courbe régulière,  $\forall t \in \mathbb{R}$ , il existe toujours une des fonction coordonnées de  $\gamma'(t)$  qui est non nulle en  $t$ . En choisissant un ouvert  $U$  tel que cette coordonnée ne s'annule pas, on montre que  $\gamma$  est bien une sous-variété en  $t$ , et par suite,  $\forall t$ .  $\square$

**4.1.23 Exemple:** La spirale logarithmique  $\gamma(t) : \mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto (e^{at} \cos(t), e^{at} \sin(t)), a > 0$  est régulière. De plus, son support (notons le  $S$ ) n'est pas une sous-variété.



**Preuve:**  $\gamma'(t) = (ae^{at}\cos(t) - e^{at}\sin(t), ae^{at}\sin(t) + e^{at}\cos(t))$ .  $\|\gamma'(t)\| = \sqrt{(a^2 + 1)e^{2at}} = e^{at}\sqrt{a^2 + 1} \neq 0, \forall t \in \mathbb{R}$ . Si la norme ne s'annule jamais, cela signifie que  $\gamma'(t)$  n'est jamais égale au vecteur nul, elle est donc bien régulière. Montrons que ce n'est pas une sous variété.

Moyen 1 : pour cela, on va montrer que le paramétrage  $\gamma$  a sa réciproque non continue au voisinage de  $(0, 0)$  :  $\lim_{t \rightarrow -\infty} \gamma(t) = (0, 0)$ . De plus,  $\gamma(t)$  est bijective sur son image (facile à montrer), donc  $\gamma^{-1}$  existe. Calculons cette réciproque :  $(x, y) \in S \Leftrightarrow \exists u \in \mathbb{R}$  tel que  $(x, y) = (e^{au}\cos(u), e^{au}\sin(u))$ . Donc :  $x^2 + y^2 = e^{2au} \Rightarrow u = \frac{\ln(x^2 + y^2)}{2a}$ . De plus, on a vérifié que  $(0, 0) \in S$  (à la limite). Or, au voisinage de  $(0, 0)$ ,  $\gamma^{-1}$  est non continue. On vient donc de montrer que, dans la topologie relative à la spirale, dans un voisinage ouvert de  $(0, 0)$  il n'existe pas de difféomorphisme rectifiant.

Moyen 2 : soit  $j : \mathbb{R} \rightarrow \mathbb{R}^2$ .  $\mathbb{R}$  est ouvert et fermé dans lui même, il en est donc de même de l'image  $j(\mathbb{R})$  dans  $\mathbb{R}^2$  (toujours dans la topologie relative). Or, cette image ne peut être fermée car le point  $(0, 0)$  n'est jamais atteint.

□

On fera donc bien attention avec l'utilisation de la régularité (le problème étant ici que  $\gamma'(t)$  tend vers  $(0, 0)$  quand  $t$  tend vers  $-\infty$ ).

## 4.2 Groupes de matrices comme sous-variétés

Remarque :ici, je calcule l'espace tangent, (que je ne définit que plus loin). Je renvoie à la définition (sous partie : espace tangent).

### 4.2.1 Exemple: $M_n(\mathbb{K})$ est une variété.

**Preuve:** Soit  $A \in M_n(\mathbb{K})$ . On pose :

$$f : M_n(\mathbb{K}) \rightarrow \mathbb{K}^{n^2}$$

$$f : A \mapsto (a_{1,1}, \dots, a_{1,n}, a_{2,1}, \dots, a_{2,n}, \dots, a_{n,1}, \dots, a_{n,n})$$

$f$  est une bijection de  $M_n(\mathbb{K})$  dans  $\mathbb{K}^{n^2}$ , et  $f$  et  $f^{-1}$  : sont clairement des homéomorphismes, donc  $f$  est un difféomorphisme. Comme on a besoin que d'une seule carte, on peut même être plus précis en écrivant :  $M_n(\mathbb{K}) \cong \mathbb{K}^{n^2}$  (les 2 espaces sont isomorphes). □

### 4.2.2 Exemple: $GL_n(\mathbb{R})$ est une sous-variété de dimension $n^2$ de $M_n(\mathbb{R})$ .

**Preuve:** Car le déterminant est une application polynomiale, donc continue, et comme  $GL_n(\mathbb{R}) = \det^{-1}(\mathbb{R}^*)$ , et que la différentielle du déterminant ne s'annule pas, alors  $GL_n(\mathbb{R})$  est une sous-variété. □

### 4.2.3 Exemple: $SL_n(\mathbb{R})$ est une sous-variété de dimension $n^2 - 1$ de $M_n(\mathbb{R})$ .

**Preuve:** Soit :

$$f : M_n(\mathbb{R}) \rightarrow \mathbb{R}$$

$A \mapsto \det(A)$ .

La différentielle de  $f$  en  $A_0$  existe, puisque les dérivées partielles de  $f$  existent et sont clairement continues. Tout d'abord, en  $A_0 = Id$ . Pour cela, on calcule les dérivées directionnelles :

$$df(Id)(X) = \lim_{n\lambda \rightarrow 0} \frac{\det(I + \lambda X) - 1}{\lambda} = \text{tr}(X)$$

En effet,

$$\det(I + \lambda X) = \sum_{\sigma \in S_n} \text{signe}(\sigma) \prod_{i=1}^n (Id_{i,\sigma(i)} + \lambda x_{i,\sigma(i)}) = 1 + \lambda \text{tr}(X) + o(\lambda)$$

, d'où la formule. Maintenant, soit  $A_0$  inversible, calculons les dérivées directionnelles :

$$\begin{aligned} df(A_0)(X) &= \lim_{n\lambda \rightarrow 0} \frac{\det(A_0 + \lambda X) - \det(A_0)}{\lambda} = \lim_{\lambda \rightarrow 0} \det(A_0) \frac{\det(I + \lambda A_0^{-1} X) - 1}{\lambda} \\ &= \det(A_0) \text{tr}(A_0^{-1} X) = \text{tr}(\det(A_0) A_0^{-1} X) = \text{tr}(\tilde{A}_0 X) \end{aligned}$$

en posant  $\tilde{A} = \det(A) A^{-1}$ .

L'application  $df : M_n(\mathbb{R}) \rightarrow \mathbb{R}^*$  est continue ( $f$  est  $C^1$  et coïncide sur  $GL_n(\mathbb{R})$  avec l'application continue définie par :  $X \mapsto \text{tr}(\tilde{A}X)$ ). La densité de  $GL_n(\mathbb{R})$  permet alors d'énoncer :  $df(A_0)(X) = \text{tr}(\tilde{A}_0 X) \forall A_0 \in M_n(\mathbb{R})$ . Enfin,  $SL_n(\mathbb{R}) = f^{-1}(1)$ , et si  $A \in GL_n(\mathbb{R})$ ,  $df(A)(x) = \text{tr}(A^{-1}X)$  : cette application est clairement surjective car c'est une forme linéaire non nulle :  $\text{tr}(XY) = 0 \forall Y \Rightarrow X = O_n$ . Le groupe  $SL_n(\mathbb{R})$  est donc une sous-variété de  $M_n(\mathbb{R})$ , de dimension  $n^2 - 1$  ( $\text{tr}X = 0$  est une équation dans linéaire dans  $M_n(\mathbb{R})$ , donc l'espace engendré est un hyperplan). Enfin, l'espace tangent en  $Id$  est l'ensemble des matrices de trace nulle (noyau de  $df(Id)$ ).  $\square$

**4.2.4 Exemple:**  $O_n(\mathbb{R})$  est une sous-variété de  $M_n(\mathbb{R})$ .

**Preuve:** On rappelle que  $O_n(\mathbb{R}) = \{X, {}^t X X = Id\}$ . On considère :

$$f : M_n(\mathbb{R}) \rightarrow S_n(\mathbb{R})$$

$$X \mapsto {}^t X X$$

où  $S_n(\mathbb{R})$  désigne l'espace des matrices symétriques réelles d'ordre  $n$ , dont la dimension est  $\frac{(n+1)n}{2}$ . Calculons  $df(X_0)(M)$  :

$${}^t(X_0 + M)(X_0 + M) = ({}^t X_0 + {}^t M)(X_0 + M) = {}^t X_0 X_0 + {}^t X_0 M + {}^t M X_0 + {}^t M M$$

Donc :  $df(X_0)(M) = {}^t X_0 M + {}^t M X_0$  car  $\|{}^t M M\| = o(\|M\|)$ . Soit  $X_0 \in O_n(\mathbb{R})$ ,  $S \in S_n(\mathbb{R})$  et  $M = \frac{1}{2} X_0 S$  :  $df(X_0)(M) = S$  et donc  $df(X_0)$  est surjective. Le groupe  $O_n(\mathbb{R})$  est donc une sous-variété de dimension  $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$ , et l'espace tangent est  $\ker(df(X_0))$ , c'est à dire l'ensemble des matrices telles que  $X_0^{-1} M + {}^t(X_0^{-1} M) = 0$ . En  $X_0 = Id$ , on obtient l'espace  $\mathfrak{g}$  des matrices antisymétriques.  $\square$

### 4.3 Variété

Avant la lecture des axiomes, je renvoie à l'annexe II, sur les espaces compacts.

#### Axiomes des variétés

Normalement on suppose qu'une variété différentiable soit un espace *de Hausdorff*, ce qui garantit que la limite d'une suite convergente soit unique, et à *base dénombrable d'ouverts*. Enfin, on suppose qu'elle est *paracompact*, (c'est d'ailleurs une conséquence des 2 axiomes précédents) ce qui garantit l'existence d'une partition de l'unité.

L'axiome de séparabilité (Hausdorff) a pour conséquence que toute variété topologique paracompacte (même sans base dénombrable) est métrisable.

Remarque : certains n'utilisent pas forcément tous ces axiomes, mais comprendre d'où viennent ces axiomes, pourquoi on les impose,... serait beaucoup trop long et sort du cadre de ce TIPE. Pour les plus curieux, voici quelques liens qui peuvent éclairer :

[https://fr.wikipedia.org/wiki/Vari%C3%A9t%C3%A9\\_topologique](https://fr.wikipedia.org/wiki/Vari%C3%A9t%C3%A9_topologique)

[https://fr.wikipedia.org/wiki/Axiome\\_de\\_s%C3%A9paration\\_\(topologie\)](https://fr.wikipedia.org/wiki/Axiome_de_s%C3%A9paration_(topologie))

Soit  $M$  un espace topologique.

**4.3.1 Définition:** Une *carte* ou *carte locale* sur  $M$  est un homéomorphisme d'un ouvert de  $M$  vers un ouvert de  $\mathbb{R}^m$ , c'est-à-dire un couple  $(U, \varphi)$  où  $U$  est un ouvert de  $M$ ,  $\varphi : U \rightarrow \mathbb{R}^m$  est un homéomorphisme, et l'image  $V = \varphi(U)$  est un ouvert de  $\mathbb{R}^m$ .

Si  $\varphi : U \rightarrow \mathbb{R}^m$  et  $\varphi' : U' \rightarrow \mathbb{R}^m$  sont deux cartes et  $U \cap U' \neq \emptyset$ , l'application  $\psi = \varphi' \circ \varphi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  s'appelle *changement de carte*. C'est évidemment un homéomorphisme de  $\varphi(U \cap U')$  vers  $\varphi'(U \cap U')$ .

**Définition 2.2.2 :** On appelle *variété différentiable de dimension  $m$*  un espace topologique  $M$  muni d'une famille de cartes  $\mathcal{A} = \{\varphi_r : U_r \rightarrow \mathbb{R}^m\}$  telle que

1. l'ensemble  $\{U_r\}$  est un recouvrement de  $M$  ;
2. les changements de cartes  $\psi_{rs} = \varphi_s \circ \varphi_r^{-1}$  sont des difféomorphismes (i.e. des applications réelles inversibles, différentiables et avec réciproque différentiable) ;
3. la famille est *maximale*, c'est-à-dire qu'elle contient toutes les cartes compatibles entre elles (dans le sens que les changements de cartes sont des difféomorphismes).

Une telle famille s'appelle *atlas (maximale)* sur  $M$ .

**4.3.2 Définition:** Une variété  $M$  s'appelle *différentiable de classe  $C^k$* , ou *de classe  $C^\infty$  (lisse)*, si les changements de cartes  $\psi_{rs}$  sont des difféomorphismes de classe  $C^k$  (i.e. différentiables  $k$  fois avec dernière dérivée continue et leurs réciproques aussi), ou bien de classe  $C^\infty$  (i.e. différentiables autant de fois qu'on veut ainsi que leurs réciproques).

**Remarque :** Au contraire, un espace topologique  $M$  doté d'un atlas de cartes dont les changements de cartes sont des simples homéomorphismes s'appelle *variété topologique*.

**4.3.3 Proposition:** *Le produit cartésien de 2 variétés  $U$  et  $V$ ,  $U \times V$  est une variété.*

**Preuve:** On a des cartes  $\varphi_{1,1}, \dots, \varphi_{1,k}$  de  $U$  et  $\varphi_{2,1}, \dots, \varphi_{2,p}$  de  $V$  sur des ouverts  $U_1, \dots, U_k$  et  $V_1, \dots, V_p$ . On prend  $p$  et  $k$  finis pour simplifier, mais, dans certains cas, on peut avoir  $p, k = \infty$  car il existe des cas où l'on peut avoir un nombre dénombrable (mais infini) de cartes. La démonstration reste valable dans ce cas. Ce sont des homéomorphismes et  $\varphi_{1,i} \circ \varphi_{1,j}^{-1}$ ,  $\varphi_{2,i} \circ \varphi_{2,j}^{-1}$  sont des difféomorphismes pour tout  $i \neq j$ . Donc, soit  $\psi_{i,j} = (\varphi_{1,i}, \varphi_{2,j})$ . C'est un homéomorphisme et  $\psi_{i,j} \circ \psi_{i',j'}^{-1}$  est un difféomorphisme. Il est clair que  $(\psi_{i,j})_{1 \leq i \leq k, 1 \leq j \leq p}$  est un recouvrement. On a donc trouvé un atlas de  $U \times V$ .  $\square$

**4.3.4 Exemple:** *Le cercle est une variété.*

Pour cela, je vais utiliser 2 méthodes. La première utilisera des cartes simples, mais nombreuses. La seconde utilisera des cartes plus compliquées, mais moins nombreuses.

**Preuve:** Moyen :1

Je vais utiliser une première méthode et nous en verrons une autre juste en dessous. Soit

$$\varphi_1(x, y) = x, y > 0$$

$$\varphi_2(x, y) = x, y < 0$$

$$\varphi_3(x, y) = y, x > 0$$

$$\varphi_4(x, y) = y, x < 0$$

Comme  $x^2 + y^2 = 1$ , on peut préciser :

$$\varphi_1(x, \sqrt{1-x^2}) = x$$



$$\varphi_2(x, -\sqrt{1-x^2}) = x$$

$$\varphi_3(\sqrt{1-y^2}, y) = y$$

$$\varphi_4(-\sqrt{1-y^2}, y) = y$$

On vérifie très simplement que ce sont des difféomorphismes sur les ouverts

$$U_1 = \{0 < x < 1\}$$

$$U_2 = \{-1 < x < 0\}$$

$$U_3 = \{0 < y < 1\}$$

$$U_4 = \{-1 < y < 0\}$$

Exprimons leur réciproque :

$$\varphi_1^{-1}(x) = (x, \sqrt{1-x^2})$$

$$\varphi_2^{-1}(x) = (x, -\sqrt{1-x^2})$$

$$\varphi_3^{-1}(y) = (\sqrt{1-y^2}, y)$$

$$\varphi_4^{-1}(y) = (-\sqrt{1-y^2}, y).$$

On va maintenant calculer la composée de  $\varphi_i \circ \varphi_j^{-1}$  pour  $i \neq j$  :

$$\varphi_1 \circ \varphi_3^{-1} = \varphi_1(\sqrt{1-y^2}, y) = \varphi_1(y', \sqrt{1-y'^2}) = y' = \sqrt{1-y^2}$$

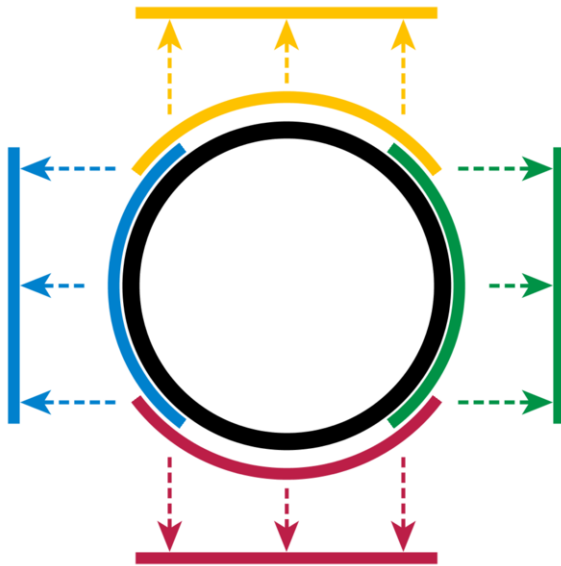
$$\varphi_1 \circ \varphi_4^{-1} = \varphi_1(-\sqrt{1-y^2}, y) = \varphi_1(y', \sqrt{1-y'^2}) = y' = -\sqrt{1-y^2}$$

$$\varphi_2 \circ \varphi_3^{-1} = \varphi_2(\sqrt{1-y^2}, y) = \varphi_2(y', -\sqrt{1-y'^2}) = y' = \sqrt{1-y^2}$$

$$\varphi_2 \circ \varphi_4^{-1} = \varphi_2(-\sqrt{1-y^2}, y) = \varphi_2(y', -\sqrt{1-y'^2}) = y' = -\sqrt{1-y^2}$$

On a des résultats relativement similaires pour le reste des composées. On vérifie également très simplement que ce sont des difféomorphismes sur  $U_i \cap U_j$  pour  $i \neq j$ , et que ces 4 cartes forment un recouvrement.

Voici une illustration des 4 cartes qu'on a choisi :



On a donc trouvé un atlas du cercle. On remarque que cet exemple a été "simple" en ce sens que les cartes choisies sont des projections canoniques. Cette méthode (utiliser des projections canoniques ou non) peut se généraliser avec un grand nombre de figures géométriques. Le seul inconvénient est que le nombre de cartes est d'autant plus élevé que la figure est "compliquée". On va donc donner un autre moyen juste après.

Moyen 2 :

Montrons que  $\mathbb{S}^1$  admet deux cartes locales. On appelle *projection stéréographique* à partir du pôle nord  $N(0, 1)$  sur l'espace  $T_S\mathbb{S}^1$  tangent au pôle sud  $S(0, -1)$ , l'application

$$\varphi_S : U_S = \mathbb{S}^1 \setminus \{N\} \longrightarrow T_S\mathbb{S}^1$$

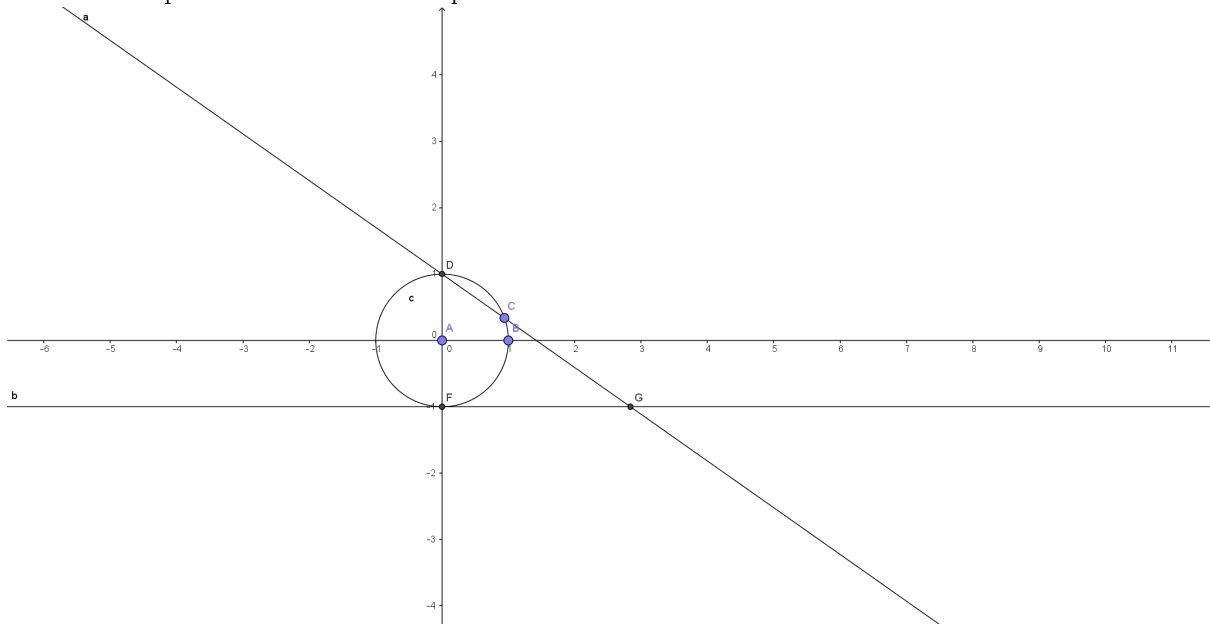
qui associe à tout point  $P \in \mathbb{S}^1, P \neq N$ , le point  $\varphi_S(P)$  obtenu comme projection de  $P$  sur  $T_S\mathbb{S}^1$  le long de la droite  $\Delta_{NP}$  qui jointe  $N$  et  $P$ .

De même, on appelle *projection stéréographique* à partir du pôle sud  $S$  sur l'espace  $T_N\mathbb{S}^1$  tangent au pôle nord, l'application

$$\varphi_N : U_N = \mathbb{S}^1 \setminus \{S\} \longrightarrow T_N\mathbb{S}^1$$

qui associe à tout point  $P \in \mathbb{S}^1, P \neq S$ , le point  $\varphi_N(P)$  obtenu comme projection de  $P$  sur  $T_N\mathbb{S}^1$  le long de la droite  $\Delta_{SP}$  qui jointe  $S$  et  $P$ .

Voici une petite illustration de ce qu'on cherche à faire :



Après calculs, on peut exprimer les deux projections stéréographiques en coordonnées cartésiennes comme suit, où l'on pose  $P(x, y) \in \mathbb{S}^1$  avec  $x^2 + y^2 = 1$  :

$$\varphi_S : U_S \longrightarrow \mathbb{R}, (x, y) \mapsto \varphi_S(x, y) = \frac{2x}{1-y}, \quad y \neq 1 \quad \varphi_N : U_N \longrightarrow \mathbb{R}, (x, y) \mapsto \varphi_N(x, y) = \frac{2x}{1+y}, \quad y \neq -1.$$

En effet :

Soit :  $C = (x, y)$  un point d'un cercle :  $x^2 + y^2 = 1$ . Soit les points  $A = (0, 1)$  et  $B = (0, -1)$ . Le but est de ramener les points du cercle à une droite. Pour avoir une idée précise de ce que l'on cherche à faire, voir le fichier géogébra. On va donner l'équation des droites  $AC$  et  $BC$ . On a  $m_1 = \frac{y-1}{x}$  et  $m_2 = \frac{y+1}{x}$ . Puis, l'équation de la droite  $(AC)$  est  $s = m_1t + p$  ( $s$  remplace le  $y$  habituel et  $t$  le  $x$ ) et celle de  $(BC)$  :  $s = m_2t + p$ . En  $t = 0, s = \pm 1$  : donc  $(AC) : s = \frac{y-1}{x}t + 1$  et  $\frac{y+1}{x}t - 1$ . On a cherché l'intersection avec les droites  $s = -1$  et  $s = 1$  respectivement. On obtient  $\varphi_N(x, y) = t_1 = \frac{2x}{1+y}$  et  $\varphi_S(x, y) = t_2 = \frac{2x}{1-y}$ . On a  $\varphi_i : \mathbb{S}^1 \longrightarrow \mathbb{R}$  continue et différentiable.

Montrons que la famille  $\{(U_S, \varphi_S), (U_N, \varphi_N)\}$  définit bien une structure de variété différentielle sur  $\mathbb{S}^1$ .

— Les ensembles

$$U_S = \mathbb{S}^1 \setminus \{N\} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y \neq 1\}$$

$$U_N = \mathbb{S}^1 \setminus \{S\} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y \neq -1\}$$

sont ouverts, et la famille  $\{U_S, U_N\}$  est bien un recouvrement de  $\mathbb{S}^1$ , i.e.  $U_S \cup U_N = \mathbb{S}^1$ .

- Les applications  $\varphi_S$  et  $\varphi_N$  sont des homéomorphismes, i.e. continues, inversibles et avec inverses continues, car admettent les applications réciproques :

$$\varphi_S^{-1} : \mathbb{R} \longrightarrow U_S, u \mapsto \varphi_S^{-1}(t) = \left( \frac{4t}{t^2+4}, \frac{t^2-4}{t^2+4} \right)$$

$$\varphi_N^{-1} : \mathbb{R} \longrightarrow U_N, v \mapsto \varphi_N^{-1}(t) = \text{a déterminer}$$

En effet :

On va maintenant calculer les application réciproques :

$$\begin{aligned} \varphi_N(x, y) = \frac{x}{1-y} = t \Rightarrow 4x^2 &= (1-y)^2 t^2 \Rightarrow 4 - 4y^2 = (1-2y+y^2)t^2 \\ &\Rightarrow y^2(4+t^2) - 2yt^2 + (t^2-4) = 0 \end{aligned}$$

c'est une équation du second degré en  $y$  :  $\Delta = 4t^4 - 4(t^2+4)(t^2-4) = 64 = 8^2$ . Donc, on a 2 solutions :  $y_1 = \frac{2t^2+8}{2(4+t^2)} = 1$ , ce qui est exclu car  $y = 1$  est une valeur interdite pour  $\varphi_N$  et

$$y_2 = \frac{2t^2-8}{2(4+t^2)} = \frac{t^2-4}{4+t^2}$$

Puis

$$x^2 = 1 - y^2 = 1 - \frac{(t^2-4)^2}{(4+t^2)^2} = \frac{(t^2+4)^2 - (t^2-4)^2}{(t^2+4)^2} = \frac{16t^2}{(t^2+4)^2}$$

Donc  $x = \pm \frac{4t}{t^2+4} = \frac{4t}{t^2+4}$  car  $t$  est une variable sur  $\mathbb{R}$ .

Puis, de la même manière :

$$\frac{2x}{y+1} = t \Rightarrow 4x^2 = 4 - 4y^2 = (y+1)^2 t^2 \Rightarrow y^2(t^2+4) + 2yt^2 + (t^2+4)$$

On vérifie par symétrie qu'on a le même  $\Delta$  que l'équation précédente. Donc

$$y_1 = \frac{-2t^2+8}{2(4+t^2)} = \frac{4-t^2}{4+t^2}$$

, l'autre solution  $y_2 = \frac{-2t^2-8}{2(4+t^2)} = -1$  est exclue. Puis

$$x^2 = 1 - y^2 = 1 - \frac{(4-t^2)^2}{(4+t^2)^2} = \frac{(4+t^2)^2 - (4-t^2)^2}{(4+t^2)^2} = \frac{16t^2}{(t^2+4)^2}$$

On a donc :  $\varphi_N^{-1} = \left( \frac{4t}{t^2+4}, \frac{t^2-4}{4+t^2} \right)$  et  $\varphi_S^{-1} = \left( \frac{4t}{t^2+4}, \frac{4-t^2}{4+t^2} \right)$

À noter qu'on a  $\varphi_S(S) = \varphi_S(0, -1) = 0$  et  $\varphi_N(N) = \varphi_N(0, 1) = 0$ .

- Les transitions de cartes  $\varphi_S \circ \varphi_N^{-1}$  et  $\varphi_N \circ \varphi_S^{-1}$  sont des difféomorphismes de classe  $C^\infty$ , ce qui garantit que le cercle  $S^1$  soit une variété lisse. En effet, les applications  $\varphi_S$  et  $\varphi_N$  restreintes à l'intersection

$$U_S \cap U_N = \mathbb{S}^1 \setminus \{N, S\} = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y \neq \pm 1 \right\}$$

sont surjectives sur l'ouvert  $\mathbb{R}^* = ]-\infty, 0[ \cup ]0, \infty[$ . On a alors

$$\varphi_N \circ \varphi_S^{-1} : \mathbb{R}^* \longrightarrow \mathbb{R}^*,$$

et après calculs, on trouve :

$$\varphi_N \circ \varphi_S^{-1} = \frac{2 \frac{4t}{t^2+4}}{1 - \frac{4-t^2}{4+t^2}} = \frac{8t}{2t^2} = \frac{4}{t}$$

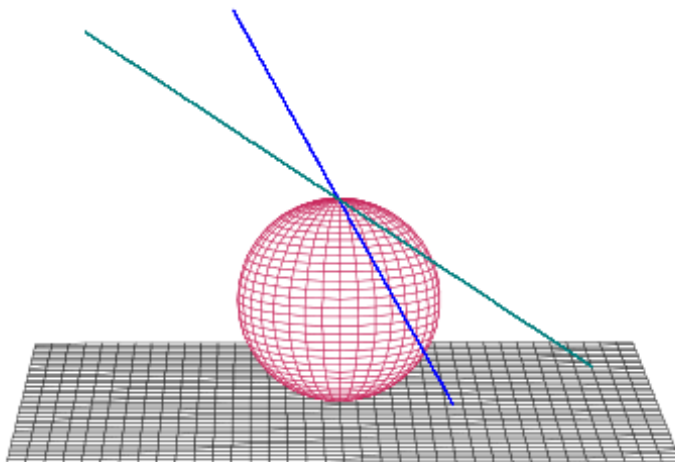
$$\varphi_N \circ \varphi_S^{-1}(t) = \frac{4}{t}, \quad u \neq 0.$$

Cette application est bien un difféomorphisme de classe  $C^\infty$ .

**4.3.5 Exemple:** *La sphère est une variété.*

**Preuve:** Analogue au cercle : on utilise encore la projection stéréographique.

Une fois de plus, voici une image :



Soit  $S = (x, y, z)$  un point de la sphère,  $A = (0, 0, 1)$  et  $B = (0, 0, -1)$ . On détermine les droites (AS) et (BS) en paramétrique.  $x(t) = xt, y(t) = yt, z_1(t) = 1 + (z - 1)t, z_2(t) = -1 + (z + 1)t$ . On détermine l'intersection de ces droites avec  $z = -1$  et  $z = 1$  respectivement. Soit : si  $z = -1, t = \frac{2}{1-z}, x(t) = \frac{2x}{1-z}, y(t) = \frac{2y}{1-z}$ . Soit  $\varphi_1(x, y, z) = (\frac{2x}{1-z}, \frac{2y}{1-z})$ . Si  $z = 1, \frac{2}{1+z}, x(t) = \frac{2x}{1+z}, y(t) = \frac{2y}{1+z}$ . Soit  $\varphi_2(x, y, z) = (\frac{2x}{1+z}, \frac{2y}{1+z})$ . De plus, rappelons que  $x^2 + y^2 + z^2 = 1$ . Déterminons les réciproques :

$$\left(\frac{2x}{1-z}, \frac{2y}{1-z}\right) = (s, t) \Rightarrow 4x^2 = s^2(1-z)^2(1), 4y^2 = t^2(1-z)^2(2) \Rightarrow (1)+(2) : 4(1-z^2) = (s^2+t^2)(1-2z+z^2)$$

$$\Rightarrow z^2(4 + s^2 + t^2) - 2z(s^2 + t^2) + (s^2 + t^2 - 4) = 0$$

Là encore, on utilise  $\Delta = 4(s^2 + t^2) - 4(4 + s^2 + t^2)(s^2 + t^2 - 4) = 64 = 8^2$ . Donc

$$z = \frac{2(s^2 + t^2) - 8}{2(s^2 + t^2 + 4)} = \frac{s^2 + t^2 - 4}{4 + s^2 + t^2}$$

et  $x = \frac{(1-z)s}{2} = \frac{4s}{4+s^2+t^2}$  et  $y = \frac{4t}{4+s^2+t^2}$ . Puis, pour  $\varphi_2$  : on trouve après calcul :  $z = \frac{4-s^2-t^2}{4+s^2+t^2}, x = \frac{(1+z)s}{2} = \frac{4s}{4+s^2+t^2}$  et  $y = \frac{4t}{4+s^2+t^2}$ . Soit

$$\varphi_1^{-1}(s, t) = \left(\frac{4s}{4 + s^2 + t^2}, \frac{4t}{4 + s^2 + t^2}, \frac{s^2 + t^2 - 4}{4 + s^2 + t^2}\right)$$

et

$$\varphi_2^{-1}(s, t) = \left(\frac{4s}{4 + s^2 + t^2}, \frac{4t}{4 + s^2 + t^2}, \frac{4 - s^2 - t^2}{4 + s^2 + t^2}\right)$$

$\varphi_1$  est un homéomorphisme,  $\varphi_2$  également, les ouverts sur lesquels sont définis (il s'agit des ouverts  $U_1 = \{(x, y, z), x^2 + y^2 + z^2 = 1, z \neq 1\}$  et  $U_2 = \{(x, y, z), x^2 + y^2 + z^2 = 1, z \neq -1\}$ )  $\varphi_1$  et  $\varphi_2$  forment un recouvrement un enfin :

$$\varphi_1 \circ \varphi_2^{-1}(s, t) = \left(\frac{2\frac{4s}{4+s^2+t^2}}{1 - \frac{4-s^2-t^2}{4+s^2+t^2}}, \frac{2\frac{4t}{4+s^2+t^2}}{1 - \frac{4-s^2-t^2}{4+s^2+t^2}}\right)$$

$$= \left( \frac{\frac{8s}{4+s^2+t^2}}{\frac{2(s^2+t^2)}{4+s^2+t^2}}, \frac{\frac{2 \cdot \frac{4t}{4+s^2+t^2}}{1 - \frac{4-s^2-t^2}{4+s^2+t^2}}}{\frac{4s}{s^2+t^2}, \frac{4t}{s^2+t^2}} \right)$$

C'est un difféomorphisme, de même que  $\varphi_2 \circ \varphi_1^{-1}(s, t)$  □

On va maintenant prouver que le tore  $T^2$  est une variété. Mais avant cela, il convient de le définir un peu plus précisément.

**4.3.6 Proposition:** *Le tore  $T^2$  est homéomorphe au produit cartésien  $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$ .*

**Preuve:**

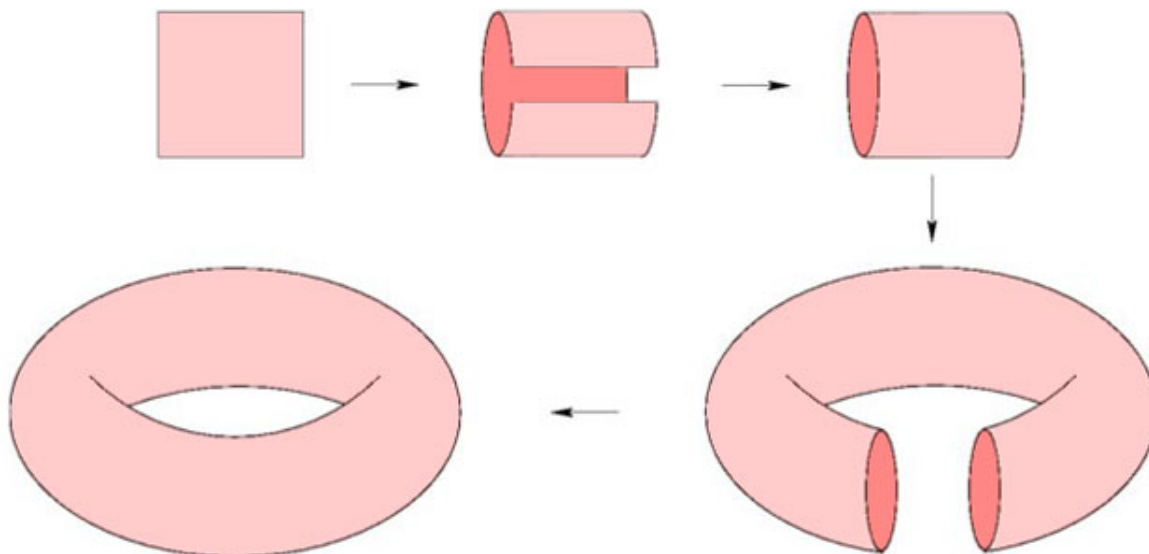
$$f(u, v) = ((r + R\cos(u))\cos(v), (r + R\cos(u))\sin(v), R\sin(u)); 0 \leq u < 2\pi; 0 \leq v < 2\pi$$

avec  $r, R > 0$ ;  $r > R$  est un paramétrage du tore. On peut écrire :

$$g(u, v) = ((r + R\cos(2\pi u))\cos(2\pi v), (r + R\cos(2\pi u))\sin(2\pi v), R\sin(2\pi u)); 0 \leq u < 1; 0 \leq v < 1$$

(qui est un paramétrage qui représente lui aussi le tore). Il est facile de vérifier que  $g$  est un homéomorphisme sur son image. Or,  $(u, v) \in \mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$ . □

**Remarque :**  $T^2$  est un donut, et  $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$  est un carré (si on le représente dans  $\mathbb{R}^2$ ). On dit que le tore de révolution est un plongement du tore carré dans  $\mathbb{R}^3$  (d'ailleurs, on l'a déjà prouvé implicitement lorsqu'on a utilisé les nappes paramétrées pour montrer que le tore est une sous variété).



**4.3.7 Proposition:**  $\mathbb{R}/\mathbb{Z}$  et  $\mathbb{S}^1$  sont homéomorphes.

**Preuve:** La fonction

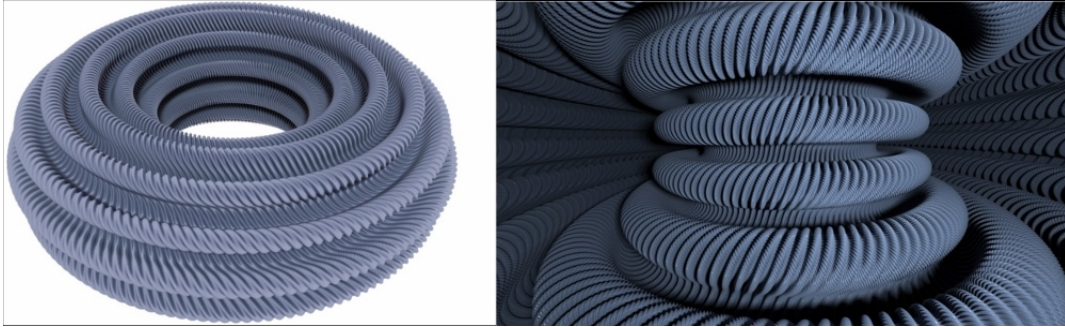
$$f : \mathbb{R}/\mathbb{Z} \longrightarrow \mathbb{S}^1, t \longmapsto (\cos(2\pi t), \sin(2\pi t))$$

est un homéomorphisme sur son image, donc  $\mathbb{R}/\mathbb{Z} \cong \mathbb{S}^1$ . □

De façon générale, on définit ainsi le tore à  $n$  trous :

**4.3.8 Définition:** Le tore à  $n$  trous est le produit cartésien :  $T^n = \mathbb{S}^1 \times \dots \times \mathbb{S}^1$  ( $n$  fois), qui est homéomorphe au produit cartésien  $\mathbb{R}/\mathbb{Z} \times \dots \times \mathbb{R}/\mathbb{Z}$  ( $n$  fois)  $\cong \mathbb{R}^n/\mathbb{Z}^n$ .

**Remarque :** Le tore de révolution  $T^2$ , en tant que nappes paramétrées est un plongement de  $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$  dans  $\mathbb{R}^3$ . Cependant, on peut trouver un plongement encore plus précis (le tore plat), dit isométrique (c'est à dire qu'il préserve les longueurs), de  $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$  dans  $\mathbb{R}^3$  :



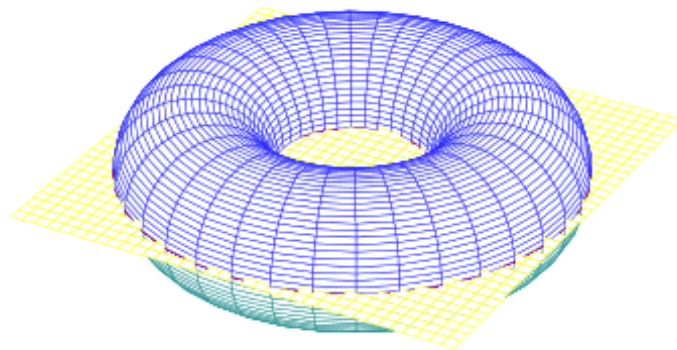
**4.3.9 Exemple:** *Le tore  $T^2$  (de révolution) est une variété. (tore de rayons  $R$  et  $r, R > r$ )*

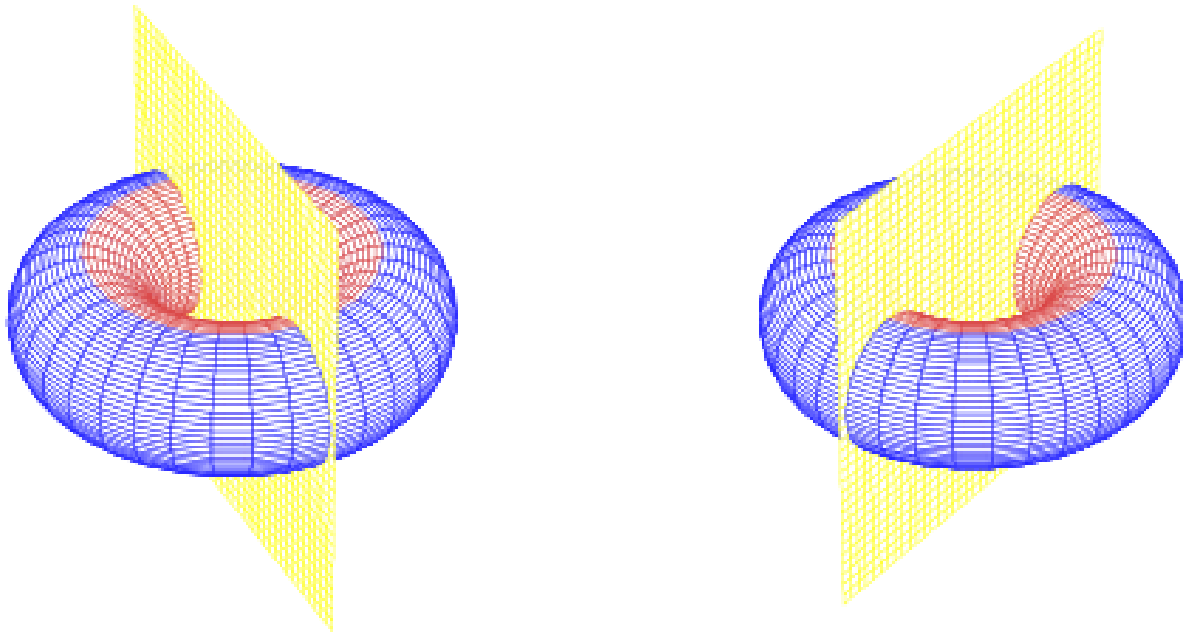
Moyen 1 :

En utilisant la paramétrisation du tore :

Petite explication pour la preuve : on coupe le tore avec le plan  $z = 0$  : on choisit alors une première carte qui envoie la partie supérieure (strictement) par rapport au plan  $z = 0$  sur le plan  $z = 0$ , puis une autre carte qui envoie la partie inférieure (strictement) par rapport au plan  $z = 0$ . Il reste alors l'intersection du tore avec le plan  $z = 0$ , 2 cercles, qui n'ont pas été recouverts. Pour pouvoir les englober, on peut choisir de projeter sur le plan  $y = 0$ . Dans ce cas, il reste alors l'intersection des 2 plans précédents ( $z = 0$  et  $y = 0$ ) avec le tore, ce qui donne 4 points. On projette enfin sur  $x = 0$ .

Voici les images :





**Preuve:** On utilise  $\varphi_i(x, y, z) = (x, y)$  ( $1 \leq i \leq 2$ ) sur les ouverts  $z > 0, z < 0$  et  $\psi_j(x, y, z) = (x, z)$  ( $1 \leq j \leq 4$ ) sur les ouverts  $x^2 + y^2 < R, y < 0, x^2 + y^2 < R, y > 0, x^2 + y^2 > R, y < 0$  et  $x^2 + y^2 > R, y > 0$ , et enfin :  $\phi_k(x, y, z) = (y, z)$  ( $1 \leq k \leq 4$ ) sur les ouverts  $x^2 + y^2 < R, y < 0, x^2 + y^2 < R, y > 0, x^2 + y^2 > R, y < 0$  et  $x^2 + y^2 > R, y > 0$

qui permettent de recouvrir l'ensemble du tore. Les  $\varphi_i, \psi_j, \phi_k$  sont des difféomorphismes, de même que leur composée. Ce qui permet de choisir ces difféomorphismes est la forme du tore. Il est également évident qu'ils forment un recouvrement.

Moyen 2 :

On a vu que le tore  $T^2$  est homéomorphe à  $\mathbb{S}^1 \times \mathbb{S}^1$ , qui est un produit cartésien de 2 variétés, alors c'est une variété. Le nombre de cartes est alors de 4 (on a un minimum de 2 cartes  $\varphi_1, \varphi_2$  pour  $\mathbb{S}^1$ , et comme on considère le produit cartésien, on a  $2 \times 2 = 4$  cartes :

$$\psi_1 = (\varphi_1, \varphi_1)$$

$$\psi_2 = (\varphi_1, \varphi_2)$$

$$\psi_3 = (\varphi_2, \varphi_1)$$

$$\psi_4 = (\varphi_2, \varphi_2)$$

□

### Problèmes de bord, théorie de la mesure

Avant de continuer, voici une remarque importante : l'intervalle ouvert  $]0, 1[$  est clairement une variété différentielle et une sous-variété de  $\mathbb{R}$ , tandis que ce n'est pas le cas de l'intervalle fermé  $[0, 1]$ , alors que celui-ci ne diffère du précédent que pour deux points, qui sont de mesure nulle dans  $\mathbb{R}$ . Où est le problème ?

**4.3.10 Lemme:** *L'intervalle  $[0, 1]$  n'est pas une variété différentielle.*

**Preuve:** Et bien, le problème est qu'il n'existe aucun recouvrement de  $[0, 1]$  formé d'ouverts. En particulier, il n'existe aucun voisinage des points de bord 0 et 1 qui soit ouvert dans  $[0, 1]$ .

Pour le vérifier, il faut d'abord préciser que l'intervalle  $[0, 1]$  hérite la topologie de  $\mathbb{R}$  : un sous-ensemble de  $[0, 1]$  est ouvert dans  $[0, 1]$  s'il est ouvert dans  $\mathbb{R}$ .

Par conséquent, un voisinage ouvert de 1 est nécessairement de la forme  $]a, b[$  avec  $a < 1$  et  $b > 1$  [ou bien choisis  $[1 - \delta, 1 + \epsilon[$  avec  $\delta, \epsilon > 0$ , mais pas le mélange  $]a, 1 + \epsilon[$  avec  $0 \leq a < 1$  et  $\epsilon > 0$ ], mais cet intervalle n'est pas contenu dans  $[0, 1]$  même si  $a > 0$  et  $b$  est aussi petit qu'on veut.  $\square$

**4.3.11 Lemme:** *L'intervalle  $[0, 1]$  n'est pas une sous-variété de dimension 1 de  $\mathbb{R}$ .*

**Preuve:** Montrons que autour du point de bord 1 il n'existe aucun ouvert  $U$  de  $\mathbb{R}$  et aucun difféomorphisme  $\varphi : U \rightarrow \mathbb{R}$  tel que

$$\varphi(U \cap [0, 1]) = \mathbb{R} \cap \varphi(U) = \varphi(U)$$

En effet, montrons qu'il y a un problème au voisinage de 0 (il en est de même en 1). Soit  $V$  un voisinage ouvert de 0.  $V$  est donc de la forme  $]a, b[$  avec  $a < 0; b > 0$ . Or,  $V \cap ]0, 1[$  n'est pas ouvert (fermé si  $b \geq 1$ , semi-ouvert si  $b < 1$ ).

Montrons par l'absurde que  $[0, 1]$  ne peut pas être une sous variété : s'il existait un difféomorphisme  $\varphi$  tel que vérifiant  $\varphi(V \cap [0, 1]) = \varphi(V)$ , alors :  $\varphi(V)$  serait ouvert, alors que  $\varphi(V \cap [0, 1]) = \varphi(V)$  ne le serait pas, car soit  $V \cap [0, 1] = [0, 1] \Rightarrow \varphi(V \cap [0, 1])$  est fermé, soit  $V \cap [0, 1] = [0, a[ \Rightarrow \varphi(V \cap [0, 1])$  n'est pas ouvert.

Cela car :

$$\varphi([0, a]) = \varphi([0, a - \alpha] \cup ]a - \alpha, a]) = \varphi(V_1 \cup V_2) = \varphi(V_1) \cup \varphi(V_2)$$

, (la dernière égalité vient du fait que  $V_1$  et  $V_2$  sont disjoints. De plus, on précise que  $0 < \alpha < a$ . Le fait que  $\varphi$  soit un difféomorphisme sur  $V$  implique alors que  $\varphi(V_1)$  et  $\varphi(V_2)$  sont disjoints, et on déduit que  $\varphi(V_1)$  est fermé tandis que  $\varphi(V_2)$  est ouvert, par suite,  $\varphi(V \cap [0, 1])$  n'est pas ouvert. Or l'égalité entre un ensemble ouvert et un ensemble non ouvert (inclus dans un espace plus grand) est impossible, ce qui permet de conclure.  $\square$

Ce problème a poussé les mathématiciens à introduire les notions de *variétés* et *sous-variétés à bord* (que je ne vais pas définir), qui s'appliquent aux exemples comme l'intervalle  $[0, 1]$ .

## 4.4 Calcul différentiel sur une variété

Avec la notion de variété, on n'est obligé de redéfinir un certain nombre de notions analytiques, qui, malgré qu'elles gardent plus ou moins les mêmes propriétés, nécessitent qu'on inclue le concept de cartes.

### Applications différentiables et difféomorphismes

Soient  $M$  et  $N$  deux variétés différentiables de dimension  $m$  et  $n$ .

**4.4.1 Définition:** Pour toute application  $\phi : M \rightarrow N$ , on appelle *expression locale* de  $\phi$  dans les cartes  $\varphi : U \rightarrow \mathbb{R}^m$  de  $M$  et  $\varphi' : U' \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$  de  $N$  la composée  $\tilde{\phi} = \varphi' \circ \phi \circ \varphi^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ .

Une application  $\phi : M \rightarrow N$  est *différentiable*, ou *de classe  $C^k$* , ou *de classe  $C^\infty$  (lisse)*, si pour toute carte sur  $M$  et toute carte sur  $N$  l'expression locale  $\tilde{\phi}$  est différentiable au sens des fonctions réelles, ou de classe  $C^k$ , ou de classe  $C^\infty$ . L'ensemble des applications de classe  $C^k$  est noté  $C^k(M, N)$ , celui des applications de classe  $C^\infty$  est noté  $C^\infty(M, N)$ .

**4.4.2 Définition:** Une application  $\phi : M \rightarrow N$  est un *difféomorphisme* si elle est différentiable, inversible et avec réciproque différentiable. Cela signifie que l'application  $\phi$  est un homéomorphisme qui, localement, est une application réelle  $\tilde{\phi}$  avec différentielle  $d\tilde{\phi}$  inversible, c'est-à-dire une matrice carrée de déterminant non nul. Dans ce cas, on dit aussi que les deux variétés  $M$  et  $N$  sont *difféomorphes*. Elles ont donc la même dimension,  $m = n$ . On note  $\text{Diff}(M, N)$  l'ensemble des difféomorphismes de  $M$  vers  $N$ .



### Fonctions réelles sur une variété

Soit  $M$  une variété de dimension  $m$ . On note  $C^\infty(M)$  l'ensemble des fonctions réelles de classe  $C^\infty$  sur  $M$ .

**4.4.3 Définition:** On dit que  $\mathcal{A}$  est une *algèbre associative* sur  $\mathbb{K}$  si :

1.  $\mathcal{A}$  est un  $\mathbb{K}$ -espace vectoriel (avec opération  $+$  et produit par un scalaire).
2.  $\mathcal{A}$  est un anneau pour la multiplication  $\times$  (associative).
3. La multiplication  $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$  est bilinéaire.

On dit que  $\mathcal{A}$  est une *algèbre associative et commutative* si la multiplication  $\times$  est commutative.

**4.4.4 Exemple:** *d'algèbres associatives*

1.  $(\mathbb{R}, +, \cdot)$  et  $(\mathbb{C}, +, \cdot)$  sont des algèbres associatives et commutatives.
2.  $\mathbb{C}^2$  est pas une algèbre associative avec le produit  $(a, b), (a', b') \in \mathbb{C}^2$  par  $(a, b) \cdot (a', b') = (aa', bb')$ .
3.  $M_n(\mathbb{K}), +, \cdot$  est une algèbre associative et non commutative.
4. L'espace des quaternions  $\mathbb{H}$ , défini comme ceci :  $(1, i, j, k)$  est une base de  $\mathbb{H}$ , de sorte que tout élément de  $x \in \mathbb{H}$  vérifie :  $x = \alpha + \beta i + \gamma j + \lambda k$  pour certains coefficients  $\alpha, \beta, \gamma, \lambda \in \mathbb{R}$ . De plus, ces éléments vérifient :

$$i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j.$$

On a alors la somme qui est définie par (si  $x = x_1 + x_2i + x_3j + x_4k, y = y_1 + y_2i + y_3j + y_4k$ )

$$x + y = x_1 + y_1 + i(x_2 + y_2) + j(x_3 + y_3) + k(x_4 + y_4)$$

et le produit par

$$\begin{aligned} x \cdot y &= (x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4) + i(x_1y_2 + x_2y_1 + x_3y_4 - x_4y_3) \\ &\quad + j(x_1y_3 + x_3y_1 + x_4y_2 - x_2y_4) + k(x_1y_4 + x_4y_1 + x_3y_4 - x_4y_3) \end{aligned}$$

Alors,  $(\mathbb{H}, +, \cdot)$  est une algèbre associative.

5.  $(\text{End}(E), +, \circ)$  est une algèbre associative non commutative.

**4.4.5 Proposition:** L'ensemble  $C^\infty(M)$  est un espace vectoriel sur  $\mathbb{R}$  (de dimension infinie) et une algèbre associative et commutative avec le produit usuel  $(f \cdot g)(x) = f(x) \cdot g(x)$ , où  $f, g \in C^\infty(M)$  et  $x \in M$ .

**4.4.6 Définition:** Si  $\{(U_r, \varphi_r)\}$  est un atlas de  $M$ , on appelle *fonctions coordonnées* ou *coordonnées locales* les fonctions  $x^i : U_r \rightarrow \mathbb{R}$  définies en tout point  $x \in U_r$  comme

$$x^i(x) := \pi^i \circ \varphi(x), \quad i = 1, \dots, m,$$

où  $\pi^i : \mathbb{R}^m \rightarrow \mathbb{R}$  est la projection sur la  $i$ -ème composante. On a alors  $\varphi(x) = (x^1(x), \dots, x^m(x))$ .

Évidemment, en tout point  $x \in M$  les coordonnées locales  $x^i$  autour de  $x$  dépendent de la carte  $\varphi$  choisie.

**4.4.7 Proposition:** Les coordonnées locales  $x^i$  sont des fonctions (locales) de classe  $C^\infty$  sur  $M$ .

**Preuve** Une fonction  $f$  sur  $M$  (ou sur tout ouvert de  $M$ ) est de classe  $C^\infty$  si elle l'est une fois lue à travers toute carte de  $M$  qui intersecte son support. Pour tout  $x \in M$ , et ayant fixé la carte  $\varphi$  pour définir  $x^i$ , choisissons alors une autre carte  $\varphi' : U' \rightarrow \mathbb{R}^m$  telle que  $U'$  contient  $x$ , et appelons  $\varphi'(x) = (y^1, \dots, y^m)$  les nouvelles coordonnées locales. Soit  $W = U \cup U'$ , et considérons  $x^i$  et  $\varphi'$  restreintes à  $W$ . La fonction  $x^i$ , dans la carte  $\varphi'$ , devient alors la fonction réelle  $\tilde{x}^i := x^i \circ (\varphi')^{-1}$ . Celle-ci est de classe  $C^\infty$ , car

$$\tilde{x}^i = x^i \circ (\varphi')^{-1} = \pi^i \circ \varphi \circ (\varphi')^{-1} = \pi^i \circ \psi,$$

où  $\psi$  est le changement de carte, qui est de classe  $C^\infty$ , et la projection  $\pi^i$  l'est évidemment aussi.

### Courbes paramétrées sur une variété

**4.4.8 Définition:** Une *courbe paramétrée* sur  $M$  est une application  $\gamma : \mathbb{R} \rightarrow M$ , de domaine  $I \subset \mathbb{R}$ . Son expression locale dans une carte  $\varphi : U \rightarrow \mathbb{R}^m$  et la courbe  $\tilde{\gamma} = \varphi \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}^m$ , avec  $\tilde{\gamma}(t) = (\tilde{\gamma}^1(t), \dots, \tilde{\gamma}^m(t)) = (x^1(t), \dots, x^m(t))$ .

La courbe  $\gamma$  est *régulière en*  $t \in I$ , ou en  $x = \gamma(t) \in M$ , si pour toute carte  $\varphi : U \rightarrow \mathbb{R}^m$  telle que  $U$  contient le point  $P$ , son expression locale  $\tilde{\gamma} = \varphi \circ \gamma$  est une courbe régulière en  $t$ , i.e.  $\dot{\tilde{\gamma}}(t) = \frac{d\tilde{\gamma}}{dt}(t) \neq \vec{0}$ . On dit que  $\gamma$  est *régulière* si elle est régulière en tout  $t \in I$ .

Si  $\gamma$  est une courbe régulière en  $x$ , on peut toujours changer sa paramétrisation pour avoir  $\gamma(0) = x$ . Localement, autour de  $x$ , on peut aussi choisir une carte  $\varphi$  telle que  $\varphi(x) = (0, \dots, 0) \equiv \vec{0} \in \mathbb{R}^m$ . On peut donc toujours avoir une courbe locale  $\tilde{\gamma} : \mathbb{R} \rightarrow \mathbb{R}^m$  régulière en  $\tilde{\gamma}(0) = \vec{0}$ .

**Remarque** La *vitesse* d'une courbe  $\gamma : \mathbb{R} \rightarrow M$  au point  $x = \gamma(0)$  où elle est régulière devrait être un *vecteur*  $\gamma'(0)$  *tangent* à la courbe en  $P$ . Mais on ne connaît pas l'espace ambiant où ce vecteur pourrait être défini. Nous allons le définir de façon indirecte, en utilisant ses propriétés comme *opérateur agissant sur les fonctions différentiables*.

Pour cela, considérons une fonction différentiable  $f : M \rightarrow \mathbb{R}$  et suivons sa valeur le long de la courbe  $\gamma$  : nous avons donc une fonction réelle  $f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$ . Pour tout choix d'une carte  $\varphi : U \rightarrow \mathbb{R}^m$  autour de  $x$  on a alors

$$f \circ \gamma = f \circ \varphi^{-1} \circ \varphi \circ \gamma = \tilde{f} \circ \tilde{\gamma},$$

où  $\tilde{f} = f \circ \varphi^{-1}$  et  $\tilde{\gamma} = \varphi \circ \gamma$  sont l'expression de  $f$  et  $\gamma$  en coordonnées locales autour de  $x$ , c'est-à-dire des fonctions réelles. Si pour tout point  $x'$  autour de  $x$  on appelle  $(x^1, \dots, x^m) = \varphi(x')$  ses coordonnées locales dans cette carte, on a  $\tilde{\gamma}(t) = (\tilde{\gamma}^1(t), \dots, \tilde{\gamma}^m(t))$ , avec  $\tilde{\gamma}^i(t) = x^i(t)$ . On peut alors dériver  $\tilde{\gamma}(t)$  et  $(\tilde{f} \circ \tilde{\gamma})(t)$  de façon usuelle par rapport à  $t$ , en utilisant la règle de la chaîne. En calculant les dérivées en  $t = 0$ , en sachant que  $\tilde{\gamma}(0) = \varphi(x)$ , on obtient :

$$\begin{aligned} \frac{d(f \circ \gamma)}{dt}(0) &= \frac{d(\tilde{f} \circ \tilde{\gamma})}{dt}(0) = \sum_{i=1}^m \frac{\partial \tilde{f}}{\partial x^i}(\tilde{\gamma}(0)) \frac{d\tilde{\gamma}^i(0)}{dt} \\ &= \left( \sum_{i=1}^m \dot{x}^i(0) \frac{\partial}{\partial x^i} \Big|_{\varphi(x)} \right) \tilde{f}. \end{aligned}$$

On voit alors que *en coordonnées locales, la dérivée de  $f \circ \gamma$  en  $t = 0$  peut être interprétée comme un opérateur différentiel dépendant seulement de  $\gamma$  et agissant sur  $\tilde{f}$* . Cela peut être énoncé indépendamment des coordonnées locales, et motive la définition suivante.

**4.4.9 Définition:** Soit  $\gamma : \mathbb{R} \rightarrow M$  une courbe régulière en  $t = 0$ , et  $C^\infty(x)$  l'ensemble des fonctions réelles sur  $M$  qui sont différentiables de classe  $C^\infty$  en  $x = \gamma(0)$ . Le *vecteur tangent* à  $\gamma$  en  $\gamma(0)$  est l'application  $\gamma'(0) : C^\infty(x) \rightarrow \mathbb{R}$  définie par

$$\gamma'(0) f := \frac{d(f \circ \gamma)}{dt}(0), \quad f \in C^\infty(x).$$

En coordonnées locales  $(x^1, \dots, x^m)$  autour de  $\gamma(0) = x$  on a donc

$$\gamma'(0) = \sum_{i=1}^m \dot{x}^i(0) \frac{\partial}{\partial x^i} \Big|_x, \quad \text{où} \quad \frac{\partial}{\partial x^i} \Big|_x := \frac{\partial}{\partial x^i} \Big|_{\varphi(x)}$$

### Espaces tangents

On suppose dorénavant que  $M^p$  est une variété de dimension  $p$ .

Lorsque  $t \mapsto c(t)$  est une courbe paramétrée, sa tangente en  $c(t_0)$  est dirigée par le vecteur  $c'(t_0)$ , si celui-ci est non nul. Pour une nappe paramétrée  $j(u, v) = (x(u, v), y(u, v), z(u, v))$ , le plan tangent en

$j(u_0, v_0)$  est engendré par les vecteurs  $\frac{\partial j}{\partial u}(u_0, v_0), \frac{\partial j}{\partial v}(u_0, v_0)$ , qui sont linéairement indépendants si  $dj$  est injective.

La notion d'espace tangent à une sous-variété généralise ces situations.

**4.4.10 Définition:** Soit  $M^p$  une sous-variété de  $\mathbb{R}^n$ , et  $x_0$  un point de  $M^p$ ,  $\varphi$  une carte en  $x_0$ . L'espace tangent (vectoriel) à  $M$  en  $x_0$  est l'espace vectoriel  $d\varphi^{-1}(0)(\mathbb{R}^p \times \{0\})$ , noté  $T_{x_0}M$ . L'espace affine tangent à  $M$  en  $x_0$  est l'unique espace affine passant par  $x_0$  et dirigé par  $T_{x_0}M$ . On le note  $T_{x_0}M$ .

**Remarque :** On peut montrer directement que l'espace tangent ne dépend pas du choix de la carte. En effet si  $\varphi, \psi$  sont deux cartes en  $x_0$ ,  $\varphi \circ \psi^{-1}$  envoie  $\mathbb{R}^p \times \{0\}$  sur lui-même et donc sa différentielle aussi. Cela signifie que

$$d\varphi(x_0)d\psi^{-1}(0)(\mathbb{R}^p \times \{0\}) = \mathbb{R}^p \times \{0\}$$

en d'autres termes que

$$d\psi^{-1}(0)(\mathbb{R}^p \times \{0\}) = d\varphi^{-1}(0)(\mathbb{R}^p \times \{0\})$$

On remarquera que la proposition suivante donne une autre démonstration de cette indépendance.

D'un point de vue pratique,

a) si  $M$  est définie localement comme l'image de  $j$  par  $\mathbb{R}^p$ , avec  $j(0) = x_0, dj(0)$  injective, on a

$$T_{x_0}M = dj(0)\mathbb{R}^p$$

b) si  $M$  est définie localement comme le graphe  $\{(z, f(z)) \mid z \in \mathbb{R}^p\}$

$$T_{x_0}M = \{(h, df(z_0)h) \mid h \in \mathbb{R}^p\}$$

où  $x_0 = (z_0, f(z_0))$

Dans ce cas  $T_{x_0}M$  est le graphe de  $df(z_0)$ .

c) si  $M$  est définie localement comme le lieu des zéros de  $F : \mathbb{R}^n \rightarrow \mathbb{R}^{n-p}$  telle que  $dF(x_0)$  soit surjective. Alors  $T_{x_0}M = \ker dF(x_0)$ .

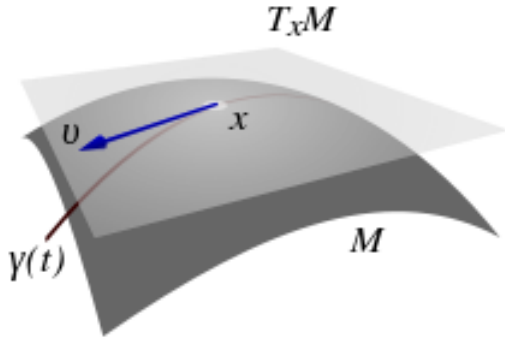
La caractérisation suivante de l'espace tangent est parfois utile.

**4.4.11 Proposition:** Soit  $M$  une sous-variété. L'espace tangent  $T_xM$  est l'ensemble des vecteurs vitesse en  $t = 0$  des chemins  $C^\infty$  tracés sur  $M$  passant par  $x$  à l'instant  $t = 0$ .

**Preuve:** Soit  $\varphi$  une carte en  $x$ , et  $c$  une courbe tracée sur  $M$  telle que  $c(0) = x$ . Alors pour tout  $t$  assez proche de 0  $\varphi(c(t))$  est définie, tracée sur  $\mathbb{R}^p \subset \mathbb{R}^n$  et passe par 0 en  $t = 0$ . Donc la tangente en 0 de  $\varphi(c(t))$  est dans  $\mathbb{R}^p$  c'est à dire que  $d\varphi(c(0))c'(0)$  est dans  $\mathbb{R}^p$ , ce qui par définition équivaut à  $c'(0) \in T_xM$ .

Inversement soit  $\zeta \in T_xM$  alors  $d\varphi(x)\zeta$  est dans  $\mathbb{R}^p$ . Soit  $t \rightarrow \gamma(t)$  une courbe paramétrée de  $\mathbb{R}^p$  tangente à  $d\varphi(x)\zeta$  en  $t = 0$ . Alors quitte à restreindre  $\gamma$  à un voisinage de 0, on peut supposer son image contenue dans  $V = \varphi(U)$ . Alors  $c(t) = \varphi^{-1}(\gamma(t))$  est une courbe tracée sur  $M$ , et le calcul précédent montre qu'elle est tangente à  $\zeta$  en  $t = 0$ .  $\square$

Voici une petite illustration :



On a défini l'espace tangent à une sous-variété et montré l'équivalence entre la définition et la proposition ci dessus. Or, pour une variété :

**4.4.12 Définition:** Soit  $M$  une variété. L'espace tangent  $T_x M$  est l'ensemble des vecteurs vitesse en  $t = 0$  des chemins  $C^\infty$  tracés sur  $M$  passant par  $x$  à l'instant  $t = 0$ .

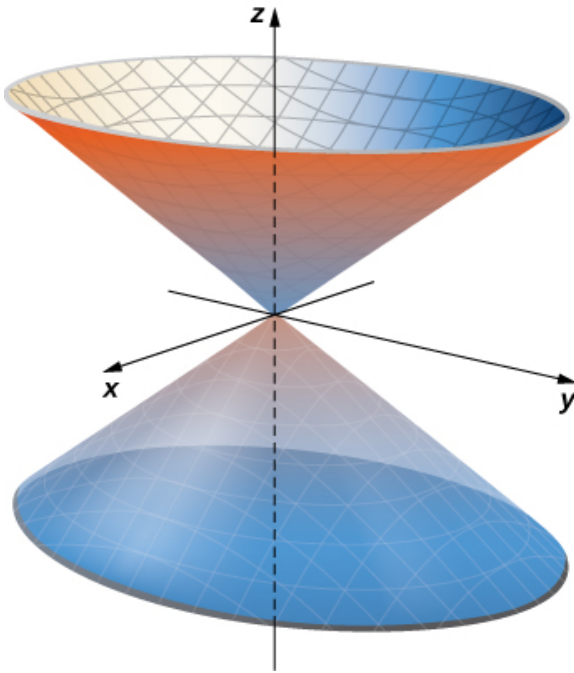
**4.4.13 Exemple:** *Espaces tangents de certaines sous-variétés.*

(A) Soit  $S^2 = \{x \in \mathbb{R}^3 \mid |x|^2 = 1\}$ . C'est une sous-variété définie par  $F(x) = 0$  où  $F(x) = |x|^2 - 1$ . Vu que  $dF(x)h = 2\langle x, h \rangle$ , l'espace tangent en  $x$  à la sphère,  $T_x M^2$  est le plan orthogonal à  $x$ , et  $T_x M^2$  le plan passant par  $x$ .

(B) Soit  $I$  un intervalle et  $\gamma : I \rightarrow \mathbb{R}^n$  une courbe paramétrée. Si  $\gamma'(t)$  ne s'annule jamais et  $\gamma$  est injective et propre,  $\Gamma = \gamma(I)$  est une courbe lisse, son espace tangent en  $\gamma(t_0)$  est donnée par  $\mathbb{R}\gamma'(t_0)$ .

(C) On va montrer de 3 façons différentes que le cône n'est pas une sous-variété. Soit  $C$  le cône  $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = 0\}$

Avant, voici une image du cône d'équation  $x^2 + y^2 = z^2$ .



1. En effet, on montre aisément que  $C$  est connexe (utiliser le fait que tout point est sur un segment de droite contenue dans  $C$  et passant par 0 pour montrer que  $C$  est connexe par des arcs), et d'après le (B) de la proposition de caractérisation des sous-variétés, un voisinage de  $(x, y, z) \neq (0, 0, 0)$  est une

sous-variété de dimension 2, vu qu'en un tel point  $dF(x, y, z)(u, v, w) = x \cdot u + y \cdot v + z \cdot w$  n'est pas identiquement nulle. Donc si  $C$  est une sous-variété, elle sera partout de dimension 2.

Mais en  $(0, 0, 0)$  les courbes  $t \rightarrow (t \cdot \cos(\theta), t \cdot \sin(\theta), t)$  tracées sur  $C$ , ont pour vecteur vitesse en 0 le vecteur  $(\cos(\theta), \sin(\theta), 1)$ . L'ensemble de ces vecteurs n'est contenu dans aucun plan vectoriel de  $\mathbb{R}^3$ . L'ensemble  $C$  ne peut donc être une sous-variété.

2.  $C$  est connexe par arcs : en effet le chemin  $(\cos(2\pi\theta)t, \sin(2\pi\theta)t, t)$  permet de relier tous les points du cône avec le point  $(0, 0, 0)$ . De plus,  $dF = 2x dx + 2y dy - 2z dz \neq 0 \forall (x, y, z) \neq 0$ . Si  $(x, y, z) = (0, 0, 0)$ ,  $dF(0, 0, 0) = (0, 0, 0)$ , donc  $F$  est non surjective en ce point. Le cône n'est pas une sous-variété (second moyen).

3. Le troisième : on vérifie que la paramétrisation donnée est injective sauf en  $(0, 0, 0)$ , le calcul des dérivées partielles donne 2 vecteurs base de l'espace tangent, donc si le cône est une sous-variété, il est de dimension 2. Mais si  $M$  est une variété connexe par arcs de dimension 2,  $M \setminus \{x\}$  l'est encore : pour montrer cela, je vais utiliser un résultat que je n'ai pas encore montré, mais qui sera prouvé par la suite : toute sous-variété est une variété. Soit un recouvrement de  $M$  par des ouverts. Soit  $U_{i,x}$  les ouverts  $U_i$  contenant le point  $x$ .  $V_i = U_i \setminus \{x\}$  est un ouvert. Par conséquent, on a toujours un recouvrement de  $M \setminus \{x\}$  par des ouverts, et les difféomorphismes associés à ces ouverts sont toujours bien définis.  $M \setminus \{x\}$  reste donc une sous-variété de dimension 2. Enfin, on remarque que le cône privé du point  $(0, 0, 0)$  n'est pas connexe :  $f(x, y, z) = z$ ,  $(x, y, z) \in C$  a des valeurs positives et négatives mais ne passe jamais par 0.

### Différentielle d'une application

Soient  $M$  et  $N$  deux variétés différentiables de dimension  $m$  et  $n$ , et  $\phi : M \rightarrow N$  une application différentiable.

**4.4.14 Définition:** Pour tout  $x \in M$ , la *différentielle de  $\phi$  en  $x$*  est l'application linéaire  $d\phi_x : T_x M \rightarrow T_{\phi(x)} N$  définie sur tout vecteur  $X_x = \gamma'(0)$  tangent à la courbe  $\gamma$  sur  $M$  en  $x = \gamma(0)$  comme le vecteur

$$d\phi_x(\gamma'(0)) := (\phi \circ \gamma)'(0).$$

Si on interprète les vecteurs tangents comme des dérivations, la différentielle de  $\phi$  en  $x$  peut être définie comme

Si on fixe des coordonnées locales  $(x^1, \dots, x^m)$  autour de  $x$ , et des coordonnées locales  $(y^1, \dots, y^n)$  autour de  $\phi(x)$ , on a  $\widetilde{\phi \circ \gamma}(t) = \tilde{\phi} \circ \tilde{\gamma}(t) = \tilde{\phi}(x^1(t), \dots, x^m(t)) = (y^1(t), \dots, y^n(t))$ , avec  $y^j = y^j(x^1, \dots, x^m)$  pour tout  $j = 1, \dots, n$ . Pour tout  $X_x = \sum_{i=1}^m X^i(x) \frac{\partial}{\partial x^i} \Big|_x$ , avec  $X^i(x) = \frac{d\tilde{\gamma}^i}{dt}(0)$ , on a donc

$$d\phi_x(X_x) = \sum_{j=1}^n \frac{d(\widetilde{\phi \circ \gamma})^j}{dt}(0) \frac{\partial}{\partial y^j} \Big|_{\phi(x)} = \sum_{j=1}^n \sum_{i=1}^m \frac{\partial \tilde{\phi}^j(\varphi(x))}{\partial x^i} \frac{d\tilde{\gamma}^i}{dt}(0) \frac{\partial}{\partial y^j} \Big|_{\phi(x)} = \sum_{j=1}^n \sum_{i=1}^m \frac{\partial y^j(\varphi(x))}{\partial x^i} X^i(x) \frac{\partial}{\partial y^j} \Big|_{\phi(x)}.$$

La différentielle  $d\phi_x$  transforme donc le vecteur tangent  $X_x$  de coefficients locaux  $X^i(x)$  en un vecteur tangent  $Y_{\phi(x)} = d\phi_x(X_x)$  de coefficients

$$Y^j(\phi(x)) = \sum_{i=1}^m \frac{\partial \tilde{\phi}^j(\varphi(x))}{\partial x^i} X^i(x).$$

La matrice  $\left( \frac{\partial \tilde{\phi}^j(\varphi(x))}{\partial x^i} \right) \in \text{Mat}_{nm}$  s'appelle *matrice Jacobienne* de  $\phi$  en  $x$ , et son déterminant  $\det \left( \frac{\partial \tilde{\phi}^j(\varphi(x))}{\partial x^i} \right)$  s'appelle *Jacobien*.

**4.4.15 Proposition:** La différentielle d'applications sur des variétés satisfait à la règle de la chaîne : si  $\phi : M \rightarrow N$  et  $\psi : N \rightarrow O$  sont deux applications différentiables, alors

$$d(\psi \circ \phi)_x = d\psi_{\phi(x)} \circ d\phi_x, \quad x \in M.$$

**4.4.16 Proposition:** Une application  $\phi : M \rightarrow N$  est un difféomorphisme si et seulement si, en tout  $x \in M$ , la différentielle  $d\phi_x : T_x M \rightarrow T_{\phi(x)} N$  est un isomorphisme.

En effet, ceci est équivalent à dire que le Jacobien de  $\phi$  en  $x$  est non-nul. On a déjà vu que ceci implique  $m = n$ .

#### Immersion, submersion, plongement sur des variétés

**4.4.17 Définition:** Soit  $f : M \rightarrow N$  une application. On dit que  $f \in C^0(M, N)$  est de classe  $C^k$  de  $M$  dans  $N$  si pour toutes cartes  $(U, \varphi)$  de  $M$  et  $(V, \psi)$  de  $N$  on a :

$$\psi \circ f \circ \varphi^{-1} : \varphi(U \cap f^{-1}(V)) \rightarrow \psi(V \cap f(U))$$

On note alors  $df$  la différentielle  $df$ .

**4.4.18 Définition:** Une application  $f$  est dite *propre* si l'image réciproque de tout compact par  $f$  est compacte.

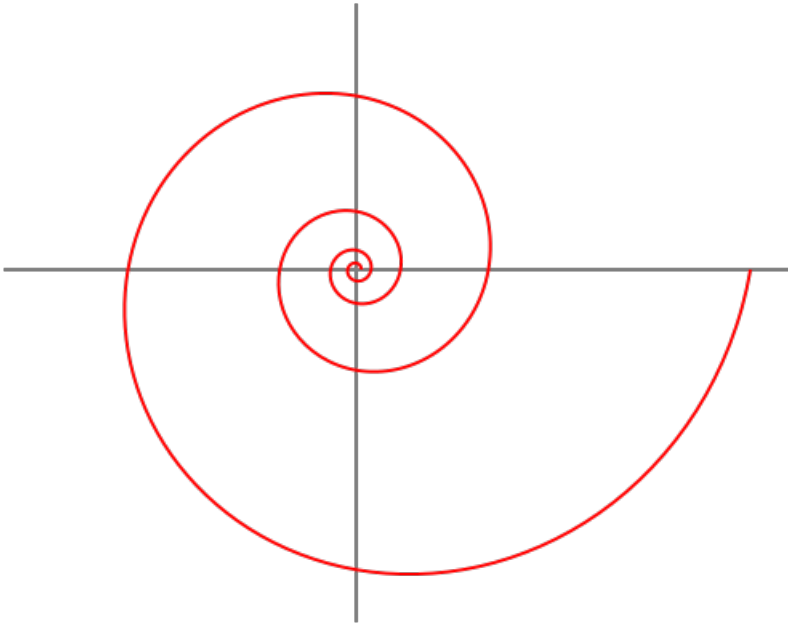
**4.4.19 Définition:** Une immersion entre 2 variétés différentielles est une application dont la différentielle est en tout point injective.

**4.4.20 Définition:** Une submersion entre 2 variétés différentielles est une application dont la différentielle est en tout point surjective.

**4.4.21 Définition:** Un plongement est une immersion injective qui est un homéomorphisme sur son image.

**4.4.22 Exemple:** *Submersions et immersions :*

En fait, on a déjà rencontré des immersions et des submersions : remarquons que, pour qu'une fonction sur une variété soit injective, la dimension de l'espace de départ doit nécessairement être inférieure à la dimension de l'espace d'arrivée. Pour une submersion, c'est l'inverse. Ainsi :  $\gamma(t) = (\cos(t), \sin(t))$  sur  $]0, 2\pi[$  est une immersion et même un plongement. Et  $f(x, y) = x^2 + y^2 - 1 = 0$  qui va du cercle dans  $\mathbb{R}$  est une submersion. Ainsi, les immersions sont des paramétrages (la réciproque est fautive) et les submersions sont un moyen de représenter un ensemble comme zéro de fonction (la réciproque est également fautive).



Cette spirale représente un plongement de la droite réelle dans le plan.

## 4.5 Théorème de plongement de Whitney dans le cas compact\*

Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$  et  $K$  un compact de  $\Omega$ . On note  $C^m(\Omega)$  l'espace des fonctions de  $C^m$  sur  $\Omega$  et par  $C^m(K)$  l'espace des fonctions  $C^m$  au voisinage de  $K$ . Enfin  $C_K^m$  est l'espace des fonctions  $C^m$  à support dans  $K$ , c'est à dire telles que  $\{x \mid f(x) \neq 0\} \subset K$ . Lorsque  $\Omega$  est omis, il faut entendre qu'il s'agit du cas  $\Omega = \mathbb{R}^n$ . Enfin  $C_0(\Omega)$  est la réunion des  $C_K^m$  pour  $K$  dans  $\Omega$ . Si  $m = \infty$ ,  $C^m$  est l'espace des fonctions  $C^\infty$ , si  $m = \omega$ , c'est l'espace des fonctions analytiques, c'est à dire des fonctions qui sont au voisinage de chaque point de la somme d'une série entière convergente.

Avant de se lancer dans la preuve du théorème, qui permet de faire le lien entre variété et sous-variétés, remarquons qu'une sous-variété est une variété :

**4.5.1 Proposition:** *Toute sous-variété  $M$  est une variété.*

**Preuve:** On a vu (je renvoie aux annexes) que  $\mathbb{R}^n$  est un ensemble à base dénombrable d'ouvert, et que tout sous-ensemble d'un ensemble à base dénombrable d'ouvert est lui-même à base dénombrable d'ouvert. Donc,  $M$  est à base dénombrable d'ouvert.

Soit  $M$  une sous variété : par définition, en chaque point  $x_0$  de  $M$  il existe un voisinage  $U_{x_0}$  de  $x_0$  et un difféomorphisme local  $\varphi$  tel que l'on ait

$$\varphi(U_{x_0} \cap M) = (\mathbb{R}^p \times \{0\}) \cap \varphi(U_{x_0})$$

On pose  $V = \bigcup_{x \in M} U_x \cap M$ , qui est un recouvrement de  $M$  (ouvert car on peut s'arranger pour prendre un voisinage ouvert  $U_x$  de  $\mathbb{R}^n$ ). Ainsi, sur chaque  $W_x = U_x \cap M$ , on définit une carte locale :  $\psi : W_x \rightarrow \mathbb{R}^p$  en posant  $\psi = \pi_1 \circ \varphi$  ( $\pi_1 : \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p} \rightarrow \mathbb{R}^p$  étant la première projection). En effet,  $\psi$  est continue car composée d'applications continues, inversible car  $\psi^{-1}(u) = \varphi^{-1}(u, 0)$  et son inverse est continue car  $\varphi$  est un difféomorphisme. Enfin, on vérifie par le calcul que  $\psi$  est un difféomorphisme (par composition). Remarque : le fait que  $U_x \cap M$  soit un ouvert de  $M$  n'est pas trivial :

$\varphi$  est continue, donc  $\varphi(U_x)$  est un ouvert puisque  $U_x$  l'est ( $\varphi$  est une application ouverte car c'est un difféomorphisme). De plus,  $\mathbb{R}^p \times \{0\}$  est un sous-espace de  $\mathbb{R}^n$  qui est fermé dans  $\mathbb{R}^n$  mais *ouvert* par rapport à la topologie que  $\mathbb{R}^n$  induit sur  $\mathbb{R}^p \times \{0\}$  par restriction car l'inclusion  $I : \mathbb{R}^p \times \{0\} \rightarrow \mathbb{R}^n$  est continue, donc la préimage (préimage d'une partie par une application = ensemble des points envoyés dans cette partie) de l'ouvert  $\mathbb{R}^n$  (qui est  $\mathbb{R}^p \times \{0\}$ ) est ouverte (par définition d'une fonction continue). Enfin, on conclut par le fait que  $\varphi(U_x \cap M) = \mathbb{R}^p \times \{0\} \cap \varphi(U_x)$ , une intersection de 2 ouverts qui est donc ouverte, donc par continuité de  $\varphi$ , l'image réciproque  $U_x \cap M$  de l'ouvert  $\mathbb{R}^p \times \{0\} \cap \varphi(U_x)$  (on a correspondance entre préimage et image réciproque car  $\varphi$  est un difféomorphisme) est un ouvert.  $\square$

**4.5.2 Proposition:** *Une immersion injective propre est un plongement. L'image d'un plongement dans  $\mathbb{R}^n$  est une sous-variété de  $\mathbb{R}^n$ . Si  $M$  est compacte, l'espace des plongements de  $M$  dans  $N$  est ouvert pour la topologie  $C^1$*

**Preuve:** De par la définition par nappes des sous-variétés. En effet, si  $f$  est un plongement,  $x \in M$ ,  $(U, \varphi)$  est une carte au voisinage de  $x$  et il existe  $V$  un voisinage de  $f(x)$  dans  $\mathbb{R}^n$  tel que  $f$  soit un homéomorphisme de  $U$  sur  $V \cap f(U)$ . Alors  $f \circ \varphi^{-1} : \varphi(U) \rightarrow V \cap f(U)$  est une application différentielle injective qui est bicontinue sur son image (par définition du plongement).  $\square$

**4.5.3 Définition:** On appelle *partition de l'unité d'un espace topologique  $X$*  une famille  $(\rho_i)_{i \in I}$  de fonctions continues, définies sur  $X$  et à valeurs dans l'intervalle  $[0, 1]$ , tel que pour tout point  $x \in X$ , les 2 conditions suivantes soient satisfaites :

1. Il existe un voisinage de  $x$  tel que toutes les fonctions  $\rho_i$  soient nulles sur ce voisinage à l'exception d'un nombre fini d'entre elles.
2.  $\sum_{i \in I} \rho_i(x) = 1, \forall x \in X$

**4.5.4 Définition:** On appelle partition de l'unité *subordonnée au recouvrement*  $(U_i)_{i \in I}$  une partition de l'unité  $(\rho_i)_{i \in I}$  au sens de la définition précédente, indexé par le même ensemble  $I$  que le recouvrement, et tel que, pour tout  $i \in I$ ,  $\text{supp} \rho_i \subseteq U_i$ .

**4.5.5 Théorème:** Soit  $M$  une variété et  $(U_\alpha)_{\alpha \in A}$  un recouvrement ouvert de  $M$ . Il existe alors une famille  $(\rho_j)_{j=1, \dots, q}$  de fonctions de classe  $C^\infty$  telles que :

1.

$$\forall j; \text{supp}(\rho_j) \subset U_j$$

2.

$$\forall x \in M, 0 \leq \rho_j(x) \leq 1$$

3.

$$\forall x \in M, \sum_{j=1}^{\infty} \rho_j(x) = 1$$

4. Chaque point  $x$  de  $M$  a un voisinage qui ne rencontre qu'un nombre fini des supports de  $\rho_j$ .

**Preuve:** On peut passer à un recouvrement plus fin, localement fini, dont chaque ouvert est contenu dans une carte. Soit  $(U_j)_{j \in \mathbb{N}}$  un tel recouvrement et  $V_j$  un sous-recouvrement tel que  $V_j \subset \overline{V_j} \subset U_j$ . Via les cartes  $\varphi_j : U_j \rightarrow \mathbb{R}^n$ , on peut trouver des fonctions  $f_j = 1$  sur  $V_j$  (voir le corollaire des partitions de l'unité). Alors, dans ce cas  $\rho_j = \frac{f_j}{\sum_{k=1}^{+\infty} f_k}$  vérifie les conditions du théorème.  $\square$

**4.5.6 Théorème:** Soit  $M$  une variété compacte de dimension  $n$ . Il existe alors un plongement de  $M$  dans un espace euclidien  $\mathbb{R}^N$ ,  $N$  n'ayant pas de lien avec  $n$ .

**Preuve:** Soient  $(U_j, \varphi_j)$  des cartes telles que les ouverts  $U_j$  forment un recouvrement de  $M$ . On considère un sous-recouvrement  $V_j$  par des ouverts tels que  $V_j \subset \overline{V_j} \subset U_j$  et une partition de l'unité  $\rho_j$  subordonnée aux  $U_j$ . On pose alors

$$F(x) = (\rho_1(x), \dots, \rho_N(x), \rho_1(x)\varphi_1(x), \dots, \rho_N(x)\varphi_N(x))$$

Cette application est bien définie, car le plongement de  $\rho_j \cdot \varphi_j$  par la fonction nulle en dehors de  $U_j$  est une fonction  $C^\infty$ . Maintenant, en chaque point  $x$ , l'un des  $\rho_j$  est non nul, et l'image de  $(d\rho_j(x), \varphi_j(x)d\rho_j(x) + \rho_j(x)d\varphi_j(x))$  est de rang  $n$  car si  $d\rho_j(x)h = 0$  et  $\varphi_j(x)d\rho_j(x)h + \rho_j(x)d\varphi_j(x)h = 0$ , alors  $d\rho_j(x)h = 0 \forall x \in V_j$ , ce qui est impossible. Elle est injective car si  $F(x) = F(y)$  on doit avoir  $\rho_j(x) = \rho_j(y)$ , donc ils sont dans la même carte  $U_j$  et  $\varphi_j(x) = \varphi_j(y)$ , donc  $x = y$ . Or, une immersion injective dans le cas compact est un plongement, puisqu'elle est propre.  $\square$



## 5 Fibré tangent

Le but de ce chapitre est d'approfondir les notions précédentes en définissant et fibré tangent et les champs de vecteurs (un cas particulier de fibré et de ses sections). Elle va permettre la transition entre la partie II et la partie III, car les champs de vecteurs sont une algèbre de Lie.

### 5.1 Fibré tangent

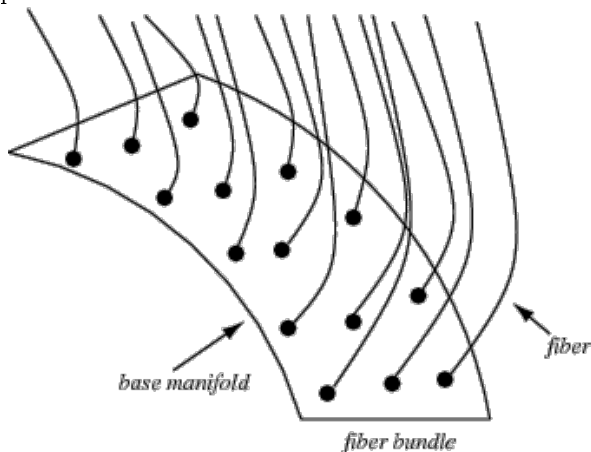
Soient  $M$  et  $N$  deux variétés différentiables de dimension  $m$  et  $n$  avec  $m < n$ , et  $\phi : N \rightarrow M$  une application différentiable.

**5.1.1 Définition:** Un fibré (différentiel) sur  $M$  de fibre une variété  $F$  est une variété  $N$  avec une application différentiable  $\pi : N \rightarrow M$  telle que

1.  $\pi$  est une submersion surjective;
2.  $N$  est localement trivial avec fibre  $F$ , c'est-à-dire que pour tout  $x \in M$  il existe un voisinage  $U$  de  $x$  dans  $M$  et un difféomorphisme  $\tau_U : \pi^{-1}(U) \rightarrow U \times F$  tel que  $\text{pr}_1 \circ \tau_U = \pi|_U$ , i.e. pour tout  $y \in \pi^{-1}(U)$  on a  $\tau_U(y) = (x, z)$ , où  $x = \pi(y)$  et  $z \in F$ .

On appelle  $N$  l'espace total,  $\pi$  la projection,  $M$  la base et  $\tau_U$  une trivialisatoin locale du fibré. On indique le fibré par  $N(M, F)$ .

Représentation visuelle d'un fibré et de sa fibre :



**5.1.2 Corollaire:** Dans tout fibré on a les propriétés suivantes :

- Pour tout ouvert  $U \subset M$ , l'ensemble  $N_U = \pi^{-1}(U)$  est un ouvert de  $N$ , qu'on appelle fibre sur  $U$ .  
 Pour tout  $x \in M$ , l'ensemble  $N_x = \pi^{-1}(x)$  est un sous-ensemble fermé de  $N$ , qu'on appelle fibre sur  $x$ ,  
 et on a  $N = \bigcup_{x \in M} N_x$ .

La restriction de  $\tau_U$  à tout point  $x \in U$  est un difféomorphisme  $\tau_U|_x : N_x \rightarrow F$  qui dépend de  $U$ .

**5.1.3 Définition:** Un fibré vectoriel de rang  $k$  sur  $M$  est un fibré  $\pi : E \rightarrow M$  de fibre un espace vectoriel  $V$  de dimension  $k$ , avec trivialisations locales  $\tau_U : \pi^{-1}(U) \rightarrow U \times V$  telles que toute restriction  $\tau_U|_x : E_x \rightarrow \{x\} \times V$  est un isomorphisme d'espaces vectoriels. On note  $E = E(M, V)$ . On peut supposer que  $V$  soit  $\mathbb{R}^k$ .

**5.1.4 Exemple:** L'espace-temps

L'espace-temps est la variété  $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$  avec coordonnées  $(t, \vec{x})$ , qui admet, en théorie, deux structures de fibré :

- l'espace sur le temps,  $\mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $(t, \vec{x}) \mapsto t$ , i.e.  $\vec{x} = \vec{x}(t)$ ;

— le temps sur l'espace,  $\mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $(t, \vec{x}) \mapsto \vec{x}$ , i.e.  $t = t(\vec{x})$ .

**5.1.5 Définition:** Un fibré vectoriel (de rang  $k$  sur  $M$ ) est dit *trivial* s'il existe un difféomorphisme  $\tau$  tel que  $\tau : \pi^{-1}(M) = M \times \mathbb{R}^k$  ou  $M \times V$ ,  $V$  étant la fibre, où  $\pi$  est la première projection :  $\pi : M \times V \rightarrow M, \pi(x, v) = x$ .

**5.1.6 Exemple:** Le fibré trivial de fibre  $F$  est la projection  $\text{pr}_1 : M \times F \rightarrow M : (x, y) \mapsto x$ . En particulier : le cylindre  $\mathbb{S}^1 \times \mathbb{R}$  et le tore  $\mathbb{S}^1 \times \mathbb{S}^1$  sont des fibrés triviaux sur  $\mathbb{S}^1$ .

Soit  $M$  une variété différentiable de dimension  $m$ .

**5.1.7 Définition:** On appelle *fibré tangent* sur  $M$  l'union disjointe des espaces tangents

$$TM = \bigcup_{x \in M} T_x M$$

**5.1.8 Proposition:** Le fibré tangent est un fibré vectoriel de rang  $m$ , avec projection  $\pi : TM \rightarrow M$ ,  $\pi(X_x) = x$ .

**Preuve:** Il faut montrer que  $TM$  est une variété différentiable, que  $\pi$  est une submersion, et qu'il y a des trivialisations locales de fibre  $\mathbb{R}^m$ . Les cartes et les trivialisations locales de  $TM$  se trouvent à partir d'un atlas de cartes sur  $M$ .

Soit  $\varphi : U \rightarrow \mathbb{R}^m$  une carte de  $M$  et appelons  $(x^1, \dots, x^m)$  les coordonnées locales dans  $\varphi(U)$ . Puisque  $\varphi : U \rightarrow \varphi(U)$  est un difféomorphisme, sa différentielle en tout point  $x \in U$  est un isomorphisme d'espaces vectoriels  $d\varphi_x : T_x U \rightarrow T_{\varphi(x)} \varphi(U)$ . Appelons  $\tau$  l'application induite sur  $TU = \bigcup_{x \in U} T_x U = TM \Big|_U$  par les différentielles  $d\varphi_x$  quand  $x$  varie dans  $U$ , composée avec l'isomorphisme canonique

$$i_x : T_{\varphi(x)} \varphi(U) \xrightarrow{\cong} \mathbb{R}^m, \quad \frac{\partial}{\partial x^i} \Big|_{\varphi(x)} \mapsto e^i, \quad i = 1, \dots, m,$$

où  $\{e^i\}$  est la base canonique de  $\mathbb{R}^m$ . On obtient une application

$$\tau : TU \rightarrow \bigcup_{x \in U} T_{\varphi(x)} \varphi(U) \cong \bigcup_{x \in U} \mathbb{R}^m \cong U \times \mathbb{R}^m, \quad (x, X_x) \mapsto (x, i_x(d\varphi_x(X_x)))$$

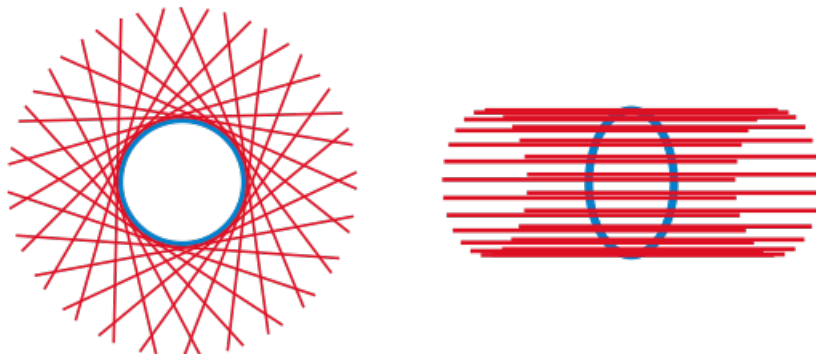
qui est évidemment un difféomorphisme et donne ainsi une trivialisation locale.

Si on compose  $\tau$  avec la carte  $\varphi : U \rightarrow \varphi(U) \subset \mathbb{R}^m$  on obtient l'application  $\bar{\varphi} : TM \Big|_U = TU \rightarrow \varphi(U) \times \mathbb{R}^m \subset \mathbb{R}^{2m}$  définie sur tout  $X_x = \sum_{i=1}^m X^i(x) \frac{\partial}{\partial x^i} \Big|_x \in T_x M$  par

$$\bar{\varphi}(x, X_x) = (\varphi(x), i_x(d\varphi_x(X_x))) = (x^1(x), \dots, x^m(x), X^1(x), \dots, X^m(x)),$$

qui donne une carte sur  $TM$ . Il est facile de montrer que les changements de cartes sont des difféomorphismes et que la projection  $\pi$  est une submersion. Le fibré  $TM$  est donc une variété différentiable de dimension  $2m$ . □

Voici maintenant une petite représentation visuelle du fibré tangent pour le cercle : 2 manières de la représenter : les espaces tangents (première image) sont regroupés de manière continue et sans se recouvrir (deuxième image) : il s'agit d'une démonstration visuelle de l'exemple qui va suivre.



**5.1.9 Exemple:** Le cercle  $\mathbb{S}^1$  est trivialisable. Autrement dit, son fibré tangent est parallélisable.

**Preuve:** Commençons par décrire le fibré tangent à  $\mathbb{S}^1$ . En tout point  $(x, y) \in \mathbb{S}^1$ , l'espace tangent à  $\mathbb{S}^1$  en  $(x, y)$  est l'ensemble des vecteurs  $\alpha'(0) \in \mathbb{R}^2$ , où

$$\alpha : I \longrightarrow \mathbb{S}^1, \quad t \mapsto \alpha(t)$$

est une courbe paramétrée (régulière) sur le cercle telle que  $\alpha(0) = (x, y)$ . On a alors forcément  $\alpha(t) = (x(t), y(t))$  avec

$$x(t)^2 + y(t)^2 = 1.$$

En dérivant, on obtient

$$0 = 2x(t)x'(t) + 2y(t)y'(t) = 2\langle (x(t), y(t)), (x'(t), y'(t)) \rangle,$$

et en posant  $t = 0$  on voit donc que le vecteur  $\alpha'(0) = (x'(0), y'(0))$  est orthogonal au vecteur position  $\alpha(0) = (x, y)$ . Dans le plan  $\mathbb{R}^2$ , il est facile de caractériser les vecteurs orthogonaux à un vecteur  $(x, y)$  donné : ce sont les multiples de  $(-y, x)$ . Par conséquent, on a

$$T_{(x,y)}\mathbb{S}^1 = \left\{ (x, y) + \lambda(-y, x) \mid \lambda \in \mathbb{R} \right\}, \quad \text{comme espace affine,}$$

$$T_{(x,y)}\mathbb{S}^1 \cong \left\{ \lambda(-y, x) \mid \lambda \in \mathbb{R} \right\}, \quad \text{comme espace vectoriel.}$$

Le fibré tangent à  $\mathbb{S}^1$  est l'union disjointe

$$T\mathbb{S}^1 = \bigcup_{(x,y) \in \mathbb{S}^1} T_{(x,y)}\mathbb{S}^1,$$

où l'on peut prendre  $T_{(x,y)}\mathbb{S}^1$  comme espace vectoriel.

Pour montrer que cet ensemble est bien un fibré vectoriel de rang 1, il suffit de trouver un recouvrement de  $\mathbb{S}^1$  par des ouverts  $U$  et des *trivialisations locales*

$$\psi_U : T_U\mathbb{S}^1 = \bigcup_{(x,y) \in U} T_{(x,y)}\mathbb{S}^1 \longrightarrow U \times \mathbb{R}.$$

Le choix est simple, prenons comme recouvrement du cercle la famille  $\{U_S, U_N\}$ , qui correspond aux deux ensembles  $U_S = \{(x, y) \in \mathbb{S}^1, y \neq -1\}$  et  $U_N = \{(x, y) \in \mathbb{S}^1, y \neq 1\}$ , dont on s'est servi lors de la projection stéréographique pour le cercle (je renvoie à la preuve que le cercle est une variété).

, et comme trivialisations locales les applications

$$\begin{aligned} \tau_S : T_{U_S}\mathbb{S}^1 &\longrightarrow U_S \times \mathbb{R}, \quad \lambda(-y, x) \mapsto ((x, y), \lambda), & y \neq 1 \\ \tau_N : T_{U_N}\mathbb{S}^1 &\longrightarrow U_N \times \mathbb{R}, \quad \lambda(-y, x) \mapsto ((x, y), \lambda), & y \neq -1. \end{aligned}$$

Il est clair que ces deux applications sont des trivialisations locales, car en tout point  $(x, y) \in U_S$  (respectivement  $(x, y) \in U_N$ ), la restriction

$$\tau_S|_{(x,y)} : T_{(x,y)}\mathbb{S}^1 \longrightarrow \mathbb{R}, \quad \lambda(-y, x) \mapsto \lambda \quad (\text{resp. } \tau_N)$$

est un isomorphisme d'espaces vectoriels.

Comme tout fibré tangent, celui du cercle est bien une variété différentielle. Pour le cercle, elle a dimension 2. Les cartes locales ont deux composantes : la première décrit une carte locale sur  $\mathbb{S}^1$ , la deuxième décrit les fibres au-dessus de l'ouvert de  $\mathbb{S}^1$  en utilisant une trivialisations locale. On peut donc prendre ;

$$\begin{aligned} \psi_S : T_{U_S}\mathbb{S}^1 &\longrightarrow \mathbb{R} \times \mathbb{R}, \quad \lambda(-y, x) \mapsto \left( \varphi_S(x, y), \tau|_{(x,y)}(\lambda(-y, x)) \right) = \left( \frac{2x}{1-y}, \lambda \right) & y \neq 1 \\ \psi_N : T_{U_N}\mathbb{S}^1 &\longrightarrow \mathbb{R} \times \mathbb{R}, \quad \lambda(-y, x) \mapsto \left( \varphi_N(x, y), \tau|_{(x,y)}(\lambda(-y, x)) \right) = \left( \frac{2x}{1+y}, \lambda \right) & y \neq -1. \end{aligned}$$

Sur l'intersection  $T_{U_S}\mathbb{S}^1 \cap T_{U_N}\mathbb{S}^1 = T_{U_S \cap U_N}\mathbb{S}^1$ , les transitions de cartes  $\psi_S \circ \psi_N^{-1}$  sont données par les applications  $\varphi_S \circ \varphi_N^{-1}$  sur  $\mathbb{S}^1$ , et par l'identité sur  $\mathbb{R}$ , donc elles sont bien des difféomorphismes.

Enfin, pour montrer que  $\mathbb{S}^1$  est parallélisable, il suffit de trouver une *trivialisat*ion globale du fibré  $T\mathbb{S}^1$ , c'est-à-dire un isomorphisme de fibrés vectoriels

$$\psi : T\mathbb{S}^1 \longrightarrow \mathbb{S}^1 \times \mathbb{R}.$$

C'est simple, il suffit de prendre

$$\psi(\lambda(-y, x)) = ((x, y), \lambda)$$

Pour tout  $(x, y) \in \mathbb{S}^1$  et pour tout  $\lambda \in \mathbb{R}$ .

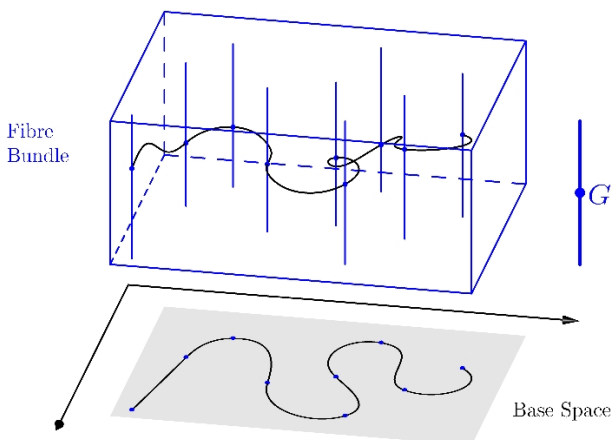
En effet, cette application respecte bien les fibres (i.e. envoie bien les fibres de  $T\mathbb{S}^1$  au-dessus du tout point  $(x, y)$  sur les fibres de  $\mathbb{S}^1 \times \mathbb{R}$  au-dessus du même point) et sa restriction à chaque fibre  $T_{(x,y)}\mathbb{S}^1$  est bien un isomorphisme d'espaces vectoriels. □

## 5.2 Champs de vecteurs

### Sections

**5.2.1 Définition:** Si  $\pi : N \longrightarrow M$  est un fibré de fibre  $F$ , une *section* est une application différentiable  $s : M \longrightarrow N$  telle que  $\pi \circ s = \text{id}_M$ , i.e.  $s(x) \in N_x$ , pour tout  $x \in M$ .

Un petit dessin pour illustrer la notion de sections :



Une *section locale* est une section définie sur un ouvert  $U \subset M$ , c'est-à-dire une application différentiable  $s : U \longrightarrow N_U$  telle que  $\pi \circ s = \text{id}_U$ .

L'ensemble des sections est noté  $\Gamma(M, N)$ , ou  $\Gamma(N)$  si la variété de base  $M$  est fixée. L'ensemble des sections locales définie sur  $U \subset M$  est noté  $\Gamma(U, N_U)$  ou  $\Gamma(N_U)$ , ou encore  $\Gamma_U(M, N)$  ou  $\Gamma_U(N)$ . Parfois on note aussi  $\Gamma_{loc}(M, N)$  ou  $\Gamma_{loc}(N)$  l'ensemble des sections locales, sans spécifier le domaine de définition.

Les sections peuvent être des applications de classe  $C^k$  ou de classe  $C^\infty$  (lisses). Cette régularité peut être indiquée comme  $\Gamma^k(M, N)$  ou  $\Gamma^\infty(M, N)$ , mais elle est très souvent sous-entendue.

Si  $\pi : N \longrightarrow M$  est un fibré de fibre quelconque, l'ensemble de sections  $\Gamma(N)$  n'a aucune structure algébrique particulière. Par contre, pour un fibré vectoriel l'ensemble des sections a une structure qui ressemble à celle de l'ensemble des fonctions réelles sur  $M$ . Considerons un fibré vectoriel  $\pi : E \longrightarrow M$ .

**5.2.2 Définition:** Soit  $\mathcal{A}$  une algèbre associative (on peut aussi définir pour  $\mathcal{A}$  comme anneau unitaire) et  $(V, +)$  un groupe abélien.  $V$  est un  $\mathcal{A}$ -module à gauche si et seulement si  $V$  est un espace vectoriel sur lequel il existe une action :  $V \times \mathcal{A} \longrightarrow V$  qui satisfait les mêmes axiomes que les espaces vectoriels. Ainsi,  $\mathcal{A}$  est le corps des scalaires et  $V$  un  $\mathcal{A}$ -espace vectoriel. On dit que  $V$  est un *bimodule* s'il existe également une seconde action  $\mathcal{A} \times V \longrightarrow V$  de sorte que  $V$  soit un  $\mathcal{A}$ -module à gauche et à droite.

**5.2.3 Proposition:** *L'ensemble  $\Gamma(E)$  est un espace vectoriel sur  $\mathbb{R}$  (de dimension infinie), et un module sur  $C^\infty(M)$ .*

**Preuve:** Pour  $s_1, s_2 \in \Gamma(E)$  et  $\alpha, \beta \in \mathbb{R}$ , on définit la combinaison linéaire  $\alpha s_1 + \beta s_2 \in \Gamma(E)$  par  
 $(\alpha s_1 + \beta s_2)(x) = \alpha s_1(x) + \beta s_2(x), \quad x \in M.$   
 Pour toute fonction  $f \in C^\infty(M)$  et toute section  $s \in \Gamma(E)$ , on définit le produit  $fs \in \Gamma(E)$  par

$$(fs)(x) = f(x) s(x), \quad x \in M$$

Il est facile de vérifier que ces deux opérations font de  $\Gamma(E)$  un module sur  $C^\infty(M)$ . □

**5.2.4 Proposition:** *Il y a une correspondance bijective entre trivialisations locales de  $E$  et bases de sections locales de  $E$ . Par conséquent, l'espace  $\Gamma(E)$  est un module localement libre de rang  $k$  sur  $C^\infty(M)$ .*

**Preuve:** En effet, les sections ont valeur dans les fibres qui sont isomorphes à  $V$ , mais l'isomorphisme n'est pas canonique, il est donné par les trivialisations locales. Plus précisément, fixons une base  $\{e_\alpha, \alpha = 1, \dots, k\}$  de  $V$  et considérons une trivialisatation  $\tau : E_U \rightarrow U \times V$ . Pour tout  $x \in U$ , l'ensemble de vecteurs

$$e_\alpha(x) := \tau^{-1}(x, e_\alpha), \quad \alpha = 1, \dots, k$$

forme alors une base de la fibre  $E_x$ . Les applications

$$e_\alpha : U \rightarrow E_U, \quad x \mapsto e_\alpha(x), \quad \alpha = 1, \dots, k$$

sont donc des sections locales (différentiables) du fibré  $E$ , qui forment une base de  $\Gamma(E_U)$  comme  $C^\infty(U)$ -module.

Cette correspondance est bijective : si pour  $\alpha = 1, \dots, k$  on a des sections locales  $\sigma_\alpha : U \rightarrow E_U$  telles qu'en tout  $x \in U$  l'ensemble  $\{\sigma_\alpha(x), \alpha = 1, \dots, k\}$  soit une base de  $E_U$ , alors on définit une trivialisatation  $\tau_U : E_U \rightarrow U \times V$  par

$$y = \sum_{\alpha=1}^k y^\alpha(x) \sigma_\alpha(x) \quad \mapsto \quad \tau_U(y) = \left( x, \sum_{\alpha=1}^k y^\alpha(x) e_\alpha \right).$$

Sur chaque ouvert  $U$  supportant une trivialisatation  $\tau_U$ , une section  $s : U \rightarrow E_U$  s'écrit donc comme combinaison  $C^\infty(U)$ -linéaire

$$s \Big|_U = \sum_{\alpha=1}^k s^\alpha e_\alpha, \quad s^\alpha \in C^\infty(U).$$

□

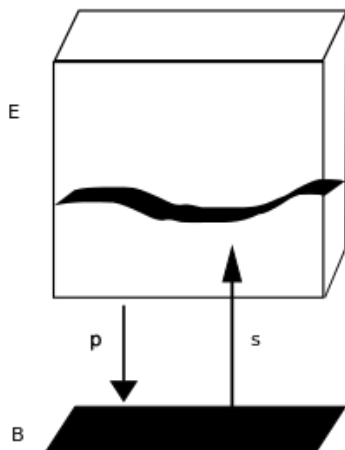
**5.2.5 Corollaire:** *Un fibré  $E$  de rang  $k$  est trivial si et seulement si le  $C^\infty(M)$ -module  $\Gamma(E)$  est libre, i.e. si  $E$  admet une base de  $k$  sections globales.*

**Remarque importante** Si on a un fibré de rang  $k$  admet une base de  $k$  sections globales, alors elles ne s'annulent jamais (auquel cas, ce n'est plus une base).

**5.2.6 Exemple:** 1. Si  $E = M \times \mathbb{R}$ , les sections de  $E$  sont les fonctions à valeurs réels sur  $M$ , i.e.  $\Gamma^\infty(E) = C^\infty(M)$ .

2. Si  $E = M \times \mathbb{R}^k$ , les sections de  $E$  sont les fonctions vectorielles sur  $M$ .

Section  $s$  d'un espace fibré  $p : E \rightarrow B$ .



### Vecteurs tangents et dérivations

**5.2.7 Définition:** Soit  $\mathcal{A}$  une algèbre associative sur  $\mathbb{R}$  et  $\mathcal{M}$  un  $\mathcal{A}$ -bimodule. Une *dérivation sur  $\mathcal{A}$  à valeurs dans  $\mathcal{M}$*  est une application  $D : \mathcal{A} \rightarrow \mathcal{M}$  telle que, pour tout  $\alpha, \beta \in \mathbb{R}$  et  $f, g \in \mathcal{A}$ , on a :

- $D(\alpha f + \beta g) = \alpha D(f) + \beta D(g)$ , i.e.  $D$  est linéaire sur  $\mathbb{R}$  ;
- $D(fg) = D(f)g + fD(g)$ , i.e.  $D$  satisfait la règle de Leibniz.

Si  $\mathcal{A}$  a une unité 1 (et donc  $\mathbb{R} \hookrightarrow \mathcal{A}$ ), on a aussi

- $D(\lambda) = 0$  pour tout  $\lambda \in \mathbb{R} \subset \mathcal{A}$  [car  $D(\lambda) = \lambda D(1) = \lambda D(1) \cdot 1 + \lambda \cdot 1 D(1) = 2 D(\lambda)$ ].

On indique avec  $\text{Der}(\mathcal{A}, \mathcal{M})$  l'ensemble de telles dérivations.

**5.2.8 Proposition:** *L'ensemble  $\text{Der}(\mathcal{A}, \mathcal{M})$  est un espace vectoriel réel.*

Autrement dit, pour tout  $D_1, D_2 \in \text{Der}(\mathcal{A}, \mathcal{M})$  et  $\alpha, \beta \in \mathbb{R}$ , la combinaison linéaire  $\alpha D_1 + \beta D_2$  est définie par

$$(\alpha D_1 + \beta D_2)(f) := \alpha D_1(f) + \beta D_2(f).$$

Soit  $C^\infty(x)$  l'algèbre des fonctions réelles sur  $M$  qui sont  $C^\infty$  en  $x$ . Considerons  $\mathbb{R}$  comme bimodule sur  $C^\infty(x)$  avec l'action d'évaluation en un point  $f \cdot \lambda := f(x)\lambda$  pour tout  $f \in C^\infty(x)$  et  $\lambda \in \mathbb{R}$ .

**5.2.9 Proposition:** *L'espace tangent à  $M$  en  $x$  est isomorphe à l'espace vectoriel des dérivations sur  $C^\infty(x)$  à valeurs dans  $\mathbb{R}$ , c'est-à-dire*

$$T_x M \cong \text{Der}(C^\infty(x), \mathbb{R}).$$

**Preuve:** Un élément de  $T_x M$  est un vecteur  $\gamma'(0)$  tangent à une courbe  $\gamma$  qui passe par  $x$  à l'instant  $t = 0$ . Celui-ci est une dérivation sur  $C^\infty(x)$ , comme on a déjà vu.

Le contraire est aussi vrai : toute dérivation sur  $C^\infty(x)$  à valeurs dans  $\mathbb{R}$  est le vecteur tangent à une courbe en un point régulier. Pour montrer ceci, d'après la Proposition, il suffit de démontrer que toute dérivation sur  $C^\infty(x)$  est une combinaison linéaire des dérivées partielles  $\left. \frac{\partial}{\partial x^i} \right|_x$ , où  $(x^1, \dots, x^m)$  sont les coordonnées locales autour de  $x$ .

Soit  $D$  une dérivation sur  $C^\infty(x)$  à valeurs dans  $\mathbb{R}$ . L'action de  $D$  sur  $f \in C^\infty(x)$  doit être lue en coordonnées locales : soit  $U$  un voisinage de  $x$  qui supporte une carte locale  $\varphi : U \rightarrow V \subset \mathbb{R}^m$ , supposons que  $\varphi(x) = (0, \dots, 0) \equiv 0$ , appelons  $(x^1, \dots, x^m)$  les coordonnées des points voisins à  $x$  dans  $U$  et  $\tilde{f} = f \circ \varphi^{-1}$  l'expression de  $f$  en coordonnées locales. Alors par  $Df$  on entend  $D\tilde{f}$ .

Calculons la valeur de  $D\tilde{f}$  en utilisant le développement de Taylor à l'ordre 1 de  $\tilde{f}$  autour de 0 :

$$\tilde{f}(x^1, \dots, x^m) = \tilde{f}(0) + \sum_{i=1}^m x^i \frac{\partial \tilde{f}(0)}{\partial x^i} + \sum_{i,j=1}^m x^i x^j \frac{\partial^2 \tilde{f}(v)}{\partial x^i \partial x^j},$$

où  $v$  est un point de  $V$  qui se trouve dans le disque centré en 0 de rayon  $\|(x^1, \dots, x^m)\|$ . En utilisant le fait que la dérivation  $D$  s'annule sur les constantes, on a alors

$$D\tilde{f} = \sum_{i=1}^m Dx^i \frac{\partial \tilde{f}(\vec{0})}{\partial x^i} + 2 \sum_{i,j=1}^m x^i(0) Dx^j \frac{\partial^2 \tilde{f}(\vec{y})}{\partial x^i \partial x^j},$$

où  $x^i(0)Dx^j = x^i \cdot Dx^j$  est l'expression en coordonnées locales de l'action de  $C^\infty(x)$  sur  $\mathbb{R}$ . Puisque  $x^i(0) = 0$  par hypothèse (car  $\varphi(x) = 0$ ), on en conclue que

$$D\tilde{f} = \sum_{i=1}^m Dx^i \frac{\partial \tilde{f}(\vec{0})}{\partial x^i} = \sum_{i=1}^m Dx^i \left. \frac{\partial}{\partial x^i} \right|_x \tilde{f}.$$

Par conséquent, tout vecteur tangent  $X_x \in T_x M$  est identifié à une dérivation sur  $C^\infty(x)$  qui agit comme  $f \mapsto X_x f$ .  $\square$

### Champs de vecteurs

**5.2.10 Définition:** Un *champ de vecteurs* sur  $M$  est une section  $X : M \longrightarrow TM$  du fibré tangent sur  $M$ .

On note  $\mathfrak{X}(M) = \Gamma(TM)$  l'ensemble des champs de vecteurs sur  $M$ . D'après la Section 3,  $\mathfrak{X}(M)$  est un espace vectoriel sur  $\mathbb{R}$  et un module localement libre de rang  $m$  sur  $C^\infty(M)$ . Une base locale de champs de vecteurs (sur un ouvert de trivialisatation  $TU$ ) est donnée par les dérivées partielles

$$\left. \frac{\partial}{\partial x^i} \right|_U : U \longrightarrow TU = TM \Big|_U, \quad x \mapsto \left. \frac{\partial}{\partial x^i} \right|_x, \quad i = 1, \dots, m.$$

Tout champ de vecteurs  $X$  sur  $M$  s'exprime alors localement comme combinaison  $C^\infty(U)$ -linéaire

$$X \Big|_U = \sum_{i=1}^m X^i \left. \frac{\partial}{\partial x^i} \right|_U, \quad X^i \in C^\infty(U).$$

**5.2.11 Proposition:** Les champs de vecteurs sur  $M$  sont isomorphes aux dérivations de  $C^\infty(M)$  à valeurs dans  $C^\infty(M)$ , i.e.

$$\mathfrak{X}(M) \cong \text{Der}(C^\infty(M), C^\infty(M)).$$

**Preuve:** Soit  $X$  un champ de vecteur, avec expression en coordonnées locales  $X \Big|_U = \sum_{i=1}^m X^i \left. \frac{\partial}{\partial x^i} \right|_U$  dans une carte  $(U, \varphi)$ . Pour tout  $f \in C^\infty(M)$ , on considère la fonction  $Xf$  sur  $M$  définie en tout point  $x \in U$  par

$$(Xf)(x) = \left( X \Big|_U f \right)(x) = \sum_{i=1}^m X^i(x) \left. \frac{\partial f}{\partial x^i} \right|_x = \sum_{i=1}^m X^i(x) \frac{\partial (f \circ \varphi^{-1})}{\partial x^i}(\varphi(x)).$$

Alors  $Xf$  est une fonction  $C^\infty$ , car sur toute carte  $(U', \varphi')$  son expression en coordonnées locales  $\widetilde{Xf} = Xf \circ \varphi'$  est  $C^\infty$  : pour tout  $(y^1, \dots, y^m) \in \varphi'(U' \cap U)$ , avec  $x \in U' \cap U$ , on a

$$\widetilde{Xf}(y^1, \dots, y^m) = \sum_{i=1}^m X^i(x) \frac{\partial (f \circ \varphi^{-1})}{\partial x^i}(\varphi(x)) = \sum_{i,j=1}^m X^i(x) \frac{\partial (f \circ (\varphi')^{-1})}{\partial y^j}(\varphi'(x)) \frac{\partial y^j}{\partial x^i}(\varphi(x)),$$

où  $(y^1, \dots, y^m) = \psi(x^1, \dots, x^m)$  est le changement de coordonnées de  $U$  à  $U'$ . Dans cette expression, tous les facteurs sont de classe  $C^\infty$ . Enfin, l'application  $f \mapsto Xf$  est évidemment une dérivation.

Inversement, comme on l'a montré pour les dérivations en un point et les vecteurs tangent, on montre que toute dérivation sur  $C^\infty(M)$  est une combinaison  $C^\infty(M)$ -linéaire des opérateurs locaux  $\left. \frac{\partial}{\partial x^i} \right|_x$ , et donc un champ de vecteurs sur  $M$ .  $\square$

### 5.2.12 Exemple: [Champs de vecteurs sur les sphères.]

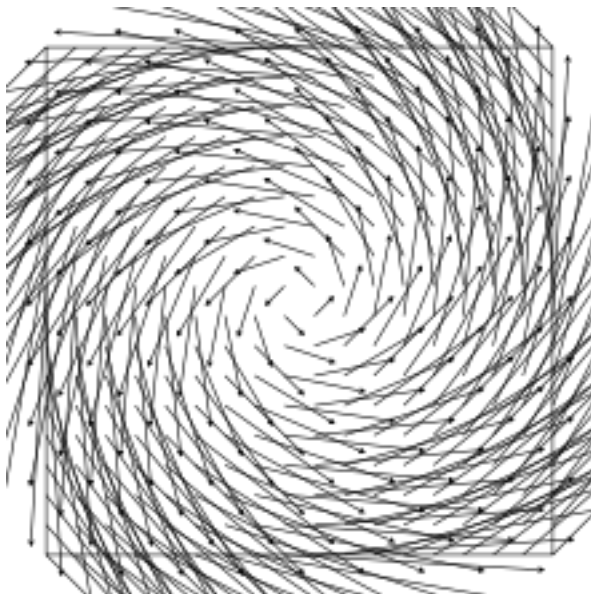
Sur  $\mathbb{S}^1 = \{e^{it}\}$ , tout champ de vecteurs est de la forme  $X_t = f(t)\frac{d}{dt}$ , avec  $f \in C^\infty(\mathbb{S}^1)$ .

Sur  $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 1\}$ , tout champ de vecteur est de la forme

$$X_{(x,y,z)} = f(x, y, z) \frac{\partial}{\partial x} + g(x, y, z) \frac{\partial}{\partial y} + h(x, y, z) \frac{\partial}{\partial z},$$

avec  $f, g, h \in C^\infty(\mathbb{S}^2)$  telles que  $x f(x, y, z) + y g(x, y, z) + z h(x, y, z) = 0$ .

Un exemple de champ de vecteurs, de la forme  $V(x, y) = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} = (-y, x)$ .



## 5.3 Fibrés trivialisables

**5.3.1 Proposition:** *S'il existe une nappe paramétrée (qui vérifie les mêmes conditions que pour la définition d'une sous variété) définie globalement sur une variété  $M$  (c'est à dire, qui paramétrise  $M$  tout entière), alors  $M$  est trivialisable.*

**Preuve:** Soit  $j : U \subseteq \mathbb{R} \rightarrow M$  une nappe paramétrée globale de  $M$ , (qui vérifie les mêmes conditions que pour la définition d'une sous variété). Alors, si  $U$  est ouvert,  $j$  est une carte globale de  $M$ . Si  $U$  est fermé, on peut trouver des ouverts  $U_i$  formant un recouvrement de  $U$  (dans le cas contraire, on se restreint à considérer  $M$  comme une variété à bord), de sorte que chaque  $j_i = j : U_i \subseteq \mathbb{R} \rightarrow M$  est une carte de  $M$ . Alors, dans ce cas, on peut trouver une base du fibré tangent en considérant les dérivées partielles de  $j$  (c'est à dire  $dj$ , la différentielle de  $j$ ), qui sont des vecteurs  $(e_1, \dots, e_n)$  (linéairement indépendants car  $dj$  est injective) et qui dépendent des coordonnées  $(x_1, x_2, \dots, x_n)$ . On prend alors comme trivialisations globales

$$\tau : TM \rightarrow M \times \mathbb{R}^n$$

$$\tau\left(\sum_{i=1}^n \lambda_i e_i\right) = (j(x_1, \dots, x_n), \lambda_1, \dots, \lambda_n)$$

□

Ce qu'on vient de faire montre également (par contraposée) que s'il n'existe pas de trivialisations globales, alors il n'existe pas de paramétrisations globales. Ce point est important car dans certains cas, il est possible de trouver une paramétrisation globale (vérifiez qu'on peut le faire pour le cercle  $\mathbb{S}^1$  et



trouvez des ouverts adaptés). Par contre, il est impossible de le faire pour la sphère  $\mathbb{S}^2$ , c'est pourquoi, dans la paramétrisation courante de la sphère, il existe deux points (opposés, qu'on appelle pôles) de non injectivité.

Dans la partie fibré tangent, on a montré avec l'exemple du cercle comment déterminer un tel fibré et comment montrer qu'il était trivialisable. On va maintenant refaire la même chose en utilisant cette fois la correspondance entre sections et champs de vecteurs.

**5.3.2 Exemple:** On va déterminer le fibré tangent du tore  $T^2$  (tore de révolution).

Pour cela, on utilise une paramétrisation du tore.

$$f(u, v) = ((a + R\cos(u))\cos(v), (a + R\cos(u))\sin(v), R\sin(u))$$

$$\frac{\partial f}{\partial u} = (-R\sin(u)\cos(v), -R\sin(u)\sin(v), R\cos(u)); \quad \frac{\partial f}{\partial v} = (-(a + R\cos(u))\sin(v), (a + R\cos(u))\cos(v), 0)$$

On voit que l'on peut simplifier l'un des vecteurs : on peut enlever le facteur  $(a + R\cos(u))$  de  $\frac{\partial f}{\partial v}$ . On pose

$$e_2 = (-\sin(v), \cos(v), 0) = \frac{1}{(a + R\cos(u))} \frac{\partial f}{\partial v}$$

$$e_3 = \frac{\partial f}{\partial u}.$$

$$e_1 = e^2 \wedge e^3 = (-R\cos(u)\cos(v), -R\cos(u)\sin(v), -\sin(u))$$

Le plan tangent est alors l'ensemble :

$$\{\lambda, \mu \in \mathbb{R}, \lambda e_2 + \mu e_3\}$$

Pour le tore : on va directement donner une trivialisatation globale. On vérifie que la paramétrisation du tore est une bijection pour  $0 \leq u, v < 2\pi$ , c'est donc une paramétrisation globale. On donne alors :

$$\phi : TT^2 \longrightarrow T^2 \times \mathbb{R}^2$$

$$\phi(\lambda e_2 + \mu e_3) = (f(u, v), \lambda, \mu)$$

qui respecte les fibres et sa restriction à chaque fibre est un isomorphisme d'espaces vectoriels.

**5.3.3 Exemple:** On va déterminer le fibré tangent de la sphère  $S^2$ .

Pour cela, on utilise une paramétrisation de la sphère unité :

$$f(u, v) = (\cos(u)\cos(v), \cos(u)\sin(v), \sin(u))$$

$$\frac{\partial f}{\partial u} = (-\sin(u)\cos(v), -\sin(u)\sin(v), \cos(u)); \quad \frac{\partial f}{\partial v} = (-\cos(u)\sin(v), \cos(u)\cos(v), 0)$$

On supprime le facteur  $\cos(u)$  du vecteur  $\frac{\partial f}{\partial v}$  et on pose :

$$e_2 = \frac{\partial f}{\partial u} = (-\sin(u)\cos(v), -\sin(u)\sin(v), \cos(u))$$

$$e_3 = (-\sin(v), \cos(v), 0) = \frac{1}{\cos(u)} \frac{\partial f}{\partial v}$$

Enfin,

$$e_1 = e_2 \wedge e_3 = (-\cos(u)\cos(v), -\cos(u)\sin(v), -\sin(u)) = -f(u, v) = e_1$$

On obtient une base de l'espace tangent  $\forall (u, v) \in \mathbb{R}^2$ . Le plan tangent est alors :

$$\{\lambda, \mu \in \mathbb{R}, \lambda e_2 + \mu e_3\}$$

On choisit :  $0 \leq v < 2\pi$  et  $-\frac{\pi}{2} \leq u \leq \frac{\pi}{2}$  pour avoir une bijection (sauf pour  $u = \pm \frac{\pi}{2}$ ) On prend alors comme trivialisations locales : ( $U_N$  et  $U_S$  étant encore les cartes aux pôles nord et sud).

$$\begin{aligned}\tau_S : T_{U_S} \mathbb{S}^2 &\longrightarrow U_S \times \mathbb{R}^2, \lambda e_2 + \mu e_3 \longmapsto (f(u, v), (\lambda, \mu), u \neq \pm \frac{\pi}{2}) \\ \tau_N : T_{U_N} \mathbb{S}^2 &\longrightarrow U_N \times \mathbb{R}^2, \lambda e_2 + \mu e_3 \longmapsto (f(u, v), (\lambda, \mu), u \neq \pm \frac{\pi}{2})\end{aligned}$$

On vérifie que la restriction

$$\tau_S |_{f(u,v)} T_{u,v} \mathbb{S}^2 \longrightarrow \mathbb{R}^2; \lambda e_2 + \mu e_3 \longmapsto (\lambda, \mu)$$

(resp  $\tau_N$ ) est bien un isomorphisme d'espaces vectoriels, donc  $\tau_S$  et  $\tau_N$  sont bien des trivialisations locales, qu'on peut "lire" à partir des cartes locales comme on l'avait fait avec le cercle.

Le problème : au pôles, la paramétrisation choisie n'est pas une bijection, mais une surjection. On doit donc définir une autre paramétrisation de la sphère permettant d'inclure ces 2 pôles, par exemple :

$$g(u, v) = (\sin(u), \cos(u)\cos(v), \cos(u), \sin(v))$$

En refaisant un travail analogue à celui-ci, on trouve 2 vecteurs  $e_5, e_6$  base de l'espace tangent. La paramétrisation n'est toujours pas définie pour  $u \neq \pm \frac{\pi}{2}$ , mais comme on a échangé  $x, y$  et  $z$ , ce n'est plus les mêmes pôles qui posent problèmes. On a donc 2 trivialisations locales pour cette paramétrisation,  $\tau'_S, \tau'_N$ , qu'on peut réduire à 2 pour la sphère en choisissant de garder  $\tau_S$  et  $\tau'_S$ . Cependant, il est alors impossible avec cela de trouver une paramétrisation globale, car on a 2 caractérisations de l'espace tangent. Il faudrait pouvoir paramétrer la sphère de façon globale pour résoudre le problème. La proposition ci dessous montre que cela n'est pas possible.

On va prouver que le fibré tangent à  $\mathbb{S}^2$  n'est pas trivialisable. Avant cela, on a besoin d'un outil géométrique bien particulier :

**5.3.4 Définition:** Un *lacet* sur  $[a, b]$  est une courbe paramétrée fermée. Soit :

$$\begin{aligned}\gamma : [a, b] &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (x(t), y(t))\end{aligned}$$

un lacet  $C^1$  dont le support ne contient pas l'origine  $O$ . Le *nombre de tours* (ou encore, *nombre d'enroulements*)  $N(\gamma, O)$  de  $\gamma$  par rapport à  $O$  est l'entier relatif

$$n = \frac{1}{2\pi} \int_a^b \frac{xy' - x'y}{x^2 + y^2} dt \in \mathbb{Z}$$

**Interprétation géométrique :** écrivons  $\gamma$  en coordonnées polaires.  $\gamma(t) = (x(t), y(t))$  et

$$x(t) = r(t)\cos(\theta(t)); y(t) = r(t)\sin(\theta(t))$$

On a alors

$$\int_a^b \frac{xy' - x'y}{x^2 + y^2} dt = \int_a^b \theta'(t) dt = \theta(b) - \theta(a)$$

Comme  $\gamma(b) = \gamma(a)$ ,  $\exists n \in \mathbb{Z}$  tel que  $\theta(b) - \theta(a) = 2\pi n$ . Donc  $N(\gamma, O) = n \in \mathbb{Z}$ , ce qui permet de vérifier que  $N(\gamma, O)$  est un entier.

**5.3.5 Définition:** Soit :  $\gamma : [a, b] \longrightarrow \mathbb{R}^2$  un lacet régulier, on appelle *indicatrice* de  $\gamma$  la courbe  $\gamma' : [a, b] \longrightarrow \mathbb{R}^2 \setminus O$ . Le nombre de tours de  $\gamma'$  par rapport à  $O$  s'appelle *l'indice de rotation* de  $\gamma$  et il se note

$$Ind(\gamma) = N(\gamma', O)$$

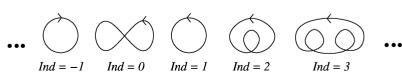
Illustration de la fonction indicatrice d'une courbe paramétrée.

CM G4 :  
Propriétés  
globales des  
courbes  
V. Borrelli

Indice de  
rotation  
Le théorème  
de Jordan  
Camille  
Jordan  
L'inégalité  
isopérimétrique  
Wilhelm  
Wirtinger  
Le théorème  
des quatre  
sommets

### Indice de rotation

**Exemples.-**



...  $Ind = -1$     $Ind = 0$     $Ind = 1$     $Ind = 2$     $Ind = 3$  ...

**Théorème.-** Soit  $\gamma : [a, b] \xrightarrow{C^1} \mathbb{R}^2$  un lacet régulier alors

$$Ind(\gamma) = -\frac{1}{2\pi} \int_a^b k_{alg}(t) \|\gamma'(t)\| dt.$$

En particulier, si  $\gamma$  est paramétrée par la l.a. alors :

$$Ind(\gamma) = -\frac{1}{2\pi} \int_a^b k_{alg}(s) ds.$$

**remarque :** On peut montrer par un simple calcul que  $Ind(\gamma) = -\frac{1}{2\pi} \int_a^b k_{alg}(t) \|\gamma'(t)\| dt$ . Avec  $k_{alg}$  la courbure algébrique.

**5.3.6 Exemple:** Le fibré tangent à  $S^2$  n'est pas trivialisable.

**Preuve:**

**Explication de la stratégie :** On choisit sur la sphère un pôle nord et un pôle sud, ainsi qu'une orientation. On peut alors parler des parallèles de la sphère et les orienter de manière continue. De plus, on peut définir un repère mobile tangent sur la sphère, en dehors des pôles nord et sud. Le premier vecteur du repère mobile est tangent au parallèle passant par ce point et orienté dans la direction de parcours des parallèles, et le second vecteur est perpendiculaire au premier, avec un sens choisi pour se conformer à l'orientation de la sphère. Tout champ de vecteurs  $v$  admet une décomposition dans le repère mobile. On va maintenant supposer (par l'absurde) que  $v$  ne s'annule jamais.

On peut donc associer à tout parallèle un nombre d'enroulement : c'est le nombre d'enroulement du champ de vecteurs dans le repère mobile le long de ce parallèle. Ce nombre est bien défini puisque le champ de vecteurs ne s'annule pas. Il dépend continûment de la latitude du parallèle, d'après les résultats standard sur la continuité du nombre d'enroulement, et il est entier. Il est donc constant.

On calcule ce nombre d'enroulement au voisinage du pôle nord. Certes, le repère mobile cesse d'être défini au pôle nord, mais pour pallier cette difficulté, on dilate un parallèle proche du pôle nord en un cercle de rayon unité, et on projette à la fois le champ de vecteurs  $v$  et le repère mobile sur le plan tangent au pôle nord. L'orientation de ce plan tangent est déduite de celle de la sphère. Par continuité, le nombre

d'enroulement ne change pas lors de cette transformation. À la limite, le repère mobile transformé tourne une fois autour du pôle, dans le sens des parallèles, et donc le champ de vecteurs tourne - une fois dans le repère mobile. Le nombre d'enroulement vaut donc  $m$  et  $m$  vaut  $+1$  ou  $-1$ , suivant le choix d'orientation des parallèles.

Avec le même raisonnement au voisinage du pôle sud, le repère mobile tourne une fois autour du pôle sud dans le sens des parallèles, mais pour maintenir une orientation cohérente avec celle de la sphère, en tant que plan dans l'espace à trois dimensions, le plan tangent au pôle sud est orienté dans le sens opposé du plan tangent au pôle nord, et donc le nombre d'enroulement du champ est  $-m$ . On a la contradiction : le nombre d'enroulement étant continu, cela n'est pas possible. Cela signifie que le champs de vecteur s'annule nécessairement, or, on a fait la remarque que pour avoir une base d'un fibré, les sections (dans ce cas, les champs de vecteurs), ne doivent jamais s'annuler.

### Version analytique :

Soit  $x$  le point courant de la sphère,  $v(x)$  un champ de vecteur quelconque. On paramétrise la sphère en coordonnées polaires, en la supposant de rayon 1 :

$$x_1 = \cos(\theta)\cos(\phi), x_2 = \sin(\theta)\cos(\phi), x_3 = \sin(\phi)$$

avec

$$0 \leq \theta \leq 2\pi; -\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}$$

On remarque que ce paramétrage ne définit pas une bijection (il est non injectif en  $\phi = \pm\frac{\pi}{2}$ ) ni, à fortiori, un difféomorphisme. En chaque point de la sphère distinct des pôles nord et sud, on introduit un repère mobile formé de 2 vecteurs tangents à la sphère :

$$e_1(\theta, \phi) = (-\sin(\theta), \cos(\theta), 0); e_2(\theta, \phi) = (-\sin(\phi)\cos(\theta), -\sin(\phi)\sin(\theta), \cos(\phi))$$

Il est évident que cette base dépend continûment de  $\theta$  et  $\phi$  pour  $0 \leq \theta \leq 2\pi; -\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}$ . Dans cette base et en dehors des pôles, le champ de vecteur  $v(x)$  admet pour coordonnées :

$$w_1(\theta, \phi) = e_1(\theta, \phi) \cdot v(x(\theta, \phi)); w_2(\theta, \phi) = e_2(\theta, \phi) \cdot v(x(\theta, \phi))$$

Ces coordonnées dépendent là encore continûment de  $\theta$  et  $\phi$  pour  $0 \leq \theta \leq 2\pi; -\frac{\pi}{2} \leq \phi < \frac{\pi}{2}$ . De plus, elles dépendent périodiquement de  $\theta$ , avec une période de  $2\pi$  et ne s'annulent pas simultanément.

Pour  $\phi$  compris strictement entre  $-\frac{\pi}{2}$  et  $\frac{\pi}{2}$ , on peut définir le nombre d'enroulement du champ de vecteurs  $(w_1(\theta, \phi), w_2(\theta, \phi))$  par rapport à la variable  $\theta$  sur le cercle unité. C'est un entier  $n$  qui est une fonction continue de  $\phi \in ]-\frac{\pi}{2}, \frac{\pi}{2}$ , donc constant. On va maintenant calculer ce nombre au voisinage des pôles nord et sud afin de parvenir à une contradiction.

Soit  $\bar{v}$  la valeur du champ de vecteur au pôle nord. Ses trois composantes sont :  $(\bar{v}_1, \bar{v}_2, 0)$ , qu'on note  $(\bar{v}_1, \bar{v}_2)$  après dilatation. Par continuité du champs, on peut trouver  $\alpha > 0$  tel que :

$$\frac{\pi}{2} - \alpha \leq \phi < \frac{\pi}{2} \Rightarrow |v(x(\theta, \phi)) - \bar{v}| \leq \frac{1}{4}$$

On peut également choisir  $\alpha$  assez petit pour avoir

$$|\bar{v}| |1 - \sin(\phi)| \leq \frac{1}{4}$$

Notons :  $\bar{z}(\theta) = \begin{pmatrix} -\bar{v}_1 \sin(\theta) + \bar{v}_2 \cos(\theta) \\ -\bar{v}_1 \cos(\theta) - \bar{v}_2 \sin(\theta) \end{pmatrix} = \begin{pmatrix} \bar{z}_1(\theta) \\ \bar{z}_2(\theta) \end{pmatrix}$  le champ de vecteur au pôle nord.

Dans ce cas :

$$w_1(\theta) = e_1 \cdot \bar{z}(\theta) = -\bar{v}_1$$

$$w_2(\theta) = e_2 \cdot \bar{z}(\theta) = \sin(\phi)(\bar{v}_1 \sin(2\theta) - \bar{v}_2 \cos(2\theta))$$

À partir de là, on déduit par le calcul :

$$|w_1(\theta) - z_1(\theta)| \leq \frac{1}{4}; |w_2(\theta) - z_2(\theta)| \leq \frac{1}{4} + |1 - \sin(\phi)| |\bar{v}| \leq \frac{1}{2}$$

Pour  $\frac{\pi}{2} - \alpha \leq \phi < \frac{\pi}{2}$ . Par conséquent,  $|w(\theta, \phi) - (\theta)| \leq \frac{\sqrt{5}}{4} \leq 1 = |\bar{z}|$ . Donc  $w(\theta, \phi)$  a le même nombre d'enroulement (en la variable  $\theta$ ) que  $\bar{z}$ , c'est donc  $-1$ .

Le même raisonnement au pôle sud montre que si l'on pose :  $\hat{z}(\theta) = \begin{pmatrix} -\hat{v}_1 \sin(\theta) + \hat{v}_2 \cos(\theta) \\ \hat{v}_1 \cos(\theta) + \hat{v}_2 \sin(\theta) \end{pmatrix} = \begin{pmatrix} \hat{z}_1(\theta) \\ \hat{z}_2(\theta) \end{pmatrix}$

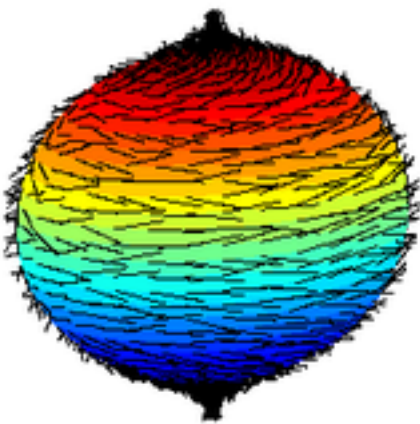
, alors pour  $\phi$  assez proche de  $-\frac{\pi}{2}$ , le nombre d'enroulement (en la variable  $\theta$ ) est le même que celui de  $\hat{z}$ , c'est donc  $1$ . On a bien la contradiction souhaitée.  $\square$

## 5.4 Théorème de la boule chevelue\*

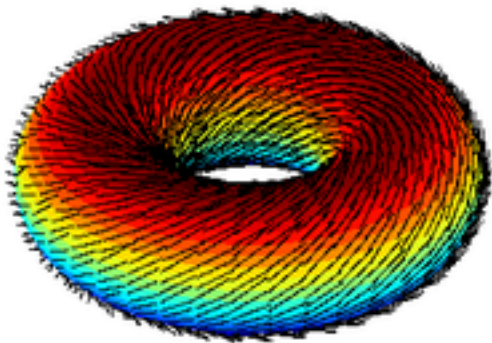
On s'intéresse à la démonstration du théorème de la boule chevelue, mais avant cela, on doit introduire un certain nombre de notions. Je vais présenter ici, non pas la démonstration la plus simple, mais la plus accessible à mon niveau. C'est une preuve alternative permettant de montrer que  $\mathbb{S}^2$ , et plus généralement, de  $\mathbb{S}^{2k}$  avec  $k$  entier, n'est pas globalement trivialisable.

**5.4.1 Théorème:** (boule chevelue) Sur une sphère réelle  $\mathbb{S}^n$  dont la dimension  $n$  est pair, tout champ de vecteurs continu  $X$  s'annule en au moins un point :  $\exists v$  (qui dépend de  $X$ ) tel que  $X(v) = 0$ .

Voici la version visuelle de ce théorème : Si un champ de vecteurs tangents sur une sphère est symbolisé par des cheveux de longueur constante, le théorème de la boule chevelue stipule que la sphère contient au moins un épi. La figure en contient deux, un sur chaque pôle.



Pour le tore, la situation est différente.



### Rappels (donc non démontrés)

**5.4.2 Définition:** Une fonction  $f : E \rightarrow \mathbb{R}$ , est dite *Lipschitzienne* si il existe un réel positif  $k$  tel que

$$\forall x, y \in E, \quad |f(x) - f(y)| \leq k |x - y|$$

.

**5.4.3 Théorème:** Soit  $f : E \rightarrow \mathbb{R}$  dérivable : S'il existe  $m, M \in \mathbb{R}$  vérifiant

$$\forall x \in I, \quad m \leq f'(x) \leq M$$

, alors  $\forall a, b \in I; a \leq b$ ,

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

De cette inégalité, on déduit :

**5.4.4 Théorème:** Soient  $U$  un ouvert de  $\mathbb{R}^n$  et  $f$  une fonction différentiable à valeurs dans  $\mathbb{R}^p$ . Soient encore  $a$  et  $b$  deux points tels que le segment d'extrémités  $a$  et  $b$  dans  $\mathbb{R}^n$ , noté  $[a, b]$ , soit contenu dans  $U$ . On suppose qu'il existe un réel positif  $k$  tel que la norme de la différentielle  $df_x$  vérifie :

$$\forall x \in [a, b], \quad \|df_x\| \leq k$$

Alors :

$$|f(b) - f(a)| \leq k |b - a|$$

**5.4.5 Corollaire:** Toute fonction continûment différentiable sur un intervalle borné  $W$  est Lipschitzienne.

**Preuve:** Pour une fonction réelle, c'est évident, pour une fonction vectorielle, il suffit de poser  $k = \sup_{x \in W} (\|df_x\|)$ .  $\square$

On sort maintenant des rappels et on introduit quelques nouveautés.

**5.4.6 Définition:** Soient  $E$  et  $F$  deux espaces vectoriels sur un même corps commutatif  $K$ . Une fonction  $f$  de  $E$  dans  $F$  est dite *homogène de degré  $\alpha$*  si

$$\forall k \in K, \forall x \in E, f(kx) = k^\alpha f(x)$$

Si  $K = \mathbb{R}$ , on dit que  $f$  est positivement homogène si

$$\forall t \geq 0, \forall x \in E, f(tx) = t^\alpha f(x)$$

**5.4.7 Proposition:** Soient  $E$  et  $F$  deux espaces vectoriels sur un même corps commutatif  $K$ . Soit  $N$  une norme  $N : F \rightarrow \mathbb{R}^+$ . Soient  $f : E \rightarrow F$  et  $g : E \rightarrow F$  deux fonctions homogène de degré  $\alpha$  vérifiant  $f(z) = g(z)$  ( $E$ ) ou  $N(f(z)) \leq N(g(z))$  ( $I$ ) pour un certain  $z \in E$ . Alors l'égalité ( $E$ ) et l'inégalité ( $I$ ) sont vérifiées pour tout élément de la forme  $kz$ ,  $k \in K$ .

**Preuve:** On a  $f(z) = g(z)$  (respectivement  $N(f(z)) \leq N(g(z))$ ). Donc  $f(kz) = k^\alpha f(z) = k^\alpha g(z) = g(kz)$  (respectivement  $N(f(kz)) = |k|^\alpha N(f(z)) \leq |k|^\alpha N(g(z)) = N(g(kz))$ ). Dans le théorème, on utilise cette propriété des fonctions homogènes.  $\square$

**5.4.8 Définition:** Une application  $k$ -contractante, ou encore, une  $k$ -contraction est une application  $k$ -Lipschitzienne avec  $0 \leq k < 1$ .

**5.4.9 Théorème:** Soient  $E$  un espace de Banach  $f$  une application  $k$ -contractante ( $k$ -Lipschitzienne) de  $E$  dans  $E$ . Il existe un unique point fixe  $x^*$  (tel que  $f(x^*) = x^*$ ). De plus, toute suite d'éléments de  $E$  vérifiant la relation de récurrence

$$x_{n+1} = f(x_n)$$

vérifie la majoration

$$d(x_n, x^*) \leq \frac{k^n}{1-k} d(x_0, x_1)$$

, et donc converge.

**Preuve:**

Existence : Soit  $x_0 \in E$  quelconque et pour tout entier naturel  $m$ ,  $x_m = f^m(x_0)$  (itérée  $m$  fois) L'application  $f$  étant  $k$ -Lipschitzienne, on a, par simple récurrence :  $d(x_m, x_{m+1}) \leq k^m d(x_0, x_1)$ . Par application successive de l'inégalité triangulaire,  $\forall n \in \mathbb{N}, \forall p \in \mathbb{N}^*$ ,

$$\begin{aligned} d(x_m, x_{m+p}) &\leq d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \dots + d(x_{m+p-1}, x_{m+p}) \\ &\leq (k^m + k^{m+1} + \dots + k^{m+p})d(x_0, x_1) = k^m \frac{1-k^{p+1}}{1-k} d(x_0, x_1) \leq \frac{k^m}{1-k} d(x_0, x_1) \end{aligned}$$

Ce majorant tend vers 0 quand  $n$  tend vers l'infini, donc  $(x_n)_{n \in \mathbb{N}}$  est une suite de Cauchy. Comme  $E$  est complet, cette suite de Cauchy converge vers une limite  $x^*$  et vérifie la majoration annoncée de l'erreur (en faisant tendre  $p$  vers l'infini :  $\lim_{p \rightarrow \infty} x_{p+n} = x^*$ ).

Unicité : Soient  $x^*$  et  $x^{**}$  deux points fixes de  $f$ . On a alors

$$d(x^*, x^{**}) = d(f(x^*), f(x^{**})) \leq kd(x^*, x^{**})$$

mais

$$0 \leq k < 1 \Rightarrow d(x^*, x^{**}) = 0 \Rightarrow x^* = x^{**}$$

$\square$

Pour la suite, on notera  $|v|$  la norme euclidienne d'un vecteur  $v$  de  $X$  et  $(v|w)$  le produit scalaire euclidien de  $v$  et  $w$  appartenant à  $X$ .

**Preuve:** du théorème de la boule chevelue

Soit  $X$  un champ de vecteurs tangents continu sur  $\mathbb{S}^n$ . Ce champs de vecteur  $X$  est une fonction définie sur la sphère  $\mathbb{S}^n$  et à valeur dans  $\mathbb{R}^{n+1}$ , avec  $n$  pair. On va démontrer par un raisonnement par l'absurde qu'il s'annule forcément en un point  $v$  de la sphère.

1/Preuve dans le cas d'un champ de vecteur continûment différentiable :

Un champ de vecteurs tangents à la sphère  $\mathbb{S}^n$  est continûment différentiable, s'il est la restriction d'un

champ continûment différentiable défini au voisinage de  $\mathbb{S}^n$ . Comme on est dans la sphère, il suffit que l'extension  $u \mapsto X(\frac{u}{|u|})$  soit continûment différentiable sur la couronne solide  $\frac{1}{2} < |u| < \frac{3}{2}$ .

Puisque le champ de vecteur est tangent, cela signifie que  $X(v)$  est orthogonal à  $v$  (pour tout  $v \in \mathbb{S}^n$ ). Supposons que  $X$  ne s'annule jamais sur la sphère  $\mathbb{S}^n$ . Comme  $X$  est continûment différentiable, sa norme  $|X|$  l'est également sur la sphère, et cette norme est non nulle. Par conséquent,  $Y = \frac{X}{|X|}$  est un champ tangent à la sphère, continûment différentiable et pour tout  $v \in \mathbb{S}^n$ ,  $Y(v)$  est de norme 1.

Comme  $v \mapsto Y(\frac{v}{|v|})$  est continûment différentiable sur la couronne  $\frac{1}{2} < |v| < \frac{3}{2}$ , il y est en particulier lipschitzien, il existe donc un nombre réel  $L$  tel que pour tout  $v, w$  dans  $\mathbb{S}^n$  l'inégalité de Lipschitz est vérifiée :

$$\forall v, w \in \mathbb{S}^n, |Y(v) - Y(w)| \leq L |v - w|$$

On définit un champ de vecteur homogène de degré 1  $Z(v) = |v| Y(\frac{v}{|v|})$ . Alors,  $Z$  est lipschitzien. En effet, supposons tout d'abord  $|v| \geq |w| = 1$ , on a alors

$$\begin{aligned} Z(v) - Z(w) &= |v| Y(\frac{v}{|v|}) - |Y(\frac{v}{|v|}) + |Y(\frac{v}{|v|}) - |w| Y(\frac{w}{|w|}) \\ &= (|v| - 1)Y(\frac{v}{|v|}) + (Y(\frac{v}{|v|}) - Y(\frac{w}{|w|})) \end{aligned}$$

(car  $|w| = 1$ ). Donc, en passant à la norme :

$$|Z(v) - Z(w)| = (|v| - 1) + L |v - w| = (|v| - |w|) + L |v - w| \leq (L + 1) |v - w|$$

Par homogénéité, on obtient le cas général (c'est à dire, quelle que soit la norme de  $|v|$  et  $|w|$ ). Soit une fonction  $f$  sur  $\mathbb{R}^{n+1} \times ]0, 1]$  définie par la formule suivante :

$$f(v, t) = v + tZ(v), v \neq 0$$

$$f(0, t) = 0$$

On voit que la fonction  $f$  est homogène de degré 1 par rapport à  $v$ . De plus, la norme  $|f(v, t)|$  est facile à calculer avec le théorème de Pythagore :

$$|f(v, t)|^2 = |v|^2 + 2(v | tZ(v)) + |tZ(v)|^2 = |v|^2 (1 + |t|^2)$$

car  $|Z(v)|^2 = 1$  et  $(v | X(v)) = 0$  et donc  $2(v | tZ(v)) = 0$  en utilisant la définition de  $Z$  à partir de  $X$ .

En d'autres termes, l'image de la sphère de rayon  $r > 0$  par  $f(v, t)$  (en considérant  $t$  comme une constante) est incluse dans la sphère de rayon  $r\sqrt{1+t^2}$ . Si  $r = 0$ , on a  $f(0, t) = 0$ . On va montrer qu'en fait cette image est toute la sphère. Par homogénéité de  $f(v, t)$  par rapport à  $v$ , il suffit de le montrer pour une valeur de  $r$ . Soit  $r = \frac{1}{\sqrt{1+t^2}}$  : on pose  $w = f(v, t)$ , de sorte que cela revient à chercher les solutions de  $w = v + tZ(v) \Leftrightarrow v = w - tZ(v)$  pour  $w$  de norme 1.

On utilise le théorème du point fixe strictement contractant : soit

$$g(v, t) = w - tZ(v)$$

La fonction  $g(v, t)$  (en considérant  $t$  comme une constante) envoie l'espace euclidien  $\mathbb{R}^{n+1}$ , qui est complet, sur lui-même. Elle est Lipschitzienne de rapport  $t(L + 1)$ . Par conséquent, si  $t < \frac{1}{L+1}$ ,  $g(v, t)$  est une contraction stricte de  $\mathbb{R}^{n+1}$  dans lui-même. Elle possède donc un unique point fixe, satisfaisant l'équation  $g(v', t) = v$ , donc  $v'$  est solution de l'équation  $f(v', t) = w$ . En particulier, d'après le calcul de la norme ci-dessus, on a :

$$|w|^2 = |f(v', t)|^2 = |v'|^2 (1 + |t|^2)$$



donc l'image de la sphère de rayon  $\frac{1}{1+t^2}$  par  $f(v, t)$  (toujours avec  $t$  constant) est la sphère de rayon 1. Par homogénéité, on déduit que  $f(v, t)$  envoie la sphère de rayon  $r$  sur la sphère de rayon  $r\sqrt{1+t^2}$ .

On calcule maintenant de 2 manières différentes le volume de la couronne solide de centre 0 et de rayon 1 et 2. D'une part, cette image est la couronne solide de rayon  $\sqrt{1+t^2}$  et  $2\sqrt{1+t^2}$ , donc son volume est donc

$$V = \omega_{n+1}(2^{n+1} - 1)(1 + t^2)^{\frac{n+1}{2}}$$

avec  $\omega_{n+1}$  le volume de la boule unité en dimension  $n + 1$ .

D'autre part,

$$Jac_v f(v, t) = I + t \cdot JacZ(v)$$

$I$  est l'application linéaire identité dans  $\mathbb{R}^{n+1}$ . Par la formule du changement de variable dans les intégrales multiples, le volume de la couronne solide est donc donné par :

$$\int_{1 \leq |v| \leq 2} det(I + t \cdot JacZ(v)) dv$$

Le déterminant est bien défini dans la formule ci-dessus car  $JacZ(v)$  est une application linéaire de  $\mathbb{R}^{n+1}$  dans lui-même, paramétrée par  $v$ . Le déterminant ne change pas de signe car  $f(v, t)$  est un difféomorphisme (en la variable  $v$ ), car  $Jac_v f(v, t) \neq 0 \forall v$ . Par continuité, il vaut 1 pour  $t = 0$ , donc il est positif. Pas besoin de valeur absolue.

Si  $n$  est pair, le volume est une expression irrationnelle en  $t$  d'après la première formule, tandis que la 2ème formule nous donne une expression polynomiale en  $t$   $det(I + t \cdot JacZ(v))$  est un polynôme en  $t$ . On parvient alors à une contradiction, ce qui montre que le champ de vecteur  $X$  doit s'annuler en un point.

2/ Cas d'un champ de vecteur continu sur la sphère  $S^n$ .

On peut approximer ce champ uniformément par une suite de champs continûment différentiables (voir lemme ci-dessous)  $(X_k)_{k \in \mathbb{N}}$ .

Mais, à priori, il n'y a pas de raison que les  $X_k$  soient tangents. On doit donc les projeter. On définit ainsi :

$$X_k(v) = \widehat{X}_k(v) - (\widehat{X}_k(v) | v)v$$

(utilisation de Gram-Schmidt). Le champ  $X_k$  est tangent à la sphère et il converge uniformément vers le champ  $v \mapsto X(v) - (X(v) | v)v$ , ce qui permet de conclure.

D'après ce qu'on a montré ci-dessous,  $\forall k, \exists v_k$  tel que  $X_k(v_k)$  s'annule. Les points  $v_k$  appartenant à l'ensemble compact  $S_n$ , on peut trouver une sous-suite convergente de la suite des  $v_k$ , ou, ce qui revient au même, une partie infinie  $K$  de  $\mathbb{N}$  telle que la suite  $(v_k)_{k \in K}$  possède une limite le long de  $K$  : notons  $\bar{v}$  cette limite. Il est évident que le long de  $K$   $X_k(v_k)$  converge vers  $X(\bar{v})$ , donc,  $X$  s'annule au point  $\bar{v}$ .  $\square$

**5.4.10 Lemme:** *Tout champ de vecteur continu est limite uniforme d'une suite de champs de vecteurs continûment différentiables.*

**Preuve:**

Soit  $\phi$  une fonction de  $\mathbb{R}^n$  dans  $\mathbb{R}^+$  ayant les propriétés suivantes :

- $\phi$  est continûment différentiable.

-Le support de  $\phi$ , c'est à dire l'adhérence de l'ensemble de tous les points où  $\phi$  n'est pas nul est inclus dans la boule de centre 0 et de rayon 1.

-L'intégrale de  $\phi$  sur  $\mathbb{R}^n$  vaut 1.

Un tel  $\phi$  est relativement simple à construire : soit une fonction  $\psi$  continûment dérivable, positive ou

nulle sur  $\mathbb{R}^+$ , valant 1 en 0 et 0 au delà de 1. Si on suppose que la dérivée de cette fonction est nulle en 0, alors on prend :  $\phi(v) = \frac{\psi(|v|)}{I}$  avec  $I = \int_{\mathbb{R}^n} \psi(|v|)dv$ .  $\psi = \exp^{-\frac{x^2}{1-x^2}}$  si  $x < 1$  et 0 si  $x > 1$  convient. On vérifie aisément que  $\phi$  a les propriétés requises.

Soit, pour  $k$  un entier positif,

$$\phi_k(v) = \phi(kv)$$

Alors, comme  $\psi(kv)$  est définie pour  $|kv| \leq 1 \Leftrightarrow |v| \leq \frac{1}{k}$ , il en est de même de  $\phi_k$ . On note  $\overline{B}_{\frac{1}{k}}(0) = \{v \in \mathbb{R}^n, |v| \leq \frac{1}{k}\}$ .

Pour tout champ de vecteur  $X : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , on pose

$$\widehat{X}_k(v) = \int_{\mathbb{R}^n} \phi_k(v-w)X(w)dw$$

On veut montrer que  $\widehat{X}_k$  est un champ de vecteurs continûment différentiable quel que soit  $k \in \mathbb{N}$  et que  $\lim_{k \rightarrow \infty} \widehat{X}_k = X$ .

Par opérations, les  $\widehat{X}_k$  sont continûments différentiables :

$$d\widehat{X}_k(v) = \int_{\mathbb{R}^n} d\phi_k(v-w)X(w)dw = \int_{\overline{B}_{\frac{1}{k}}(0)} d\phi_k(v-w)X(w)dw$$

$d\phi$  étant continue, de même que  $X, d\widehat{X}_k$  est continue car on intègre sur un compact.

Montrons maintenant la convergence uniforme vers  $X$  :

$$\phi_k(v-w) = \phi(kv-w) = \frac{1}{I} \psi(|kv-w|)$$

Donc, si

$$k|v-w| \geq 1 \Leftrightarrow |v-w| \geq \frac{1}{k}$$

alors  $\phi_k(v-w) = 0$ . Ainsi, quand  $k \rightarrow \infty$ ,

$$\begin{aligned} \widehat{X}_k &= \int_{(v-w) \in \overline{B}_{\frac{1}{k}}(0)} \phi_k(v-w)X(w)dw = \int_{\overline{B}_{\frac{1}{k}}(0)} \phi_k(-u)X(u+v)du \\ &= \int_{\overline{B}_{\frac{1}{k}}(0)} \frac{\psi(|ku|)}{I} X(u+v)du \end{aligned}$$

par le changement de variable  $u = w - v$ . On a tout d'abord,

$$\begin{aligned} I &= \int_{\mathbb{R}^n} \psi(|kv|)dv = \int_{\overline{B}_{\frac{1}{k}}(0)} \psi(|kv|)dv \\ &= \frac{1}{k^n} \int_{\overline{B}_1(0)} \psi(|w|)dw \end{aligned}$$

par le changement de variable  $w = kv$ , ce qui entraîne  $Jac \ w = kId$ . Puis, par le même changement de variable  $Jac \ w = kId$ ,

$$\int_{\overline{B}_{\frac{1}{k}}(0)} \frac{\psi(|ku|)}{I} X(u+v)du = \int_{\overline{B}_1(0)} \frac{1}{k^n} \frac{\psi(|w|)}{I} X\left(\frac{w}{k} + v\right)dw$$

Or, il est facile de vérifier que cette dernière expression vaut

$$A_k(v) = \int_{\overline{B}_1(0)} \frac{\psi(|w|)}{\int_{\overline{B}_1(0)} \psi(|u|)du} X\left(\frac{w}{k} + v\right)dw$$

On pose, pour simplifier l'écriture

$$g : \mathbb{R}^n \rightarrow \mathbb{R}, g(w) = \frac{\psi(|w|)}{\int_{\overline{B_1(0)}} \psi(|u|) du}$$

Or, la convergence uniforme de

$$Z_k(v, w) = \frac{\psi(|w|)}{\int_{\overline{B_1(0)}} \psi(|u|) du} X\left(\frac{w}{k} + v\right)$$

vers  $g(w)X(v)$ , implique que l'on puisse passer la limite dans l'intégrale,

$$\lim_{k \rightarrow +\infty} A_k(v) = \int_{\overline{B_1(0)}} \lim_{k \rightarrow +\infty} \frac{\psi(|w|)}{\int_{\overline{B_1(0)}} \psi(|u|) du} X\left(\frac{w}{k} + v\right) dw = \int_{\overline{B_1(0)}} g(w)X(v) = X(v)$$

On a donc la convergence simple, la convergence uniforme se montre également.

Pour une idée de preuve rigoureuse : on peut utiliser le fait que  $X$  est continu. Montrons que  $Z_k(v, w)$  converge uniformément vers  $g(w)X(v)$  pour tout  $w \in \overline{B_1(0)}$ .

$$|g(w)X\left(\frac{w}{k} + v\right) - g(w)X(v)| \leq |g(w)| |X\left(\frac{w}{k} + v\right) - X(v)|$$

Comme  $w \in \overline{B_1(0)}$ ,  $|g(w)|$  est majoré par un certain  $\alpha \in \mathbb{R}$ . On pose  $\beta_k(v) = \sup_{w \in \overline{B_1(0)}} |X\left(\frac{w}{k} + v\right) - X(v)|$ . De sorte que  $|g(w)X\left(\frac{w}{k} + v\right) - g(w)X(v)| \leq \alpha\beta_k(v)$ . Donc, quand  $k \rightarrow \infty$ ,

$$\lim_{k \rightarrow \infty} \sup |Z(v, w) - Z_k(v, w)| = \lim_{k \rightarrow \infty} \alpha\beta_k(v) = 0$$

, ce qui montre la convergence uniforme de  $Z_k(v, w)$  vers  $g(w)X(v)$ . De plus, cela est vrai  $\forall v \in \mathbb{R}^n$ , par continuité de  $X(v)$ . La convergence simple de  $A_k$  est donc établie.

Montrons maintenant la convergence uniforme de la suite  $A_k(v)$  :

$$\begin{aligned} |A_k(v) - X(v)| &= \left| A_k - \int_{\overline{B_1(0)}} \frac{\psi(|w|)}{\int_{\overline{B_1(0)}} \psi(|u|) du} X(v) dw \right| = \left| \int_{\overline{B_1(0)}} \frac{\psi(|w|)}{\int_{\overline{B_1(0)}} \psi(|u|) du} (X\left(\frac{w}{k} + v\right) - X(v)) dw \right| \\ &\leq \int_{\overline{B_1(0)}} |g(w)(X\left(\frac{w}{k} + v\right) - X(v))| dw \leq \int_{\overline{B_1(0)}} \alpha\beta_k(v) = \alpha\beta_k(v) \end{aligned}$$

Donc,  $\lim_{k \rightarrow \infty} |A_k(v) - X(v)| = 0$ , de plus cela est vrai  $\forall v \in \mathbb{R}^n$ , par continuité de  $X(v)$ . On a bien la convergence uniforme de  $A_k$ .

□

# PARTIE III : Groupes de Lie et algèbre de Lie

Le but de cette partie est de donner une introduction à la théorie des groupes de Lie.

## 6 Groupes et Algèbre de Lie

Le but de ce chapitre est de définir l'algèbre de Lie d'un groupe de Lie, après avoir défini au préalable ce qu'est un groupe de Lie et ce qu'est une algèbre de Lie. Les outils utilisés dans la partie précédente nous permettront de faire ce lien.

### 6.1 Groupes de Lie et premiers exemples

**6.1.1 Définition:** Un *groupe de Lie* est un groupe dont la loi de groupe et l'inverse sont des applications différentiables. C'est aussi une variété.

**6.1.2 Proposition:**  $Heis_3$  (groupe de Heisenberg) est une variété, et donc un groupe de Lie.

**Preuve:**

$\mathbb{R}^2$  est une variété et  $\mathbb{R}/\mathbb{Z} = ]0; 1[$  est également une variété : en effet, on considère les fonctions

$$\begin{aligned} \varphi_1 : ]0, 1[ &\longrightarrow ]0, 1[, \quad x \longmapsto x \\ \varphi_2 &\left] \frac{1}{2}, \frac{3}{2} \right[ \longrightarrow ]0, 1[, \quad x \longmapsto x - \frac{1}{2}. \end{aligned}$$

On vérifie que ce sont des homéomorphismes, de même que les composées sur leur domaine de définition. Ces 2 cartes forment un recouvrement. Donc, comme  $H_3(\mathbb{R}) = \mathbb{R}^2 \times \mathbb{R}/\mathbb{Z}$ , c'est une variété par produit cartésien de 2 variétés.  $\square$

De plus, d'après tout ce que l'on a déjà fait :

- Le cercle  $\mathbb{S}^1$  et la sphère  $\mathbb{S}^3$  sont des groupes de Lie.
- La sphère  $\mathbb{S}^2$  n'est pas un groupe de Lie (ce n'est pas un groupe).
- Le groupe de Heisenberg est un groupe de Lie (mais pas un groupe matriciel).
- Les groupes  $GL_n(\mathbb{R})$ ,  $SL_n(\mathbb{R})$  et  $O_n(\mathbb{R})$  sont des groupes de Lie.
- Le tore  $T^2$  en tant que produit cartésien  $\mathbb{S}^1 \times \mathbb{S}^1$  muni de la loi de groupe  $(x, y) \cdot (x', y') = (xx', yy')$  est un groupe de Lie.

**Dimension topologique d'un groupe.** Un groupe de Lie étant une variété, on peut parler de sa dimension en tant que variété. Par exemple,  $\dim U(1) = 1$  et  $\dim SU(2) = \dim SO(3) = 3$  (en raison des homéomorphismes entre  $\mathbb{S}^1$  et  $U(1)$ , puis  $\mathbb{S}^3$  et  $SU(2)$ ; pour  $SO(3)$ , on le verra un peu plus tard).

**6.1.3 Définition:** La *dimension* d'un de Lie réel  $G$  est  $\dim G = \dim T_e G$ ,  $e$  étant le neutre du groupe, et  $T_e G$  l'espace tangent en l'identité.

### 6.2 Algèbre de Lie

**6.2.1 Définition:** Une  $\mathbb{K}$ -algèbre de Lie sur  $\mathbb{K}$  est un espace vectoriel  $E$  sur un corps  $\mathbb{K}$ , qui est  $\mathbb{K}$  bilinéaire :  $E \times E \longrightarrow E$ , est appelée le crochet de Lie, tel que :

$$[x, y] = -[y, x] \text{ (antisymétrie)}$$

$$[x, [y, z]] + [y, [x, z]] + [z, [x, y]] = 0 \text{ (Identité de Jacobi)}$$

$\mathbb{K}$  bilinéaire signifie :  $\forall x_1, x_2, x, y_1, y_2, y \in E$  et  $r_1, r_2, r, s_1, s_2, s \in \mathbb{K}$  :

$$[r_1 x_1 + r_2 x_2, y] = r_1 [x_1, y] + r_2 [x_2, y]$$

$$[x, s_1 y_1 + s_2 y_2] = s_1 [x, y_1] + s_2 [x, y_2]$$

**6.2.2 Proposition:** *Le produit vectoriel réel sur  $\mathbb{R}^3 \wedge$  avec  $x = (x_1, x_2, x_3)$  et  $y = (y_1, y_2, y_3)$  tel que :  $x \wedge y = (x_2y_3 - y_2x_3, x_3y_1 - y_3x_1, x_1y_2 - x_2y_1)$ . est une algèbre de Lie.*

**Preuve:**  $x \wedge y = -y \wedge x$ . On vérifie la bilinéarité très simplement. Enfin, l'identité de Jacobi. Rappel : formule de Gibbs :  $(x \wedge y) \wedge z = (x \cdot z)y - (y \cdot z)x$  avec  $\cdot$ . (je ne la démontrerai pas, elle se vérifie de plusieurs manières, par un calcul direct par exemple).

$$(x \wedge y) \wedge z + (y \wedge z) \wedge x + (z \wedge x) \wedge y = (x \cdot z)y - (y \cdot z)x + (x \cdot y)z - (x \cdot z)y + (z \cdot y)x - (x \cdot y)z = 0. \quad \square$$

**6.2.3 Proposition:** *Les champs de vecteurs  $\mathfrak{X}(M)$  sur une variété forment une algèbre de Lie, avec crochet*

$$[X, Y]f = X(Yf) - Y(Xf), \quad f \in C^\infty(M),$$

qui vaut, en coordonnées locales,

$$[X, Y] \Big|_U = \sum_{i,j=1}^m \left( X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}.$$

**Preuve:** Le crochet de Lie sur  $\mathfrak{X}(M)$  est induit par le crochet de Lie naturel sur les dérivations  $\text{Der}(C^\infty(M), C^\infty(M))$ , donné par  $[D_1, D_2] = D_1 \circ D_2 - D_2 \circ D_1$ . Montrons que  $[D_1, D_2]$  est encore une dérivation :

$$\begin{aligned} (D_1 D_2 - D_2 D_1)(fg) &= D_1 D_2(fg) - D_2 D_1(fg) = D_1(D_2(f)g) + D_1(fD_2(g)) - D_2(D_1(f)g) - D_2(fD_1(g)) \\ &= D_1 D_2(f)g + D_2(f)D_1(g) + D_1(f)D_2(g) + fD_1 D_2(g) - D_2 D_1(f)g - D_1(f)D_2(g) - D_2(f)D_1(g) - fD_2 D_1(g) \\ &= [D_1, D_2](f)g + f[D_1, D_2](g) \end{aligned}$$

Puis, on a également :

$$[D_1, D_2](\alpha f + \beta g) = [D_1, D_2]\alpha f + [D_1, D_2]\beta g$$

donc  $[D_1, D_2]$  est bilinéaire. Par conséquent,  $[D_1, D_2] \in \text{Der}(C^\infty(M))$ . Reste à prouver que  $[,]$  est un crochet de Lie, cela découle du fait que  $\text{Der} \subset (\text{End}, \circ)$  et  $\circ$  est associative.  $\square$

Remarque : ce résultat nous sera très utile, pour introduire le flot.

**6.2.4 Proposition:** *Le crochet de Lie pour les matrices, et plus généralement pour tout espace vectoriel, est une algèbre de Lie ( $[x, y] = xy - yx$ ).*

**Preuve:** La bilinéarité et l'antisymétrie sont évidentes. Identité de Jacobi :

$$[[x, y], z] + [[y, z], x] + [[z, x], y] = (xy - yx)z - z(xy - yx) + (yz - zy)x - x(yz - zy) + (zx - xz)y - y(zx - xz) = xyz - yxz - zxy + zyx + yzx - zyx - xyz + xzy + zxy - xzy - yzx + yxz = 0. \quad \square$$

**6.2.5 Définition:** Si  $E$  est une  $\mathbb{K}$ -algèbre et un  $\mathbb{K}$ -sev  $V \subseteq A$  est une  $\mathbb{K}$  sous algèbre de Lie de  $E$  si :  $\forall x, y \in V, [x, y] \in V$ . On écrit  $V \leq E$

**6.2.6 Proposition:** *Les matrices antisymétriques de taille  $n$  sont une sous-algèbre de Lie de  $GL_n(\mathbb{K})$ .*

**Preuve:** Soit  $A, B \in \text{ASym}_n(\mathbb{K})$   $[A, B] = AB - BA$ .  ${}^t[A, B] = {}^t(AB - BA) = {}^t(AB) - {}^t(BA) = {}^t B {}^t A - {}^t A {}^t B = BA - AB = -[A, B]$ . Donc  $[A, B]$  est antisymétrique.  $\square$

**6.2.7 Proposition:** *Soit  $\Phi : E \rightarrow V$  un homomorphisme surjectif de  $\mathbb{K}$ -algèbre de Lie.*

*$V$  est un abélien si et seulement si  $E' \subseteq \ker \Phi$  (abélien : algèbre de Lie dont tous les crochets sont triviaux  $\forall x, y \in E : [x, y] = 0$ )*

**Preuve:**

(i) Si  $V$  est abélien : Supposons qu'il existe par l'absurde l'existence d'éléments  $u, v \in E$  tels que  $\Phi([u, v]) \neq 0 \Leftrightarrow [\Phi(u), \Phi(v)] \neq 0$  et donc,  $V$  ne serait pas abélien car  $\Phi(u), \Phi(v) \in V$ . On parvient donc à une contradiction.

(ii) Si  $E' \subseteq \ker \Phi$  : toujours par l'absurde, supposons qu'il existe  $x, y \in V$  tels que  $xy \neq yx \forall x, y \in V$ . Par surjectivité,  $\exists x', y' \in E$  tels que  $\Phi(x')\Phi(y') \neq \Phi(y')\Phi(x') \Leftrightarrow [\Phi(x'), \Phi(y')] = \Phi([x', y']) \neq 0 \Leftrightarrow [x', y'] \notin \ker \Phi$  et cela contredit le fait que  $E' \subseteq \ker \Phi$ .  $\square$

**6.2.8 Proposition:** Soit  $\Phi : E \rightarrow V$  un homomorphisme de  $\mathbb{K}$ -algèbre de Lie.  $\ker \Phi \triangleleft E$  et il existe un isomorphisme d'algèbre de Lie :  $\text{im} \Phi \rightarrow E/\ker \Phi$ . (attention : rappel  $\triangleleft$  signifie que  $\ker \Phi$  est normal)

**Preuve:** Il suffit d'appliquer le premier théorème d'isomorphisme dans la partie quotient de groupes, en considérant  $E$  et  $V$  comme des groupes additifs.  $\square$

### 6.3 Algèbre de Lie d'un groupe de Lie

On veut définir une structure d'algèbre de Lie sur  $\mathfrak{g}T_e G$ , on le fait en utilisant la représentation adjointe de  $G$  sur  $\mathfrak{g}$  qui est très utile car elle permet d'étudier les autres représentations de  $G$

La représentation adjointe est la principale représentation qu'on utilise pour ensuite trouver d'autres représentations.

Dans toute cette partie, soit  $G$  un groupe de Lie de dimension  $n$  (comme variété différentielle). On note  $\mathfrak{g} = T_e G$  l'espace tangent en l'identité de  $G$ , qui est bien défini car  $G$  est une variété différentielle et est isomorphe à  $\mathbb{R}^n$ .

**6.3.1 Definition:** On appelle *action par conjugaison* de  $G$  sur lui-même l'application définie, pour tout  $x \in G$ , par

$$\begin{aligned} \text{aut}_x : G &\longrightarrow G \\ y &\mapsto \text{aut}_x(y) := xyx^{-1}. \end{aligned}$$

Il s'agit bien d'une action, car pour tout  $x, y, z \in G$  on a

$$\text{aut}_{xy}(z) = \text{aut}_x(\text{aut}_y(z)).$$

De plus, pour tout  $x$  fixé, l'application  $\text{aut}_x$  est un morphisme de groupes, car

$$\text{aut}_x(yz) = xyzx^{-1} = xyx^{-1}zx^{-1} = \text{aut}_x(y)\text{aut}_x(z).$$

**6.3.2 Definition:** L'*action adjointe* de  $G$  sur  $\mathfrak{g}$  est définie, pour tout  $x \in G$ , par l'application

$$\text{Ad}_x := d_e \text{aut}_x : \mathfrak{g} \longrightarrow \mathfrak{g}$$

donnée par la différentielle en l'élément neutre  $e$  de la conjugaison  $\text{aut}_x$ .

Plus précisément, pour tout vecteur  $v$  tangent à  $G$  en  $e$ , i.e.  $v \in \mathfrak{g}$ , soit  $\gamma : \mathbb{R} \rightarrow G$  une courbe différentiable telle que  $\gamma(0) = e$  et  $\gamma'(0) = v$ . On a alors

$$\text{Ad}_x(v) := d_e \text{aut}_x(v) = \left. \frac{d}{dt} \right|_{t=0} (x\gamma(t)x^{-1}).$$

À noter qu'on ne peut pas écrire  $\text{Ad}_x(v) = x\gamma'(0)x^{-1} = xv x^{-1}$ , car en général on ne peut pas multiplier des éléments  $x \in G$  par des vecteurs  $v \in \mathfrak{g}$ . L'application  $\text{Ad}_x$  est justement ce qui remplace la conjugaison comme action de  $G$  sur  $\mathfrak{g}$ .

**6.3.3 Proposition:** Si  $G \subset GL_n(\mathbb{K})$  est un groupe de matrices, on a que  $\mathfrak{g} \subset M_n(\mathbb{K})$  est un sous-espace vectoriel de matrices, donc l'action adjointe est exactement la conjugaison définie par le produit de matrices,

$$Ad_A(V) = AVA^{-1}, \quad A \in G, \quad V \in \mathfrak{g}.$$

**Preuve:** Dans ce cas, la loi de groupe est le produit de matrices, et la dérivée par rapport à  $t$  de matrices qui dépendent de  $t$  satisfait la règle de Leibniz. Par conséquent, on a

$$\begin{aligned} Ad_A(V) &= \frac{d}{dt}(A\gamma(t)A^{-1})\Big|_{t=0} \\ &= \frac{d}{dt}(A)\Big|_{t=0}\gamma(t)A^{-1} + A\frac{d}{dt}(\gamma(t))\Big|_{t=0}A^{-1} + A\gamma(t)\frac{d}{dt}(A^{-1})\Big|_{t=0} \\ &= A\gamma'(0)A^{-1} = AVA^{-1}. \end{aligned}$$

□

**6.3.4 Proposition:** L'action adjointe est une représentation linéaire de  $G$  sur  $\mathfrak{g}$ , c'est-à-dire qu'elle vérifie les identités suivantes :

$$\begin{aligned} Ad_{xy}(v) &= Ad_x(Ad_y(v)) \\ Ad_x(tv + su) &= tAd_x(v) + sAd_x(u), \end{aligned}$$

pour tout  $x, y \in G$ ,  $u, v \in \mathfrak{g}$  et  $t, s \in \mathbb{R}$ . Par conséquent, elle détermine un morphisme de groupes

$$\overline{Ad} : G \longrightarrow GL(\mathfrak{g}), \quad x \mapsto Ad_x,$$

où  $GL(\mathfrak{g}) \cong GL_n(\mathbb{R})$  est vu comme sous-groupe de  $End(\mathfrak{g}) \cong M_n(\mathbb{R})$ .

**Preuve:** Pour tout  $x \in G$  fixé, l'application  $Ad_x$  est clairement linéaire car elle est définie comme une différentielle. Montrons que  $x \mapsto Ad_x$  est bien une action.

Considérons  $x, y \in G$ . Pour tout  $v \in \mathfrak{g}$ , soit  $\gamma : \mathbb{R} \longrightarrow G$  une courbe différentiable telle que  $\gamma(0) = e$  et  $\gamma'(0) = v$ . On a alors

$$\begin{aligned} Ad_{xy}(v) &= \frac{d}{dt}((xy)\gamma(t)(xy)^{-1})\Big|_{t=0} \\ &= x\frac{d}{dt}(y\gamma(t)y^{-1})\Big|_{t=0}x^{-1} \\ &= Ad_x(Ad_y(v)). \end{aligned}$$

□

**6.3.5 Definition:** On remarque que  $\overline{Ad}(e) = I \in GL(\mathfrak{g})$  est l'identité (élément neutre) du groupe  $GL(\mathfrak{g})$ , et qu'on peut identifier l'espace tangent à  $GL(\mathfrak{g})$  en  $I$  avec l'espace vectoriel  $End(\mathfrak{g})$ . La différentielle en  $e \in G$  de l'application  $\overline{Ad} : G \longrightarrow GL(\mathfrak{g})$  est donc une application linéaire

$$ad := d_e \overline{Ad} : \mathfrak{g} \longrightarrow End(\mathfrak{g})$$

qui s'appelle *action adjointe* de  $\mathfrak{g}$  sur  $\mathfrak{g}$ . Plus précisément, pour tout  $u \in \mathfrak{g}$ , soit  $\gamma : \mathbb{R} \longrightarrow G$  une courbe différentiable telle que  $\gamma(0) = e$  et  $\gamma'(0) = u$ . Pour tout  $v \in \mathfrak{g}$  on a alors :

$$ad_u(v) := (d_e \overline{Ad}(u))(v) = \frac{d}{dt}Ad_{\gamma(t)}(v)\Big|_{t=0}.$$

**6.3.6 Proposition:** L'opération  $[\ , \ ] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$  définie par

$$[u, v] := ad_u(v), \quad u, v \in \mathfrak{g}$$

est un crochet de Lie sur  $\mathfrak{g}$ .

**Preuve:** antisymétrie + Jacobi □

**6.3.7 Définition:** L'algèbre de Lie du groupe de Lie  $G$  est l'espace vectoriel  $\mathfrak{g}$  muni du crochet de Lie  $[u, v] = ad_u(v)$ .

**6.3.8 Proposition:** Si  $G \subset GL_n(\mathbb{K})$  est un groupe de matrices, le crochet de Lie sur  $\mathfrak{g} \subset M_n(\mathbb{K})$  est donné par le commutateur du produit de matrices :

$$[U, V] = UV - VU.$$

**Preuve:** Calculons  $[U, V] = ad_U(V)$  dans ce cas, en sachant que pour tout  $A \in G$  et tout  $V \in \mathfrak{g}$  on a  $Ad_A(V) = AVA^{-1}$ .

Soit  $\gamma : \mathbb{R} \rightarrow G$  une courbe telle que  $\gamma(0) = I$  et  $\gamma'(0) = U \in \mathfrak{g}$ . Alors on a

$$\begin{aligned} [U, V] &= ad_U(V) = \left. \frac{d}{dt} Ad_{\gamma(t)}(V) \right|_{t=0} = \left. \frac{d}{dt} (\gamma(t)V\gamma(t)^{-1}) \right|_{t=0} \\ &= \gamma'(0)V\gamma(0)^{-1} + \gamma(0)V \left. \frac{d}{dt} \gamma(t)^{-1} \right|_{t=0}. \end{aligned}$$

Pour calculer  $\left. \frac{d}{dt} \gamma(t)^{-1} \right|_{t=0}$  on dérive  $\gamma(t)\gamma(t)^{-1} = I$  : on a

$$\gamma'(t)\gamma(t)^{-1} + \gamma(t)(\gamma(t)^{-1})' = 0,$$

donc

$$\gamma(t)(\gamma(t)^{-1})' = -\gamma'(t)\gamma(t)^{-1},$$

d'où suit

$$(\gamma(t)^{-1})' = -\gamma(t)^{-1}\gamma'(t)\gamma(t)^{-1}.$$

Finalement, en sachant que  $\gamma(0) = I$  et  $\gamma'(0) = U$ , on obtient

$$[U, V] = UVI^{-1} - IVI^{-1}UI^{-1} = UV - VU.$$

□

**6.3.9 Remarque:** La formule explicite  $[u, v] = uv - vu$  du crochet de Lie n'est valable que pour les groupes de matrices, car pour un groupe de Lie général l'algèbre de Lie  $\mathfrak{g} = T_eG$  est un espace vectoriel sur lequel il n'y a aucun produit disponible pour pouvoir multiplier deux vecteurs tangents  $u$  et  $v$ .

Cependant, l'espace vectoriel  $\mathfrak{g} = T_eG$  peut être identifié à un sous-espace de l'ensemble des champs de vecteurs sur  $G$ , sur lequel on a déjà vu qu'il existe un crochet de Lie obtenu comme commutateur de l'opération de composition d'applications linéaires sur les fonctions sur  $G$ .

On veut donc maintenant montrer comment identifier  $\mathfrak{g} = T_eG$  à certains champs de vecteurs sur  $G$ .

## 6.4 Champs de vecteurs invariants à gauche

**6.4.1 Définition:** Pour tout  $x \in G$ , on appelle *translation par  $x$  à gauche et à droite* les applications  $L_x, R_x : G \rightarrow G$  définies par

$$L_x(y) = xy \quad \text{et} \quad R_x(y) = yx$$

pour tout  $y \in G$ .

**6.4.2 Proposition:** Les translations à gauche et à droite sont des applications  $C^\infty$  avec les propriétés suivantes :

—  $L$  et  $R$  sont des actions (respectivement à gauche et à droite) de  $G$  sur  $G$ , c'est-à-dire :

$$L_{xy} = L_x \circ L_y \quad \text{et} \quad R_{xy} = R_y \circ R_x$$

pour tout  $x, y \in G$ .



— Pour tout  $x \in G$ ,  $L_x$  et  $R_x$  sont donc des applications inversibles sur  $G$ , d'inverses

$$(L_x)^{-1} = L_{x^{-1}} \quad \text{et} \quad (R_x)^{-1} = R_{x^{-1}}.$$

—  $L$  et  $R$  commutent, dans le sens que

$$L_x \circ R_y = R_y \circ L_x$$

pour tout  $x, y \in G$ .

**Preuve:**

□

**6.4.3 Corollaire:** L'action par conjugaison de  $G$  sur  $G$  s'écrit alors en termes des translations comme :

$$\text{aut}_x = L_x \circ (R_x)^{-1}.$$

Par conséquent, l'action adjointe de  $G$  sur  $\mathfrak{g}$  s'écrit en termes des différentielles de  $L$  et  $R$  :

$$\text{Ad}_x = d_{x^{-1}}L_x \circ d_e(R_x)^{-1} : T_eG \longrightarrow T_{x^{-1}}G \longrightarrow T_eG.$$

**6.4.4 Définition:** Soit  $X$  un champ de vecteurs au moins  $C^1$  sur  $G$ . En général, si  $x$  et  $y$  sont deux points distincts de  $G$ , les vecteurs  $X_x \in T_xG$  et  $X_y \in T_yG$  ne coïncident pas. Le mieux qu'on puisse dire est que le passage de  $X_x$  à  $X_y$  est de classe  $C^1$ , par hypothèse sur  $X$ .

On dit que  $X$  est *invariant à gauche* si, pour tout  $x, y \in G$ , on a :

$$X_{L_x(y)} = d_yL_x(X_y),$$

où  $L_x : G \longrightarrow G$  amène  $y$  en  $xy$  donc sa différentielle en  $y$  est une application

$$d_yL_x : T_yG \longrightarrow T_{xy}G = T_{L_x(y)}G.$$

Cela signifie qu'on peut trouver la valeur de  $X$  en tout point  $xy \in G$  à partir de la valeur de  $X$  en  $y$  par "translation", en prenant garde à maintenir les vecteurs bien tangents à  $G$  pendant cette translation. En fait, le vecteur  $X_{xy}$  est le même vecteur  $X_y$ , par rapport au repère donné sur l'espace tangent en  $xy$  au lieu que  $y$ .

On note  $\mathfrak{X}^L(G) \subset \mathfrak{X}(G)$  l'ensemble des champs de vecteurs sur  $G$  qui sont invariants à gauche.

**6.4.5 Exemple:** Si  $G$  est le groupe additif  $\mathbb{R}$ , un champ de vecteurs lisse est une section  $X$  de  $T\mathbb{R} \cong \mathbb{R} \times \mathbb{R}$  de la forme

$$X_x = \left( x, f(x) \frac{d}{dx} \Big|_x \right) \equiv f(x) \frac{d}{dx} \Big|_x, \quad x \in \mathbb{R},$$

où  $f$  est une fonction  $C^\infty$  sur  $\mathbb{R}$ . Évidemment, en général un tel champ n'est pas invariant à gauche, où  $L_x(y) = x + y$  dans ce cas, car

$$X_{x+y} = f(x+y) \frac{d}{dx} \Big|_{x+y},$$

alors que

$$d_yL_x(X_y) = f(y) \frac{d}{dx} \Big|_{x+y}.$$

L'ensemble des champs invariants à gauche sur  $\mathbb{R}$  est donc

$$\mathfrak{X}^L(G) = \left\{ x \mapsto X_x = f(x) \frac{d}{dx} \Big|_x, f(x+y) = f(x) \text{ pour tout } x, y \in \mathbb{R} \right\}.$$

**6.4.6 Proposition:** L'espace vectoriel  $\mathfrak{X}^L(G)$  des champs invariants à gauche est une sous-algèbre de Lie de  $\mathfrak{X}(G)$ .

De plus,  $\mathfrak{X}^L(G)$  est isomorphe à l'algèbre de Lie  $\mathfrak{g}$  de  $G$ , et l'isomorphisme est

$$\mathfrak{X}^L(G) \longrightarrow \mathfrak{g}, X \longmapsto X_e.$$

**Preuve:** 1)  $\mathfrak{X}^L(G)$  est une sous-algèbre de Lie de  $\mathfrak{X}(G)$  ssi  $\mathfrak{X}^L(G)$  est fermé pour le crochet de Lie des champs :

$$X, Y \in \mathfrak{X}^L(G) \implies [X, Y] \in \mathfrak{X}^L(G).$$

On le montre par le calcul.

2)  $\mathfrak{X}^L(G)$  est isomorphe à  $\mathfrak{g}$  comme espace vectoriel, car l'application

$$\mathfrak{X}^L(G) \longrightarrow \mathfrak{g}, X \longmapsto X_e$$

est un isomorphisme : c'est clair qu'elle est linéaire sur  $\mathbb{R}$ , et c'est inversible car tout  $X \in \mathfrak{X}^L(G)$  est déterminé par sa valeur en  $e$  : pour tout  $x \in G$  on a  $x = xe = L_x(e)$  donc

$$X_x = X_{L_x(e)} = d_e L_x(X_e).$$

3)  $\mathfrak{X}^L(G)$  est isomorphe à  $\mathfrak{g}$  comme algèbre de Lie, car l'isomorphisme  $X \mapsto X_e$  preserve les crochets de Lie, c'est-à-dire que

$$[X_e, Y_e]_{\mathfrak{g}} = ([X, Y]_{\mathfrak{X}(G)})_e.$$

On le montre par le calcul. □

**6.4.7 Corollaire:** Avec un abus de notation, le crochet de Lie dans  $\mathfrak{g}$  peut s'exprimer comme un commutateur, i.e.

$$[u, v] = u \circ v - v \circ u, \quad \text{pour tout } u, v \in \mathfrak{g},$$

où l'on a identifié les vecteurs tangents en  $e$  à des champs de vecteurs invariants à gauche, et on comprend que le symbol  $\circ$  est la projection sur  $\mathfrak{g}$  d'une opération qui n'est bien définie que sur  $\mathfrak{X}^L(G)$ .

**Preuve:** Notons

$$\mathfrak{g} \longrightarrow \mathfrak{X}^L(G), v \longmapsto X^v$$

la réciproque de l'isomorphisme  $X \mapsto X_e$ . Autrement dit,  $X^v$  est l'unique champ de vecteurs invariant à gauche sur  $G$  tel que  $(X^v)_e = v \in \mathfrak{g}$ . Alors on a

$$[u, v] = [X^u, X^v]_e = (X^u \circ X^v - X^v \circ X^u)_e,$$

et par abus de notation on indique ce résultat par

$$X_e^u \circ X_e^v - X_e^v \circ X_e^u = u \circ v - v \circ u.$$

□

Avant de terminer, un théorème très utile : (qui requiert une proposition)

**6.4.8 Proposition:** Pour un groupe de Lie  $G$ ,  $T_g(G)$  est isomorphe à  $T_e(G)$  de façon canonique.

**Preuve:** Soit  $a \in T_e(G) : \exists \gamma : ]-\epsilon, \epsilon[ \rightarrow G$  telle que  $\gamma'(0) = a$  et  $\gamma(0) = e$  Donc pour tout  $t$ , on définit :

$$\alpha(t) = g\gamma(t), \alpha(0) = g$$

de sorte que  $\alpha'(0) = ga \in T_g(G)$  Réciproquement : soit  $b \in T_g(G) : \exists \alpha \in G$  tel que  $\alpha'(0) = b$  et  $\alpha(0) = g$ . Donc pour tout  $t$ , on pose :

$$\gamma(t) = g^{-1}\alpha(t), \gamma(0) = e$$

de sorte que  $\gamma'(0) = g^{-1}b$ . □

**6.4.9 Théorème:** Le fibré tangent d'un groupe de Lie est toujours parallélisable.

**Preuve:** Les deux applications

$$L'_{g^{-1}} : TG \longrightarrow G \times T_e G, L'(g, v) = (g, (TL_g)^{-1}(v))$$

et

$$L'_g : G \times T_e G \longrightarrow TG, L'(g, v) = (g, (TL_g)(v))$$

sont des isomorphismes de fibrés vectoriels inverses l'un de l'autre. □

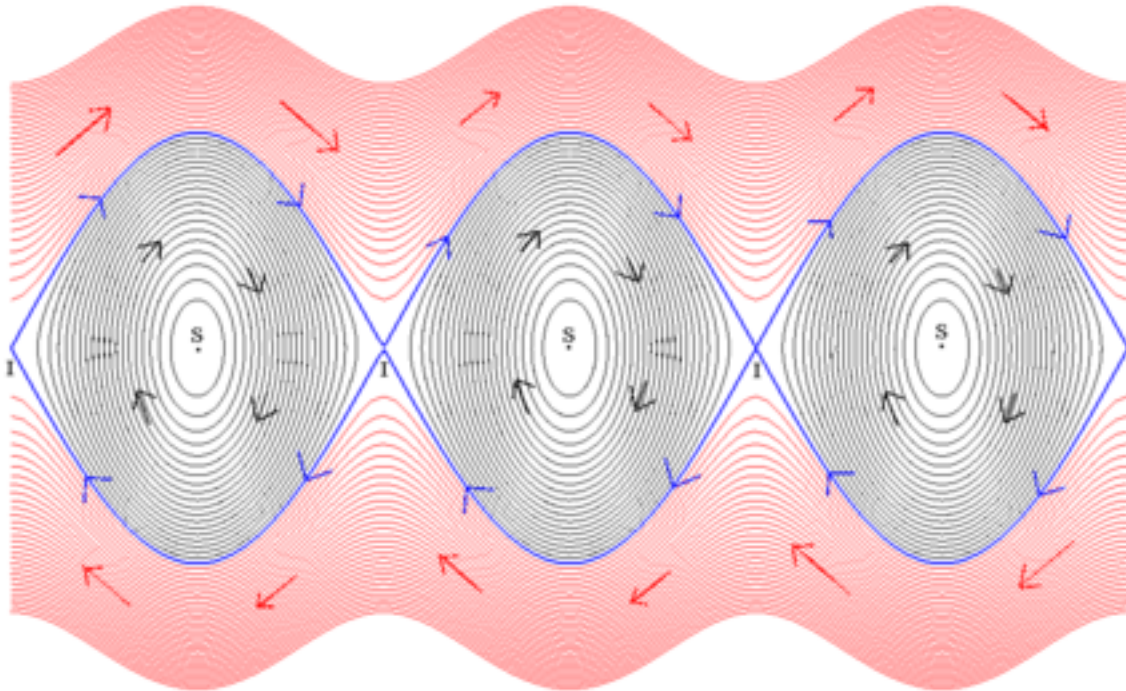
## 6.5 Flot

Dans cette section, on utilise des outils très analytiques. On a donc besoin d'un certain nombre de rappels car l'analyse n'est pas forcément la plus grande partie de ce TIPE. On verra par la suite à quoi nous servent ces outils.

**6.5.1 Définition:** Un champ de vecteur sur  $M$  est une application de classe  $C^k$  tel que  $v : M \rightarrow TM$  et  $v(x) \in T_x M$  où  $T_x M$  désigne l'espace tangent en  $M$ . Une courbe solution de  $v$  est une application différentiable  $\gamma : I \rightarrow M$  avec  $I$  un intervalle ouvert de  $\mathbb{R}$  tel que

$$\frac{d\gamma}{dt} = v(\gamma(t))$$

Courbes intégrales pour l'équation du pendule simple dans l'espace des phases  $(y, y')$



**6.5.2 Théorème:** Soit  $v$  un champ de vecteur sur  $M$  de classe  $C^1$ . Si  $\gamma_1 : I_1 \rightarrow M$  et  $\gamma_2 : I_2 \rightarrow M$  sont des courbes solution de  $v$  si  $t_0 \in I_1 \cap I_2$  et  $\gamma_1(t_0) = \gamma_2(t_0)$ , alors il existe un intervalle ouvert  $I_0$  autour de  $t_0$  dans  $I_1 \cap I_2$  tel que  $\gamma_1(t) = \gamma_2(t) \forall t \in I_0$ .

**6.5.3 Corollaire:** Si  $\gamma_1 : I_1 \rightarrow M$  et  $\gamma_2 : I_2 \rightarrow M$  sont des courbes solution de  $v \in C^1$  et  $\gamma_1(s) = \gamma_2(s)$  pour un certain  $s \in I_1 \cap I_2$ , alors  $\gamma_1(t) = \gamma_2(t) \forall t \in I_1 \cap I_2$ .

**Preuve:** Si  $\gamma_1 : I_1 \rightarrow M$  et  $\gamma_2 : I_2 \rightarrow M$  sont des courbes solution arbitraires de  $v$ , cela implique que  $J = \{t \in I_1 \cap I_2 \mid \gamma_1(t) = \gamma_2(t)\}$  est un sous-ensemble ouvert de  $I_1 \cap I_2$ . Mais  $J$  est également fermé de par la continuité de  $\gamma_1$  et  $\gamma_2$ . Cela implique que  $J = \emptyset$  ou alors  $J = I_1 \cap I_2$  car les seuls sous-ensembles à la fois ouverts et fermés d'un ensemble  $X$  sont  $X$  et  $\emptyset$ . Mais comme cet ensemble est non vide,  $J = I_1 \cap I_2$ . On remarque de plus que  $I_1 \cap I_2$  est connexe.

Pour démontrer le théorème, on va utiliser un certain nombre de lemmes et définir le flot. □

**6.5.4 Définition:** Pour  $s \in \mathbb{R}$  et  $x \in M$ , on note  $I(s, x, v)$  la réunion de tous les domaines de définition  $I$  des courbes solutions  $\gamma : I \rightarrow M$  vérifiant :  $s \in I, \gamma(s) = x$ .

Si  $t \in I(s, x, v)$ , il existe une courbe solution  $\gamma : I \rightarrow M$  et on note  $\Phi_v^{t,s}(x) = \gamma(t)$ . Cette définition ne dépend pas du choix de la courbe solution de  $v$  qui satisfait  $s, t \in I, \gamma(s) = x$ , en raison du corollaire 1. Donc  $\Phi_v^{s,s}(x) = \gamma(s) = x$  et la courbe  $\Phi_v^{t,s}(x)$  est la courbe solution maximale de valeur initiale  $x$  pour  $t = s$ . Soit  $M_v^{t,s} = \{x \in M \mid t \in I(s, t, x)\}$ . Alors :

$$\Phi_v^{s,t} : x \mapsto \Phi_v^{t,s}(x); M_v^{t,s} \rightarrow M$$

est le flot de  $v$  allant de  $s$  à  $t$ .

**6.5.5 Proposition:** Si  $x \in M_v^{s,u}(x)$ ,  $t \mapsto \Phi_v^{t,s}(\Phi_v^{s,u}(x)) = \Phi_v^{t,u}(x)$  avec  $\Phi_v^{s,u}(x) \in M_v^{t,s}$ .

**Preuve:** En effet,  $\Phi_v^{s,u}$  est la courbe solution  $\gamma(s)$  telle que  $\gamma(u) = x$  vérifiant  $\frac{d\gamma(s)}{ds} = v(\gamma(s))$  et  $\Phi_v^{t,s}$  est la courbe solution  $\gamma(t)$  vérifiant  $\frac{d\gamma(t)}{dt} = v(\gamma(t))$ .

De plus, en utilisant les conditions initiales,  $\Phi_v^{t,u}(x) = \gamma(t) = x$  et  $\Phi_v^{s,s}(\Phi_v^{s,u}(x)) = \gamma(s) = \Phi_v^{s,s}(x)$ , cette courbe vérifiant elle-même  $\gamma(u) = x$ . D'où la proposition.  $\square$

**6.5.6 Proposition:**  $\Phi_v^{t,s}(x) = \Phi_v^{t-\tau, s-\tau}(x)$  pour tout  $\tau \in \mathbb{R}$ .

**Preuve:** En effet  $\Phi_v^{t-\tau, s-\tau}(x) = \gamma(t-\tau)$ , et en  $s-\tau = t-\tau \Rightarrow t = s$ ,  $\gamma(t) = \gamma(s) = x$ .

Donc pour  $\tau = s$ , on peut noter

$$M_v^t := M_v^{t,0}, \Phi_v^t := \Phi_v^{t,0} : M_v^t \rightarrow M$$

On dit alors simplement que  $\Phi_v^t$  est le temps  $t$  du flot de  $v$ ; et on a alors :  $\Phi_v^{t,s} = \Phi_v^{t-s}$ .

Si on réutilise l'avant dernière proposition, on déduit alors que si  $x \in M_v^p$ ,  $\Phi_v^p(x) \in M_v^q$ , alors  $x \in M_v^{p+q}$  et  $\Phi_v^{p+q}(x) = \Phi_v^q(\Phi_v^p(x))$ . Donc  $t \mapsto \Phi_v^t$  est un homéomorphisme de  $(\mathbb{R}, +)$ .

Soit  $\chi^k(M)$  l'espace des champs de vecteurs  $C^k$  sur  $M$ . On peut maintenant énoncer le principal théorème de cette sous-partie :  $\square$

**6.5.7 Théorème:** On a les résultats suivants :

(i)  $\Omega^1 = \{(t, x, v) \in \mathbb{R} \times M \times \chi^1(M) \mid x \in M_v^t\}$  est un sous-ensemble ouvert de  $\mathbb{R} \times M \times \chi^1(M)$  contenant  $\{0\} \times M \times \chi^1(M)$ .

(ii)  $(t, x, v) \mapsto \Phi_v^t$  est une application (ou carte)  $C^k$  allant de  $\Omega^k = \Omega^1 \cap (\mathbb{R} \times M \times \chi^k(M)) \subseteq \mathbb{R} \times M \times \chi^k(M)$  dans  $M$ .

(iii) Si  $\epsilon \mapsto v_\epsilon$  est différentiable :  $J \rightarrow \chi^1(M)$ , avec  $J$  un intervalle de  $\mathbb{R}$ , on a la formule variationnelle :

$$\frac{\partial}{\partial \epsilon} \Phi_{v_\epsilon}^t(x) = \int_0^t T_{\Phi_{v_\epsilon}^s(x)}(\Phi_{v_\epsilon^{t-s}}) \frac{\partial v_\epsilon}{\partial \epsilon}(\Phi_{v_\epsilon^s}(x)) ds \in T_{\Phi_{v_\epsilon^t}(x)} M$$

(formule n1) (iv) Si  $T$  est une frontière finie de  $I(s, x, v)$ , alors il existe, pour tout compact  $K$  de  $M$ , un voisinage  $I$  de  $T$  dans  $\mathbb{R}$  tel que  $\Phi_v^t(x) \notin K$  pour tout  $t \in I \cap I(s, x, v)$ . En particulier, si  $M$  est compact, alors  $\Omega^1 = \mathbb{R} \times M \times \chi^1(M)$ .

On va maintenant s'attaquer à la démonstration des théorèmes.

**6.5.8 Lemme:** Soit  $x_0 \in \mathbb{R}^n$ ,  $r > 0$ ,  $B = \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\}$  et  $f_0 \in C^1(B, \mathbb{R}^n)$ . Pour une fonction arbitraire  $f \in C^1(B, \mathbb{R}^n)$ , notons :

$$M_0(f) = \sup \{\|f(x)\| \mid x \in B\}; M_1(f) = \sup \{\|Df(x)\| \mid x \in B\}$$

Choisissons également  $\delta > 0$  tel que  $\delta M_0(f_0) < r$  et  $\delta M_1(f_0) < 1$ . Alors pour chaque  $x \in \mathbb{R}^n$  et  $f \in C^1(B, \mathbb{R}^n)$  tel que  $\|x - x_0\| + \delta M_0(f) < r$  et  $\delta M_1(f) < 1$ , il existe une unique solution  $\gamma : ]-\delta, \delta[ \rightarrow \mathbb{R}^n$  de l'équation :

$$\frac{d\gamma}{dt}(t) = f(\gamma(t)) \quad (|t| < \delta) \quad (1)$$

de valeur initiale  $\gamma(0) = x$ .

Si on restreint  $f$  à une fonction  $C^k(B, \mathbb{R}^n)$ , alors l'application  $\gamma$  associée à ce  $x$  et à cette fonction  $f$  est de classe  $C^k$ . Finalement, si  $f$  dépend du paramètre  $\epsilon$ , alors

$$\frac{\partial}{\partial \epsilon} \gamma(t, x, \epsilon) = \int_0^t \frac{\partial \gamma}{\partial x}(t-s, \gamma(s, x, \epsilon), \epsilon) \frac{\partial f}{\partial \epsilon}(\gamma(s, x, \epsilon), \epsilon) ds \quad (2)$$

**Preuve:** En intégrant l'équation (1), on obtient

$$\gamma(t) = x + \int_0^t f(\gamma(s)) ds := F(\gamma, x, f)(t), \quad (|t| < \delta) \quad (3)$$

Si  $\gamma : ]-\delta, \delta[ \rightarrow \mathbb{R}^n$  est continue, alors elle est différentiable (car l'intégrale est une fonction différentiable en  $t$ ). De plus, on a bien  $\gamma(0) = x$ , et si on applique la formule n1 à l'équation (3), on obtient l'équation (2).

Soit  $\Gamma$  l'espace de Banach des courbes  $\gamma : ]-\delta, \delta[ \rightarrow \mathbb{R}^n$  qui sont bornées et continues, muni de la norme  $\|\cdot\|$ , qui est la norme sup. On note  $\underline{x}_0$  la courbe sur  $]-\delta, \delta[$  qui est constante et égale à  $x_0$ .

$$B := \{\gamma \in \Gamma \mid \|\gamma - \underline{x}_0\| \leq r\}$$

est un sous-ensemble convexe et fermé de  $\Gamma$ , car  $F(\gamma, x, f)$  est définie pour tout  $\gamma \in B$ ,  $x \in \mathbb{R}^n$  et  $f \in C^1(B, \mathbb{R}^n)$ . On a donc, par suite :

$$\|F(\gamma, x, f) - \underline{x}_0\| \leq \|x - x_0\| + \delta M_0(f)$$

et

$$\|F(\gamma, x, f) - F(\gamma', x, f)\| \leq \delta M_1(f) \|\gamma - \gamma'\|; \quad (\gamma, \gamma' \in B) \quad (4)$$

(théorème des accroissements finis). Comme  $F$  est une application  $k$ -contractante ( $k$ -lipschitzienne), c'est une contraction de  $B$  dans  $B$ . En appliquant le théorème du point fixe strictement contractant (voir encore une fois le théorème de la boule chevelue, dans les pré-requis), on déduit qu'il existe une unique courbe  $\gamma$  solution, telle que  $\gamma = F(\gamma, x, f)$ .

De par l'équation (1) toujours, on déduit que  $F$  est une application  $C^k$  de  $B \times \mathbb{R}^n \times C^k(B, \mathbb{R}^n)$  dans  $\Gamma$ . Puis, par l'équation (4), on obtient que la norme de  $\frac{\partial F}{\partial \gamma}(\gamma, x, f)$  de  $\gamma \mapsto F(\gamma, x, f) : B \rightarrow \Gamma$  est dominée par  $\delta M_1(f)$ , et  $\frac{\partial}{\partial \gamma}(\gamma - F(\gamma, x, f)) = I - \frac{\partial F}{\partial \gamma}(\gamma, x, f)$  est inversible tant que  $\delta M_1(f) < 1$ . On conclut, d'après le théorème des fonctions implicites sur les espaces de Banach, qu'il y a une solution de  $\gamma - F(\gamma, x, f) = 0$ , qui est  $C^k$  et dépend des paramètres  $(x, f)$ . On peut maintenant dériver en  $\epsilon$  l'équation (1) en remplaçant  $\gamma$  par  $t \mapsto \gamma(t, x, \epsilon)$  pour obtenir :

$$\frac{\partial \gamma}{\partial \epsilon}(t, x, \epsilon) = \int_0^t \left( \frac{\partial f}{\partial x}(\gamma(s, x, \epsilon), \epsilon) \frac{\partial \gamma}{\partial \epsilon}(s, x, \epsilon) + \frac{\partial f}{\partial \epsilon}(\gamma(s, x, \epsilon), \epsilon) \right) ds$$

En dérivant par rapport à  $t$ , on obtient :

$$\frac{d}{dt} \frac{\partial \gamma}{\partial \epsilon}(t, x, \epsilon) = \frac{\partial f}{\partial x}(\gamma(s, x, \epsilon), \epsilon) \frac{\partial \gamma}{\partial \epsilon}(s, x, \epsilon) + \frac{\partial f}{\partial \epsilon}(\gamma(s, x, \epsilon), \epsilon), \quad \frac{\partial \gamma}{\partial \epsilon}(0, x, \epsilon) = 0 \quad (5)$$

Or, en dérivant l'équation (3) par rapport à  $x$ , puis à  $t$ , on a :

$$\frac{d}{dt} \frac{\partial}{\partial x} \gamma(t, x, \epsilon) = \frac{\partial f}{\partial x}(\gamma(t, x, \epsilon), \epsilon) \frac{\partial \gamma}{\partial x}(t, x, \epsilon)$$

Donc  $\frac{\partial \gamma}{\partial x}$  est solution homogène de (5), par conséquent, on déduit par le calcul (méthode de la variation de constante) la formule (2) du lemme, qu'on vient maintenant de prouver.  $\square$

**Preuve:** du premier théorème de cette section

Le lemme nous assure l'unicité de la solution. On peut choisir  $t_0 = 0, \gamma(0) = x_0 = \gamma'(0)$  (sans pertes de généralités). En choisissant  $I_0$  suffisamment petit,  $I_0 = ]-\delta, \delta[$  ( $\delta$  du lemme), (qui reste valide si on le prend arbitrairement petit). Donc on prend  $\delta$  proche de 0 tel que  $\gamma(I_0) \subset B, \gamma'(I_0) \subset B$ , donc  $\gamma|_{I_0}$  et  $\gamma'|_{I_0}$  sont solutions comme dans le lemme.  $\square$

**Preuve:** du dernier théorème de cette section Cette partie est très calculatoire. Comme on n'a pas besoin de ce théorème (qui devait servir à la preuve du troisième théorème de Lie), je renvoie au livre Lie groups : J.J Duistermaat et J.A.C Kolk , à l'appendice B.  $\square$

## 7 Exponentielle

Le but de ce chapitre est d'introduire l'application exponentielle, définie de  $\mathfrak{g}$  dans  $G$ , qui permet de faire le lien entre un groupe de Lie et une algèbre de Lie. Pour cela, on a besoin de la notion de flot.

### 7.1 Définition à partir du flot

**7.1.1 Définition:** *Sous-groupe à un paramètre* de  $G =$  toute courbe  $\gamma : \mathbb{R} \rightarrow G$  qui soit un morphisme du groupe additif  $(\mathbb{R}, +)$  vers  $G$ , etc

**7.1.2 Définition:** Soit  $X$  un champ de vecteurs invariant à gauche sur  $G$ . Le *flot* de  $X$  est l'application  $\phi_X : I \rightarrow G$  de domaine maximale  $I \subset \mathbb{R}$  qui soit solution de l'équation différentielle (problème de Cauchy)

$$\frac{d}{dt} \phi_X(t) = X_{\phi_X(t)}$$

avec condition initiale  $\phi_X(0) = e$ . On note aussi  $\phi_X(t) = \phi(t, X)$ .

**7.1.3 Proposition:** *Pour tout  $X \in \mathfrak{X}^L(G)$ , le flot de  $X$  est un sous-groupe à un paramètre de  $G$ .*

**7.1.4 Définition:** Un champ de vecteur  $v$  sur  $G$  est dit invariant à gauche si

$$v_{L(x)y} = T_y L(x) v_y$$

et invariant à droite si

$$v_{R(x)y} = T_y R(x) v_y$$

De sorte que si  $v_1 = X, v_x = T_1 L(x) X, v_x = T_1 R(x) X$ . On note alors  $X_x^L$  et  $X_x^R$  les champs de vecteurs à gauche et à droite.

**7.1.5 Proposition:** *Soit  $\Phi^t$  le flot d'un champ de vecteur  $v$  invariant à gauche à droite sur  $G$ . C'est une application  $C^2, G \rightarrow G$  qui vérifie :*

$$\Phi^t = R(\Phi^t(1))$$

$$\Phi^t = L(\Phi^t(1))$$

la première formule pour  $v$  invariant à gauche, la seconde pour  $v$  invariant à droite.

**Preuve:** Soit  $v$  invariant à gauche (l'invariance à droite se montre de la même manière) :

$$x(t) = R(\Phi^t(1))(x) = x\Phi^t(1) = L(x)(\Phi^t(1))$$

Alors  $x(0) = x$  et

$$\frac{dx}{dt}(t) = T_{\Phi^t(1)}L(x)(v(\Phi^t(1))) = v(L(x)(\Phi^t(1))) = v(x(t))$$

On vient simplement de montrer que  $L(x)(\Phi^t(1))$  est également le flot du champ de vecteur  $v$  mais de valeur initiale différente. Plus exactement,  $x(t) = \Phi^t(x)$ .  $\square$

**7.1.6 Théorème:** Pour tout  $X \in \mathfrak{g}$ , il existe un unique homomorphisme

$h = h_X : (\mathbb{R}, +) \rightarrow (G, \cdot)$  différentiable en  $t = 0$  et vérifie  $\frac{dh}{dt}(0) = X$ . Cet homomorphisme est égal à la courbe solution d'un champ de vecteur invariant  $X^L, X^R$ , qui vérifie  $h(0) = 1$ , 1 étant l'élément identité du groupe  $G$ , et le flot des champs de vecteurs  $X^L, X^R$  est globalement définie pour tout  $t \in \mathbb{R}$ .

**Preuve:** Soit  $\Phi^t$  le flot de  $X^L$ . Par définition :

$$\Phi^{t+s}(1) = \Phi^s(\Phi^t(1)) = \Phi^t(1) \cdot \Phi^s(1)$$

(en utilisant la proposition précédente). Cela montre qu'il existe un  $\epsilon > 0$  tel que, si  $\Phi^t(1)$  est définie, alors  $\Phi^{t'}(1)$  est définie pour tout  $t' \in ]t - \epsilon, t + \epsilon[$ . Comme  $t \mapsto \Phi^t(1)$  est définie pour tout  $t \in \mathbb{R}$ , et c'est un homomorphisme  $C^1$ ,  $h : (\mathbb{R}, +) \rightarrow (G, \cdot)$ , car  $\frac{dh}{dt}(0) = X$ . On déduit que  $\Phi^t : G \rightarrow G$  existe pour tout  $t \in \mathbb{R}$ . On obtient la même chose pour  $X^R$ .

On a alors, comme  $h$  est un homomorphisme :

$$h(t+s) = h(t)h(s) = h(s)h(t) \Rightarrow \frac{dh}{ds}(s+t) = h'(s+t) = h'(s)h(t); (s=0); = Xh(t) = h'(t)$$

en dérivant puis en appliquant en  $s = 0$ , ou plus exactement :

$$\frac{dh}{dt}(t) = T_1L(h(t))(X) = X^L(h(t))$$

donc  $h(t)$  est courbe solution pour  $X^L$  et  $X^R$  (pour  $X^R$ , preuve analogue), avec  $h(0) = 1$ . De plus,  $h$  est uniquement déterminée par  $X \in \mathfrak{g}$ .  $\square$

**7.1.7 Définition:** Pour chaque  $X \in \mathfrak{g}$ ,  $\exp(X) \in G$  est définie par  $h(1)$ ,  $h$  étant l'homomorphisme différentiable tel que  $\frac{dh}{dt}(0) = X$ .  $\exp : X \mapsto \exp(x) : \mathfrak{g} \rightarrow H$  est l'exponentielle de  $\mathfrak{g}$  dans  $G$ .

Dans ce cas,  $\exp(X)$  est la courbe solution (le flot) du champ de vecteur  $X^L$  au temps  $t = 1$  qui part de  $t = 0$  et qui vaut l'élément identité du groupe  $G$  en  $t = 0$ . De plus,

$$h : s \mapsto h_X(st)$$

est un homomorphisme différentiable tel que  $\frac{dh}{dt}(0) = tX : h_X(st) = h_{tX}(s)$ , soit en  $s = 1$ ,

$$h_X(t) = \exp(tX), (X \in \mathfrak{g}, t \in \mathbb{R})$$

ce qui implique que  $t \mapsto \exp(tX)$  est un homomorphisme. En calculant la différentielle selon  $t$  en  $t = 0$ ,  $X = T_0(\exp)(X)$ .

## 7.2 Exponentielle matricielle : Définition, Convergence, Propriétés

Ici,  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$

Remarque : le développement en série de Taylor de l'exponentielle est valable pour tout groupe de Lie  $G$  dans un voisinage de 0 de son algèbre de Lie  $\mathfrak{g}$ .

**7.2.1 Définition:** On définit ainsi l'exponentielle matricielle :

$$\exp : M_n(\mathbb{K}) \longrightarrow M_n(\mathbb{R})$$

$$\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!}$$

On définit également le logarithme népérien dans un voisinage de  $O_n$  de  $M_n(\mathbb{K})$  ( $\neq$  de la fonction logarithme sur les matrices, qu'on définira un peu plus tard) :

$$\text{Log}(A + I_n) = \sum_{n=0}^{+\infty} \frac{(-1)^{n+1} \cdot A^n}{n}$$

(Log avec une majuscule).

**7.2.2 Proposition:** La série exponentielle converge normalement pour tout  $A \in M_n(\mathbb{K})$  et la série Logarithme converge normalement pour tout  $A \in M_n(\mathbb{K})$  telle que  $\|A\| < 1$ . Pour cela, on doit introduire la notion de normes (qui sera détaillée plus loin dans la partie topologie sur les matrices. Rappel :  $\|\cdot\|$  est une norme d'algèbre si :

-  $\|A\| = 0 \Leftrightarrow A = 0$ .

- Pour tout  $A \in M_n(\mathbb{K}), \forall \lambda \in \mathbb{K}, \|\lambda \cdot A\| = |\lambda| \cdot \|A\|$

- Pour tout  $A, B \in M_n(\mathbb{K}) \ \|A + B\| \leq \|A\| + \|B\|$

- Pour tout  $A, B \in M_n(\mathbb{K}) \ \|AB\| \leq C \|A\| \cdot \|B\|$  ( $C =$  constante)

Remarque : les 3 premières propriétés sont la définition d'une norme, la quatrième est une condition supplémentaire pour qu'on puisse parler de normes d'algèbre.

**Preuve:** Pour  $\exp$  :

$$\|\exp(A)\| = \left\| \sum_{n=0}^{+\infty} \frac{A^n}{n!} \right\| \leq \sum_{n=0}^{+\infty} \left\| \frac{A^n}{n!} \right\| = \exp(\|A\|)$$

Or, l'utilisation des séries numériques montre très clairement que cette série converge. De même pour Log :

$$\|\text{Log}(A)\| = \left\| \sum_{n=0}^{+\infty} \frac{(-1)^{n-1} \cdot A^n}{n} \right\| \leq \sum_{n=0}^{+\infty} \left\| \frac{(-1)^{n-1} \cdot A^n}{n} \right\| = \text{Log}(\|A\|)$$

Là encore, les séries numériques montrent que cette série converge si  $\|A\| < 1$ . Ici, on a montré la convergence mais il suffit d'enlever une étape de la démonstration (on passe directement à la norme) pour avoir la convergence normale. □

**7.2.3 Proposition:**  $\exp((u + v)A) = \exp(uA)\exp(vA)$  pour tout  $A \in M_n(\mathbb{K})$  et  $u, v \in \mathbb{K}$ . De plus,  $\exp(A) \in GL_n(\mathbb{K})$  et  $\exp(A)^{-1} = \exp(-A)$ .

**Preuve:**

$$\begin{aligned} \exp((u + v)A) &= \sum_{n=0}^{+\infty} (u + v)^n \frac{A^n}{n!}; \exp(uA)\exp(vA) = \sum_{n=0}^{+\infty} u^n \frac{A^n}{n!} \sum_{n=0}^{+\infty} v^n \frac{A^n}{n!} = \\ &= \sum_{r=0}^{+\infty} \sum_{s=0}^{+\infty} u^r v^s \cdot \frac{A^{r+s}}{(r + s)!} = \sum_{n=0}^{+\infty} \sum_{r=0}^n \frac{u^r}{r!} \frac{v^{n-r}}{(n - r)!} A^n = \sum_{n=0}^{+\infty} \frac{1}{n!} \left( \sum_{r=0}^n \binom{n}{r} u^r \cdot v^{n-r} \right) A^n = \end{aligned}$$



$$\sum_{n=0}^{+\infty} (u+v)^n \cdot \frac{A^n}{n!} = \exp((u+v)A).$$

Donc :  $I_n = \exp(O_n) = \exp(A \cdot 1 + (-1) \cdot A) = \exp(A) \cdot \exp(-A)$ . □

**7.2.4 Proposition:** Si  $A$  et  $B$  commutent :  $\exp(A+B) = \exp(A)\exp(B)$ .

**Preuve:**

$$\begin{aligned} \exp(A)\exp(B) &= \sum_{r=0}^{+\infty} \frac{A^r}{r!} \sum_{s=0}^{+\infty} \frac{B^s}{s!} = \sum_{r=0, s=0}^{+\infty} \frac{A^r}{r!} \frac{B^s}{s!} = \sum_{n=0}^{+\infty} \sum_{r=0}^n \frac{A^r}{r!} \frac{B^{n-r}}{(n-r)!} = \\ &= \sum_{n=0}^{+\infty} \frac{1}{n!} \left( \sum_{r=0}^n \binom{n}{r} A^r B^{n-r} \right) = \sum_{n=0}^{+\infty} \frac{(A+B)^n}{n!} = \exp(A+B). \end{aligned}$$

Remarque : on peut utiliser le binôme de Newton uniquement car  $A$  et  $B$  commutent. □

**7.2.5 Definition:** On définit l'application logarithme de  $M_n(\mathbb{K})$  par

$$\log(A) = \text{Log}(A - I_n) = \sum_{n=0}^{+\infty} \frac{(-1)^{n+1} \cdot (A - I_n)^n}{n}$$

définie dans un voisinage de  $I_n$ .

**7.2.6 Proposition:** Si  $\|A\| < 1$ ,  $\exp(\log(A)) = A$  et si  $\|\exp(B)\| < 1$ ,  $\log(\exp(B)) = B$ .

**Preuve:** Immédiat d'après les propriétés établies pour le Log. □

Soit

$$F(X) = \sum_{k=0}^{+\infty} \frac{X^k}{(k+1)!}$$

$$F(X)X + 1 = \sum_{k=0}^{+\infty} \frac{X^{k+1}}{(k+1)!} + 1 = \exp(X)$$

$\Leftrightarrow F(X)^n = \frac{\exp(X)^n - 1}{X}$  On pose :  $\text{Ad}_A(C) = AC - CA$  :

$$F(\text{Ad}_A(C)) = \sum_{k=0}^{+\infty} \frac{\text{Ad}_A^k(C)}{(k+1)!}$$

**7.2.7 Proposition:**  $\left(\frac{d}{dt}\right)_{t=0}(\exp(A+tB)) = F(\text{Ad}_A(B))\exp(A)$ .

**Preuve:** Soit  $D = \frac{d}{ds}$  et  $f(s)$  une fonction d'une variable réelle.

$$F(D) \Big|_{s=0} f(s) = \sum_{k=0}^{+\infty} \left(\frac{d}{ds}\right)_{s=0} f(s) \sum_{k=0}^{+\infty} \frac{f^k(0)}{(k+1)!}$$

Soit  $f$  une fonction telle que son développement limité en 0 converge normalement vers  $f$  (pour un certain rayon  $R \geq 1$ ).

$$\begin{aligned} \int_0^1 f(x) dx &= \int_0^1 \left( \sum_{k=0}^{+\infty} \frac{f^k(0)}{k!} x^k \right) dx = \sum_{k=0}^{+\infty} \frac{f^k(0)}{k!} \int_0^1 x^k dx \\ &= \sum_{k=0}^{+\infty} \frac{f^k(0)}{(k+1)!} F(D) \Big|_{s=0} f(s) \end{aligned}$$

Maintenant, posons

$$\phi(s) = \exp(sA)B\exp((1-s)A) = \exp(sA)B\exp(-sA)\exp(A)$$

. On va avoir besoin d'un lemme :  $Ad_{e^X}() = e^{Ad(X)()}$  (on le démontre plus tard). Donc :

$$\phi(s) = \exp(sAd_A)(B)\exp(A)$$

. Donc, en dérivant à chaque fois et en appliquant en  $s=0$

$$F(D) |_{s=0} \phi(s) = \left( \sum_{k=0}^{+\infty} \frac{Ad_A^k}{(k+1)!} \right) (B)\exp(A)F(Ad_A(B))\exp(A)$$

De plus,

$$\left( \frac{d}{dt} \right) |_{t=0} \exp(A+tB) = \left( \frac{d}{dt} \right) |_{t=0} \left( \sum_{k=0}^{+\infty} \frac{(A+tB)^k}{k!} \right) = \sum_{k=0}^{+\infty} \frac{1}{k!} \sum_{i+j=k-1} A^i B A^j = \sum_{k=0}^{+\infty} \sum_{i+j=k-1} \frac{1}{k!} A^i B A^j$$

(en fait, comme il s'agit de la dérivée d'un polynôme évaluée en 0, on s'arrange pour ne garder que le second terme du binôme en t, car si  $P(X) = \sum_{k=0}^n a_k X^k$ , alors  $P'(0) = a_1$  mais on ne voit pas apparaître le second terme  $\binom{k}{1} = k$  du binôme à cause de la non commutativité de A et B). Enfin, on utilise le fait que

$$\int_0^1 \frac{s^m (1-s)^n ds}{m!n!} = \frac{1}{(m+n+1)!}$$

(se démontre par intégration par partie successives, voir juste en dessous). En utilisant cette égalité pour  $\frac{1}{k!}$  :

$$\frac{1}{k!} = \int_0^1 \frac{s^i (1-s)^j ds}{i!j!} \quad (i+j = k-1)$$

, on obtient :

$$\left( \frac{d}{dt} \right) |_{t=0} \exp(A+tB) = \int_0^1 \sum_{k=0}^{+\infty} \sum_{i+j=k-1} \frac{s^i A^i}{i!} B \frac{A^j (1-s)^j ds}{j!} = \int_0^1 \phi(s) ds$$

(formule du produit de séries infinie, et on identifie très clairement  $\exp(sA)$  et  $\exp((1-s)A)$ ). Donc, on a bien l'égalité de la proposition.  $\square$

### 7.2.8 Proposition:

$$\int_0^1 \frac{s^m (1-s)^n ds}{m!n!} = \frac{1}{(m+n+1)!}$$

**Preuve:** Par intégration par partie, soit :

$$\begin{aligned} u(s) = s^m, v'(s) = (1-s)^n &\Rightarrow u'(s) = ms^{m-1} \text{ et } v(s) = -\frac{(1-s)^{n+1}}{n+1} : \int_0^1 s^m (1-s)^n ds \\ &= \\ & \left[ -\frac{s^m (1-s)^{n+1}}{n+1} \right]_0^1 + \frac{m}{n+1} \int_0^1 s^{m-1} (1-s)^{n+1} ds = \\ & \frac{m}{n+1} \int_0^1 s^m (1-s)^n ds \text{ car } \left[ -\frac{s^m (1-s)^{n+1}}{n+1} \right]_0^1 = 0. \forall m, n > 0 \end{aligned}$$

Notons :

$$I_{m,n} = \int_0^1 \frac{s^m (1-s)^n ds}{m!n!}$$

. On a alors, par ce qu'on vient de faire :

$$I_{m,n} \frac{m-1}{n+1} I_{m,n}$$

. Finalement, par récurrence :

$$I_{m,n} = \frac{m!}{(n+1)(n+2)\dots(n+m)} I_{0,m+n}$$

avec

$$I_{0,m+n} \int_0^1 \frac{(1-s)^{n+m} ds}{m!n!} = \frac{1}{m!n!(m+n+1)}$$

Donc,

$$I_{m,n} = \frac{m!}{(n+1)(n+2)\dots(n+m)} \cdot \frac{1}{m!n!(m+n+1)} \frac{1}{(m+n+1)!}$$

□

**7.2.9 Definition:** Une courbe différentiable dans  $M_n(\mathbb{K})$  est une fonction  $\alpha : [a; b] \rightarrow M_n(\mathbb{K})$  pour laquelle la dérivée  $\alpha'(t)$  existe pour tout  $t \in [a; b]$  et telle que  $\alpha'(t) = \lim_{s \rightarrow t} \frac{\alpha(s) - \alpha(t)}{s - t}$ .

**7.2.10 Proposition:** Soit  $A \in M_n(\mathbb{K})$ . Il existe une unique solution de l'équation  $\alpha'(t) = \alpha(t)A$ ,  $\alpha(0) = C \in M_n(\mathbb{K})$ .

**Preuve:** Soit

$$\begin{aligned} \alpha(t) = \exp(tA) &= \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!} : \alpha'(t) = \exp(tA) = \sum_{k=k}^{+\infty} \frac{k t^{k-1} A^k}{k!} = A \left( \sum_{k=1}^{+\infty} \frac{t^{k-1} A^{k-1}}{(k-1)!} \right) \\ &= A \left( \sum_{k=0}^{+\infty} \frac{t^k A^k}{k!} \right) = \alpha(t)A \end{aligned}$$

Soit  $\beta$  une autre solution de l'équation et

$$\gamma(t) = \beta(t) \exp(-tA) : \gamma'(t) = \beta'(t) \exp(-tA) - \beta(t) A \exp(-tA) = \beta A \exp(-tA) - \beta(t) A \exp(-tA) = O_n$$

Donc  $\gamma(t) = C$  (constante) et  $\beta(t) = C \exp(tA)$ , qui vérifie  $\beta(0) = C$ . □

**7.2.11 Definition:** Un sous-groupe à un paramètre dans  $G$  est l'image d'une fonction  $\gamma : [-\epsilon; \epsilon] \rightarrow G$  qui est différentiable en 0 et qui vérifie  $\gamma(s+t) = \gamma(s)\gamma(t)$  pour tout  $s, t, s+t \in [-\epsilon; \epsilon]$ .  $\gamma$  est un morphisme de groupe  $]-\epsilon, \epsilon[$  vers  $G$ . Remarque :  $\gamma(s) = \gamma(s+0) = \gamma(s)\gamma(0)$  pour tout  $s \in [-\epsilon; \epsilon] \Rightarrow \gamma(0) = I$ .

**7.2.12 Proposition:** Soit  $\gamma : [-\epsilon; \epsilon] \rightarrow G$  un sous-groupe à un paramètre dans  $G$  : Alors, pour tout  $t \in [-\epsilon; \epsilon]$ ,  $\gamma$  est une courbe différentiable en  $t$  et  $\gamma'(t) = \gamma'(0)\gamma(t) = \gamma(t)\gamma'(0)$

**Preuve:** Soit  $h \in \mathbb{R}$  suffisamment petit :

$$\gamma(t)\gamma(h) = \gamma(t+h) = \gamma(h)\gamma(t)$$

(de par la commutativité de l'addition).

$$\gamma'(t) = \lim_{h \rightarrow 0} \frac{\gamma(t+h) - \gamma(t)}{h} = \gamma(t) \cdot \lim_{h \rightarrow 0} \frac{\gamma(h) - I}{h} = \gamma'(0)\gamma(t)$$

□

**7.2.13 Proposition:** Soit  $\gamma : [-\epsilon; \epsilon] \rightarrow G$  un sous-groupe d'un paramètre dans  $G$ . Il existe une unique extension de  $\gamma$  à tout  $\mathbb{R}$ ,  $\tilde{\gamma} : \mathbb{R} \rightarrow G$ , c'est à dire, une fonction  $\tilde{\gamma}$  telle que  $\tilde{\gamma}(t) = \gamma(t)$  pour tout  $t \in [-\epsilon; \epsilon]$ .

**Preuve:** Soit  $t \in \mathbb{R} : \exists m \in \mathbb{N}$  tel que  $\frac{t}{m} \in [-\epsilon; \epsilon]$ . Donc  $\gamma(\frac{t}{m})$  et  $\gamma(\frac{t}{m})^m \in G$ . De même  $\exists n \in \mathbb{N}$  tel que  $\frac{t}{n} \in [-\epsilon; \epsilon]$  et donc  $\gamma(\frac{t}{n})$  et  $\gamma(\frac{t}{n})^n \in G$ . Supposons  $m, n \neq 0 : mn \geq n, m$ . Donc  $\frac{t}{mn} \in [-\epsilon; \epsilon]$  et  $\gamma(\frac{t}{n})^n = \gamma(\frac{t}{mn})^{mn} = \gamma(\frac{t}{nm})^{mn} = \gamma(\frac{tn}{nm})^m = \gamma(\frac{t}{m})^m$ . Donc  $\gamma(\frac{t}{n})^n = \gamma(\frac{t}{m})^m$ . Cela définit  $\tilde{\gamma} : \mathbb{R} \rightarrow G$  tel que  $\tilde{\gamma} = \gamma(\frac{t}{n})^n$  pour  $n$  assez grand.  $\square$

**7.2.14 Théorème:** Soit  $\gamma : \mathbb{R} \rightarrow G$  un groupe à un paramètre dans  $G$ . Alors on a  $\gamma(t) = \exp(tA)$  pour une certaine matrice  $A \in M_n(\mathbb{K})$

**Preuve:**  $\gamma'(t) = \gamma'(0)\gamma(t) \Rightarrow \gamma(t) = \exp(tA)$  (car  $\gamma(0) = I$ ).  $\square$

### 7.3 Théorème de Cartan\*

Dans cette section on veut caractériser les sous-groupes d'un groupe de Lie qui sont des sous-groupes de Lie, c'est-à-dire qu'ils sont aussi des sous-variétés avec loi de groupe et inversion lisses.

Tout d'abord, on remarque qu'un ouvert de  $\mathbb{R}^n$  ne peut pas être de "dimension" inférieure à  $n$ . Par exemple, un segment ouvert de  $\mathbb{R}$  n'est pas un ouvert de  $\mathbb{R}^2$ .

**7.3.1 Lemme:** Soit  $M$  une variété différentielle de dimension  $n$ , et soit  $N \subset M$  un sous-ensemble ouvert de  $M$ . Alors  $N$  est aussi une variété différentielle, de même dimension  $n$ .

**Preuve:** Sur  $M$ , on a un recouvrement  $\{U_i\}_i$  d'ouverts qui supportent des cartes locales  $\varphi_i : U_i \rightarrow \mathbb{R}^n$  (homéomorphismes) telles que tout changement de carte  $\varphi_i \circ \varphi_j^{-1}$  soit un difféomorphisme.

Les intersections  $V_i = N \cap U_i$  sont ouvertes et l'ensemble  $\{V_i\}_i$  est évidemment un recouvrement de  $N$ . En restreignant les cartes  $\varphi_i$  à  $V_i$ , on trouve bien un atlas de  $N$ . Par conséquent  $N$  est une variété différentielle.

Cela montre également que  $\dim N = \dim M$ .  $\square$

**7.3.2 Remarque:** Si on applique le lemme précédent à un groupe de Lie  $G$ , on serait tentés de déduire que tout sous-groupe ouvert  $H$  de  $G$  est une sous-variété réelle de  $G$ , et par conséquent un sous-groupe de Lie de  $G$ . Or, cela est faux, comme le montre le contre-exemple donné à la fin de cette section.

Au contraire, la condition sur un sous-groupe  $H$  de  $G$  qui garantit que  $H$  est un sous-groupe de Lie est donnée par le théorème de Cartan suivant.

**7.3.3 Théorème:** Soit  $G$  un groupe de Lie. Tout sous-groupe fermé de  $G$  est une sous-variété, et par conséquent un sous-groupe de Lie de  $G$ .

Cette partie est consacrée à la démonstration du théorème de Cartan (qui est en fait due à Von Neuman).

L'idée est d'abord de montrer qu'un sous-groupe fermé détermine une sous-algèbre de Lie de l'algèbre de Lie  $\mathfrak{g}$  de  $G$ . Ensuite, qu'il est l'image de cette algèbre de Lie par l'exponentielle.

Pour cela, on a besoin de plusieurs résultats sur l'application exponentielle.

**7.3.4 Théorème:** (admis) Soit  $G$  un groupe de Lie et  $\mathfrak{g}$  son algèbre de Lie. Il existe un voisinage ouvert  $U$  de 0 dans  $\mathfrak{g}$  et un voisinage ouvert  $V$  de  $e$  dans  $G$  tels que l'application exponentielle  $\exp : \mathfrak{g} \rightarrow G$  restreinte à  $U$  soit un difféomorphisme sur  $V$ , développable en série entière (son développement de Taylor).

De plus, le développement de Taylor de  $\exp$  est formellement le même que celui de la fonction exponentielle réelle (et matricielle également).

**7.3.5 Lemme:** Soit  $G$  un groupe de Lie. Soient  $X$  et  $Y$  deux éléments de  $\mathfrak{g}$ , alors :

$$\lim_{n \rightarrow \infty} \left[ \exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right) \right]^n = \exp(X + Y)$$

$$\lim_{n \rightarrow \infty} \left[ \exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right) \exp\left(\frac{-X}{n}\right) \exp\left(\frac{-Y}{n}\right) \right]^{n^2} = \exp(XY - YX).$$

**Preuve:** D'après le théorème précédent, on peut utiliser le développement en série entière de  $\exp$  si  $n$  est suffisamment grand, ce qui est le cas puisque  $n \rightarrow +\infty$ . Puisque

$$\exp\left(\frac{X}{n}\right) = Id + \frac{X}{n} + O\left(\frac{1}{n^2}\right)$$

on a

$$\exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right) = \frac{X + Y}{n} + O\left(\frac{1}{n^2}\right),$$

et donc

$$n \text{Log}\left(\exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right)\right) = X + Y + O\left(\frac{1}{n^2}\right)$$

d'où suit

$$\lim_{n \rightarrow \infty} n \text{Log}\left(\exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right)\right) = X + Y$$

et

$$\lim_{n \rightarrow \infty} \left[ \exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right) \right]^n = \exp(X + Y).$$

De la même façon, on trouve

$$\begin{aligned} \left[ \exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right) \exp\left(\frac{-X}{n}\right) \exp\left(\frac{-Y}{n}\right) \right]^{n^2} &= \left( Id + \frac{X}{n} + \frac{X^2}{2n^2} + O\left(\frac{1}{n^3}\right) \right) \left( Id + \frac{Y}{n} + \frac{Y^2}{2n^2} + O\left(\frac{1}{n^3}\right) \right) \\ &\quad \times \left( Id - \frac{X}{n} + \frac{X^2}{2n^2} + O\left(\frac{1}{n^3}\right) \right) \left( Id - \frac{Y}{n} + \frac{Y^2}{2n^2} + O\left(\frac{1}{n^3}\right) \right) \\ &= \left( Id + \frac{X + Y}{n} + \frac{X^2 + 2XY + Y^2}{2n^2} \right) \left( Id - \frac{X + Y}{n} + \frac{X^2 + 2XY + Y^2}{2n^2} \right) + O\left(\frac{1}{n^3}\right) \\ &= Id + \frac{X^2 + 2XY + Y^2}{n^2} - \frac{(X + Y)^2}{n^2} + O\left(\frac{1}{n^3}\right) = Id + \frac{XY - YX}{n^2} + O\left(\frac{1}{n^3}\right). \end{aligned}$$

Donc

$$n^2 \text{Log}\left(\exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right) \exp\left(\frac{-X}{n}\right) \exp\left(\frac{-Y}{n}\right)\right) = XY - YX,$$

et finalement :

$$\lim_{n \rightarrow \infty} \left[ \exp\left(\frac{X}{n}\right) \exp\left(\frac{Y}{n}\right) \exp\left(\frac{-X}{n}\right) \exp\left(\frac{-Y}{n}\right) \right]^{n^2} = \exp(XY - YX).$$

□

**7.3.6 Proposition:** Soit  $H$  un sous groupe fermé de  $G$ , alors l'ensemble  $W = \{X \in \mathfrak{g}, \exp(tX) \in H, \forall t \in \mathbb{R}\}$  est un sous-espace vectoriel de  $\mathfrak{g}$  stable par l'application

$$(X, Y) \longrightarrow XY - YX.$$

Autrement dit, c'est une sous-algèbre de Lie de  $\mathfrak{g}$ .

**Preuve:** Soit  $X, Y \in W, \lambda \in \mathbb{R}$ . Soit  $Z = X + \lambda Y$  :

$$\exp(tZ) = \exp(tX + t\lambda Y) = \lim_{n \rightarrow \infty} \left[ \exp\left(\frac{tX}{n}\right) \exp\left(\frac{t\lambda Y}{n}\right) \right]^n \in W$$

car  $W$  est un sous-groupe fermé. De même, la stabilité par le crochet de Lie est assurée par l'application est assurée par la deuxième formule du lemme précédent. □

Donc  $W$  est bien une sous-algèbre de Lie de  $\mathfrak{g}$ , notons-la  $\mathfrak{h}$ .

**Remarque :** Il est clair que 0 (le neutre de  $\mathfrak{h}$ ) est dans  $\mathfrak{h}$ . Mais on ne sait pas s'il y a dans  $\mathfrak{h}$  des éléments non nuls. Le lemme suivant nous garantit leur existence.

**7.3.7 Lemme:** Soit  $h_n$  une suite d'éléments de  $H$  tendant vers  $Id$ , avec  $h_n \neq Id$  (ce qui est possible puisque  $H$  n'est pas discret). Alors  $\text{Log}(h_n)$  est une suite de vecteurs dans  $\mathfrak{h}$ , et on peut considérer leur norme euclidienne.

Si  $h$  est une valeur d'adhérence de la suite  $\frac{\text{Log}(h_n)}{\|\text{Log}(h_n)\|}$  (suite bien définie pour  $n$  suffisamment grand), alors  $h \in \mathfrak{h}$ .

De plus,  $h$  est alors forcément non nul, car tout élément  $\frac{\text{Log}(h_n)}{\|\text{Log}(h_n)\|}$  appartient à la sphère unité de  $\mathfrak{g}$  qui est compacte. Par conséquent,  $\mathfrak{h} \neq \{0\}$ .

**Preuve:** On peut supposer que  $\frac{\text{Log}(h_n)}{\|\text{Log}(h_n)\|}$  converge vers  $h$ . On doit calculer  $\exp(th)$  :

$$\exp(th) = \lim_{n \rightarrow \infty} \exp\left(t \frac{\text{Log}(h_n)}{\|\text{Log}(h_n)\|}\right).$$

Si  $\frac{t}{\|\text{Log}(h_n)\|} = k_n$  était un entier, cela ne poserait pas de problèmes car

$$\exp(k_n \text{Log}(h_n)) = (\exp(\text{Log}(h_n)))^{k_n} = h_n^{k_n}$$

serait une limite d'éléments de  $H$ , donc élément de  $H$  puisque  $H$  est fermé.

Sinon, on pose tout simplement :  $\frac{t}{\|\text{Log}(h_n)\|} = P_n + V_n$  où  $P_n$  est la partie entière de  $\frac{t}{\|\text{Log}(h_n)\|}$  et  $V_n \in [0; 1[$ . Alors

$$\exp\left(t \frac{\text{Log}(h_n)}{\|\text{Log}(h_n)\|}\right) = h_n^{P_n} \times \exp(V_n \text{Log}(h_n)).$$

Enfin, on a  $\|V_n \text{Log}(h_n)\| \leq \|\text{Log}(h_n)\|$ , qui tend vers 0, donc  $\exp(V_n \text{Log}(h_n)) \rightarrow Id$  et  $\exp(th) = \lim_{n \rightarrow \infty} h_n^{P_n}$  est une limite d'éléments de  $H$ , donc un élément de  $H$  et  $h \in \mathfrak{h}$ .  $\square$

**7.3.8 Théorème:** Soit  $G$  un groupe de Lie. Si  $H$  est un sous-groupe fermé de  $G$ , il existe un voisinage  $V$  de 0 dans  $\mathfrak{h}$  et un voisinage  $W$  de  $e$  dans  $H$ , tels que l'exponentielle réalise un homéomorphisme entre  $V$  et  $W$ .

**Preuve:** D'après le théorème que l'on a admis, il existe un voisinage  $V_0$  de 0 dans  $\mathfrak{g}$  et un voisinage  $W_0$  de  $e$  dans  $G$  tel que l'exponentielle réalise un homéomorphisme entre  $V_0$  et  $W_0$ . L'enjeu est de montrer qu'il en est de même pour  $H$ .

Soit  $V = V_0 \cap \mathfrak{h}$  et  $W = W_0 \cap G$ . Ce sont également des ouverts, mais rien ne nous permet d'affirmer la surjectivité dans  $\mathfrak{h}$  : si  $Y$  est proche de  $e$  et  $Y \in H$ , alors  $\text{Log} Y$  n'appartient pas a priori à  $\mathfrak{h}$ , et rien ne permet de dire que  $\exp(V)$  est un voisinage de 0 dans  $W$ .

Dans la suite, on suppose  $H$  non discret (non réduit à un singleton), connexe (s'il n'est pas connexe, on se ramène à ce cas en considérant une composante connexe).

Soit  $\mathfrak{h}'$  un supplémentaire de  $\mathfrak{h}$  dans  $\mathfrak{g}$ . On va montrer qu'il existe un voisinage  $V'$  de 0 dans  $\mathfrak{h}'$  tel que  $\exp(V') \cap H = \{e\}$ . En effet, dans le cas contraire, il existerait un voisinage  $W_0$  de 0 dans  $\mathfrak{h}'$ , que l'on pourrait choisir compact et étoilé par rapport à 0 (étoilé signifie que tout point peut être relié par un chemin rectiligne), et tel que  $\exp(\frac{W_0}{n}) \cap H \neq \{e\}$  pour tout  $n$ . Soit alors  $h_n = \exp(\frac{X'_n}{n}) \in H$ , avec  $X'_n \in W_0$ . La suite  $(h_n)$  tend vers  $e$  puisque, pour  $n$  assez grand, on a que  $\frac{\text{Log}(h_n)}{\|\text{Log}(h_n)\|}$  tend vers 0 ( $X'_n$  est bornée). Par conséquent, pour  $n$  assez grand, la suite  $\frac{\text{Log}(h_n)}{\|\text{Log}(h_n)\|} = \frac{X'_n}{\|X'_n\|}$  possède une sous-suite convergeant vers  $h$  qui, d'une part, appartient à la boule unité de  $\mathfrak{g}'$  et, d'après le lemme précédent,  $h \in \mathfrak{h}$ , ce qui est absurde.

Pour finir, soit  $\phi$  l'application sur  $\mathfrak{h} \times \mathfrak{h}'$  à valeurs dans  $G$ , définie en posant  $\phi(X, X') = \exp(X) \exp(X')$ . Alors :  $\phi(X, X') = e + X + X' + O(\sqrt{\|X\|^2 + \|X'\|^2})$ , lorsque  $(X, X')$  tend vers  $(0, 0)$ . Par conséquent, la

différentielle de  $\phi$  en  $(0, 0)$  est l'application  $(X, X') \mapsto X + X'$ , qui est bijective. Le théorème d'inversion locale garantit alors l'existence d'un voisinage  $U$  de 0 dans  $\mathfrak{h}$ , d'un voisinage  $U'$  de 0 dans  $\mathfrak{h}'$  et enfin d'un voisinage  $V$  de 0 dans  $\mathfrak{h}$ , tels que  $\phi$  réalise un difféomorphisme entre  $U \times U'$  et  $V$ . Nous allons voir que  $\exp(U)$  est un voisinage de  $e$  dans  $H$ . En effet, l'intersection  $U_1 = U' \cap V'$  (avec  $V'$  défini précédemment) est un voisinage de 0 dans  $\mathfrak{g}'$ , et l'intersection  $V_1 = \phi(U \times U_1)$  est un voisinage de  $e$  dans  $G$ . L'application  $\phi$  réalise un difféomorphisme entre  $U \times U_1$  et  $V_1$ . Il en résulte que  $V_1 \cap H = \exp(U)$  : l'inclusion  $\exp(U) \subseteq V_1 \cap H$  est évidente, l'inclusion réciproque vient du fait que  $\exp(V') \cap H = \{e\}$ , donc  $\exp(U)$  est un voisinage de  $e$  dans  $G$ .  $\square$

Nous pouvons enfin montrer le théorème de Cartan : tout sous-groupe fermé de  $G$  est une sous-variété, et par conséquent un sous-groupe de Lie de  $G$ .

**Preuve:** Avec les notations du théorème précédent, il existe en effet un voisinage  $V_1$  de  $e$  dans  $G$ , un voisinage  $U \times U_1$  de 0 dans  $\mathfrak{g}$  et un difféomorphisme  $\phi$  entre  $U \times U_1$  et un espace vectoriel, à savoir  $\mathfrak{h}$ , tels que :

$$\phi(\mathfrak{h} \cap (U \times U_1)) = V_1 \cap H (= \exp(U)).$$

Ceci est exactement la définition d'une sous-variété (définition (A)). On a établi ceci pour  $x = e$ , mais  $X \mapsto PX$  est un difféomorphisme, pour tout  $P$  inversible et, par conséquent,  $H$  est une sous-variété de  $G$  de dimension égale à  $\dim \mathfrak{h}$ .  $\square$

On va maintenant montrer qu'il existe des sous-groupes ouverts d'un groupe de Lie qui ne sont pas des sous-variétés.

**7.3.9 Exemple:** Soit  $G$  le tore "géométrique"  $T^2$ , qu'on a défini au chapitre 6 comme la surface paramétrée par

$$f(t, s) = ((r + R \cos(2\pi t)) \cos(2\pi s), (r + R \cos(2\pi t)) \sin(2\pi s), R \sin(2\pi t)), \quad t, s \in \mathbb{R}$$

et qu'on a identifié au groupe abélien  $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z} = [0, 1[ \times [0, 1[ \text{ mod}(1, 1)$ , avec loi de groupe additive

$$(t, s) \cdot (t', s') = (t + t', s + s').$$

Pour tout  $w \in \mathbb{R}$ , on considère le sous-ensemble

$$H_w = \{ (t, wt), t \in \mathbb{R} \}$$

de  $T^2$ , qui est clairement un sous-groupe, parce-que

$$(t, wt) \cdot (t', wt') = (t + t', wt + wt') = (t + t', w(t + t')).$$

Le sous-groupe  $H_w$  peut être visualisé comme une courbe tracée sur le tore géométrique, ou bien comme une courbe "périodique" tracée sur le carré  $[0, 1[ \times [0, 1[$ , d'équation  $s = wt$ , formée d'une série de segments transversaux.

La forme de cette courbe dépend de  $w$ , et on a la caractérisation suivante.

**7.3.10 Proposition:** 1. *Le sous-groupe  $H_w$  est dense dans le tore si et seulement si  $w$  est irrationnel.*  
2. *Si  $w$  est rationnel, alors  $H_w$  est un sous-groupe fermé.*

**Preuve:** Pour montrer la première assertion, on rappelle un théorème et une proposition montrés dans la partie I :

**7.3.11 Théorème:** *Si  $H$  est un sous-groupe additif de  $\mathbb{R}$  non réduit à  $\{0\}$ , alors soit  $H$  est dense dans  $\mathbb{R}$ , soit il existe  $a \in \mathbb{R}$  tel que  $H = a\mathbb{Z}$ .*

**7.3.12 Proposition:** *Soit  $w$  un irrationnel. Alors l'ensemble  $G = \{nw + p, n, p \in \mathbb{Z}\}$  est dense dans  $\mathbb{R}$ .*

Donc  $G/Z$  est dense dans  $[0, 1[$ . Lien avec le problème :  $(t, wt)$  est dense dans  $[0, 1[ \times [0, 1[$  si et seulement si  $(n, wn)$  est dense dans  $[0, 1[$ , c'est à dire  $w$  est irrationnel.

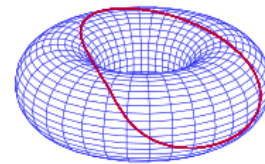
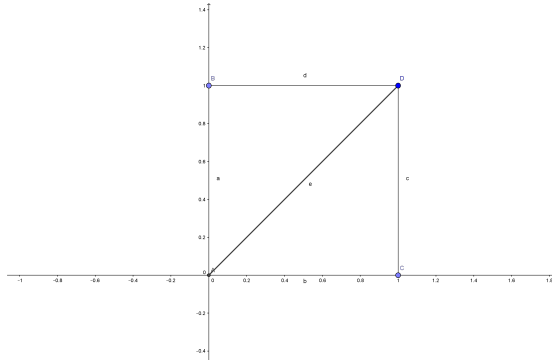
Pour montrer la deuxième assertion, il suffit de montrer que la courbe  $\gamma(t) = (t, wt)$  est périodique (donc fermée) dans  $[0, 1[ \times [0, 1[ \text{mod}(1, 1)$ .

En effet, supposons qu'il existe deux entiers  $p, q$  tels que  $w = \frac{p}{q}$ . En  $u = 0$ , on a  $\gamma(0) = (0, w0) = (0, 0)$ . En  $u = q$ , on a  $\gamma(q) = (q, wq) = (q, p) \equiv (0, 0) \pmod{1, 1}$  car on est dans le tore carré "périodique". Donc  $\gamma$  est bien une courbe fermée.

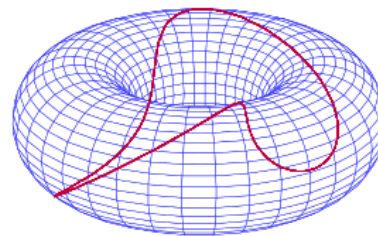
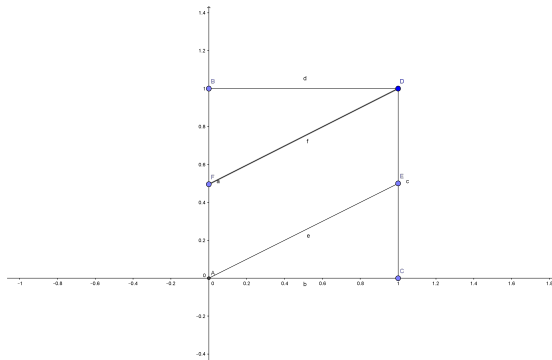
En particulier, si  $\frac{p}{q}$  est irréductible, la courbe  $\gamma$  est  $q$ -périodique (on retombe sur  $(0, 0)$  pour tout multiple de  $q$ ). □

Voici des illustrations des sous-groupes  $H_w$ , pour un tore tel que  $r = 2, R = 1$  :

Cas  $w = 1$  :

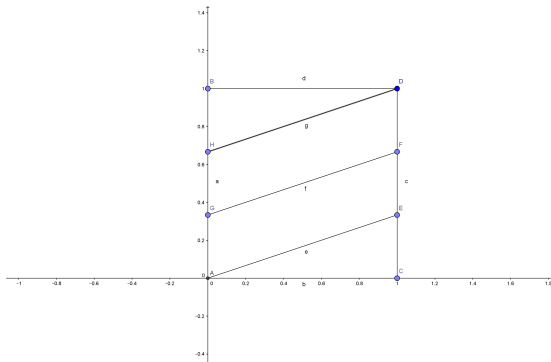


Cas  $w = \frac{1}{2}$  :

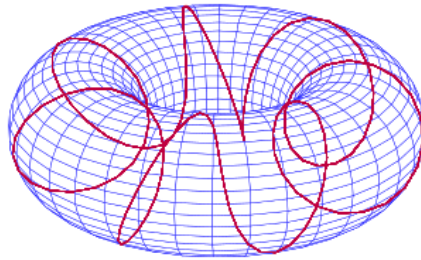
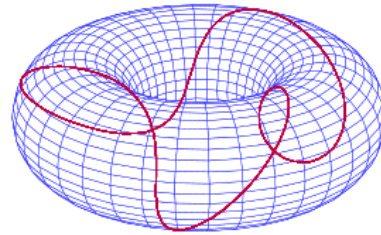


Cas  $w = \frac{1}{3}$  :

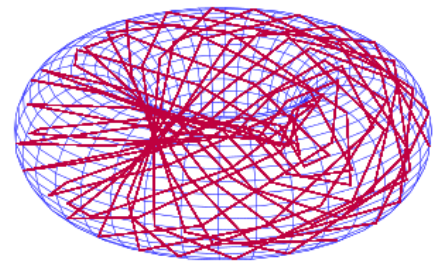
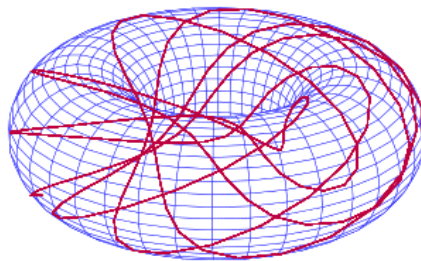




Cas  $w = \frac{1}{7}$  :



Cas  $w = \frac{\sqrt{2}}{2}$  pour  $t$  variant sur  $[0, 50]$ , puis  $[0, 150]$  :



**7.3.13 Proposition:** *Le sous groupe  $H_w$  du tore  $T^2$  est une sous-variété si  $w$  est rationnel, et ce ne l'est pas si  $w$  est irrationnel.*

**Preuve:** On regarde  $H_w$  comme image de la courbe paramétrée par  $\gamma(t) = (t, wt)$ , avec  $t \in \mathbb{R}$ . Alors  $H_w$  est une sous-variété si  $\gamma$  est un plongement, c'est-à-dire un homéomorphisme sur son image. La paramétrisation  $\gamma$  est clairement continue, injective, et elle est une immersion (car comme courbe elle est régulière partout). Il reste à vérifier qu'elle soit une application ouverte (ou fermée).

Si  $w$  est rationnel, on a montré que  $H_w$  est fermé, et également l'image  $\gamma(I) \subset H_w$  est fermée pour tout intervalle fermé  $I \subset \mathbb{R}$ . Donc  $\gamma$  est bien un plongement, et  $H_w$  est une sous-variété de  $T^2$ .

Si  $w$  est irrationnel, on a montré que  $H_w$  est dense dans  $T^2$  : dans la topologie relative induite par celle de  $T^2$  sur le sous-ensemble  $H_w$ , cela signifie que  $H_w$  est ouvert. Or,  $\mathbb{R}$  est fermé dans lui-même, et  $\gamma(\mathbb{R}) = H_w$  est ouvert : cela montre que  $\gamma$  n'est pas un plongement, et par conséquent  $H_w$  n'est pas une sous-variété de  $T^2$ .  $\square$

Revenons sur un corollaire intéressant du théorème qu'on a admis :

**7.3.14 Théorème:** (admis) Soit  $G$  et  $H$  2 groupes de Lie. Si  $H$  est un sous-groupe connexe de  $G$ ,  $\exp(\mathfrak{g})$  engendre  $G^\circ$ , la composante connexe de  $G$  contenant l'élément neutre.

**7.3.15 Corollaire:**  $\exp(\mathfrak{sl}_n(\mathbb{K})), \exp(\mathfrak{o}_n(\mathbb{R})), \exp(\mathfrak{so}_n(\mathbb{R})), \exp(\mathfrak{u}_n(\mathbb{C}))$  et  $\exp(\mathfrak{su}_n(\mathbb{C}))$  engendrent les différents groupes  $SL_n(\mathbb{K}), SO_n(\mathbb{R}), SO_n(\mathbb{R}), U_n(\mathbb{C})$  et  $SU_n(\mathbb{C})$  (respectivement). On verra en fait que ce résultat est réellement utile uniquement pour  $SL_n(\mathbb{K})$ . On verra également pourquoi  $\exp(\mathfrak{o}_n(\mathbb{R}))$  ne peut engendrer que  $SO_n(\mathbb{R})$ .

Le théorème ci-dessus permet de justifier que l'on s'intéresse ici aux groupes de matrices plus en détail, puisque ce sont des groupes de Lie.

## 7.4 Formule de Baker-Campbell-Hausdorff \*

Soit  $G$  un groupe de Lie. La formule de Baker Campbell Hausdorff permet de donner, pour  $X, Y \in \mathfrak{g}$  une expression de  $Z = \log(\exp(X)\exp(Y))$ , à partir du moment où  $X$  et  $Y$  sont définis dans un voisinage de 0 de  $\mathfrak{g}$  (car le développement série de Taylor de l'exponentielle est valable dans ce cas).

### Lemme 1

**7.4.1 Lemme:** On note :  $Ad_A(X) = AXA^{-1}$  et  $ad_A(X) = AX - XA$ . On a :  $Ad_{e^X}(\cdot) = e^{ad_X}(\cdot)$ .

**Preuve:**

Soit

$$f(s)Y = e^{sX}Ye^{-sX} : \frac{d}{ds}(f(s)Y) = \frac{d}{ds}(e^{sX}Ye^{-sX}) = Xe^{sX}Ye^{-sX} - e^{sX}YXe^{-sX}$$

=

$$X(e^{sX}Ye^{-sX}) - (e^{sX}Ye^{-sX})X$$

(rappel :  $X$  et  $e^{kX}$  commutent car  $e^{kX}$  est un polynôme en  $X$ )  $= ad_X(e^{sX}Ye^{-sX})$ .

(On remarque également que  $Xe^{sX}Ye^{-sX} - e^{sX}YXe^{-sX} = e^{sX}(XY - YX)e^{-sX} = e^{sX}ad_X(Y)e^{-sX}$  : second moyen de l'écrire mais ce n'est pas celui qui nous intéresse). Donc  $f'(s) = ad_X(f(s))(\cdot)$  et  $f(0) = 1$ . La solution de cette équation différentielle est  $f(s) = e^{ad_X}(\cdot)$ .  $\square$

### Lemme 2

**7.4.2 Lemme:**  $\frac{d}{dt}e^{X(t)} = e^{X(t)} \frac{1 - e^{-ad_X}}{ad_X} \frac{dX}{dt}$

**Preuve:** On pose :  $\Gamma(s, t) = e^{-sX(t)} \frac{\partial}{\partial t} e^{sX(t)}$ .

$$\begin{aligned} \frac{\partial \Gamma(s, t)}{\partial s} &= e^{-sX(t)} (-X(t)) \frac{\partial}{\partial t} e^{sX(t)} + e^{-sX(t)} \frac{\partial}{\partial t} (X(t) e^{sX(t)}) \\ &= e^{-sX(t)} (-X(t)) e^{sX(t)} + e^{-sX(t)} \frac{dX(t)}{dt} + e^{sX(t)} (X(t)) \frac{\partial}{\partial t} e^{-sX(t)} = e^{-sX(t)} \frac{dX(t)}{dt} e^{sX(t)} \end{aligned}$$

(On utilise ici la dérivée d'un produit  $(uv)' = u'v + uv'$ ). Donc :

$$\frac{\partial \Gamma(s, t)}{\partial s} = Ad_{e^{-sX}}(X') = e^{-ad_s(X)}(X')$$

, d'après le lemme 1 (on enlève les  $t$  pour alléger la notation). Puis :

$$\begin{aligned}\Gamma(1, t) &= e^{-X(t)} \frac{\partial}{\partial t} e^{X(t)} = \int_0^1 \frac{\partial \Gamma(s, t)}{\partial s} ds = \int_0^1 e^{-ad_s(X)} (X') ds \\ &= \int_0^1 \sum_{k=0}^{\infty} \frac{(-1)^k s^k}{k!} (ad_X)^k \frac{dX(t)}{dt} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} (ad_X)^k \frac{dX(t)}{dt} = \frac{1 - e^{-ad_X}}{ad_X} \frac{dX(t)}{dt}\end{aligned}$$

(voir le lien entre cette fonction et la fonction  $F$  définie précédemment. Donc :  $e^{-X(t)} \frac{\partial}{\partial t} e^{X(t)} = \frac{1 - e^{-ad_X}}{ad_X} \frac{dX(t)}{dt}$   
 $\Rightarrow \frac{\partial}{\partial t} e^{X(t)} = \frac{d}{dt} e^{X(t)} = e^X \frac{1 - e^{-ad_X}}{ad_X} \frac{dX(t)}{dt}$ . Le lemme est démontré.  $\square$

## Formule intégrale

### 7.4.3 Proposition:

$$\log(e^X e^Y) = X + \int_0^1 \frac{d}{dt} \psi(e^{ad_X} e^{tad_Y})(Y)$$

où :  $\varphi(z) = \frac{e^z - 1}{z}$  et  $\psi(z) = \frac{z \ln(z)}{z-1}$ ,  $\varphi$  et  $\psi$  étant liées par la relation :

$$\begin{aligned}\varphi(-\ln(z)) &= \frac{e^{-\ln(z)} - 1}{-\ln(z)} = \frac{\frac{1}{z} - 1}{-\ln(z)} = \\ &= \frac{1 - z}{-z \ln(z)} = \frac{z - 1}{z \ln(z)} = \frac{1}{\psi(z)}\end{aligned}$$

**Preuve:** Soit  $Z(t) = \log(e^X e^{tY})$  :

$$e^{Z(t)} = e^X e^{tY} \Rightarrow e^{-Z(t)} = e^{-tY} e^{-X}. \forall W, e^{Z(t)} W e^{-Z(t)} = e^X e^{tY} W e^{-tY} e^{-X}$$

par le lemme 1 :

$$e^{ad_Z} (W) = e^{Z(t)} W e^{-Z(t)} = e^X e^{tY} W e^{-tY} e^{-X} = e^{ad_X} e^{ad_{tY}} (W)$$

Ainsi :

$$\begin{aligned}e^{ad_Z} = e^{ad_X} e^{ad_{tY}} &\Rightarrow ad_Z = \log(e^{ad_X} e^{ad_{tY}}) \Rightarrow e^{-Z(t)(t)} \frac{\partial}{\partial t} e^{Z(t)} \\ &= \varphi(-ad_{Z(t)}) \frac{\partial Z(t)}{\partial t} = Y\end{aligned}$$

, la première égalité se déduit du lemme 2 et la dernière du fait que :

$$e^{-Z(t)} = e^{-tY} e^{-X} \text{ et } \frac{de^{Z(t)}}{dt} = \frac{d}{dt} (e^X e^{tY}) = e^X Y e^{tY} :$$

$$e^{-Z(t)} \frac{de^{Z(t)}}{dt} = e^{-tY} e^{-X} e^X Y e^{tY} = e^{-tY} Y e^{tY} = Y e^{-tY} e^{tY} = Y$$

Finalement :

$$\varphi(-ad_{Z(t)}) \frac{\partial Z(t)}{\partial t} = \varphi(-\log(e^{ad_X} e^{ad_{tY}})) \frac{\partial Z(t)}{\partial t} = Y \Rightarrow \frac{dZ(t)}{dt} = \psi(e^{ad_X} e^{ad_{tY}})(Y)$$

(en multipliant par  $\psi(e^{ad_X} e^{ad_{tY}})$  des 2 côtés, on fait disparaître  $\varphi$  d'un côté et apparaître  $\psi$  de l'autre).  
 En intégrant :  $Z(0) = \log(e^X) = X$  et  $Z(1) = \log(e^X e^Y)$ . Donc

$$Z(x) = Z(0) + \int_0^x \psi(e^{ad_X} e^{ad_{tY}})(Y) dt \Rightarrow Z(1) = X + \int_0^1 \psi(e^{ad_X} e^{ad_{tY}})(Y) dt = \log(e^X e^Y)$$

$\square$

### Formule de Dynkin

Il s'agit d'une formule explicite de  $\log(e^X e^Y)$ . accrochez vous, car c'est une partie très calculatoire, certainement la plus difficile de ce TIPE.

On pose cette fois :  $Z(t) = \log(e^{tX} e^{tY}) \Rightarrow e^{Z(t)} = e^{tX} e^{tY}$  :

$$\begin{aligned} e^{-Z(t)} \frac{de^{Z(t)}}{dt} &= e^{-tY} e^{-tX} (X e^{tX} e^{tY} + e^{tX} Y e^{tY}) = e^{-tY} e^{-tX} X e^{tX} e^{tY} + e^{-tY} e^{-tX} e^{tX} Y e^{tY} \\ &= e^{-tY} X e^{tY} + Y = Ad_{e^{-tY}}(X) + Y = e^{-ad_{tY}}(X) + Y \end{aligned}$$

Or,

$$\begin{aligned} e^{-Z(t)} \frac{de^{Z(t)}}{dt} &= \frac{1 - e^{-ad_Z}}{ad_Z} \frac{dZ}{dt} = e^{-ad_{tY}}(X) + Y \Rightarrow Z'(t) = \frac{ad_Z}{1 - e^{-ad_Z}} (Y + e^{-tad_Y}(X)) = \\ &= \frac{ad_Z}{e^{ad_Z} - 1} (X + e^{tad_X}(Y)) \end{aligned}$$

(on peut le vérifier, par symétrie). Or :

$$ad_Z = \log(\exp(ad_Z)) = \text{Log}(\exp(ad_Z) - 1) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (\exp(ad_Z) - 1)^n, \|ad_Z\| < \log(2)$$

Donc :

$$Z'(t) = \frac{ad_Z}{e^{ad_Z} - 1} (X + e^{tad_X}(Y)) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (\exp(ad_Z) - 1)^{k+1} (X + e^{tad_X}(Y))$$

En utilisant :  $e^{ad_Z} = e^{ad_{tX}} e^{ad_{tY}}$  et en développant :

$$Z'(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} [(e^{ad_{tX}} e^{ad_{tY}} - 1)^k X + (e^{ad_{tX}} e^{ad_{tY}} - 1)^k e^{ad_{tX}Y}] = \log(e^{ad_{tX}} e^{ad_{tY}} - 1) X + \log(e^{ad_{tX}} e^{ad_{tY}} - 1) e^{ad_{tX}Y}$$

**7.4.4 Lemme:** Pour des réels, dans un corps non commutatif,

$$\log(\exp(x)\exp(y)) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{r_i+s_i>0, 1 \leq i \leq k} \frac{x^{r_1} y^{s_1} x^{r_2} y^{s_2} \dots x^{r_k} y^{s_k}}{r_1! s_1! \dots r_k! s_k!}$$

(à condition que  $\|\exp(x)\exp(y) - 1\| < 1$ , question de développement limité et de convergence).

**Preuve:**

$$\begin{aligned} \log(\exp(x)\exp(y)) &= \log(1 + (e^x e^y - 1)) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (e^x e^y - 1)^k = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left[ \left( \sum_{k=0}^{\infty} \frac{x^k}{k!} \right) \left( \sum_{k=0}^{\infty} \frac{y^k}{k!} \right) - 1 \right]^k \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left( \sum_{i+j \geq 1} \frac{x^i y^j}{i! j!} \right) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{r_i+s_i>0, 1 \leq i \leq k} \frac{x^{r_1} y^{s_1} x^{r_2} y^{s_2} \dots x^{r_k} y^{s_k}}{r_1! s_1! \dots r_k! s_k!} \end{aligned}$$

Revenons à  $Z'(t)$  : à la place de  $e^X e^Y - 1$ , on met  $e^{ad_{tX}} e^{ad_{tY}} - 1$ . On a donc :

$$\begin{aligned} Z'(t) &= \log(e^{ad_{tX}} e^{ad_{tY}} - 1) X + \log(e^{ad_{tX}} e^{ad_{tY}} - 1) e^{ad_{tX}Y} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s \in S_k, i_{k+1} \geq 0} t^{i_1+j_1+\dots+i_k+j_k} \frac{ad_X^{i_1} ad_Y^{j_1} \dots ad_X^{i_k} ad_Y^{j_k}}{i_1! j_1! \dots i_k! j_k!} X + \\ &= t^{i_1+j_1+\dots+i_k+j_k+i_{k+1}} \frac{ad_X^{i_1} ad_Y^{j_1} \dots ad_X^{i_k} ad_Y^{j_k}}{i_1! j_1! \dots i_k! j_k! i_{k+1}!} X^{i_{k+1}} Y \end{aligned}$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s \in S_k, i_{k+1} \geq 0} t^{i_1+j_1+\dots+i_k+j_k} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}X]}{i_1!j_1!\dots i_k!j_k!} +$$

$$t^{i_1+j_1+\dots+i_k+j_k+i_{k+1}} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}X^{i_{k+1}}Y]}{i_1!j_1!\dots i_k!j_k!i_{k+1}!} (i_r, j_r \geq 0, i_r + j_r > 0, 1 \leq r \leq k)$$

(les crochets  $[\ ]$  permettent de simplifier la notation). On a donc :

$$Z(1) = \int_0^1 Z'(t)dt = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s \in S_k, i_{k+1} \geq 0} \frac{1}{i_1 + j_1 + \dots + i_k + j_k + 1} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}X]}{i_1!j_1!\dots i_k!j_k!} +$$

$$\frac{1}{i_1 + j_1 + \dots + i_k + j_k + i_{k+1} + 1} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}X^{i_{k+1}}Y]}{i_1!j_1!\dots i_k!j_k!i_{k+1}!} (i_r, j_r \geq 0, i_r + j_r > 0, 1 \leq r \leq k)$$

On peut écrire cette somme sous la forme :

$$Z(1) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s \in S_k, i_{k+1} \geq 0} \frac{1}{i_1 + j_1 + \dots + i_k + j_k + (i_{k+1} = 1) + (j_{k+1} = 0)} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}X^{(i_{k+1}=1)}Y^{(j_{k+1}=0)}]}{i_1!j_1!\dots i_k!j_k!(i_{k+1}=1)!(j_{k+1}=0)!} +$$

$$\frac{1}{i_1 + j_1 + \dots + i_k + j_k + i_{k+1} + (j_{k+1} = 1)} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}X^{i_{k+1}}Y^{(j_{k+1}=1)}]}{i_1!j_1!\dots i_k!j_k!i_{k+1}!(j_{k+1}=1)!} (i_r, j_r \geq 0, i_r + j_r > 0, 1 \leq r \leq k)$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \sum_{s \in S_{k+1}} \frac{1}{\sum_{r=1}^{k+1} i_r + j_r} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}X^{i_{k+1}}Y^{j_{k+1}}]}{i_1!j_1!\dots i_k!j_k!i_{k+1}!j_{k+1}!}$$

$$(i_r, j_r, i_{k+1} \geq 0, i_r + j_r > 0, 1 \leq r \leq k+1)$$

En effet, on a  $[T, T] = 0 \forall T$ , donc  $j_{k+1} = 0$  ou  $1$ . Si  $j_{k+1} = 0$ ,  $i_{k+1} = 1$  et si  $i_{k+1} = 0$ ,  $j_{k+1} = 1$ . Finalement : formule de Dynkin :

$$= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{s \in S_k} \frac{1}{\sum_{r=1}^k i_r + j_r} \frac{[X^{i_1}Y^{j_1}\dots X^{i_k}Y^{j_k}]}{i_1!j_1!\dots i_k!j_k!}$$

$$(i_r, j_r \geq 0, i_r + j_r > 0, 1 \leq r \leq k)$$

□

## 7.5 3ème théorème de Lie\*

### Problème

L'enjeu de ce chapitre est de répondre à cette question : à toute algèbre de Lie  $\mathfrak{g}$ , peut-on associer un groupe de Lie? En effet, on a prouvé dans le chapitre précédent que tout groupe de Lie possède une algèbre de Lie, déterminée uniquement par l'espace tangent du groupe  $G$  en l'identité. Le problème est l'unicité.

Notion de simple connexité : il s'agit d'un raffinement de la connexité : je ne vais pas en parler beaucoup, mais je renvoie à la toute fin, dans la conclusion.

Voici comment les théorèmes de Lie résolvent le problème : (ils s'énoncent ici pour des algèbres de Lie de dimension finie).

**7.5.1 Théorème:** Soit  $G$  un groupe de Lie d'algèbre de Lie  $\mathfrak{g}$  : il existe un unique groupe de Lie simplement connexe d'algèbre de Lie  $\mathfrak{g}$  (à isomorphisme près).

**7.5.2 Théorème:** Soit  $G$  et  $H$  2 groupes de Lie d'algèbre de Lie  $\mathfrak{g}$  et  $\mathfrak{h}$ ,  $G$  étant simplement connexe. Soit  $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$  un homomorphisme d'algèbre de Lie. Il existe un unique homomorphisme de groupes de Lie  $\Phi : G \rightarrow H$ , tel que  $\phi$  soit la représentation adjointe de  $\Phi$ , comme définie dans la partie sur l'algèbre de Lie, c'est à dire :  $\Phi^* = \phi$ .

**7.5.3 Théorème:** *Pour toute algèbre de Lie  $\mathfrak{g}$ , il existe un groupe de Lie  $G$  d'algèbre de Lie  $\mathfrak{g}$ .*

Un peu d'histoire : Sophus Lie a démontré ces théorèmes de façon locale, en terme de transformation de groupes. Les versions globales de ces 3 théorèmes ont été démontrés plus tard par Élie Cartan et Herman Weyl. Les premier et deuxième théorèmes de Lie ne sont pas les plus durs à démontrer, mais la preuve du dernier théorème a été faite par Cartan, en utilisant des outils algébriques (en 1936) (en utilisant le théorème d'Ado) puis par Van Est, en utilisant des outils géométriques (en 1953).

Je propose ici de prouver le troisième théorème de Lie localement.

On va, avant cela, donner quelques exemples.

**7.5.4 Exemple:** *Soit  $\mathbb{R}$  : on sait que c'est une algèbre de Lie. On vérifie que  $U_1(\mathbb{C})$  est un groupe de Lie d'algèbre de Lie  $i\mathbb{R}$ , donc isomorphe à  $\mathbb{R}$ . Pourtant, on peut également montrer que l'algèbre de Lie de  $(\mathbb{R}, +)$  est  $\mathbb{R}$  : en effet, soit  $\alpha(t) \in \mathbb{R}, \alpha(0) = 0$ . On pose par exemple  $\alpha(t) = ta, a \in \mathbb{R} : \alpha'(0) = a, \forall a \in \mathbb{R}$ . Pourtant,  $U_1(\mathbb{C}) \cong \mathbb{S}^1$  n'est pas homéomorphe à  $\mathbb{R}$ . Donc il y a plusieurs possibilités de groupes ayant même algèbre de Lie. Par contre, seul  $\mathbb{R}$  est simplement connexe.*

**7.5.5 Exemple:** *De la même façon, on peut montrer que l'algèbre de Lie du tore  $T^2 = \mathbb{S}^1 \times \mathbb{S}^1$  est isomorphe à l'algèbre de Lie groupe matriciel (additif)  $G = \left\{ \begin{pmatrix} 1 & x & 0 \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, x, y \in \mathbb{R} \right\}$*

**Démonstration**

Malheureusement, et malgré mon envie de prouver ce théorème, je n'ai pas pu le faire dans le temps imparti (pas le temps d'écrire). Toutefois, j'ai compris une grande majorité de la preuve que donne le livre Lie groups : J.J Duistermaat et J.A.C Kolk , au Chapitre 1, excepté quelques petits détails. C'est pourquoi je renvoi à ce livre pour la preuve..

## 8 Groupes matriciels

Le but de ce chapitre est de montrer l'utilité de ce qu'on a fait précédemment en l'appliquant aux groupes matriciels. On verra également un problème intéressant qui est la surjectivité de l'exponentielle d'une algèbre de Lie dans un groupe de Lie (pour des groupe matriciels).

### 8.1 Algèbre de Lie des principaux groupes de matrice

Soit  $G \subset GL_n(\mathbb{K})$  un groupe matriciel. On a les mêmes propriétés que pour un groupe de Lie quelconque, à la différence que l'exponentielle est toujours développable en série entière.

**8.1.1 Définition:** La *dimension* d'un de Lie matriciel  $G$  est  $dim G = dim T_I G$ . On utilise la notation  $\mathfrak{g} = T_I G$  pour cet espace vectoriel réel de  $M_n(\mathbb{K})$ .

**8.1.2 Lemme:**  $\frac{d(det(\alpha(t)))}{dt} \Big|_{t=0} = tr(\alpha'(0))$ .

**Preuve:** On a déjà vu que  $d(det)(Id)(X) = tr(X)$ , dans la partie sur les sous-variétés. Donc : par la règle de composition :

$$\frac{d(det(\alpha(t)))}{dt} \Big|_{t=0} = (d(det)(\alpha(0))) \circ \frac{d}{dt} \Big|_{t=0} (\alpha(t)) = tr(\alpha'(0))$$

Le résultat utilisé est le théorème de dérivation des fonctions composées dans le cas général. □

**8.1.3 Lemme:**  $det(exp(A)) = e^{tr(A)}$  (quelles que soit  $A \in M_n(\mathbb{C})$ )

**Preuve:** Un résultat bien connu utilise la décomposition de Dunford-Jordan. On donne ici une autre méthode. Soit  $\gamma(t) = \det(\exp(tA))$  :  $\gamma(0) = 1$ .

$$\gamma'(t) = \lim_{h \rightarrow 0} \frac{\det(\exp((t+h)A)) - \det(\exp(tA))}{h} = \lim_{h \rightarrow 0} \frac{\det(\exp(tA)\exp(hA)) - \det(\exp(tA))}{h}$$

=

$$\det(\exp(tA)) \lim_{h \rightarrow 0} \frac{\det(\exp(hA)) - I_n}{h} = \det(\exp(tA)) \operatorname{tr}(A) = \gamma(t) \operatorname{tr}(A)$$

car  $\gamma'(0) = \operatorname{tr}(A \exp(O_n)) = \operatorname{tr}(A)$  d'après le lemme précédent. Donc :  $\alpha(t) = \exp(\operatorname{tr}(A)t)$  est l'unique solution de l'équation différentielle  $\gamma'(t) = \gamma(t) \operatorname{tr}(A)$ ,  $\gamma(0) = 1$ . En  $t = 1$ ,  $\alpha(1) = \exp^{\operatorname{tr}(A)} = \det(\exp(A))$ .

□

**8.1.4 Proposition:** *L'espace tangent en l'identité (l'algèbre de Lie) de  $GL_n(\mathbb{K})$  est l'ensemble des matrices de  $\mathbb{K}$ .*

**Preuve:** Soit  $\alpha(t) = I_n + tA$  une courbe différentiable avec  $A \in GL_n(\mathbb{K})$ .  $\alpha(0) = I_n$  et  $\alpha'(t) = A = \alpha'(0)$ . On vérifie que, comme  $A$  n'admet qu'un nombre fini de valeurs propres non nulles (car elle est inversible), alors  $\alpha(t) \in GL_n(\mathbb{K})$  dans un voisinage de 0, car  $\det(I_n) = 1 \Rightarrow \exists \epsilon$  tel que  $\alpha(t)$  inversible  $\forall t \in ]-\epsilon; \epsilon[$ . De plus, si  $A \in M_n(\mathbb{K})$ ,  $\exp(tA) \in GL_n(\mathbb{K})$ .

L'espace tangent est donc l'ensemble des matrices de  $M_n(\mathbb{K})$ . □

**8.1.5 Proposition:** *L'espace algèbre de Lie de  $A \in T_n^+(\mathbb{K})$ ,  $\det(A) \neq 0$  est l'ensemble des matrices triangulaire  $T_n^+(\mathbb{K})$ .*

**Preuve:** Soit  $\alpha(t) = I_n + tA$ , avec  $A \in T_n^+(\mathbb{K})$ ,  $\det(A) \neq 0$   $\alpha(0) = I_n$ ,  $\Rightarrow \exists \epsilon$  tel que  $\alpha(t)$  inversible  $\forall t \in ]-\epsilon; \epsilon[$ .  $\alpha'(t) = A = \alpha'(0)$ . De plus, si  $A$  est triangulaire supérieure, il en est de même de  $\exp(A)$ .

L'espace tangent est donc l'ensemble des matrices de  $T_n^+(\mathbb{K})$ . □

**8.1.6 Proposition:** *L'algèbre de Lie de  $SL_n(\mathbb{K})$  est l'ensemble des matrices de  $\mathbb{K}$  de trace nulle.*

**Preuve:** Soit  $\alpha(t) \in SL_n(\mathbb{K})$  une courbe différentiable :

$$\det(\alpha(t)) = 1 \Rightarrow \operatorname{tr}(\alpha'(0)) = 0$$

L'espace tangent est donc l'ensemble des matrices de  $M_n(\mathbb{K})$  de trace nulle  $n$ . De plus, si  $A \in M_n(\mathbb{K})$  et que  $\operatorname{tr}(A) = 0$ ,  $\det(e^{tA}) = \exp(\operatorname{tr}(tA)) = \exp(nt \cdot \operatorname{tr}(A)) = \exp(0) = 1$ , donc  $\exp(tA) \in GL_n(\mathbb{K})$ .

L'espace tangent est donc l'ensemble des matrices de  $M_n(\mathbb{K})$  de trace nulle. □

**8.1.7 Proposition:** *L'algèbre de Lie de  $O_n(\mathbb{R})$  est l'ensemble des matrices antisymétriques de taille  $n$ .*

**Preuve:** Soit  $\alpha(t) \in O_n(\mathbb{R})$ , une courbe différentiable telle que  $\alpha(0) = I_n$  :

$$\alpha(t)^T \alpha(t) = I_n \Rightarrow \alpha'(t)^T \alpha(t) + \alpha(t)^T \alpha'(t) = O_n$$

En  $t = 0$ , on obtient

$$\alpha'(0)^T \alpha(0) + \alpha(0)^T \alpha'(0) = O_n \Rightarrow \alpha'(0)^T = -\alpha'(0)$$

De plus, si  $A \in ASym_n(\mathbb{R})$ ,  $e^{tA} \in O_n(\mathbb{R})$ . L'espace tangent est donc l'ensemble des matrices antisymétriques de taille  $n$ . On vérifie assez simplement que  $O_n(\mathbb{R})$  et  $SO_n(\mathbb{R})$  ont même espace tangent : en effet, on a vu avec  $SL_n(\mathbb{K})$  que la condition  $\det(A) = 1$  imposait  $\operatorname{tr}(B) = 0$  pour toute matrice  $B$  dans l'espace tangent. Or, une matrice antisymétrique réelle vérifie déjà cette condition. □

**8.1.8 Proposition:** *L'algèbre de Lie de  $U_n(\mathbb{C})$  est l'ensemble des matrices anti-hermitiennes de taille  $n$ .*

**Preuve:** Soit  $\alpha(t) \in O_n(\mathbb{R})$ , une courbe différentiable telle que  $\alpha(0) = I_n$  :

$$\alpha(t)^* \alpha(t) = I_n \Rightarrow \alpha'(t)^* \alpha(t) + \alpha(t)^* \alpha'(t) = O_n$$

En  $t = 0$ , on obtient

$$\alpha'(0)^* \alpha(0) + \alpha(0)^* \alpha'(0) = O_n \Rightarrow \alpha'(0)^* = -\alpha'(0)$$

De plus, si  $A \in AHerm_n(\mathbb{R})$ ,  $e^{tA} \in U_n(\mathbb{C})$ .

L'espace tangent est donc l'ensemble des matrices anti-hermitiennes de taille  $n$ .  $\square$

**8.1.9 Proposition:** *L'algèbre de Lie de  $SU_n(\mathbb{C})$  est l'ensemble des matrices anti-hermitiennes de taille  $n$  et de trace nulle.*

**Preuve:** On reprend le raisonnement précédent et on ajoute la condition  $tr(\alpha'(0)) = 0$ .  $\square$

Soit le tore à  $n$  trous  $T^n$ . C'est un groupe de Lie (car produit cartésien de groupes matriciels) qu'on peut représenter comme :

$$\begin{pmatrix} e^{2i\pi\theta_1} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & e^{2i\pi\theta_n} \end{pmatrix}$$

Avec  $\theta_1, \dots, \theta_n \in \mathbb{R}/\mathbb{Z}$ . C'est une matrice unitaire et diagonale.

**8.1.10 Proposition:** *L'algèbre de Lie du tore  $T^n$  est  $i\mathbb{R}^n$ .*

**Preuve:** Soit  $D(t) \in T^n$  tel que

$$D(t)D^*(t) = I_n, D(0) = D^*(0) = I_n$$

:

$$D'(t)D^* + D(t)D'^*(t) = O_n \Rightarrow D'^*(0) = -D(0)$$

$D$  est donc diagonale et anti-hermitienne, donc de la forme :

$$\begin{pmatrix} i\alpha_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & i\alpha_n \end{pmatrix}$$

Donc cet espace tangent est isomorphe à  $i\mathbb{R}^n$ , donc comme on s'y attendait,  $dim\mathfrak{T}^n = n$ .  $\square$

Dimension des groupes matriciels : par définition, cette dimension est exactement celle de l'algèbre de Lie (dimension dans le corps des réels et non des complexes), qui est aussi l'espace tangent. Comme on note  $\mathfrak{g}$  l'algèbre de Lie, on utilisera également cette écriture pour (lettres celtiques) les algèbres de Lie des groupes matriciel.

$$\mathfrak{gl}_n(\mathbb{R}) = M_n(\mathbb{R}) \Rightarrow dimGL_n(\mathbb{R}) = n^2$$

$$\mathfrak{sl}_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}), tr(A) = 0\} \Rightarrow dimSL_n(\mathbb{R}) = n^2 - 1$$

$$\mathfrak{gl}_n(\mathbb{C}) = M_n(\mathbb{C}) \Rightarrow dimGL_n(\mathbb{C}) = 2n^2$$

$$\mathfrak{sl}_n(\mathbb{C}) = \{A \in M_n(\mathbb{C}), tr(A) = 0\} \Rightarrow dimSL_n(\mathbb{C}) = 2n^2 - 2$$

$$\mathfrak{o}_n(\mathbb{R}) = ASym_n(\mathbb{R}) \Rightarrow dimO_n(\mathbb{R}) = \frac{n(n-1)}{2} = dimSO_n(\mathbb{R})$$

$$\mathfrak{u}_n(\mathbb{C}) = AHerm_n(\mathbb{C}) \Rightarrow dimU_n(\mathbb{C}) = n(n-1) + n = n^2$$

$$\mathfrak{su}_n(\mathbb{C}) = \{AHerm_n(\mathbb{C}), tr(A) = 0\} \Rightarrow dimSU_n(\mathbb{C}) = n^2 - 1$$

$$\mathfrak{t}_n^+(\mathbb{R}) = t_n^+(\mathbb{R}) \Rightarrow dimT_n^+(\mathbb{R}) = \frac{n(n+1)}{2}$$

$$\mathfrak{t}_n^+(\mathbb{C}) = t_n^+(\mathbb{C}) \Rightarrow dimT_n^+(\mathbb{C}) = n(n+1) = n^2 + n$$

**Remarques :**

-Le groupe engendré par  $exp(\mathfrak{gl}_n(\mathbb{R}))$  est  $GL_n(\mathbb{R})^+$ , c'est à dire l'ensemble des matrices inversibles de déterminant positif, car  $det(exp(A)) = e^{tr(A)} > 0$ . Cela vient du fait que  $GL_n(\mathbb{R})$  n'est pas connexe, mais  $GL_n(\mathbb{R})^+$  est l'une de ces 2 composantes connexes, et que ce 2 groupes ont même algèbre de Lie.



-Le groupe engendré par  $\exp(\mathfrak{o}_n(\mathbb{R}))$  est  $SO_n(\mathbb{R})$ , c'est à dire l'ensemble des matrices orthogonales de déterminant 1, car  $\det(\exp(A)) = e^{\text{tr}(A)} = 1$  si  $A$  est antisymétrique. Cela vient du fait que  $O_n(\mathbb{R})$  n'est pas connexe, mais  $SO_n(\mathbb{R})$  est l'une de ces 2 composantes connexes, et que ces 2 groupes ont même algèbre de Lie.

## 8.2 Surjectivité de l'exponentielle matricielle

Rappel : On a vu que pour un groupe  $G$ ,  $\mathfrak{g} = \{X, \exp(tX) \in G\}$ .

On a vu également que si  $G$  est un sous-groupe fermé et connexe de  $GL_n(\mathbb{K})$ ,  $\exp(\mathfrak{g})$  engendre  $G$ .

Enfin, on a vu que pour les groupes matriciels, l'exponentielle est le seul sous groupe à un paramètre, ce qui signifie que, pour vérifier que l'espace tangent en 0 de  $G$  qu'on a déterminé est correct, il faut vérifier que l'exponentielle d'un élément de cet espace est bien dans  $G$ . Il semble alors naturel de se demander : est-ce que, pour tout élément  $Y$  de  $G$ , il existe un élément de  $X \in \mathfrak{g}$  tel que  $\exp(X) = Y$ . On voit que c'est un résultat plus précis que d'engendrer un groupe. Ce qu'on demande ici, c'est la surjectivité de l'exponentielle. Le théorème, les 2 remarques ci-dessous et l'exemple répondent de façon précise à cette question.

**8.2.1 Théorème:** *Si  $\exp_G(\mathfrak{g}) \rightarrow G$  est surjective, alors  $G$  est connexe (par arcs).*

**Preuve:** On va montrer la connexité par arcs. Soit  $X, Y \in G$  : Alors, il existe  $X', Y' \in \mathfrak{g}$  telles que  $X = \exp_G(X'), Y = \exp_G(Y')$ . De plus,  $\exp(tX'), \exp(tY') \in G$  car  $X', Y' \in \mathfrak{g}$ . Donc, soit  $\gamma(t) = Y \exp(-tY') \exp(tX')$  :  $\gamma(0) = Y$  et  $\gamma(1) = X$ . On a donc bien un chemin dans  $[0, 1]$  reliant  $X$  et  $Y$ .  $\square$

La réciproque est fautive !

Cependant, on aurait pu augmenter les hypothèses et prouver que si  $G$  est un groupe de Lie connexe et compact, alors  $\exp_G(\mathfrak{g}) \rightarrow G$  est surjective, mais je ne le ferai pas dans ce TIPE (utilise des notions de topologies).

**8.2.2 Exemple:** (et proposition) *Considérons l'exponentielle :  $\exp_{SL_2(\mathbb{R})} : \mathfrak{sl}_2(\mathbb{R}) \rightarrow SL_2(\mathbb{R})$ . Alors, pour  $\delta \geq 0$ , la matrice :  $B = \begin{pmatrix} -(2+\delta) & 0 \\ 0 & \frac{-1}{2+\delta} \end{pmatrix} \in SL_2(\mathbb{R})$  n'est pas dans l'image de  $\exp_{SL_2(\mathbb{R})}$ .*

**Preuve:** Soit

$$A \in \mathfrak{sl}_2(\mathbb{R}) = \{A \in M_2(\mathbb{R}) : \text{tr}(A) = 0\}.$$

Par le théorème de Cayley-Hamilton :

$$A^2 + \det(A)I_2 = 0$$

Si  $\det(A) = 0$ , la seule valeur propre possible de  $A$  est 0, et on a de plus  $\forall t \in \mathbb{R}$

$$A^2 = 0 \Rightarrow \exp(tA) = I_2 + tA$$

et  $\text{tr}(\exp(tA)) = 2$ . Or, on vérifie que  $\text{tr}(B) = -(2+\delta) - \frac{1}{2+\delta} < -2 < 2$ , donc on ne peut avoir  $\exp(tA) = B$

Si  $\det(A) \neq 0$ ,  $A$  a 2 valeurs propres distinctes, et selon le signe de  $\det(A)$ , on a comme valeurs propres :  $\pm i\sqrt{\det(A)}$ , si  $\det(A) > 0$  et  $\pm\sqrt{-\det(A)}$ , si  $\det(A) < 0$

Dans tous les cas, comme on a 2 valeurs propres distinctes, la matrice  $A$  est diagonalisable et il existe une matrice  $P \in GL_2(\mathbb{C})$  telle que  $P^{-1}AP = \begin{pmatrix} i\sqrt{\det(A)} & 0 \\ 0 & -i\sqrt{\det(A)} \end{pmatrix}$  si  $\det(A) > 0$  et

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} \sqrt{-\det(A)} & 0 \\ 0 & \sqrt{-\det(A)} \end{pmatrix} \text{ si } \det(A) < 0. \text{ Donc : } \forall t \in \mathbb{R} \ P^{-1}\exp(tA)P = \begin{pmatrix} e^{it\sqrt{\det(A)}} & 0 \\ 0 & e^{-it\sqrt{\det(A)}} \end{pmatrix} \\ &= \cos(t\sqrt{\det(A)})I_2 + i\sin(t\sqrt{\det(A)})J_2 \end{aligned}$$

(en utilisant la formule  $e^{ix} = \cos(x) + i\sin(x)$ ) si  $\det(A) > 0$  avec

$J_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . On remarque alors que  $i\sqrt{\det(A)}J_2 = P^{-1}AP$ . En réorganisant cela :  $A = i\sqrt{\det(A)}PJ_2P^{-1}$  et donc :

$$P^{-1}\exp(tA)P = \cos(t\sqrt{\det(A)})I_2 + \frac{\sin(t\sqrt{\det(A)})}{\sqrt{\det(A)}} \times i\sqrt{\det(A)}J_2$$

et en multipliant par  $P$  et  $P^{-1}$  :

$$\exp(tA) = \cos(t\sqrt{\det(A)})I_2 + \frac{\sin(t\sqrt{\det(A)})}{\sqrt{\det(A)}}A$$

si  $\det(A) < 0$  :

$$\begin{aligned} P^{-1}\exp(tA)P &= \begin{pmatrix} e^{t\sqrt{-\det(A)}} & 0 \\ 0 & e^{-t\sqrt{-\det(A)}} \end{pmatrix} \\ &= \operatorname{ch}(t\sqrt{-\det(A)})I_2 + \operatorname{sh}(t\sqrt{-\det(A)})J_2 \end{aligned}$$

(en utilisant la formule  $e^x = \operatorname{ch}(x) + \operatorname{sh}(x)$ ). On remarque alors que  $\sqrt{-\det(A)}J_2 = P^{-1}AP$ . En réorganisant cela :  $A = \sqrt{-\det(A)}PJ_2P^{-1}$  et donc :

$$P^{-1}\exp(tA)P = \operatorname{ch}(t\sqrt{-\det(A)})I_2 + \frac{\operatorname{sh}(t\sqrt{-\det(A)})}{\sqrt{-\det(A)}} \times \sqrt{-\det(A)}J_2$$

et en multipliant par  $P$  et  $P^{-1}$  :

$$\exp(tA) = \operatorname{ch}(t\sqrt{-\det(A)})I_2 + \frac{\operatorname{sh}(t\sqrt{-\det(A)})}{\sqrt{-\det(A)}}A$$

Dans les 2 cas, comme  $\operatorname{tr}(A) = 0$ ,  $\operatorname{tr}(\exp(tA)) = 2\cos(t\sqrt{\det(A)})$  si  $\det(A) > 0$  et  $2\operatorname{ch}(t\sqrt{-\det(A)})$  si  $\det(A) < 0$ . Dans tous les cas,  $\operatorname{tr}(\exp(tA)) > -2$  alors que  $\operatorname{tr}(A) < -2$ . Fin de la preuve.  $\square$

### Dunford-Jordan

Rappels sur Dunford Jordan : Soit  $A \in \mathbb{K}$  avec  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ .

Il existe un unique couple de matrices  $D$  et  $N$  telles que :

$$A = D + N$$

$D$  est diagonalisable

$N$  est nilpotente

$N$  et  $D$  commutent :  $DN = ND$

C'est ce qu'on appelle la décomposition de Dunford-Jordan.

On va maintenant exprimer cette décomposition pour l'exponentielle :

$$\exp(A) = \exp(D + N) = \exp(D)\exp(N) = \exp(D) + \exp(D)(\exp(N) - I).$$

On a :

$$\exp(N) - I = \sum_{k=1}^{+\infty} \frac{N^k}{k!} = N \sum_{k=1}^{+\infty} \frac{N^{k-1}}{k!} = NB$$

avec  $B = \exp(N) - I$ , un polynôme en  $N$ . Soit  $k$  l'indice de nilpotence de  $N$  :  $N^k = 0$ . Donc  $NB^k = N^k B^k = 0$  ( $N$  et  $B$  commutent). Puis :  $\exp(D) = \sum_{k=0}^{+\infty} \frac{D^k}{k!}$ .  $D$  est diagonalisable, on peut donc écrire  $D = PD'P^{-1}$  pour une certaine matrice  $P$  inversible et  $D'$  diagonale.  $\exp(D) = \sum_{k=0}^{+\infty} \frac{(PD'P^{-1})^k}{k!} = P\exp(D')P^{-1}$ . Pour finir :  $\exp(D') = \sum_{k=0}^{+\infty} \frac{D'^k}{k!}$

Soit  $D' = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix}$  Donc :  $\exp(D') = \begin{pmatrix} \sum_{k=0}^{+\infty} \frac{d_1^k}{k!} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sum_{k=0}^{+\infty} \frac{d_n^k}{k!} \end{pmatrix}$   
 $= \begin{pmatrix} e^{d_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{d_n} \end{pmatrix}$  Finalement,  $\exp(D')$  est diagonale et donc  $\exp(D)$  est diagonalisable. On a donc bien trouvé la décomposition de Dunford-Jordan de l'exponentielle.

### Matrices diagonales semblables

**8.2.3 Proposition:** 2 matrices diagonales sont semblables si et seulement si elles ont mêmes valeurs propres, avec même multiplicité.

**Preuve:** On s'intéresse à l'équation (en X) :  $D_1 = X D_2 X^{-1} \Leftrightarrow D_1 X = X D_2$  où  $D_1$  et  $D_2$  sont 2 matrices diagonales. On pose :

$$X = (x_{i,j})_{1 \leq i,j \leq n}, D_1 = (d_{1,i,j})_{1 \leq i,j \leq n}$$

et

$$D_2 = (d_{2,i,j})_{1 \leq i,j \leq n}. \text{ Soit } C = D_1 X \text{ et } E = X D_2. C_{i,j} = \sum_{k=1}^n d_{1,ik} x_{k,j} = d_{1,ii} x_{i,j}$$

car  $d_{1,ik} = 0$  si  $i \neq k$ . De même, on trouve  $e_{i,j} = x_{i,j} d_{2,jj}$ . L'équation  $C=E$  donne  $e_{i,j} = c^{ij} \forall i,j \Leftrightarrow x_{ij} d_{1,ii} = x_{ij} d_{2,jj}$ .

On a plusieurs cas de figures :  $d_{1,ii} = d_{2,jj} \Rightarrow x_{i,j}$  quelconque. Sinon,  $x_{i,j} = 0$ . On constate alors que pour que 2 matrices diagonales puissent être semblables, elles doivent avoir même valeurs propres (dans le cas contraire, une ligne de X est nulle et X n'est pas inversible). Dans ce cas, on déduit que les matrices X possibles sont les matrices semblables aux matrices de permutations si toutes les valeurs propres sont distinctes (facile à vérifier) et elles sont plus générales si les valeurs propres sont non distinctes (du fait que la matrice diagonale se "rapproche" de l'identité).  $\square$

On note :  $D_n(\mathbb{C})$  l'ensemble des matrices diagonalisables dans  $\mathbb{C}$  et  $D'_n(\mathbb{C})$  l'ensemble des matrices diagonalisables inversibles dans  $\mathbb{C}$ . Remarque : on est obligé d'imposer que l'ensemble image de l'exponentielle soit un ensemble où les matrices sont inversibles puisque l'on a vu précédemment que  $\forall A \in M_n(\mathbb{K}), \exp(A)$  est inversible d'inverse  $\exp(-A)$ .

### Surjectivité de $\exp : D_n(\mathbb{C}) \rightarrow D'_n(\mathbb{C})$

**8.2.4 Proposition:**  $\exp : D_n(\mathbb{C}) \rightarrow D'_n(\mathbb{C})$  est surjective.

**Preuve:** Soit  $M \in D_n(\mathbb{C})$  et  $K \in D'_n(\mathbb{C})$ . Donc :  $\exists P, Q$  telles que :  $M = P M' P^{-1}$  et  $K = Q K' Q^{-1}$  avec  $M', P'$  diagonales.  $\exp(M) = K \Leftrightarrow P \exp(M') P^{-1} = Q K' Q^{-1} \Leftrightarrow \exp(M') = (P^{-1} Q) K' (P^{-1} Q)^{-1}$ . Soit  $R = (P^{-1} Q) : \exp(M') = R K' R^{-1}$ . Donc :  $\exp(M')$  et  $K'$  sont 2 matrices diagonales semblables. Ce qui nous intéresse, pour prouver la surjectivité est de trouver une solution  $M_0$  de  $\exp(M) = K$ . On pose :  $P = Q$ , de sorte que  $R = I$  et donc :  $\exp(M') = K'$ . Soit  $(k_{11}, \dots, k_{nn})$  les coefficients de  $K'$  et  $(m_{11}, \dots, m_{nn})$  les coefficients de  $M'$ . L'équation précédente revient à :  $e^{m_{ii}} = k_{ii}$ . Cette équation a une solution dans  $\mathbb{C}$  (partie réelle unique et partie imaginaire modulo  $2\pi$ ) tant que  $k_{ii} \neq 0 \Leftrightarrow K$  inversible, ce qui est le cas.  $\square$

### Surjectivité de $\exp : M_n(\mathbb{C}) \rightarrow GL_n(\mathbb{C})$

**8.2.5 Proposition:**  $\exp : M_n(\mathbb{C}) \rightarrow GL_n(\mathbb{C})$  est surjective.

**Preuve** On va utiliser Dunford-Jordan : On cherche à résoudre  $\exp(A) = B$  avec  $A \in M_n(\mathbb{C})$  et

$B \in GL_n(\mathbb{C})$ . Soit  $A = D_1 + N_1$   $B = D_2 + N_2$ . Comme  $B$  est inversible,  $D_2$  l'est également : (je ne vais pas le prouver, c'est un résultat un peu plus précis de la décomposition de Dunford-Jordan qui stipule que les valeurs propres de la matrice  $B$  sont celles de la matrice  $D_2$ , donc :  $0$  est valeur propre de  $B \Leftrightarrow \det(B) = 0$  (ou  $\ker B \neq 0$ )  $\Leftrightarrow \det(D_2) = 0$ ). Donc, en utilisant Dunford Jordan pour  $\exp(A)$ , on obtient :  $\exp(D_1) + \exp(D_1)(\exp(N_1) - I_n) = D_2 + N_2$ . Comme cette décomposition (Dunford-Jordan) est unique, on obtient 2 équations :

$\exp(D_1) = D_2$  Or  $D_1$  et  $D_2$  étant diagonalisables avec  $D_2$  inversible, on a déjà montré que cette équation avait une solution. On peut donc trouver  $D_1$

(1)  $\exp(D_1)(\exp(N_1) - I_n) = N_2$ . On va donc chercher une solution  $N_1$  de cette équation :

(2)  $\exp(D_1)(\exp(N_1) - I_n) = N_2 \Leftrightarrow D_2(\exp(N_1) - I_n) = N_2$  (d'après l'équation (1)). Donc :  $\exp(N_1) = D_2^{-1}N_2 + I_n$ . On va maintenant montrer que  $D_2^{-1}N_2$  est nilpotente : on remarque tout d'abord que  $D_2^{-1}N_2 =$

$D_2^{-1}N_2D_2D_2^{-1}$  ( $D_2$  est inversible)  $= D_2^{-1}D_2N_2D_2^{-1}$  (car  $N_2$  et  $D_2$  commutent)  $= N_2D_2^{-1}$ . Donc  $N_2$  et  $D_2^{-1}$  commutent et comme  $N_2$  est nilpotente, on pose  $N = D_2^{-1}N_2$ , nilpotente également (on a montré précédemment, dans une preuve, que si une matrice  $C$  est le produit d'une matrice nilpotente et d'une autre qui commute avec cette matrice nilpotente, alors  $C$  est également nilpotente). On arrive à  $\exp(N_1) = N + I_n$ .

La dernière étape est de montrer qu'il existe une solution  $N_1$  de cette équation. On pose  $S = \log(I_n + N) = \sum_{k=1}^{n-1} \frac{(-1)^{k-1} N^k}{k}$  (rappel : une matrice nilpotente est semblable à une matrice triangulaire supérieure stricte, ce qui fait que si une matrice de taille  $n$  est nilpotente,  $N^n = 0$ ).

Soit

$$\begin{aligned} f(t) &= \sum_{k=1}^{n-1} \frac{(-1)^{(k-1)} N^k t^k}{k} \cdot f'(t) = \sum_{k=1}^{n-1} (-1)^{k-1} N^k t^{k-1} \Rightarrow (I_n + tN)f'(t) = f'(t) + tNf'(t) = \\ & \sum_{k=1}^{n-1} (-1)^{k-1} N^k t^{k-1} + \sum_{k=1}^{n-1} (-1)^{k-1} N^{k+1} t^k = \\ & \left( \sum_{k=1}^{n-1} (-1)^{k-1} N^k t^{k-1} \right) + \left( \sum_{k=1}^{n-2} (-1)^{k-1} N^{k+1} t^k \right) + t^{n-1} N^n (= 0) = \\ & \left( \sum_{k=2}^{n-1} (-1)^{k-1} N^k t^{k-1} \right) + \left( \sum_{k=1}^{n-2} (-1)^{k-1} N^{k+1} t^k + N \right) = \\ & \left( \sum_{k=1}^{n-2} (-1)^k N^{k+1} t^k \right) + \\ & \left( \sum_{k=1}^{n-2} (-1)^{k-1} N^k t^k \right) + N = \left( \sum_{k=1}^{n-2} (-1)^k (t^k N^{k+1} - t^k N^{k+1}) \right) + N = N \end{aligned}$$

Donc :  $(I_n + tN)f'(t)\exp(f(t)) = N\exp(f(t))$ . De plus :  $f(t)$  et  $f'(t)$  commutent car ce sont des polynômes en  $N$ . Pour finir : soit  $S(t) = \exp(f(t))$ .  $S'(t) = D'(t)\exp(f(t))$ . On obtient alors :  $(I_n + tN)S'(t) = NS'(t)$ . On redérive :  $S''(t) + NS'(t) + tNS''(t) = NS'(t)$  puis  $(I_n + N)S''(t) = O_n$ . Mais  $(I_n + N)$  est inversible car  $N = PTP^{-1}$  avec  $T$  triangulaire supérieure stricte donc  $(I_n + N) = P(I_n + T)P^{-1}$  et donc :  $\det(I_n + N) = \det(I_n + T) = 1$  car  $T$  est triangulaire supérieure stricte. On déduit enfin :  $S''(t) = 0_n \Rightarrow S(t) = S(0) + tS'(0)$ . Or,  $f(0) = O_n \Rightarrow S(0) = I_n$  et  $S'(0) = N \Rightarrow S(t) = I_n + tN$ . Puis  $S(1) = \exp(f(1)) = I_n + N = \exp(N_1)$ ,  $N_1$  étant la matrice cherchée au début de ce paragraphe.

Remarque : par la méthode qu'on vient d'utiliser, on vient en fait de prouver un résultat plus précis :  $f(1)$  est un polynôme en  $N$ , qui vérifie  $N = D_2^{-1}N_2$ , où les matrices  $N_2$  et  $D_2$  sont les matrices de la décomposition de Dunford Jordan telles que  $B = N_2 + D_2$  (on rappelle que le but était de résoudre  $\exp(A) = B$ ). Or, toujours d'après Dunford-Jordan,  $D_2$  et  $N_2$  sont des polynômes en  $B$ , il en est donc de même de  $N$ . Par conséquent, la solution  $N_1$  qu'on a trouvé est un polynôme en  $B$  (car c'est un polynôme

en  $N$ ). Enfin, la solution  $D_1$  trouvée est un polynôme en  $D_2$ , donc en  $B$  (preuve : car elles sont toutes les 2 diagonales dans une certaine base, par conséquent, il existe un polynôme complexe  $Q$  tel que  $P(D_2') = D_1' : D_1'$  et  $D_2'$  étant les matrices diagonales dans leur base commune de diagonalisation. Il suffit de prendre le polynôme de degré  $n$  ou moins vérifiant  $P(d_{1,i}) = d_{2,i} \forall i = 1, 2, \dots, n$  ( $d_{1,i}$  et  $d_{2,i}$  étant les valeurs propres de  $D_1'$  et  $D_2'$ ). Finalement  $A = D_1 + N_1 = R(B)$ ,  $R$  étant un polynôme en  $B$ .

**8.2.6 Corollaire:** *Pour toute matrice  $B \in GL_n(\mathbb{C})$ ,  $\exists P \in \mathbb{C}[X]$  tel que  $\exp(P(B)) = B$ .*

### Cas réel et applications

**8.2.7 Proposition:** *Soit  $A \in GL_n(\mathbb{R})$ . Il existe une matrice réelle  $M$  telle que  $\exp(M) = A$  si et seulement si il existe une matrice  $B \in GL_n(\mathbb{R})$  telle que  $A = B^2$ .*

**Preuve:** Si une telle matrice  $M$  existe, on pose  $B = \exp(\frac{M}{2})$ , de sorte que  $B^2 = A$ . Maintenant, si  $B$  existe : par le corollaire précédent, il existe  $Q \in \mathbb{C}[X]$  telle que  $B = \exp(Q(B))$ . Mais  $B$  est réelle, donc  $B = \overline{B} = \overline{\exp(Q(B))} = \exp(\overline{Q(B)}) = \exp(\overline{Q}(B))$ . Donc  $A = B^2 = B\overline{B} = \exp(Q(B))\exp(\overline{Q}(B)) = \exp(Q(B) + \overline{Q}(B)) = \exp((Q + \overline{Q})(B))$ . Or, par la définition du conjugué, on a  $Q + \overline{Q} \in \mathbb{R}[X]$ , et  $B$  étant réelle, on a bien le résultat voulu, avec comme point supplémentaire que  $M$  est un polynôme en  $B$ .  $\square$

Remarque : une condition nécessaire pour l'existence de  $M$  est donc que  $\det(A) = \det(B^2) = \det(B)^2 > 0$ . Mais elle n'est pas suffisante :

$\exp : M_n(\mathbb{R}) \rightarrow GL_n^+(\mathbb{R})$  n'est pas surjective. Prendre par exemple :  $C = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix}$  avec  $a \neq b, a, b \in \mathbb{R}_+^*$ , elle ne peut être exprimée sous la forme  $\exp(M)$ ,  $M \in M_n(\mathbb{R})$  car il est facile de montrer, en prenant une matrice  $D$  de  $M_2(\mathbb{R})$  avec 4 coefficients inconnus, qu'il n'est pas possible de résoudre  $D^2 = C$ .

On peut se demander si, néanmoins, elle engendre le groupe. Pour cela, il faudrait que  $GL_n^+(\mathbb{R})$  soit fermé (pour appliquer le dernier théorème juste après celui de Cartan). Néanmoins, on peut montrer que  $\exp : M_n(\mathbb{R})$  engendre  $GL_n^+(\mathbb{R})$ . Pour cela, on va tout simplement réutiliser un résultat qu'on connaît déjà (je renvoie à la partie sur les générateurs du groupe linéaire).

Les matrices  $A$  de  $GL_n^+(\mathbb{R})$  peuvent s'exprimer comme produit de transvections et d'une matrice de dilatation  $D_n(\alpha)$  (au milieu des transvections) de coefficient  $\alpha = \det(A) > 0$  (ce détail a de l'importance). Soit  $T_{ij}(\lambda)$  une transvection :  $\exp(\lambda E_{ij}) = I_n + \lambda E_{ij} + O_n = T_{ij}(\lambda)$  car  $E_{ij}^2 = O_n \Rightarrow E_{ij}^k = O_n \forall k \geq 2$ . Donc, les transvections sont dans l'image de  $\exp : M_n(\mathbb{R})$ . Enfin,  $\exp(\ln(\alpha)E_{nn}) = D_n(\alpha)$  (facile à vérifier), donc toutes les matrices permettant d'engendrer  $GL_n^+(\mathbb{R})$  sont dans l'image de  $\exp(M_n(\mathbb{R}))$ , par conséquent  $\exp : M_n(\mathbb{R})$  engendre  $GL_n^+(\mathbb{R})$ .

Une application de ce qu'on a fait jusqu'à maintenant :

**8.2.8 Proposition:** *Il n'existe pas de sous-groupes arbitrairement petits de  $GL_n(\mathbb{C})$ .*

**Preuve:** On va montrer pour cela qu'il existe un voisinage  $V$  de  $I_n$  tel que le seul sous-groupe contenu dans  $V$  soit  $\{I_n\}$ .

Soit  $V$  un voisinage de  $I_n$  dans  $GL_n(\mathbb{C})$  et  $U$  un voisinage de  $O_n$  dans  $M_n(\mathbb{C})$  tel que  $\exp$  réalise un difféomorphisme de  $U$  sur  $V$ . On pose  $U' = \frac{U}{2}$  et  $V' = \exp(U)$ .  $V'$  est ouvert, c'est un voisinage de  $I_n$ . Soit  $M \in V'$ . On va montrer qu'il existe  $k \in \mathbb{N}$  tel que  $M^k \notin V'$  : on peut écrire  $M = \exp(A)$  avec  $A \in U'$ . Il existe  $k \in \mathbb{N}$  tel que  $kA \in U \setminus U'$ .  $\exp(kA) = M^k \in V \setminus V'$ , donc  $M^k \notin V'$ , qui n'est donc pas un sous-groupe.  $\square$

**Surjectivité de  $\exp : AHerm_n(\mathbb{C}) \rightarrow U_n(\mathbb{C})$**

**8.2.9 Proposition:**  *$\exp : AHerm_n(\mathbb{C}) \rightarrow U_n(\mathbb{C})$  est surjective.*

**Preuve** Soit  $\exp(A) = B$  avec  $B$  unitaire : Si  $B = PDP^{-1}$ , on peut écrire :  $\exp(A) = PD'P^{-1}$  ( $P$  est unitaire) soit  $A = PD'P^{-1}$ . On a l'équation  $\exp(D') = D$ . Mais rappelons qu'on a montré (partie 2) qu'une matrice unitaire était diagonalisable dans une base unitaire, avec des valeurs propres **de module 1**. Par conséquent, la résolution de cette équation coefficient par coefficient donne :  $e^{d'_{kk}} = d'_{kk} = e^{i\theta_k}$ , donc

$d'_{kk} = i\theta_k$ . Donc  $D'$  est de la forme : 
$$\begin{pmatrix} i\theta_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & i\theta_n \end{pmatrix}$$
 On vérifie très clairement que c'est une matrice

anti-hermitienne. De plus,  $P$  est unitaire  $\Rightarrow P^{-1} = P^*$ . (Pour la signification de  $*$ , je vous renvoie à la partie 2). Donc :  $A^* = (PD'P^{-1})^* = {}^t(PD'P^{-1})^* = {}^t(PD'{}^tP) = {}^t(P) {}^tD' {}^tP$  (avec  ${}^tD' = D'$ )  $= P\overline{D'}P^* = -PD'P^*$  (car  $\overline{D'} = -D'$ )  $= -A$ . Donc  $A$  est anti-hermitienne, et s'écrit même sous forme  $iH$  avec  $H$  hermitienne (facile à vérifier).

**Surjectivité de  $\exp : ASym_n(\mathbb{C}) \rightarrow O_n(\mathbb{R})$**

**8.2.10 Proposition:**  $\exp : ASym_n(\mathbb{C}) \rightarrow O_n(\mathbb{R})$  est surjective.

**Preuve** Soit  $\exp(A) = B$  avec  $B$  orthogonale (**réelle**) : Si  $B = PDP^{-1}$ , on peut écrire :  $\exp(A) = PD'P^{-1}$  ( $P$  est orthogonale) soit  $A = PD'P^{-1}$ . On a l'équation  $\exp(D') = D$ . Mais rappelons qu'on a montré (partie 2) qu'une matrice orthogonale était diagonalisable par blocs dans une base orthogonale. Donc : si les blocs sont de taille 1, les 2 équations qu'on peut avoir sont  $e^{d_{ii}} = 1$  ou  $e^{d_{ii}} = -1$ , donc  $d_{ii} = 0$  ou  $d_{ii} = i\pi$  (modulo  $2\pi$ ). Si les blocs sont de taille 2, la matrice  $2*2$  est une matrice de rotation :  $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$  On s'intéresse alors à l'équation  $\exp(C) = A$ . Pour cela : on calcule le polynôme caractéristique de  $A$  :

$$\det\left(\begin{pmatrix} X - \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & X - \cos(\theta) \end{pmatrix}\right)$$

$= (X - \cos(\theta))^2 + \sin(\theta)^2 = X^2 - 2X\cos(\theta) + 1$ .  $\Delta = 4\cos^2(\theta) - 4 = -4\sin^2(\theta) = (\pm 2i\sin(\theta))^2$ . Donc les 2 solutions possibles sont  $x_1 = \frac{1}{2}(2\cos(\theta) + 2i\sin(\theta)) = e^{i\theta}$  et  $x_2 = \frac{1}{2}(2\cos(\theta) - 2i\sin(\theta)) = e^{-i\theta}$ .

Calcul des vecteurs propres :  $A - e^{i\theta}I_2 = \begin{pmatrix} -i\sin(\theta) & -\sin(\theta) \\ \sin(\theta) & -i\sin(\theta) \end{pmatrix}$

Le calcul du ker donne le système :  $-i\sin(\theta)x_1 - \sin(\theta)x_2 \Rightarrow x_2 = -ix_1$  Le premier vecteur propre est  $v_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ , et c'est une base de  $\ker(A - e^{i\theta}I_2)$ .

De même, on trouve le second vecteur propre : (base de  $\ker(A - e^{-i\theta}I_2)$ )  $v_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$

La matrice de passage est donc :  $P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$

Soit  $D = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & -e^{-i\theta} \end{pmatrix}$  On vérifie que  $A = PDP^{-1}$ .

Enfin, on pose  $D' = \begin{pmatrix} i\theta & 0 \\ 0 & -i\theta \end{pmatrix}$

De sorte que  $C = PD'P^{-1}$  vérifie  $\exp(C) = A$  On vérifie également par un calcul très simple que  $C = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}$

Donc, pour finir, si on revient à l'équation tout en haut :  $\exp(A) = B$  avec  $B$  orthogonale, on déduit que  $A = PD'P$  sera diagonalisable par blocs dans une base orthogonale avec des blocs de taille au plus 2, de la forme 0,  $i\pi$ , ou de la forme  $\begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}$

On vérifie très simplement, que la matrice diagonalisable par blocs est anti-symétrique de par sa forme, et donc que  ${}^tA = {}^t(PD'P) = {}^tP {}^tD' P = -PD'P = -A$ . Donc  $A$  est antisymétrique.

**Surjectivité de  $\exp : ASym_n(\mathbb{R}) \longrightarrow SO_n(\mathbb{R})$**

**8.2.11 Proposition:**  $\exp : ASym_n(\mathbb{R}) \longrightarrow SO_n(\mathbb{R})$  est surjective.

**Preuve:** La démonstration de la surjectivité dans le cas complexe nous donne la forme de cette matrice :  $Q = RDR^{-1}$  où  $R$  est orthogonale réelle et  $D$  anti-symétrique complexe, diagonale avec des blocs  $2 \times 2$  réels et des coefficients diagonaux égaux à 0 ou  $i\pi$ , donc qui vérifie :  $\exp(D) \in M_n(\mathbb{R})$ . Mais dans ce cas, on peut trouver un moyen de supprimer les coefficients diagonaux complexes. En effet, comme  $P \in SO_n(\mathbb{R})$ , il y a un nombre pair de (-1) sur la diagonale. De ce fait, la matrice  $A =$

$$\begin{pmatrix} 0 & \pi \\ -\pi & 0 \end{pmatrix} \text{ remplace la matrice } B = \begin{pmatrix} i\pi & 0 \\ 0 & i\pi \end{pmatrix} \text{ Car les 2 matrices vérifient : } \exp(A) = \exp(B) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Ainsi, en changeant  $D$  en  $D'$ , c'est à dire  $Q$  (anti-symétrique complexe) en  $Q'$  (anti-symétrique réelle), on vient de prouver la surjectivité de  $\exp : ASym_n(\mathbb{R}) \longrightarrow SO_n(\mathbb{R})$ .  $\square$

**Surjectivité de  $\exp : \mathfrak{su}_n(\mathbb{C}) \longrightarrow SU_n(\mathbb{C})$  et  $\exp : \mathfrak{sl}_n(\mathbb{C}) \longrightarrow SL_n(\mathbb{C})$**

**8.2.12 Proposition:**  $\exp : \mathfrak{su}_n(\mathbb{C}) \longrightarrow SU_n(\mathbb{C})$  et  $\exp : \mathfrak{sl}_n(\mathbb{C}) \longrightarrow SL_n(\mathbb{C})$  est surjective

**Preuve:** On sait que si  $A \in SU_n(\mathbb{C})$  (respectivement  $:SL_n(\mathbb{C})$ ), alors il existe une matrice complexe  $M \in U_n(\mathbb{C})$  (respectivement  $:M_n(\mathbb{C})$ ) telle que  $\exp(M) = A \in SU_n(\mathbb{C})$  (respectivement  $:SL_n(\mathbb{C})$ ). Cependant, on sait que  $\det(A) = 1 = \det(\exp(M)) = e^{\text{tr}(M)} \Rightarrow \text{tr}(M) = 0$ . Dans les 2 cas, cela prouve la proposition (par la définition de l'algèbre de Lie de  $SL_n(\mathbb{C})$  et  $U_n(\mathbb{C})$ ).  $\square$

### 8.3 Homomorphisme entre $SU_2(\mathbb{C})$ et $SO_3(\mathbb{R})$ \*

On va montrer qu'il existe un homomorphisme surjectif entre  $SU_2$  et  $SO_3$ . Pour cela, on passe par leurs algèbres de Lie. On va donner une base de leurs algèbres de Lie :  $\mathfrak{so}_3$  : soit :  $P = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$$Q = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

On vérifie très facilement que c'est une base des matrices antisymétriques de taille 3. Puis :  $\mathfrak{su}_2$   $H = \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$

$$E = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad F = \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Là encore, on vérifie que c'est une base des matrices antihermitiennes de taille 2.

Par un calcul direct, on trouve comme crochets de Lie :

$$[P, Q] = R, [Q, R] = P, [R, P] = Q$$

$$[H, E] = F, [E, F] = H, [F, H] = E$$

Cela implique l'existence d'un isomorphisme  $\varphi : \mathfrak{su}_2 \longrightarrow \mathfrak{so}_3$ , car ce sont 2 espaces vectoriels de même dimension. On peut définir :

$$\varphi(xH + yE + zF) = xP + yQ + zR; (x, y, z) \in \mathbb{R}^3$$

et il satisfait  $\varphi([U, V]) = [\varphi(U), \varphi(V)]$  (on le vérifie en le faisant pour  $[H, E], [E, F], [F, H]$ , car  $P, Q$  et  $R$  forment une base. On peut également remarquer que ces 2 espaces sont donc liés à  $\mathbb{R}^3$ , et, par exemple, il existe un isomorphisme

$$\theta_0 : \mathbb{R}^3 \longrightarrow \mathfrak{so}_3; \theta_0(xe_1 + ye_2 + ze_3) = xP + yQ + zR$$

, où  $(e_1, e_2, e_3)$  est la base canonique de  $\mathbb{R}^3$ , qui est un isomorphisme de  $\mathbb{R}$  algèbre de Lie.

On va maintenant construire un homomorphisme de Lie :  $SU_2 \longrightarrow SO_3$  de dérivée  $\varphi$ .

Pour cela, on utilise l'action adjointe de  $SU_2$  sur  $\mathfrak{su}_2$ , qu'on a déjà définie avant mais qu'on rappelle :  $Ad_A = AUA^{-1} = AUA^* : A \in SU_2, U \in \mathfrak{su}_2$ .

Remarque :  $Ad_A$  est une application linéaire. On vérifie très facilement l'injectivité de  $Ad_A$  de  $\mathfrak{su}_2$  dans lui même. La surjectivité est tout aussi évidente (les 2 démonstrations s'appuient sur le fait que  $\mathfrak{su}_2$  est stable par  $Ad_A$  :

$$Ad_A(U)^* = (AUA^*)^* = (A^*)^*U^*A^* = A(-U)A^* = -AUA^* = -Ad_A(U)$$

) Donc c'est un isomorphisme de  $\mathbb{R}$  algèbre de Lie :

$$Ad_A : \mathfrak{su}_2 \longrightarrow \mathfrak{su}_2$$

quelle que soit la matrice  $A$  dans  $SU_2$ .

On peut définir le produit réel intérieur (|) sur  $\mathfrak{su}_2$  par  $(X | Y) = -tr(XY)$ ,  $X, Y \in \mathfrak{su}_2$ . C'est l'équivalent du produit scalaire euclidien sur  $\mathbb{R}^3$ .

C'est une application bilinéaire symétrique (se vérifie) et les éléments :

$\widehat{H} = \sqrt{2}H$ ,  $\widehat{E} = \sqrt{2}E$  et  $\widehat{F} = \sqrt{2}F$  forment une base orthogonale pour le produit réel intérieur. On peut définir un  $\mathbb{R}$ -isomorphisme linéaire

$$\theta : \mathbb{R}^3 \longrightarrow \mathfrak{su}_2, \theta(xe_1 + ye_2 + ze_3) = x\widehat{H} + y\widehat{E} + z\widehat{F}$$

, qui est également une isométrie :

$$(\theta(x) | \theta(y)) = x.y, x, y \in \mathbb{R}^3$$

**8.3.1 Proposition:** (|) est une forme bilinéaire symétrique sur  $\mathfrak{su}_2$  qui est définie positive. C'est un invariant dans le sens où :

$$\forall X, Y, Z \in \mathfrak{su}_2, ([Z, X] | Y) + (X | [Y, Z]) = 0$$

**Preuve:** On a déjà expliqué que la bilinéarité et la symétrie étaient triviales (propriétés de la trace). Pour ce qui est de la forme définie positive :  $(x\widehat{H} + y\widehat{E} + z\widehat{F} | x\widehat{H} + y\widehat{E} + z\widehat{F}) = x^2 + y^2 + z^2 > 0$  si  $(x, y, z) \neq (0, 0, 0)$  et cela vaut 0 si  $x, y$  et  $z$  sont tous nuls. Pour l'invariance : on le vérifie par un calcul direct :

$$tr((ZX - XZ)Y) + tr(X(YZ - ZY)) = tr(ZXY) - tr(XZY) + tr(XYZ) - tr(XZY) = 0$$

car  $tr(AB) = tr(BA)$  pour toutes matrices carrées de même taille. □

**Remarque :** Soit  $A \in SU_2$  et  $X, Y \in \mathfrak{su}_2$  :

$$(AXA^* | AYA^*) = -tr(AXA^*AYA^*) = -tr(AXYA^{-1}) = -tr(XY) = (X | Y)$$

$Ad_A$  est donc une transformation linéaire orthogonale car elle préserve le produit scalaire euclidien (produit réel intérieur).

On définit

$$\overline{Ad} : SU_2 \longrightarrow SO_3; \overline{Ad}(A) = Ad_A$$

**8.3.2 Théorème:** L'homomorphisme dérivé  $d\overline{Ad} : \mathfrak{su}_2 \longrightarrow \mathfrak{so}_3$  est un isomorphisme de  $\mathbb{R}$  algèbres de Lie.



**Preuve:** Soit  $B \in \mathfrak{su}_2$  et la courbe :

$$\beta : \mathbb{R} \longrightarrow SU_2; \beta(t) = \exp(Bt)$$

Soit

$$\bar{\beta} : \mathbb{R} \longrightarrow SO_3; \bar{\beta}(t) = \overline{Ad}_{\beta(t)}$$

On peut dériver  $\bar{\beta}$  en  $t = 0$  :

$$\bar{\beta}'(0)(X) = \frac{d}{dt} \exp(tB) X \exp(-tB) |_{t=0} = BX - XB = [B, X] \in \mathfrak{so}_3$$

On a (en reprenant les matrices  $H, E, F$  définies au début :

$$[H, H] = 0, [H, E] = F, [H, F] = -E; [E, H] = -F, [E, E] = 0, [E, F] = H; [F, H] = E, [F, E] = -H, [F, F] = 0$$

On peut donc représenter l'action de ces 3 matrices ( $H, E, F$ ) sur  $\mathfrak{su}_2$  dans la base ( $H, E, F$ ) :

$$\begin{aligned} \text{Pour } H : R &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\ \text{Pour } E : Q &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \\ \text{Pour } F : Q &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Donc

$$d\overline{Ad} : \mathfrak{su}_2 \longrightarrow \mathfrak{so}_3; d\overline{Ad}(xH + yE + zF) = xR + yQ + zP$$

est bien un isomorphisme.  $\square$

Avant de commencer la preuve de la proposition suivante, on va avoir besoin d'un petit lemme :

On rappelle au préalable que la composante connexe d'un élément  $x \in U$  ( $U$  est un ensemble) est le plus grand ensemble connexe contenant  $X$ .

**8.3.3 Lemme:** Soit  $G_0$  la composante connexe de l'élément identité  $e$  d'un groupe topologique  $G$ .  $G_0$  est un sous groupe normal de  $G$ , fermé.

**Preuve:** Avant de commencer, une petite remarque : la continuité des application  $f, g$  et  $w$  qu'on considère en dessous vient du fait que  $G$  est un groupe topologique.

Soit  $f : G_0 \times G_0 \longrightarrow G_0, (g, h) \longmapsto gh$ .  $f$  est une application continue et  $G_0 \times G_0$  est connexe, par conséquent  $Im f \subseteq G_0$ . Comme il est évident que  $G_0 \subseteq Im f$ , en prenant  $g = e$  et  $h$  quelconque, on a  $Im f = G_0$ . Ceci montre que si  $g, h \in G_0, gh \in G_0$ , donc  $G_0$  est stable par la loi du groupe  $G$ . De plus,  $e \in G_0$ . On peut considérer  $f : G_0 \times G_0 \longrightarrow G_0, (g, h) \longmapsto gh^{-1}$ , également continue. Un raisonnement analogue permet de montrer que  $Im f = G_0$ , donc  $G_0$  est un sous-groupe de  $G$ . Montrons qu'il est normal : même principe. Soit  $g \in G$  et  $w : G_0 \longrightarrow G_0, h \longmapsto ghg^{-1}$ , une application continue. avec un raisonnement similaire à celui fait pour  $f$ ,  $Im w \subset G_0 \forall g \in G$  et par conséquent,  $G_0$  est normal dans  $G$ .  $\square$

**8.3.4 Proposition:**  $\overline{Ad}$  tel que définit ci-dessus est un homomorphisme continue de groupes matriciels, lisse, surjectif et son noyau est  $\{\pm I_2\}$ .

**Preuve:** D'après la remarque précédente :

$$\overline{Ad} : SU_2 \longrightarrow O_3$$

car elle préserve le produit scalaire.  
C'est un homomorphisme :

$$\begin{aligned}\overline{Ad}_{AB}(U) &= (AB)U(AB)^* = (AB)U(B^*A^*) = A(BUB^*)A^* = \overline{Ad}_A \circ \overline{Ad}_B(U) \\ \overline{Ad}_{I_2}(U) &= I_2UI_2^* = U\end{aligned}$$

De plus :  $\overline{Ad}$  est clairement continue et  $SU_2$  est connexe. Donc  $Im(\overline{Ad})$  est connexe et contient l'élément neutre de  $O_3$ ,  $I_3$  (car c'est un homomorphisme). Elle est également fermée car  $SU_2$  est compact. Or, on a vu (voir la partie topologie sur les matrices) que  $O_3$  n'est pas connexe, mais qu'il possède 2 composantes connexes  $O_3 = SO_3 \cup O_3^-$ . La composante connexe qui contient  $I_3$  est  $SO_3$ , par conséquent  $Im(\overline{Ad}) \subseteq SO_3$ . De plus,  $dim SU_2 = dim SO_3$ , ce qui donne  $Im \overline{Ad} = SO_3$ .

Remarque : si on est pas convaincu par l'argument de la dimension, on peut imaginer que  $Im(Ad)$  soit un sous-ensemble connexe de  $SO_3$ , autrement dit,  $Im Ad$  serait la composante connexe d'un sous-groupe  $H$  de  $SO_3$ . Mais le lemme affirme que  $Im Ad$  serait alors normal dans ce sous-groupe  $H$ , et on peut également montrer en considérant encore  $w : G_0 \rightarrow G_0, h \mapsto ghg^{-1}$  (avec  $g \in SO_3$ ) que  $Im Ad$  serait alors normal dans  $SO_3$ . Par conséquent,  $Im Ad = SO_3$  (car  $Im Ad \neq I_3$ ) en utilisant le fait que  $SO_3$  est simple.

Calcul du noyau (c'est le plus simple). Simple rappel : attention, on s'intéresse à des homomorphismes, le neutre est donc l'identité comme application linéaire et non  $I_2$  ou encore  $O_2$ . Donc

$$Ad_A(X) = Id = X \forall X \in \mathfrak{su}_2 \Leftrightarrow AXA^{-1} = X \Leftrightarrow AX = XA$$

$$\text{On considère : } X_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad X_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

$$\text{On trouve alors comme équations, après calculs, si : } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$b = c = 0; a = d; a = \pm 1 \Rightarrow A = \pm I_2$$

, en effet, les 2 premières équations imposent que  $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$  Mais  $A \in SU_2 \Rightarrow a^2 = 1$ . Donc  $A = \pm I_2$ .

Par conséquent, cet homomorphisme n'est pas injectif (on rappelle que si  $f$  est un homomorphisme,  $ker f = \{e\} \Leftrightarrow f$  est injective).  $\square$

**8.3.5 Corollaire:**  $SU_2(\mathbb{C}) / \{\pm I_2\}$  est isomorphe à  $SO_3(\mathbb{R})$ .

**Preuve:**  $\overline{Ad}$  est un homomorphisme de  $SU_2$  dans  $SO_3$ . Il existe donc un isomorphisme de  $SU_2 / ker(\overline{Ad})$  dans  $Im \overline{Ad} = SO_3$  (théorème sur les groupes);  $\square$

**Remarque :** Il existe une méthode plus directe, mais plus calculatoire pour trouver un homomorphisme entre  $SU_2$  et  $SO_3$ . l'avantage de cette preuve est qu'elle permet de voir ce qu'il se passe et montre un lien particulier entre groupes et algèbres de Lie pour  $SU_2$  et  $SO_3$ .

La voici : [https://en.wikipedia.org/wiki/Rotation\\_group\\_SO\(3\)](https://en.wikipedia.org/wiki/Rotation_group_SO(3)) .

**Remarque :** Cette partie permet ensuite de s'intéresser à la généralisation de ce qu'on a vu à d'autres groupes, avec les spineurs ( $Spin_n$ ) (ce qui sort du cadre de ce TIPE).

## 8.4 Caractérisation de $SL_2(\mathbb{R})$ \*

On va maintenant utiliser l'algèbre de Lie pour étudier  $SL_2(\mathbb{R})$ .

On rappelle que  $SL_2(\mathbb{R})$  est l'ensemble des matrices réelles de la forme

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  telles que  $\det(A) = ad - bc = 1$ . Son algèbre de Lie est (on l'a vu précédemment) l'ensemble des matrices réelles de trace nulle :  $\{B \in M_2(\mathbb{R}), \text{tr}(B) = 0\}$  qu'on peut écrire sous la forme :

$$B = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}$$

Pour étudier la structure topologique de  $SL_2(\mathbb{R})$ , on peut écrire :

$$a = p + q, d = p - q, b = r + s, c = r - s$$

de sorte que

$$ad - bc = p^2 - q^2 - (r^2 - s^2) = 1 \Leftrightarrow p^2 + s^2 = q^2 + r^2 + 1$$

Donc, pour  $(q, r) \in \mathbb{R}^2$  fixés, le point  $(p, s)$  parcourt le cercle dans  $\mathbb{R}^2$ , de rayon  $(q^2 + r^2 + 1)^{\frac{1}{2}}$ . Donc  $SL_2(\mathbb{R})$  est difféomorphe au produit cartésien d'un cercle et d'un plan. Pour se le représenter, on peut utiliser le difféomorphisme (facile à vérifier) :

$$\psi : \mathbb{R}^3 \longrightarrow SL_2(\mathbb{R})$$

$$(\theta, (u, v)) \longmapsto (1 - u^2 - v^2)^{-\frac{1}{2}} \begin{pmatrix} u + \cos(\theta) & -\sin(\theta) + v \\ \sin(\theta) + v & \cos(\theta) - u \end{pmatrix}$$

En effet, un simple calcul montre que  $\det(\psi(\theta, u, v)) = 1$ . Soit  $A \in SL_2(\mathbb{R})$  de coefficients  $a, b, c$  et  $d$  comme ci-dessus. Soit  $r = (1 - u^2 - v^2)^{\frac{1}{2}}$

On a :

$$a = \frac{u + \cos(\theta)}{r}, b = \frac{-\sin(\theta) + v}{r}, c = \frac{\sin(\theta) + v}{r}, d = \frac{\cos(\theta) - u}{r}$$

Donc

$$\tan(\theta) = \frac{c - b}{a + d} \Rightarrow \theta = \text{Arctan}\left(\frac{c - b}{a + d}\right), \text{ ou } : \pi + \text{Arctan}\left(\frac{c - b}{a + d}\right)$$

Ceci nous permet de déterminer  $\cos(\theta)$  et  $\sin(\theta)$  au signe près car l'équation  $\tan(y) = z$  d'inconnue  $y$  possède 2 solutions : rappel de trigonométrie :

$$\cos(\arctan(x)) = \frac{1}{\sqrt{1 + x^2}}, \sin(\arctan(x)) = \frac{x}{\sqrt{1 + x^2}}$$

Donc

$$\cos(\theta) = \pm \frac{|a + d|}{\sqrt{1 + (c - b)^2}}, \sin(\theta) = \pm \frac{|a + d|}{a + d} \frac{c - b}{\sqrt{1 + (c - b)^2}}$$

(le signe est + pour les 2 égalités si  $\theta = \text{Arctan}\left(\frac{c-b}{a+d}\right)$ , - pour les 2 égalités si  $\theta = \pi + \text{Arctan}\left(\frac{c-b}{a+d}\right)$ )

On peut alors déterminer la valeur de  $r$  car

$$a + d = \frac{2\cos(\theta)}{r}, c - b = \frac{2\sin(\theta)}{r} \Rightarrow r = \frac{2\cos(\theta)}{a + d} = \frac{2\sin(\theta)}{c - b} = \pm \frac{|a + d|}{a + d} \frac{2}{\sqrt{1 + (c - b)^2}} = \frac{2}{\sqrt{1 + (c - b)^2}}$$

car  $r > 0$ . On peut alors en déduire la valeur de  $\theta$  car : si  $\theta = \text{Arctan}\left(\frac{c-b}{a+d}\right)$ , on a

$$\frac{|a + d|}{a + d} = 1 \Rightarrow a + d \geq 0$$

si  $\theta = \text{Arctan}\left(\frac{c-b}{a+d}\right) + \pi$ , alors  $a + d \leq 0$ , la réciproque étant vraie, on peut connaître  $\theta$ .

Puis on termine :

$$u = ar - \cos(\theta) \cos(\theta) - dr, v = br + \sin(\theta) = cr - \sin(\theta)$$

De plus, on vérifie que  $u, v \in D = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 1\}$ . On a bien un difféomorphisme.

Déterminons les valeurs propres d'une matrice  $A \in SL_2(\mathbb{R})$  : le polynôme caractéristique est, d'après le théorème de Cayley-Hamilton :

$$X^2 - \text{tr}(A)X + \det(A) = X^2 - \text{tr}(A)X + 1$$

, ce qui donne, si  $tr(A) \geq 2$  ou  $tr(A) \leq -2$  :

$$\delta = tr(A)^2 - 4 \Rightarrow X = \lambda_{1,2} = \frac{tr(A) \pm \sqrt{\delta}}{2}; tr(A) = \frac{2\cos(\theta)}{\sqrt{1-u^2-v^2}} = p$$

Si  $tr(A) = \pm 2$ ,  $\lambda_1 = \lambda_2 = \pm 1$ . Les éléments de  $SL_2(\mathbb{R})$  autres que l'identité  $I_2$  tels que  $tr(A) = 2$  sont unipotents car  $(A - I_2)^2 = O_2$ . De même,  $(A + I_2)^2 = O_2$  si  $tr(A) = -2$ . On dit que ce sont les éléments paraboliques de  $SL_2(\mathbb{R})$ .

Donc, si  $|tr(A)| > 2$ , on rappelle que

$$det(A) = \lambda_1 \lambda_2 = 1 \Rightarrow \lambda_2 = \lambda_1^{-1} = \lambda^{-1} (\lambda = \lambda_1)$$

De plus, dans ce cas, si  $\lambda = \frac{tr(A) + \sqrt{\delta}}{2}$  et  $tr(A) > 2$ ,  $\lambda > 1$ , et si  $\lambda = \frac{tr(A) - \sqrt{\delta}}{2}$  et  $tr(A) < -2$ ,  $\lambda < -1$ . Ce sont les éléments hyperboliques de  $SL_2(\mathbb{R})$ . Dans les 2 cas, la matrice  $A$  est diagonalisable (car ses valeurs propres sont différentes).

Si  $|tr(A)| < 2$ , dans ce cas  $\exists \varphi$  tel que

$$tr(A) = p = 2\cos(\varphi); p^2 - 4 = -(4 - p^2) = 4\sin^2(\varphi) \Rightarrow \lambda_1 = \cos(\varphi) + i\sin(\varphi), \lambda_2 = \cos(\varphi) - i\sin(\varphi)$$

De sorte que  $A$  est semblable à une matrice de rotation, donc elle l'est (d'angle  $\varphi$ ) (car, voir la surjectivité de l'exponentielle : une matrice de rotation a pour valeur propre  $e^{i\theta}$  et  $e^{-i\theta}$ ). Ce sont ici les éléments appelés elliptiques de  $SL_2(\mathbb{R})$ .

On remarque alors que la caractérisation qu'on a faite de  $SL_2(\mathbb{R})$  est complète :

On a les classes conjuguées suivantes :

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \quad \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \quad \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

Et de plus, aucune de ces classes ne s'intersectent (facile à vérifier). Remarque : les 2 premières classes sont valables avec  $|\lambda| > 1$  et  $\theta \in [0, 2\pi[$ . Les 4 dernières ne sont autres que les moyens de représenter les éléments unipotents et ceux qui vérifient  $(A + Id)^2 = 0$ , c'est à dire les éléments de forme :

$$A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix}$$

Avec  $a, b \in \mathbb{R}^*$ , en distinguant les cas positifs et négatifs. En effet, pour déterminer un représentant d'une classe de conjugaison : (on va le faire pour les matrices du premier type, avec  $a > 0$ ) : on a clairement que 1 est une valeur propre. Déterminons  $ker(A - I_2)$

$$\begin{aligned} ker \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= 0 \\ \Leftrightarrow ay = 0 &\Leftrightarrow y = 0. \end{aligned}$$

Un vecteur propre est donc de la forme

$$\begin{pmatrix} \alpha \\ 0 \end{pmatrix} = 0$$

( $\alpha \neq 0$ ) Comme la matrice n'est pas diagonalisable, on doit utiliser un procédé similaire à la trigonalisation. Les possibilités pour le second vecteur propre sont donc

$$\begin{pmatrix} \beta \\ \alpha^{-1} \end{pmatrix} = 0 \text{ car la matrice de passage est alors } P = \begin{pmatrix} \alpha & \beta \\ 0 & \alpha^{-1} \end{pmatrix} \in SL_2(\mathbb{R}) \text{ (la condition } \in SL_2(\mathbb{R})$$

est requise pour la détermination des classes de conjugaison). Donc :

$$\begin{aligned} PAP^{-1} &= \begin{pmatrix} \alpha & \beta \\ 0 & \alpha^{-1} \end{pmatrix} \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha^{-1} & -\beta \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} \alpha & a\alpha + \beta \\ 0 & \alpha^{-1} \end{pmatrix} \\ \begin{pmatrix} \alpha^{-1} & -\beta \\ 0 & \alpha \end{pmatrix} &= \begin{pmatrix} 1 & -\alpha\beta + \alpha(a\alpha + \beta) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \alpha^2 a \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Donc on peut poser  $\alpha = \sqrt{\left(\frac{1}{a}\right)}$ , de façon à ce que le représentant de la classe de conjugaison soit bien le troisième annoncé au dessus. On procède de même pour les cas  $a < 0, b < 0$  et  $b < 0$ , et on retrouve bien ce qui a été annoncé. On a donc "classifié"  $SL_2(\mathbb{R})$ .

Déterminons maintenant la représentation adjointes de  $SL_2(\mathbb{R})$ . Soit  $A \in SL_2(\mathbb{R})$

$$Ad_A : \mathfrak{sl}_2 \longrightarrow \mathfrak{sl}_2, Ad_A(U) = AUA^{-1}$$

(En effet,  $tr(Ad_A)(U) = tr(U) = 0$  si  $U \in \mathfrak{sl}_2$ ) De la même manière que pour  $SU_2$  et  $SO_3$ , on peut définir un produit scalaire (forme de Killing). Soit  $\overline{Ad} : SL_2(\mathbb{R}) : c$ 'est une application linéaire orthogonale pour ce produit scalaire.

**8.4.1 Proposition:** *Le noyau de l'homomorphisme  $\overline{Ad}$  est  $\{I_2, -I_2\}$  et l'image de  $Ad(SL_2(\mathbb{R}))$  est par conséquent isomorphe à  $PSL_2(\mathbb{R})$  (voir le dernier théorème de la partie algèbre de Lie).*

**Preuve:** Soit une base de  $\mathfrak{sl}_2$  : les matrices  $P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  ;  $Q = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  ;  $R = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$  conviennent.

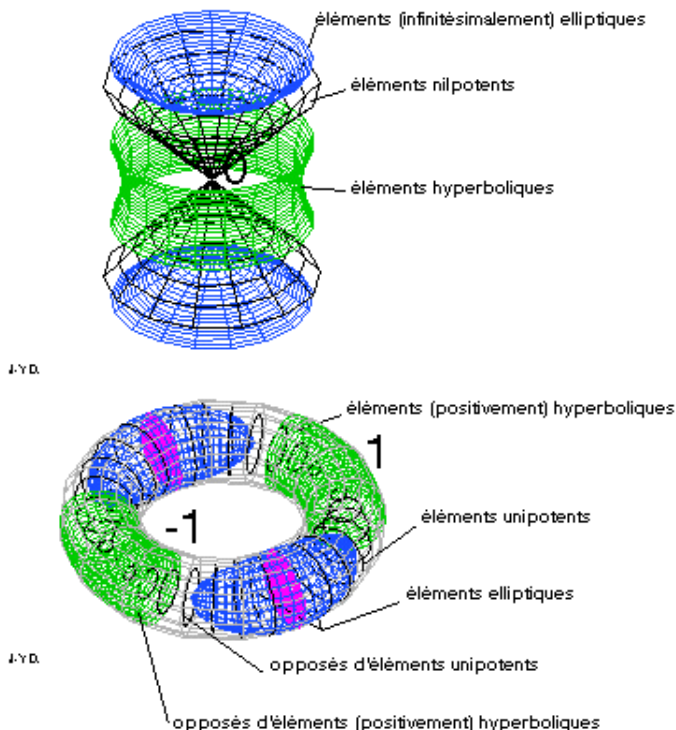
Montrons que le noyau est  $\{I_2, -I_2\}$  : Il s'agit de résoudre  $\forall U \in \mathfrak{sl}_2(\mathbb{R}) AUA^{-1} = Id(U) = U$  (c'est à dire, trouver tous les  $A \in SL_2(\mathbb{R})$  possibles) : soit  $AU = UA$ . Soit  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$U_1 = Q = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ et } U_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

On trouve après calcul,  $b = c = 0$ ;  $a = d$ . Puis, comme  $A \in SL_2(\mathbb{R}), det(A) = a^2 = 1 \Rightarrow a = \pm 1 \Rightarrow A = \pm I_2$ . □

Voici une 2 images qui illustrent cela :

Orbites adjointes dans  $\mathfrak{sl}_2(\mathbb{R})$  et classes de conjugaison dans  $SL_2(\mathbb{R})$  :



Explication du dessin : on a déterminé l'ensembles des classes de conjugaisons de  $SL_2(\mathbb{R})$ . En reprenant la partie sur la non connexité de  $exp(\mathfrak{sl}_2(\mathbb{R}))$  : on avait montré que si une matrice  $A \in \mathfrak{sl}_2(\mathbb{R})$  :

1.  $det(A) = 0 \Rightarrow tr(exp(tA)) = 2$
2.  $det(A) < 0 \Rightarrow tr(exp(tA)) > 2$

$$3. \det(A) > 0 \Rightarrow -2 < \operatorname{tr}(\exp(tA)) < 2$$

Par conséquent, on en déduit en reprenant la caractérisation précédente que l'image de  $\exp(\mathfrak{sl}_2)$  est l'ensemble des éléments elliptiques ( $\det(A) > 0$ ), paraboliques de trace égale à 2 (c'est à dire, unipotents) ( $\det(A) = 0$ ) ou hyperboliques de trace supérieure à 2 ( $\det(A) > 0$ ).

Voici la page web où l'on trouve ces jolis dessins :

[https://webusers.imj-prg.fr/~jean-yves.ducloux/sl2/exp\\_sl2.html](https://webusers.imj-prg.fr/~jean-yves.ducloux/sl2/exp_sl2.html)

**Généralisation :** Pour trouver l'image de l'adjoint en général, pour n'importe quel groupe de Lie  $G$ , on utilise toujours le même procédé : on détermine le  $\ker$  de l'adjoint, car c'est le plus simple, et on utilise le fait que le quotient du groupe  $G/\ker \operatorname{Ad}$  est isomorphe à  $\operatorname{Im} \operatorname{Ad}$ , c'est en fait le premier théorème d'isomorphisme sur le quotient de groupes. On obtient à chaque fois l'équation  $AX = XA$  avec  $A \in G$  et  $X \in \mathfrak{g}$ . Une fois le  $\ker$  déterminé, on a plus rien à faire. De plus, le gros avantage (par rapport à la détermination du centre d'un groupe), c'est que  $\mathfrak{g}$  est un espace vectoriel, il suffit donc de prendre une base et de résoudre l'équation pour chaque vecteurs de la base et on trouve la solution. Pourquoi est-ce si intéressant ? Parce que les représentations, qui sont une théorie vraiment abstraite, peuvent être déterminées (on explicite l'image et le noyau) par un procédé simple.

$SL_2(\mathbb{R})$  est un exemple de groupe de Lie non compact et non commutatif de dimension 3, mais il en existe aussi en dimension 2 : par exemple, le groupe (notons le  $T_2$  des matrices triangulaires supérieures ayant des valeurs propres positives :  $U = \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix}$  avec  $a > 0, b \in \mathbb{R}$ ). On peut montrer par le calcul que les différentes classes de conjugaison sont :

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

$$U = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} (b > 0)$$

$$U = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} (b < 0)$$

$$U = \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} (b \in \mathbb{R}, a > 0, a \neq 1)$$

## ANNEXES ET CONCLUSION

### 9 Annexe I : Groupes de matrices

Le but de cette annexe est de définir un certain nombre de groupes matriciels, et d'étudier leur propriétés.

.Remarque : dans tous les groupes matriciels multiplicatifs que nous allons considérer, la matrice identité  $I_n$  sera dans le groupe et dans tous les groupes additifs, la matrice nulle  $O_n$  sera dans le groupe. De plus, comme on sait que l'addition et la multiplication matricielle sont associatives, on ne vérifiera que la première et la quatrième propriété sur les groupes.

#### 9.1 Groupes élémentaires et représentation matricielle

Représentation matricielle de  $\mathbb{R}^n$ .

On cherche ici à représenter  $\mathbb{R}^n$  (en tant que groupe additif) comme un groupe de matrice.

Commençons par  $\mathbb{R}^2$  : On considère la matrice :  $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$  Donc  $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \lambda + \mu \\ 0 & 1 \end{pmatrix}$

Pour  $\mathbb{R}^n$ , on considère :

$$\begin{pmatrix} 1 & x_1 & \cdots & x_n \\ 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\text{Donc} \begin{pmatrix} 1 & x_1 & \cdots & x_n \\ 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & y_1 & \cdots & y_n \\ 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} 1 & x_1 + y_1 & \cdots & x_n + y_n \\ 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Un raisonnement analogue vaut pour  $\mathbb{C}^n$

Maintenant, on va voir comment représenter la multiplication de matrices de  $M_n(\mathbb{C})$  dans  $M_{2n}(\mathbb{R})$ , car  $\dim_{\mathbb{R}} M_n(\mathbb{C}) = 2n$ . L'intérêt est que l'on peut voir  $\mathbb{C}^n$  comme un groupe matriciel réel plutôt que complexe, ce qui sera utile par la suite (car les groupes de Lie nécessitent de trouver la dimension d'un groupe dans le corps des réels).

Là encore, on commence par  $\mathbb{C}$ . Pour cela : on sait que  $i^2 = -1$  dans  $\mathbb{C}$ . Cherchons donc une matrice  $J$  telle que  $J^2 = -I_2$  :  $J = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$   $J^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

On a les équations :  $b(a+d) = 0$  et  $c(a+d) = 0$ . Si  $b = c = 0$ ,  $a^2 = -1$  et  $d^2 = -1$ . On exclut ce cas car  $a$  et  $d$  ne peuvent pas être réels. Donc :  $a+d = 0$ . On a donc  $a = -d$  et  $a^2 + bc = 0$ . Donc on a une infinité de possibilités (on a le choix de  $a$  et  $b$  par exemple, et  $c$  est directement déterminé). Le cas le plus simple est  $a = d = 0$  et  $c = 1 \Rightarrow b = -1$ . Donc  $J = J_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Comme pour les complexes, on écrit :  $z \in \mathbb{C} = xI_2 + yJ_2$  et  $z' \in \mathbb{C} = x'I_2 + y'J_2$ . Donc  $zz' = (xI_2 + yJ_2)(x'I_2 + y'J_2) = (xx' - yy')I_2 + (xy' + x'y)J_2$ . On a donc (comme on pouvait s'en douter) représenté la multiplication dans  $\mathbb{C}$ .

Pour  $\mathbb{C}^n$  : On a deux méthodes : la première consiste à écrire  $A \in M_n(\mathbb{C})$  :  $A = X + iY$ ,  $X \in M_n(\mathbb{R})$ ,  $Y \in M_n(\mathbb{R})$  (remarque : cette méthode est intéressante car, quand on a 2 groupes (exemple pour 2 groupes additifs :  $\mathbb{R}^2$  et  $M_n(\mathbb{R})$ , et on cherche à représenter  $M_n(\mathbb{R}^2)$ ) on choisit la façon la plus simple pour représenter : au lieu d'écrire  $A \in M_n(\mathbb{R}^2)$ , on écrit  $A = (A_1, A_2)$ ,  $A_1, A_2 \in M_n(\mathbb{R})$ ). Donc, ici, on écrit :

$$\text{si } A = X + iY \quad A = A_X I_{2n} + A_Y J_{2n} \quad \text{avec : } J_{2n} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \quad A_X = \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix} \quad A_Y = \begin{pmatrix} Y & 0 \\ 0 & Y \end{pmatrix}$$

Un deuxième moyen consiste tout simplement de représenter chaque coefficient complexe de la matrice  $A$  par des blocs  $2 \times 2$  :  $z_{ij} = z_{1,ij}I_2 + z_{2,ij}J_2$  : on a encore une matrice  $A' \in M_{2n}(\mathbb{R})$ , et on vérifie assez simplement qu'elle permet la multiplication et l'addition dans  $M_n(\mathbb{C})$ .

## 9.2 Groupe Linéaire

**9.2.1 Définition:** Le groupe linéaire  $GL_n(\mathbb{K})$  est l'ensemble des matrices inversibles de  $M_n(\mathbb{K})$  ( $\mathbb{K}$  étant un corps quelconque), c'est à dire l'ensemble des matrices  $A$  de  $M_n(\mathbb{K})$  telles que  $\det(A) \neq 0$ .

**9.2.2 Proposition:**  $GL_n(\mathbb{K})$  est un groupe.

**Preuve:**

$\det(A) \neq 0$  signifie également qu'il existe une matrice  $B$ , notée  $A^{-1}$  telle que  $AB=BA=I_n$ . De plus, si  $A, B \in GL_n(\mathbb{K})$ ,  $AB \in GL_n(\mathbb{K})$  car  $\det(AB) = \det(A)\det(B) \neq 0$  puisque  $\det(A) \neq 0$  et  $\det(B) \neq 0$ . Donc  $GL_n(\mathbb{K})$  est un groupe.  $\square$

## 9.3 Groupe Orthogonal

**9.3.1 Définition:** Les matrices *orthogonales*  $A$  de  $M_n(\mathbb{R})$  sont l'ensemble des matrices telles que

$${}^t AA = A^t A = I_n$$

,  ${}^t$  étant la transposée. On note  $O_n(\mathbb{R})$  cet ensemble.

**9.3.2 Proposition:**  $O_n(\mathbb{R})$  est un groupe.

**Preuve:**  $A^t A = I_n \Rightarrow {}^t A = A^{-1}$  donc l'inverse de A est également dans  $O_n(\mathbb{R})$  car  ${}^t({}^t A) = A$ . Enfin, si  $A$  et  $B \in O_n(\mathbb{R})$ ,

$$(AB) \cdot {}^t(AB) = (AB)({}^t B^t A) = A(B^t B)^t A = A^t A = I_n$$

Donc  $O_n(\mathbb{R})$  est un groupe. □

Nous allons maintenant nous intéresser à des propriétés générales de ce groupe :

**9.3.3 Proposition:** -Le déterminant d'une matrice orthogonale est égal à 1 ou -1.

**Preuve:**

$$A^t A = I_n \Rightarrow \det(A^t A) = \det(A)\det({}^t A) = \det(A)^2 = \det(I_n) = 1$$

□

**9.3.4 Proposition:** -Les vecteurs colonnes d'une matrice orthogonale sont orthogonaux entre eux et de norme 1.

**Preuve:** On utilise le produit de matrices : Soit  $C=A^t A$ . On pose

$$A = (a_{ij})_{1 \leq i, j \leq n} \text{ et } {}^t A = (a_{ji})_{1 \leq i, j \leq n}$$

Donc :  $c_{ij} = \sum_{k=1}^n a_{ik} \cdot a_{jk} = 1$  si  $i=j$  et 0 si  $i \neq j$ .

On pose  $v_i = \begin{pmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,n} \end{pmatrix}$

On a :  $c_{ij} = {}^t v_i \cdot v_j = v_i \cdot v_j$  où  $\cdot$  désigne le produit scalaire. Or, si  $i \neq j$ , ce produit scalaire est nul et si  $i=j$ ,  $(v_i)^2 = 1 = \|v_i\|^2 \Rightarrow \|v_i\| = 1$  où  $\|\cdot\|$  est la norme euclidienne. Remarque : la réciproque de cette propriété est immédiate. De plus, on a également cette propriété pour les lignes de la matrice (la démonstration se fait de la même manière). □

**9.3.5 Proposition:** -La multiplication d'un vecteur par une matrice orthogonale préserve la norme.

**Preuve:** Soit  $A \in O_n(\mathbb{R})$  et  $X \in M_{n,1}(\mathbb{R})$ . Soit  $Y=AX$  :

$$\begin{aligned} y_i &= \sum_{k=1}^n a_{ik} \cdot x_k \cdot \|Y\|^2 = \sum_{p=1}^n (y_p)^2 = \sum_{p=1}^n \left( \sum_{k=1}^n a_{pk} \cdot x_k \right)^2 = \sum_{p=1}^n \left( \sum_{k=1}^n a_{pk}^2 x_k^2 + 2 \sum_{k \neq i} a_{pk} a_{pi} x_k x_i \right) \\ &= \sum_{p=1}^n \left( \sum_{k=1}^n a_{pk}^2 x_k^2 + 2 \sum_{p=1}^n \sum_{k \neq i} a_{pk} a_{pi} x_k x_i \right) = \sum_{k=1}^n \left( \sum_{p=1}^n a_{pk}^2 \right) x_k^2 + 2 \sum_{k \neq i} \left( \sum_{p=1}^n a_{pk} a_{pi} \right) x_k x_i \\ &= \sum_{k=1}^n (x_k)^2 = \|X\|^2 \end{aligned}$$

En effet, on peut mettre les parenthèses car  $x_k^2$  et  $x_k x_i$  ne dépendent pas de p. Enfin, la dernière égalité se déduit en identifiant  $\sum_{p=1}^n a_{pk}^2$  comme la norme au carrée d'un vecteur correspondant à une colonne de la matrice A, donc égale à 1 puisque la matrice est orthogonale; et en identifiant  $\sum_{p=1}^n (a_{pk} a_{pi})$  comme le produit scalaire de deux vecteurs colonne différents de la matrice A, donc ce produit est nul puisque A est orthogonale. Finalement, comme  $\|Y\| \geq 0, \|X\| \geq 0, \|Y\| = \|X\|$ . □

**9.3.6 Proposition:** -Les 2 seules valeurs propres possibles d'une matrice orthogonale sont -1 et 1.



**Preuve:**

$$AX = \lambda X \Rightarrow \|AX\| = \|\lambda X\| = |\lambda| \|X\| \text{ Donc } |\lambda| = \frac{\|AX\|}{\|X\|} = 1$$

car d'après la propriété précédente,

$$\|AX\| = \|X\|$$

□

**9.3.7 Théorème:** -Pour finir : toute matrice orthogonale est diagonalisable par blocs d'une taille d'au plus 2. Autrement dit, on peut écrire une matrice  $B$  orthogonale sous la forme  $B = {}^t P A P$  avec  $P$  orthogonale et :  $A =$

$$\begin{pmatrix} R_1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & R_k & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & \ddots & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & -1 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & -1 \end{pmatrix}$$

Où  $R_1, \dots, R_k$  sont des matrices de rotation de taille 2 :  $R_i = \begin{pmatrix} \cos(\theta_i) & \sin(\theta_i) \\ -\sin(\theta_i) & \cos(\theta_i) \end{pmatrix}$

On va démontrer ce résultat en 3 étapes. Avant de commencer, je préciserai que ces résultats sont valables plus généralement pour les isométries vectorielles, mais je n'en parlerai pas ici (il y aura simplement un lien dans la bibliographie). Je donne simplement la définition : On appelle *isométrie vectorielle* de  $E$  tout endomorphisme  $u \in L(E)$  conservant la norme : Pour tout  $x \in E$ ,  $\|u(x)\| = \|x\|$ .

**9.3.8 Lemme:** Soit  $u$  une application linéaire dont la matrice associée est orthogonale : si  $F$  est un sous-espace vectoriel stable par  $u$ , alors  $F^\perp$  l'est aussi.

**Preuve:**  $F$  est stable par  $u$  donc  $u(F) \subset F$ . Or  $\det(u) = \pm 1$ , donc  $u$  est bijective et conserve la dimension. Par conséquent  $u(F) = F$  car les 2 espaces ont même dimension. Soit  $x \in F^\perp$ .  $\forall y \in F$ ,  $\exists a \in F$  tel que  $y = u(a)$ . Donc,

$$(u(x) | y) = (u(x) | u(a)) = (x | a) = 0$$

Ici  $(e | f)$  est le produit scalaire de  $e$  et  $f$ . Pour justifier l'avant dernière égalité, on va montrer un résultat intermédiaire : □

**9.3.9 Lemme:** Si  $u$  est une application linéaire dont la matrice est orthogonale,  $(u(x) | u(y)) = (x | y) \forall x, y \in \mathbb{R}^n$ . Remarque : la réciproque est également vraie (admis : voir un cours d'algèbre bilinéaire pour la preuve).

**Preuve:**

$$\|u(x+y)\|^2 = \|u(x) + u(y)\|^2 = \|u(x)\|^2 + 2(u(x) | u(y)) + \|u(y)\|^2 = \|x+y\|^2 = \|x\|^2 + 2(x | y) + \|y\|^2$$

Mais rappelons nous que  $u$  préserve la norme, ce qui a pour conséquence :

$$\forall t, \|u(t)\|^2 = \|t\|^2, (u(x) | u(y)) = (x | y).$$

□

**9.3.10 Lemme:** *Si  $u$  est un endomorphisme d'un  $\mathbb{R}$ -espace vectoriel réel de dimension finie non nulle alors il existe au moins une droite vectorielle ou un plan stable par  $u$ . Si  $u$  est un endomorphisme d'un  $\mathbb{C}$ -espace vectoriel complexe de dimension finie non nulle alors il existe au moins une droite vectorielle stable par  $u$ .*

**Preuve:** Soit  $P \in \mathbb{R}[X]$  (respectivement :  $\mathbb{C}[X]$ ) un polynôme unitaire annulateur de  $u$  (par exemple, son polynôme caractéristique ou minimal). On peut écrire  $P = P_1 P_2 \dots P_m$  avec  $P_k$  polynômes unitaires irréductibles de  $\mathbb{R}[X]$  (resp :  $\mathbb{C}[X]$ ). Comme  $P(u)=0$ ,  $P_1(u) \circ P_2(u) \circ \dots \circ P_m(u)=0$ , et donc, au moins l'un des endomorphismes composés n'est pas injectif (rappel :  $u$  est injectif  $\Leftrightarrow \ker u = \{0\}$ ). Supposons que ce soit celui d'indice  $k$  : qu'on soit dans  $\mathbb{R}$  ou  $\mathbb{C}$ , une forme possible est  $P_k(X) = X - \lambda$ , ce qui signifie :  $\exists x \neq 0$  tel que  $P_k(u(x))=0 \Leftrightarrow u(x) = \lambda x$ .  $\lambda$  est alors valeur propre de  $u$  et tout vecteur propre associé engendre une droite vectorielle stable. Si on est dans  $\mathbb{R}$ , il existe une autre forme possible :  $P_k(X) = X^2 + pX + q$  avec  $\Delta = p^2 - 4q < 0$ . Soit  $x \in \ker P_k(u) : u(x)^2 + pu(x) + qx = 0$  et donc  $F = \text{Vect}(x, u(x))$  est stable par  $u$  car dans ce cas,  $u(x)^2 = -pu(x) - qx \in F$ . □

**Preuve:** du théorème On va le prouver par récurrence sur la dimension de  $E$  (espace vectoriel). Cas  $n=1$  :  $u$  est une isométrie d'une droite et peut donc être représentée en base orthonormale par (1) ou (-1). Cas  $n=2$  :  $u$  est une isométrie du plan et peut donc être représentée en base orthonormale par :  $R(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$  ou  $R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Supposons la propriété établie jusqu'au rang  $n$  avec  $n \geq 2$ . Soit  $E$  un espace euclidien de dimension  $n+1$  et  $u$  un endomorphisme dont la matrice est orthogonale. Il existe une droite ou un plan  $F$  stable par  $u$  et  $F^\perp$  est alors aussi stable par  $u$ . Par hypothèse de récurrence, il existe une base orthonormale de  $F^\perp$  telle que la matrice de  $u$  dans celle-ci soit de la forme voulue. Par l'étude initiale, il existe une base orthonormale de  $F$  telle que la matrice de  $u$  dans celle-ci soit de la forme voulue. En accolant ces deux, on forme une base orthonormale de  $E$  comme voulue. Récurrence établie. Le seul point qui reste à prouver est la forme d'une isométrie vectorielle dans les cas  $n=1$  (dont on pouvait se douter) et  $n=2$ . Cas  $n=1$  :

$${}^tAA = A^tA = I_n \text{ où } A \text{ est de taille } 1 \Rightarrow A = (a) \text{ et donc : } a^2 = 1, \text{ soit } a = \pm 1 \text{ Cas } n=2 : A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\text{Donc } {}^tAA = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Donc :

$$a^2 + c^2 = 1 = b^2 + d^2, ab + cd = 0; \det(A) = \pm 1$$

$\exists \theta_1, \theta_2$  tels que  $a = \cos(\theta_1), c = \sin(\theta_1),$

$$b = \cos(\theta_2), d = \sin(\theta_2).$$

$$ab + cd = \cos(\theta_1 - \theta_2) = 0 = \cos(\theta_2 - \theta_1)$$

et

$$ad - bc = \sin(\theta_2 - \theta_1) = \pm 1 \cdot \cos(\alpha) = 0 \text{ et } \sin(\alpha) = \pm 1 \Rightarrow \alpha = \pm \frac{\pi}{2}$$

Donc : - si  $\det(A) = 1, \theta_2 = \theta_1 + \frac{\pi}{2}$  et donc, avec les propriétés de  $\cos$  et  $\sin$  et en posant  $\theta_1 = \theta$  :

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

- si  $\det(A) = -1, \theta_2 = \theta_1 - \frac{\pi}{2}$  et donc, avec les propriétés de  $\cos$  et  $\sin$  et en posant  $\theta_1 = \theta$  :

$$A = S(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$$

Dans ce cas, On a  $S(\theta) = S(0)R(\theta) = S(0)R(\frac{\theta}{2})R(\frac{\theta}{2}) = R(-\frac{\theta}{2})S(0)R(\frac{\theta}{2})$  donc  $S(\theta)$  est semblable à  $S(0)$  par le biais d'une matrice de passage orthogonale. Ainsi, une isométrie négative représentée initialement dans une base orthonormale par  $S(\theta)$  peut aussi être représentée dans une base orthonormale par  $S(0)$ . On reconnaît alors une symétrie orthogonale.

Et voilà pour les matrices orthogonales. □

## 9.4 Groupe Unitaire

Le groupe unitaire est assez semblable au groupe orthogonal, mais pour les complexes. C'est pourquoi je ne vais pas démontrer l'intégralité des propriétés que je vais énoncer car les preuves sont très similaires au cas réel.

**9.4.1 Définition:** Une matrice  $A$  unitaire de taille  $n$  est une matrice telle que  $A^*A=I_n$  avec  $A^*=\overline{A}^t$  où  $\overline{A}$  est la matrice conjuguée de  $A$  :  $\overline{A}=(\overline{a_{ij}})_{(1 \leq i,j \leq n)}$ . Propriétés :

**9.4.2 Proposition:** -Les vecteurs colonnes d'une matrice unitaire sont orthogonaux entre eux (pour le produit scalaire complexe  $(u | v)=u \cdot \overline{v}$  avec  $\cdot$  qui est le produit scalaire réel) et de norme 1.

**9.4.3 Proposition:** -La multiplication d'un vecteur par une matrice unitaire préserve la norme.

Les 3 propriétés précédentes sont très similaires à celles des matrices orthogonales et se démontrent de la même façon en manipulant le conjugué. Voici maintenant les petites différences avec le cas réel :

**9.4.4 Proposition:** -Le déterminant d'une matrice unitaire est de module 1 :

**Preuve:** Soit  $A$  une matrice unitaire.

$$\det(\overline{A}^t) = \det(\overline{A}) = \sum_{\sigma \in S_n} \text{signe}(\sigma) \left( \prod_{i=1}^n \overline{a_{i,\sigma(i)}} \right) = \overline{\sum_{\sigma \in S_n} \text{signe}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)}} = \overline{\det(A)}$$

(Remarque :  $S_n$  est l'ensemble des permutations). Donc :

$$\det(A^*A) = \det(A^*)\det(A) = \det(A)\overline{\det(A)} = |\det(A)|^2 = \det(I_n) = 1, \text{ soit } |\det(A)| = 1.$$

□

**9.4.5 Théorème:** Toute matrice unitaire est diagonalisable (dans une base orthonormale), semblable à une matrice diagonale unitaire dont les valeurs propres sont de module 1.

Tout d'abord, le lemme suivant reste valable dans  $\mathbb{C}$  :

**9.4.6 Lemme:** Soit  $u$  une application linéaire dont la matrice associée est unitaire : si  $F$  est un sous-espace vectoriel stable par  $u$ , alors  $F^\perp$  l'est aussi. Cela vient du fait (je vous renvoie à la démonstration dans  $\mathbb{R}$ ) que  $|\det(A)| = 1$  pour une matrice unitaire, et donc  $\det(A) \neq 0$ .

**Preuve:** Il suffit de reprendre la dernière démonstration pour les matrices orthogonales. On procède par récurrence sur  $n=\dim E$ . Dans  $\mathbb{C}$ , pour  $n=1$ , les isométries vectorielles sont telles que :  $A=(a)$  et  $(a)(\overline{a})=1=|a|^2$  donc  $a$  est de module 1. Supposons la propriété établie jusqu'au rang  $n$  avec  $n \geq 1$ . Soit  $E$  un espace euclidien de dimension  $n+1$  et  $u$  un endomorphisme dont la matrice est unitaire. Il existe une droite  $F$  stable par  $u$  et  $F^\perp$  est alors aussi stable par  $u$ . Par hypothèse de récurrence, il existe une base orthonormale de  $F^\perp$  telle que la matrice de  $u$  dans celle-ci soit de la forme voulue. Par l'étude initiale, il existe une base orthonormale de  $F$  telle que la matrice de  $u$  dans celle-ci soit de la forme voulue. En accolant ces deux, on forme une base orthonormale de  $E$  comme voulue. Récurrence établie.

Autrement dit, on peut écrire une matrice unitaire  $A$  sous la forme  $PDP^*$  avec  $P$  unitaire et  $D=$

$$\begin{pmatrix} e^{i\theta_1} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & e^{i\theta_n} \end{pmatrix}$$

□

## 9.5 Groupes Spécial Linéaire, Spécial Orthogonal et Spécial Unitaire

C'est très simple : on notes ces différents groupes  $SL_n(\mathbb{K})$ ,  $SO_n(\mathbb{R})$  et  $SU_n(\mathbb{C})$ .

$$SL_n(\mathbb{K}) = \{A \in GL_n(\mathbb{K}), \det(A) = 1\}$$

$$SO_n(\mathbb{R}) = \{A \in O_n(\mathbb{R}), \det(A) = 1\}$$

$$SU_n(\mathbb{C}) = \{A \in U_n(\mathbb{C}), \det(A) = 1\}$$

**9.5.1 Proposition:** *Ce sont des groupes.*

**Preuve:**  $\det(A) = 1$  et  $\det(B) = 1$ ,  $\det(AB) = \det(A)\det(B) = 1$ . C'est la seule chose à vérifier. Après ce qu'on a déjà vu, c'est tout ce qu'il y a à dire sur le sujet.  $\square$

## 9.6 Matrices triangulaires supérieures inversibles

**9.6.1 Definition:** Si  $T$  est triangulaire supérieure, Pour tout  $i > j$ ,  $t_{ij} = 0$ .

**9.6.2 Proposition:** *Cet ensemble est un groupe.*

**Preuve:** Soit  $T_1, T_2$  2 matrices triangulaires supérieures : soit  $A = T_1 T_2$  :

$$a_{i,j} = \sum_{k=1}^n t_{1,ik} t_{2,kj}$$

avec :  $t_{1,ik} = 0 \forall k$  tel que  $i > k$  et  $t_{2,kj} = 0 \forall k$  tel que  $k > j$ . Si  $i > j$ , alors soit  $i > k$  et  $t_{1,ik} = 0$ , soit  $k \geq i > j$  et  $t_{2,kj} = 0$ , donc  $a_{i,j} = 0 \forall i > j$  et donc :  $T_1 T_2$  est triangulaire supérieure.

Montrons que l'inverse d'une matrice triangulaire supérieure inversible est triangulaire supérieure : pour cela on utilise le mineur et la comatrice. On note  $T = (t_{k,l})_{1 \leq k, l \leq n}$ . Exprimons le mineur :  $T_{ij} = (s_{k,l})_{1 \leq k, l \leq n}$  avec :

-Si  $k < i$  et  $l < j$  :  $s_{k,l} = t_{k,l}$

-Si  $k \geq i$  et  $l < j$  :  $s_{k,l} = t_{k+1,l}$

-Si  $k < i$  et  $l \geq j$  :  $s_{k,l} = t_{k,l+1}$

-Si  $k \geq i$  et  $l \geq j$  :  $s_{k,l} = t_{k+1,l+1}$

Pour s'en convaincre, il suffit d'écrire le mineur (et revoir sa définition si besoin) : je renvoie ici à un cours d'algèbre de première année qui explique en détail les choses à savoir sur le déterminant (par exemple, le site mp ddl, cours sup, déterminant, développement selon une rangée). Le cofacteur est donc :

$$|T_{i,j}| = \sum_{\sigma \in S_{n-1}} \text{signe}(\sigma) \prod_{p=1}^n s_{p,\sigma(p)}$$

Si  $i < j$  :  $\exists m$  tel que  $s_{m,\sigma(m)} = 0$ , sauf si  $\sigma = \text{Id}$ . En effet :  $\forall \sigma \in S_{n-1}$ ,  $\exists m$  tel que  $\sigma(m) < m$ . On le prouve par l'absurde : Si

$$\sigma(k) \geq k \forall k, \sigma(n-1) \geq n-1 \Rightarrow \sigma(n-1) = n-1, \text{ puis } \sigma(n-2) \geq n-2 \Rightarrow \sigma(n-2) = n-2$$

, puis finalement,  $\sigma = \text{Id}$  (on peut le démontrer rigoureusement par récurrence).

Donc : l'existence de  $m$  étant prouvée : supposons  $i < j$ , afin de prouver que dans ce cas, le mineur est nul (et donc que la comatrice est triangulaire inférieure).

-Si  $m < i$ ,  $\sigma(m) < m < i < j$  et  $s_{m,\sigma(m)} = t_{m,\sigma(m)} = 0$

-Si  $j > m \geq i$  :  $\sigma(m) < j$  donc  $s_{m,\sigma(m)} = t_{m+1,\sigma(m)} = 0$  car  $m+1 \geq \sigma(m)$

-Si  $\sigma(m) \geq j$ ,  $m > j$  car  $m \geq \sigma(m)$ . Donc  $s_{m,\sigma(m)} = t_{m+1,\sigma(m)+1} = 0$

-Si  $\sigma = \text{Id}$  :  $\sigma(i) = i < j$  et  $s_{i,\sigma(i)} = s_{i,i} = t_{i+1,i} = 0$  car  $i+1 > i$ .

Donc  $\text{com}(A)$  est triangulaire inférieure. Pour finir, on utilise la formule  $A^{-1} = \frac{t_{\text{com}}(A)}{\det(A)}$ ,  $t_{\text{com}}(A)$  est triangulaire supérieure et  $\det(A) \neq 0$  ce qui termine la preuve.

Pour l'instant, voici les groupes matriciels que j'ai trouvé utile d'introduire dès le début, car ils sont essentiels. Dans la suite de ce TIPE, nous croiserons tout de même d'autres groupes. Avant de terminer cette section, je vais faire un bref rappel : On appelle matrice symétrique une matrice telle que  ${}^t A = A$ , antisymétrique une matrice telle que  ${}^t A = -A$ , hermitienne une matrice telle que  $A^* = A$  et antihermitienne une matrice telle que  $A^* = -A$ . On vérifie aisément qu'en notant respectivement  $Sym_n$ ,  $ASym_n$ ,  $Herm_n$ ,  $AHerm_n$  l'ensemble de ces matrices de taille fixée (dans  $M_n(\mathbb{K})$  avec  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$ ) : chacun de ces ensembles est un espace vectoriel. Comme c'est très simple, je vais juste donner les 2 propriétés principales à utiliser :  ${}^t(A + B) = {}^t A + {}^t B$  et  $\overline{(A + B)} = \overline{A} + \overline{B}$ .  $\square$

## 10 Annexe II : Topologie

Le but de cette annexe est de donner des rappels de topologie qui sont nécessaires : dans la définition rigoureuse d'une variété, dans la preuve du théorème de plongement de Whitney, dans l'étude topologique des groupes, et également, parfois, dans certains résultats où utiliser ces notions rend une preuve plus simple. Certaines notions sont niveau L2, tandis que d'autres sont un peu plus précises (on introduit de manière plus rigoureuse les compacts).

### 10.1 Espaces topologiques

**10.1.1 Définition:** (espaces topologiques) Soit  $X$  un ensemble, et désignons par  $\mathcal{P}(X)$  l'ensemble de ses parties. Une *topologie* sur  $X$  est un sous-ensemble  $\tau \subset \mathcal{P}(X)$  ( $\tau$  est un certain ensemble de parties de  $X$ ) qui vérifie :

$$(1) \emptyset, X \in \tau$$

$$(2) \text{Si } \{U_i\}_{i \in I} \subset \tau, \text{ alors } \bigcup_{i \in I} U_i \in \tau.$$

$$(3) \text{Si } U_1, \dots, U_n \in \tau, \text{ alors } \bigcup_{j=1}^n U_j \in \tau.$$

Les éléments  $\tau$  sont appelés les ouverts de la topologie. Les conditions (1), (2) et (3) sont appelés les axiomes de topologie.

**10.1.2 Exemple:** (1)  $X$  un ensemble quelconque,  $\tau = \{\emptyset, X\}$ . On l'appelle *topologie grossière*, elle contient le minimum possible d'ouverts.

(2)  $X$  un ensemble quelconque,  $\tau = \mathcal{P}(X)$ . On l'appelle *topologie discrète*, elle contient le maximum possible d'ouverts.

(3) Soit  $(X, \tau_X)$  un espace topologique et  $A \subset X$  un sous-ensemble. On définit la topologie induite sur un sous-ensemble  $\tau_A$  sur  $A$  en posant

$$\tau_A = \{U \cap A \mid U \in \tau_X\}$$

Autrement dit, on prend comme ouverts de  $A$  les intersections d'ouverts de  $X$  avec  $A$ .

**10.1.3 Définition:** Un ensemble  $B$  de parties de  $X$  est une base de topologie si et seulement si il vérifie les 2 conditions suivantes :

$B$  est un recouvrement de  $X$

Soient  $B_1$  et  $B_2$  deux éléments de  $B$  et  $I$  leur intersection. Pour tout  $x \in I, \exists B_3 \in B$  tel que  $x \in B_3$  et  $B_3 \subset I$ .

**10.1.4 Exemple:**

(1)  $\mathbb{R}^n$  est à base dénombrables d'ouverts.

**Preuve:** Montrons le d'abord pour  $\mathbb{R}$ . Soit  $U$  l'ensemble des ouvert de  $\mathbb{R}$  d'extrémités rationnelles. C'est clairement un recouvrement. De plus, l'intersection de deux ouverts d'extrémités rationnelles est lui-même un ouvert d'extrémités rationnelles (ou l'ensemble vide), ce qui montre la seconde condition. Par produit cartésien, on montre de la même façon que  $\mathbb{R}^n$  est à base dénombrable.  $\square$

**10.1.5 Proposition:** *Tout sous-ensemble  $Y$  d'un ensemble  $X$  à base dénombrable d'ouvert est lui-même à base dénombrable d'ouvert.*

**Preuve:** Soit  $B$  une base dénombrable d'ouverts de  $X$ , et  $\{B_i\}_{i \in I}$  ses éléments. Si  $Y$  est ouvert,  $B' = \{B_i \cap Y\}_{i \in I}$  est une base dénombrable d'ouverts de  $Y$ . Si  $Y$  n'est pas ouvert, on peut avoir des problèmes de bord (voir cette remarque dans la sous-partie sur les variétés). On se restreint alors à  $\text{int}(Y)$ , qui est un ouvert, et  $B' = \{B_i \cap \text{int}(Y)\}_{i \in I}$  est bien une base dénombrable d'ouverts de  $\text{int}(Y)$ .  $\square$

**10.1.6 Définition:** Soit  $(X, \tau)$  un espace topologique,  $x \in X$ . On dit que  $V \subset X$  est un *voisinage de  $x$*  s'il existe un ouvert  $U_x$  tel que  $x \in U_x \subset V$ . On note  $V_x$  l'ensemble des voisinages de  $x$ .

Rappel sur  $\setminus$  : on rappelle que les ensembles si  $X$  et  $Y$  sont deux sous-ensembles d'un ensemble  $Z$ , on a par définition :  $X \setminus Y = X \cap C_Z Y$ ,  $C_Z Y$  étant le complémentaire dans  $Z$  de  $Y$ . Il ne faut surtout pas le confondre avec le quotient.

## 10.2 Fonctions sur des espaces topologiques

**10.2.1 Définition:** Soit  $(X, \tau_X), (Y, \tau_Y)$  2 espaces topologique,  $f : X \rightarrow Y$  une application et  $x \in X$ ,  $y = f(x)$ . On dit que  $f$  est *continue au point  $x$*  si

$$\forall V \in \tau_Y, y \in V, \exists U \in \tau_X, x \in U; t.q : f(U) \subset V$$

Autrement dit, pour tout ouvert  $V$  de  $Y$  contenant  $y$ , on peut trouver un ouvert  $U$  de  $X$  contenant  $x$ , dont l'image par  $f$  est contenue dans  $V$ .

**10.2.2 Proposition:** *Soient  $X$  et  $Y$  des espaces topologiques et  $f : X \rightarrow Y$  une application. Alors :  $f$  est continue  $\Leftrightarrow \forall V \in \tau_Y, f^{-1}(V) \in \tau_X$ .*

**Preuve:** Si  $f : X \rightarrow Y$  est continue et  $V \subset Y$  est ouvert, soit  $U = f^{-1}(V)$  et  $x \in U$ . Puisque  $f$  est continue au point  $x$ ,  $\exists U_x$  ouvert dans  $X, x \in U_x$  tel que  $f(U_x) \subset V$ , soit  $U_x \subset f^{-1}(V) = U$ . Alors  $f^{-1}(V) = U = \bigcup_{x \in U} U_x$  est une réunion d'ouverts, donc ouvert. Réciproquement, soit  $x \in X$  et  $V$  un ouvert de  $Y$  contenant  $f(x)$ , alors par hypothèse  $U = f^{-1}(V)$  est un ouvert de  $X$ , et  $x \in U, f(U) \subset V$ .  $f$  est donc continue en  $x$ .  $\square$

**10.2.3 Proposition:** *Soient  $f : X \rightarrow Y$  et  $g : Y \rightarrow Z$  des applications continues.*

(1) *La composition  $g \circ f : X \rightarrow Z$  est continue.*

(2) *Soit  $A \subset X$ , si on muni  $A$  de la topologie induite par celle de  $X$ , la restriction  $f|_A$  de  $f$  à  $A$  est continue.*

(3) *Soit  $B \subset Y$  tel que  $f(X) \subset B$ . Si on munit  $B$  de la topologie induite,  $f : X \rightarrow B$  est continue.*

**Preuve:** (1) Si  $W \in \tau_Z, g^{-1}(W) \in \tau_Y$  car  $g$  est continue, donc  $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$  est dans  $\tau_X$  car  $f$  est continue.

(2) Si  $V \in \tau_Y, f|_A^{-1}(V) = f^{-1}(V) \cap A$ , et  $f^{-1}(V)$  est ouvert dans  $X$  car  $f$  est continue et  $f^{-1}(V) \cap A$  est ouvert dans  $A$  par définition de la topologie induite.

(3) Il suffit de remarquer que, puisque  $f(X) \subset B$ , si  $V \in \tau_Y, f^{-1}(V) = f^{-1}(B \cap V)$ .  $\square$

**10.2.4 Definition:** Soient  $X$  et  $Y$  des espaces topologiques. on dit que  $f : X \longrightarrow Y$  est un homéomorphisme si :

- (1)  $f$  est continue.
- (2)  $f$  est une bijection, dont l'inverse est notée  $f^{-1} : Y \longrightarrow X$ .
- (3) L'application  $f^{-1} : Y \longrightarrow X$  est continue.

**10.2.5 Definition:** Soient  $X$  et  $Y$  deux espaces topologiques. On dit que l'application  $f : X \longrightarrow Y$  une *ouverte* si :

$\forall U \subset X$  ouvert,  $f(U)$  est ouvert dans  $Y$ .

On dit que  $f$  est *fermée* si

$\forall V \subset X$  fermé,  $f(V)$  est fermé dans  $Y$

**10.2.6 Proposition:** Soient  $X$  et  $Y$  deux espaces topologiques et  $f : X \longrightarrow Y$  un application. Les conditions suivantes sont équivalentes :

- (1)  $f$  est un homéomorphisme.
- (2)  $f$  est bijective continue et ouverte
- (3)  $f$  est bijective continue et fermée.

**Preuve:** Si  $f$  est bijective, désignons par  $g : Y \longrightarrow X$  son inverse. Alors, si  $A \subset X, g^{-1}(A) = f(A)$ . Il suffit d'appliquer les 2 propositions précédentes sur la continuité (à  $g$ ) (en prenant  $A$  ouvert ou  $A$  fermé selon les cas).  $\square$

## 10.3 Espaces compacts

### Motivation

Ici, on va avoir besoin de notions topologiques plus générales, on ne se limitera pas aux espaces métriques, mais à tous les espace topologiques. On va donc redéfinir un ouvert (de manière plus générale) (et on admettra que les propriétés sur la réunion et l'intersection d'ouverts sont encore valables). En fait, elles le sont, car dans un espace topologique, ces propriétés des ouverts sont en fait des axiomes et on déduit alors les mêmes propriétés pour les fermés par passage au complémentaire.

Mais on verra ensuite que se restreindre aux espaces métriques est suffisant concernant ce que l'on veut étudier : les variétés (voir le théorème énoncé à la fin). Donc pourquoi introduire des notions plus générales ?

Pour au moins 2 raisons :

- 1/ Les définitions dans les notions générales sont plus "fortes" que celles restreintes aux espaces métriques, plus faciles à utiliser. Elles prouvent plus simplement un résultat.
- 2/ Parce qu'en maths, il arrive souvent de tomber sur des ensembles "bizarres", qui ne vérifient pas les propriétés de régularité auxquelles on est habitué. Par exemple, un espace métrique est ce qu'on pourrait appeler un espace relativement "régulier", "simple" (voir définition plus bas), mais certains espaces ne le sont pas. Ainsi, dans les variétés, on définit 3 axiomes pour éviter de se casser la tête sur des notions de topologie trop compliquées. Mais, en pratique, ces notions sont intéressantes (car elles permettent de donner des propriétés des espaces moins "réguliers" que des espaces métriques, grâce aux nombreuses et incalculables définitions (et propositions) de la topologie).

### Espaces séparés

**10.3.1 Définition:** Soit  $X$  un espace topologique, on dit qu'il est *séparé* si, pour tout  $x, y \in X$  avec  $x \neq y$ , il existe 2 ouverts  $U_x$  et  $U_y$  tels que  $x \in U_x, y \in U_y$  et  $U_x \cap U_y = \emptyset$ . On dit que  $U_x$  et  $U_y$  séparent les points  $x$  et  $y$ .

Dans l'approche de F.Hausdorff de la notion d'espace topologique, la condition d'être séparé faisait partie de la définition. C'est pourquoi on dit parfois espace de Hausdorff pour un espace séparé.

**10.3.2 Exemple:** 1/ Tout espace métrique  $(X, d)$  est séparé : si  $x, y \in X; x \neq y$ , soit  $r = d(x, y)$ , on pose alors  $U_x = B(x, \frac{r}{2})$  et  $U_y = B(y, \frac{r}{2})$  (boules ouvertes).

2/ L'espace de Sierpinski  $S = \{0, 1\}$  muni de la topologie  $\tau = \{\{0\}, \{0, 1\}, \emptyset\}$  n'est pas séparé : tout ouvert contenant 1 contient aussi 0.

**10.3.3 Proposition:** Voici des propriétés des espaces séparés.

(1) Si  $X$  est séparé et  $A \subseteq X$  muni de la topologie induite, alors  $A$  est séparé.

(2) Soit  $\{X_i\}_i$  une famille d'espaces. Alors

$\prod_{i \in I} X_i$  est séparé  $\Leftrightarrow X_i$  est séparé  $\forall i \in I$ .

(3) Soit  $\{X_i\}_i$  une famille d'espaces non vides. Alors  $\prod_{i \in I} X_i$  est séparé  $\Leftrightarrow X_i$  est séparé  $\forall i \in I$ .

(4) Si  $\sim$  est une relation d'équivalence sur l'espace topologique  $X$ , alors :

$X/\sim$  est séparé  $\Leftrightarrow \forall x, y \in X$ , tels que  $x$  et  $y$  ne soient pas dans la même classe d'équivalence, il existe des ouverts saturés  $U_x, U_y \subset X$  tels que  $U_x \cap U_y = \emptyset$ .

**Preuve:** (1) Si  $x, y \in A$ ,  $x \neq y$  et  $U_x, U_y \subset X$  sont des ouverts de  $X$  et  $Y$  qui séparent  $x$  et  $y$ , alors  $U_x \cap A$  et  $U_y \cap A$  sont des ouverts de  $A$  qui séparent  $x$  et  $y$ .

Je ne vais pas prouver le reste, car il faudrait pour cela que j'ai défini  $\prod \prod$  et la notion de saturation. En fait (1) sera le plus utile pour la suite.  $\square$

### Espaces compacts

**10.3.4 Définition:** Soit  $X$  un espace topologique. Un recouvrement de  $X$  par des ouverts est une famille  $\{U_i\}_{i \in I}$  d'ouverts de  $X$  tels que

$$X = \bigcup_{i \in I} U_i$$

**10.3.5 Définition:** On dit que l'espace  $X$  est *compact* s'il est séparé et de plus satisfait la condition suivante, qu'on appelle *condition de compacité* : pour tout recouvrement  $\{U_i\}_{i \in I}$  de  $X$  par des ouverts, on peut trouver un sous-ensemble fini  $J \subset I$  tel que

$$X \subset \bigcup_{i \in J} U_i$$

On dit que  $\{U_i\}_{i \in J}$  est un *sous-recouvrement extrait* du recouvrement  $\{U_i\}_{i \in I}$ .

On peut donc reformuler la condition précédente en disant que de tout recouvrement de  $X$  par des ouverts, on peut extraire un sous-recouvrement fini.

En passant aux complémentaires, on peut reformuler cette condition en termes de fermés : si  $\{F_i\}_{i \in I}$  est une famille de fermés de  $X$  telle que  $\bigcap_{i \in I} F_i = \emptyset$ , alors il existe un sous-ensemble fini  $J \subset I$  tel que



$\bigcap_{i \in J} F_i = \emptyset$ . Autrement dit, si l'intersection d'une famille de fermés est vide, il existe une sous-famille dont l'intersection est vide.

Dans la pratique, on se trouve souvent dans la situation où  $X$  est un sous-espace d'un espace  $Y$ , et que le recouvrement de  $X$  est donné par une famille  $\{V_i\}_{i \in I}$  d'ouverts de  $Y$ ; strictement parlant, le recouvrement de  $X$  est la famille  $\{V_i \cap X\}_{i \in I}$ . La condition de compacité s'exprime alors en disant qu'il existe  $J \subset I$  fini tel que  $X \subset \{V_i\}_{i \in J}$ . En termes de fermés, si  $\{F_i\}_{i \in I}$  est une famille de fermés de  $Y$ , la compacité de  $X$  s'exprime en disant que si  $X \cap (\bigcap_{i \in I} F_i) = \emptyset$ , alors il existe  $J \subset I$  fini tel que  $X \cap (\bigcap_{i \in J} F_i) = \emptyset$ .

On va maintenant définir une notion liée : la paracompacité, qui est une notion capitale dans les variétés :

Un *raffinement* de  $\{U_r\}$  est un recouvrement  $\{V_s\}$  tel que chaque ouvert  $V_s$  est contenu dans un ouvert  $U_r$ .

Un recouvrement est *localement fini* si tout point est contenu dans un nombre fini d'ouverts du recouvrement.

L'espace  $M$  est *paracompact* s'il est de Hausdorff et tout recouvrement admet un raffinement localement fini.

Un recouvrement est à *base dénombrable* s'il existe un recouvrement contenant un nombre dénombrable d'ouverts. Dans ce cas, tous les recouvrements admettent un raffinement localement compact. Donc un espace de Hausdorff et à base dénombrable est paracompact.

Voici quelques propriétés des espaces compacts :

### 10.3.6 Proposition:

(1) Soit  $X$  un espace compact et soit  $F \subset X$  un sous-espace fermé, alors  $F$  est compact.

(2) Soit  $X \subset Y$ ; si  $X$  est compact et  $Y$  est séparé, alors pour tout  $y \in Y \setminus X$  il existe des ouverts  $V_y$  et  $U_X$  de  $Y$  tels que  $y \in V_y$ ,  $X \subset U_X$  et  $V_y \cap U_X = \emptyset$ . En particulier,  $X$  est fermé dans  $Y$ .

**Preuve:** (1)  $F$  est séparé d'après la proposition précédente (le (1)). Si  $\{U_i\}_{i \in I}$  est une famille d'ouverts de  $X$  tels que  $F \subset \bigcup_{i \in I} U_i$ , alors si on ajoute l'ouvert  $X \setminus F$  à la famille  $\{U_i\}_{i \in I}$ , on obtient un recouvrement

de  $X$  par des ouverts, dont on peut extraire un sous-recouvrement fini, de la forme  $\{U_i\}_{i \in J}$ ,  $J \subset I$ , plus éventuellement l'ouvert  $X \setminus F$ . Dans tous les cas,  $\{U_i\}_{i \in J}$ ,  $J \subset I$  est un recouvrement fini de  $F$ .

(2) Soit  $y \in Y \setminus X$ . Comme  $Y$  est séparé, pour tout  $x$ , on peut trouver des ouverts  $U_x$  et  $V_y^x$ ,  $x \in U_x$ ,  $y \in V_y^x$  tels que  $U_x \cap V_y^x = \emptyset$ . Évidemment, la famille d'ouverts  $\{U_x\}_{x \in X}$  recouvre  $X$ ; on peut donc extraire un sous-recouvrement fini  $\{U_{x_1}, \dots, U_{x_k}\}$  de  $X$ . Posons  $U_X = U_{x_1} \cup \dots \cup U_{x_k}$  et  $V_y = \bigcap_{h=1, \dots, k} V_y^{x_h}$ ; alors  $U$  et  $V$  sont des ouverts,  $X \subset U_X$ ,  $y \in V_y$  et  $U_X \cap V_y = \emptyset$ , car si  $z \in U_x$ ,  $\exists h \in \mathbb{N}$  tel que  $z \in U_{x_h}$  et comme  $U_{x_h} \cap V_y^{x_h} = \emptyset$ ,  $z \notin V_y$ . En particulier,  $V_y \cap X = \emptyset$ , donc  $Y \setminus X = \bigcup_{y \in Y \setminus X} V_y$  est un ouvert.  $\square$

Remarque : on vient de prouver (le point (2)) que, sous les hypothèses de la proposition, on peut séparer  $y \in Y \setminus X$  et le fermé  $X$  par les ouverts  $V_y$  et  $U_X$ , utilisons donc ce résultat pour prouver :

### 10.3.7 Proposition: Soit $X$ un espace compact, alors :

(1)  $\forall x \in X$  et  $F \subset X$  fermé, il existe des ouverts  $U_x, U_F \subset X$  tels que  $x \in U_x$ ,  $F \subset U_F$  et  $U_x \cap U_F = \emptyset$ .

(2) Soient  $F_1, F_2 \subset X$  2 fermés. Il existe des ouverts  $U_1, U_2 \subset X$  tels que  $F_1 \subset U_1$ ,  $F_2 \subset U_2$  et  $U_1 \cap U_2 = \emptyset$ .

**Preuve:** (1) est un cas particulier du point (2) de la proposition précédente, en posant  $X = F$  et en remplaçant l'espace  $Y$  par  $X$ . Montrons (2) : soit  $y \in F_2$ , d'après (1), il existe des ouverts  $U_{F_1}^y$  et  $U_y$  tels

que  $F_1 \subset U_{F_1}^y$  et  $y \in U_y$  avec  $U_{F_1}^y \cap U_y$ . D'après le (1) de la proposition précédente,  $F_2$  est compact, et les ouverts  $\{U_y\}_{y \in Y}$  forment un recouvrement de  $F_2$ , on peut donc extraire un recouvrement fini  $U_{y_1}, \dots, U_{y_n}$ . Posons  $U_1 = \bigcap_{h=1, \dots, k} U_{F_1}^{y_h}$  et  $U_2 = U_{y_1} \cup \dots \cup U_{y_n}$ . Alors,  $F_1 \subset U_1, F_2 \subset U_2$  et  $U_1 \cap U_2 = \emptyset$ .  $\square$

**10.3.8 Définition:** Un espace topologique séparé vérifiant la propriété que l'on peut séparer un point et un espace fermé est appelé *régulier*. Un espace topologique dans lequel on peut séparer deux fermés disjoints est appelé *normal*. Ainsi, on peut résumer en la proposition en disant que tout espace compact est normal.

**10.3.9 Théorème:** Soit  $f : X \rightarrow Y$  une application continue. Si  $X$  est compact et  $Y$  séparé, alors  $f(X)$  est compact.

**Preuve:** Puisque  $Y$  est séparé,  $f(X)$  l'est aussi, par le (1) de la proposition sur les espaces séparés. Condition de compacité : supposons que  $f(X) \subset \bigcup_{i \in I} U_i$ , où  $\{U_i\}_{i \in I}$  est une famille d'ouverts de  $Y$ . Puisque  $f$  est continue,  $f^{-1}(U_i)$  est ouvert dans  $X, \forall i \in I$ . Donc  $\{f^{-1}(U_i)\}_{i \in I}$  est un recouvrement de  $X$  par des ouverts, dont on peut extraire un sous-recouvrement fini :  $X = \bigcup_{i \in J} f^{-1}(U_i)$ , où  $J \subset I$  est un sous-ensemble fini. Donc les  $\{U_i\}_{i \in J}$  forment un recouvrement fini de  $f(X)$ .  $\square$

**10.3.10 Corollaire:** Soit  $f : X \rightarrow Y$  une application continue, où  $X$  est compact et  $Y$  séparé. Alors  $f$  est fermée.

**Preuve:** Prenons un fermé  $F \subset X$  : d'après l'une des propositions précédentes, il est compact, et d'après la proposition juste au-dessus,  $f(F)$  est compact, donc fermé dans  $Y$ .  $\square$

**10.3.11 Corollaire:** Soit  $f : X \rightarrow Y$  une application continue, bijective, où  $X$  est compact et  $Y$  séparé. Alors  $f$  est un homéomorphisme.

**Preuve:**  $f$  étant fermée, la proposition sur le lien entre applications ouvertes, fermées et homéomorphismes permet de conclure.  $\square$

### Espaces localement compacts

**10.3.12 Définition:** On dit que l'espace topologique séparé  $X$  est *localement compact* si tout point  $x \in X$  possède un voisinage compact. En symboles :

$$\forall x \in X, \exists V \in V_x$$

avec  $V$  compact et  $V_x$  un voisinage quelconque de  $x$ .

Par exemple,  $\mathbb{R}^n$  est localement compact, car  $\forall a \in \mathbb{R}^n$ , la boule fermée  $\{a \in \mathbb{R}^n \mid \|x - a\| \leq 1\}$  est compacte. Aussi, si  $X$  est compact, il est voisinage de chacun de ses points, donc  $X$  est localement compact.

Dans la définition de localement compact, on exige l'existence d'un voisinage compact pour tout point, mais cela implique en fait l'existence de beaucoup de voisinages compacts de chaque point (c'est la proposition suivante, qu'on va montrer en utilisant 2 lemmes).

**10.3.13 Lemme:** Soit  $X$  un espace compact,  $x \in X$  et  $V \in V_x$  un voisinage de  $x$ . alors il existe un voisinage compact  $V' \subset V$  de  $x$ .

**Preuve:** Soit  $V_O \subset V$  un ouvert tel que  $x \in V_O$ . Alors  $X \setminus V_O$  est fermé. Donc, d'après la proposition, il existe des ouverts disjoints  $U_V$  et  $U_x$ ,  $U_V \subset X \setminus V_O$ ,  $x \in U_x$ . Posons  $V' = \overline{U_x}$  (l'adhérence); c'est un voisinage fermé, donc compact, de  $x$ , et  $V' \cap (X \setminus V_O) = \emptyset$ , ce qui fait que  $V' \subset V_O \subset V$ .  $\square$

Illustration des différents ensembles qui apparaissent dans le lemme.

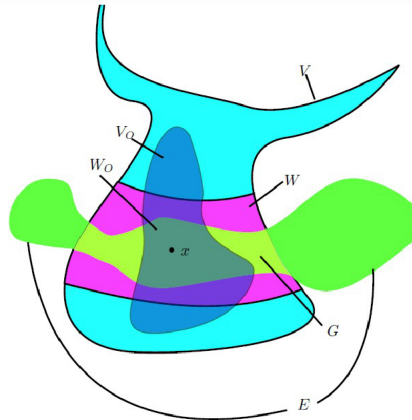


Figure III.1: Les divers ensembles qui apparaissent dans le lemme III.4.5

**Remarque :** Si  $X$  est compact et  $a \in X$ , il suit du lemme précédent que  $X \setminus a$  est localement compact. En effet, si  $x \in X \setminus a$ , l'ensemble  $X \setminus a$  est un voisinage de  $x$  dans  $X$ , qui contient donc un voisinage compact  $V$  de  $x$  dans  $X$ . Comme  $X \setminus a$  est ouvert dans  $X$ ,  $V$  est aussi un voisinage de  $x$  dans  $X \setminus a$ , et il est compact.

La construction du compactifié d'Alexandrov, un peu plus loin, montre que tous les espaces localement compacts sont de cette forme.

**10.3.14 Lemme:** Soit  $X$  un espace topologique,  $x \in X$ ,  $V$  un voisinage de  $x$  dans  $X$  et  $W \subset V$  un voisinage de  $x$  dans  $V$ . Alors  $W$  est un voisinage de  $x$  dans  $X$ .

**Preuve:** Soit  $V_O \subset V$  un ouvert de  $X$  tel que  $x \in V_O$ , et soit  $G \subset V$  un ouvert de  $V$  tel que  $x \in G \subset W$ . Il existe un ouvert  $E$  de  $X$  tel que  $E \cap V = G$ . Posons  $W_O = E \cap V_O$ , c'est un ouvert de  $X$ , et  $x \in W_O \subset G \subset W$ , ce qui prouve que  $W$  est un voisinage de  $x$ .  $\square$

**10.3.15 Proposition:** Soit  $X$  un espace topologique localement compact. Alors, pour tout  $x \in X$  et tout voisinage  $V$  de  $x$ , il existe un voisinage compact  $V'$  de  $x$  tel que  $V' \subset V$ .

**Preuve:** Soit  $V$  un voisinage de  $x$  et  $K$  un voisinage compact de  $x$  dans  $V$ .  $V \cap K$  est un voisinage de  $x$  dans  $K$ , il suit du premier lemme qu'il existe un voisinage  $V' \subset V \cap K$  de  $x$  dans  $K$ . Et d'après le second lemme,  $V'$  est un voisinage de  $x$  dans  $X$ .  $\square$

Voici une proposition qui nous dit que, dans un espace localement compact, dans une certaine mesure, la topologie est déterminée par les sous-espaces compacts :

**10.3.16 Proposition:** Soit  $X$  un espace localement compact. Alors  $F \subset X$  est fermé si et seulement si pour tout  $K \subset X$  compact,  $F \cap K$  est fermé.

**Preuve:** Si  $F$  est fermé,  $F \cap K$  est fermé car c'est l'intersection de 2 fermés de  $X$ . Réciproquement, supposons que  $F \cap K$  soit fermé pour tout compact  $K \subset X$  et soit  $x \in X \setminus F$ . Soit  $V$  un voisinage compact de  $x$ , alors  $V \cap F$  est fermé et il existe un voisinage  $W$  de  $x$  dans  $V$  tel que  $W \cap (F \cap V) = \emptyset$ . Comme on l'a vu avec le second lemme,  $W$  est aussi un voisinage de  $x$  dans  $X$ , et  $W \cap F = \emptyset$ , ce qui montre que  $F$  est fermé dans  $X$ .  $\square$

### Applications propres

**10.3.17 Définition:** Soient  $X$  et  $Y$  des espaces localement compacts et  $f : X \rightarrow Y$  une application. On dit que  $f$  est *propre* si elle est continue et de plus :

$\forall K \subset Y$  compact,  $f^{-1}(K) \subset X$  est compact.

**10.3.18 Proposition:** Soient  $(X, d_X)$  et  $(Y, d_Y)$  des espaces métriques localement compacts et  $f : X \rightarrow Y$  une application continue.  $f$  est propre si et seulement si la condition suivante est vérifiée :  
Pour toute suite  $\{x_n\}_{n \in \mathbb{N}} \subset X$  qui n'admet pas de valeur d'adhérence dans  $X$ , la suite  $\{f(x_n)\}_{n \in \mathbb{N}}$  n'admet pas de valeur d'adhérence.

**Preuve:** Supposons que  $f$  soit propre, soit  $\{x_n\}_{n \in \mathbb{N}} \subset X$  une suite et supposons qu'on puisse extraire une suite  $\{f(x_{n_k})\}$  qui converge vers  $y \in Y$ . alors  $K = \{f(x_{n_k}), k \in \mathbb{N}\} \cup \{y\}$  est compact, donc  $f^{-1}(K)$  aussi, mais  $f^{-1}(K) \subset \{x_{n_k}, k \in \mathbb{N}\}$ , ce qui implique qu'on peut extraire une sous-suite convergente de la suite  $\{x_{n_k}\}$ , donc de la suite  $\{x_n\}$ .

Par contraposée, on obtient le résultat annoncé.  $\square$

**10.3.19 Corollaire:** Soit  $f : \mathbb{R}^k \rightarrow \mathbb{R}^p$  une application continue. Alors  $f$  est propre si et seulement si pour toute suite  $\{x_n\} \subset \mathbb{R}^k$  qui tend vers l'infini, la suite  $\{f(x_n)\}_{n \in \mathbb{N}} \subset \mathbb{R}^p$  tend vers l'infini.

**Preuve:** Il suffit de remarquer que dans  $\mathbb{R}^k$ , dire qu'une suite tend vers l'infini équivaut à dire qu'elle n'a pas de valeur d'adhérence, ce qui équivaut encore à dire qu'on ne peut en extraire une suite qui converge. Réciproquement, si une suite  $\{x_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^k$  n'a pas de valeur d'adhérence, pour tout  $R > 0$  l'ensemble  $\{n \in \mathbb{N} \mid \|x_n\| \leq R\}$  est fini, sinon, toute boule fermée  $\overline{B}(0, R)$  étant compacte, on pourrait extraire  $\{x_n\}_{n \in \mathbb{N}}$  une suite convergente dans cette boule. On pose  $N_R = \sup \{n \in \mathbb{N} \mid \|x_n\| \leq R\} + 1$ , et alors  $\|x_n\| > R$  si  $n \geq N_R$ , c'est à dire :  $\{x_n\}$  tend vers l'infini.  $\square$

**Exemples :** (1) Soit  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = \sum_{i=0}^d a_i x^i$ ,  $a_i \in \mathbb{R}$  une application polynomiale non constante, on peut supposer  $a_d \neq 0$  et  $d \geq 1$ . Alors elle est propre car  $f(x) = x^d (\sum_{i=0}^d a_i x^{i-d})$ , de sorte que  $\lim_{x \rightarrow \infty} |f(x)| = \lim_{x \rightarrow \infty} |a_d x^d| = \infty$ , par conséquent, il en est de même pour toute suite  $x_n$  de limite  $\pm \infty$  (par continuité de  $f$ ).

(2) On utiliserait le même argument pour montrer qu'une fonction polynomiale complexe est propre.

(3) L'application  $\sin(x) : \mathbb{R} \rightarrow [-1, 1]$  n'est pas propre, car l'image réciproque de  $\{0\}$  est  $\{k\pi, k \in \mathbb{Z}\}$ , qui n'est pas compact.

(4) L'exponentielle  $e^x : \mathbb{R} \rightarrow \mathbb{R}$  n'est pas propre : la suite  $x_n = -n$  tend vers  $-\infty$ , mais  $\lim_{n \rightarrow +\infty} e^{x_n} = 0$ .

(5) L'application  $f : \mathbb{R} \rightarrow \mathbb{R}_+^*$  est propre, car c'est un homéomorphisme, l'inverse étant le logarithme népérien.

(6) L'application  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(t_1, \dots, t_n) = t_1^2 + \dots + t_n^2 = \|t\|^2$  est propre, car si  $\{x_n\} \subset \mathbb{R}^n$  est une suite qui tend vers l'infini,  $f(x_k) = \|x_k\|^2$  tend vers l'infini.

(7) L'application  $(x, y) \rightarrow x^2 - y^2$  n'est pas propre car la suite  $x_n = (n, n)$ , quand  $n$  tend vers  $\infty$  a une norme qui tend vers  $\infty$  alors que  $f(x_n) = 0 \forall n \in \mathbb{N}$ .

Voici une bonne raison de s'intéresser à ces applications :

**10.3.20 Théorème:** Soient  $X$  et  $Y$  des espaces localement compact et  $f : X \rightarrow Y$  propre. Alors  $f$  est fermée.

**Preuve:** Soit  $F \subset X$  fermé et  $K \subset Y$  compact. Alors  $f^{-1}(K)$  est compact, donc  $F \cap f^{-1}(K)$  est aussi compact. Mais alors  $f(F) \cap K = f(F \cap f^{-1}(K))$  est compact, donc fermé dans  $Y$ . Il suit alors que  $f(F)$  est fermé dans  $Y$ .  $\square$

## 10.4 Partitions de l'unité

**10.4.1 Définition:** On appelle partition de l'unité d'un espace topologique  $X$  une famille  $(\rho_i)_{i \in I}$  de fonctions continues, définies sur  $X$  et à valeurs dans l'intervalle  $[0, 1]$ , tel que pour tout point  $x \in X$ , les 2 conditions suivantes soient satisfaites :

1. Il existe un voisinage de  $x$  tel que toutes les fonctions  $\rho_i$  soient nulles sur ce voisinage à l'exception d'un nombre fini d'entre elles.
2.  $\sum_{i \in I} \rho_i(x) = 1, \forall x \in X$

**10.4.2 Définition:** On appelle partition de l'unité *subordonnée au recouvrement*  $(U_i)_{i \in I}$  une partition de l'unité  $(\rho_i)_{i \in I}$  au sens de la définition précédente, indexé par le même ensemble  $I$  que le recouvrement, et tel que, pour tout  $i \in I$ ,  $\text{supp} \rho_i \subseteq U_i$ .

Le théorème énoncé à la fin de cette section permet dans de nombreux cas de remplacer une question globale par une question locale. Il joue surtout un rôle crucial dans l'étude des variétés et des sous-variétés.

Tout d'abord il est important de remarquer qu'il existe des fonctions  $C^m$  à support compacte (ici  $m \leq \infty$  mais  $m \neq \omega$ ). En effet,

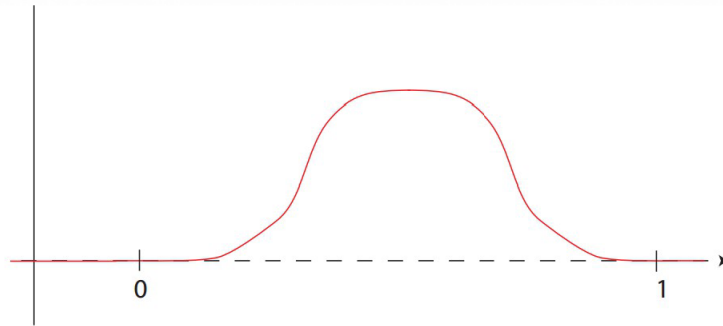
**10.4.3 Lemme:** La fonction définie sur  $\mathbb{R}$  par  $f(t) = e^{-1/t}$  pour  $t \geq 0$  et par  $f(t) = 0$  pour  $t < 0$  est  $C^\infty$ .

**Preuve:** On montre par récurrence que pour  $t > 0$ ,  $f^{(p)}(t)$  est de la forme  $\frac{Q(t)}{t^{2p}e^{-1/t}}$  où  $Q$  est un polynôme. En effet, cela est vrai pour  $p = 0$  et la dérivée de  $\frac{Q(t)}{t^p e^{-1/t}}$  est

$$\frac{tQ'(t) - 2pQ(t)}{t^{2p+1}e^{-1/t}} - \frac{Q(t)}{t^{2p+2}e^{-1/t}}$$

Il en résulte que  $f^{(p)}(t)$  est bien définie pour  $t > 0$  et que  $\lim_{t \rightarrow 0^+} f^{(p)}(t) = 0$ . On en déduit que  $f$  est  $p$ -fois dérivable en 0, et que sa  $p$ -ième dérivée est continue. Donc  $f$  est de classe  $C^p$  pour tout  $p$ , et elle est finalement de classe  $C^\infty$ .  $\square$

On peut alors construire une fonction  $g$  telle que  $g(t) = 0$  pour  $t \leq 0$  et  $g(t) = 1$  pour  $t \geq 1$ . Il suffit de poser  $I = \int_{\mathbb{R}} f(t)f(1-t)dt$ . Alors  $I > 0$  et  $g(t) = \frac{1}{I} \int_{-\infty}^t f(s)f(1-s)ds$ .

FIGURE 1.1 – La fonction  $f(s)f(1-s)$ .

**10.4.4 Proposition:** Soit  $K$  un compact contenu dans  $\Omega$ . Il existe une fonction  $\varphi$  de  $C_0^\infty(\Omega)$  égale à 1 au voisinage de  $K$ .

**Preuve:** Soit  $\delta$  tel que  $B(x, \delta) \subset \Omega$  pour tout  $x$  de  $K$  (qui existe car  $d(x, \mathbb{C}\Omega)$  atteint son minimum  $> 0$  sur le compact  $K$ ). Alors posant

$$\psi_x(y) = 2g\left(1 - \frac{|y-x|^2}{\delta^2}\right)$$

et  $W_x = \{y \mid \psi_x(y) > 1\}$ , on a  $K \subset \bigcup_{x \in K} W_x$  et en extrayant un sous-recouvrement fini par les  $W_{x_j}$ , on pose  $\varphi(y) = g(\sum_j \psi_{x_j}(y))$ . Alors  $\varphi = 1$  sur  $K$  et  $\varphi = 0$  hors d'un  $\delta$ -voisinage de  $K$ .  $\square$

**10.4.5 Corollaire:** Si  $F$  est un fermé contenu dans un ouvert  $U$ , il existe une fonction  $C^\infty$  à support dans  $U$  égale à 1 sur  $F$ .

**Preuve:** On écrit  $F_j = F \cap \overline{B(0, j) / B(0, j-2)}$  et on voit que  $F$  est réunion des compacts  $F_j$ . La proposition précédente associe à  $F_j$  une fonction  $\varphi_j$  à support dans  $U_j = U \cap (B(0, j+1) / B(0, j-3))$ , et la somme  $\sum_j \varphi_j$  étant localement finie<sup>1</sup>,  $\varphi = g(\sum_j \varphi_j)$  répond à la question.  $\square$

**10.4.6 Théorème:** Soit  $X$  un sous-ensemble fermé de  $\mathbb{R}^n$  et  $(U_j)_{j \in \mathbb{N}}$  un recouvrement localement fini de  $X$ . Il existe alors des fonctions  $\chi_j$  dans  $C_0^\infty(U_j)$ , à valeurs dans  $[0, 1]$  vérifiant

$$\sum_j \chi_j(x) = 1 \text{ sur } X$$

**Preuve:** Il existe un recouvrement de  $X$  par des ouverts  $V_j$  tels que  $\overline{V_j} \subset U_j$  (recouvrement qui est automatiquement localement fini). Supposons en effet avoir défini  $V_1, \dots, V_m$  tels que

$$X \subset \bigcup_{1 \leq j \leq m} V_j \cup \bigcup_{k \geq m+1} U_k$$

On pose  $F_{m+1} = X / (\bigcup_{1 \leq j \leq m} V_j \cup \bigcup_{k \geq m+2} U_k)$ . On a par hypothèse  $F_{m+1} \subset U_{m+1}$ , on peut alors trouver un ouvert  $V_{m+1}$  tel que  $\overline{V_{m+1}} \subset U_{m+1}$ . On a alors

$$X \subset \bigcup_{1 \leq j \leq m+1} V_j \cup \bigcup_{k \geq m+2} U_k$$

On prend alors  $\varphi_j$  valant 1 sur  $\overline{V_j}$  et s'annulant hors de  $U_j$ . Alors  $\sum_j \varphi_j > 0$  sur  $\bigcup_j V_j$  et si  $\psi(x) = 1$  sur  $X$  et vaut 0 hors de  $\bigcup_j V_j$ , on a sur  $X$  l'égalité  $\sum_j \frac{\varphi_j(x)}{(1-\psi(x)) + \sum_j \varphi_j(x)} = 1$ . Les fonctions  $\chi_j(x) = \frac{\varphi_j(x)}{(1-\psi(x)) + \sum_j \varphi_j(x)}$  répondent donc à la question.  $\square$

**10.4.7 Théorème:** Soit  $K$  un compact de  $\mathbb{R}^n$  contenu dans la réunion d'ouverts  $\bigcup_\alpha U_\alpha$ . Il existe alors des fonction  $\chi_\alpha$  dans  $C_0^\infty(U_\alpha)$ , à valeurs dans  $[0, 1]$  vérifiant

$$\sum_\alpha \chi_\alpha(x) = 1, \forall x \in K$$

La somme étant supposée localement finie.

**Preuve:** On rappelle le théorème de Lebesgue : pour tout recouvrement de  $K$  il existe un  $\delta$  tel que les boules  $B(x, \delta)$  pour  $x$  dans  $K$  sont toutes contenues dans un  $U_\alpha$ . Posant  $\psi_x(y) = 2f(\delta^2 - |y - x|^2)$  et  $W_x = \{y \mid \psi_x(y) > 1\}$ , les  $W_x$  pour  $x$  dans  $K$  forment un recouvrement de  $K$  dont on extrait un sous-recouvrement fini par les  $W_{x_j}$ . On pose  $\psi_j = \psi_{x_j}$ . Alors  $\sum_j \varphi_j(y) > 1$  sur  $K$  et si  $\varphi$  est la fonction associée par la proposition précédente à  $K$ , les fonctions

$$\chi_j = \frac{\varphi_j}{(1 - \varphi) + \sum_j \varphi_j}$$

répondent au problème posé.  $\square$

## 10.5 Topologie dans les espaces métriques

### Espaces de Banach

**10.5.1 Définition:** Une *distance* est une application  $d : E \times E \rightarrow \mathbb{R}^+$  telle que :

$$\begin{aligned} \forall (a, b) \in E^2, d(a, b) &= d(b, a) \\ \forall (a, b) \in E^2, d(a, b) &= 0 \Leftrightarrow a = b \\ \forall (a, b, c) \in E^3, d(a, c) &\leq d(a, b) + d(b, c) \end{aligned}$$

**10.5.2 Définition:** Un espace  $M$  est un *espace métrique* si c'est un ensemble dans lequel on a défini une distance.

**10.5.3 Définition:** Un espace métrique est dit *complet* si toute suite de Cauchy de  $M$  a une limite dans  $M$ . Cette propriété dépend de la distance choisie. Il est donc important de toujours préciser celle-ci.

**10.5.4 Définition:** Un *espace de Banach* est un espace vectoriel normé sur un sous-corps  $\mathbb{K}$  de  $\mathbb{C}$ , complet pour la distance issue de sa norme ( $d(x, y) = \|x - y\|$  est toujours une distance si  $\|\cdot\|$  est une norme, la réciproque étant fautive, ce qui veut dire que la notion de distance est plus générale que celle de norme).

Pour un espace euclidien, la distance est  $d(x, y) = \|x - y\|$  avec  $\|\cdot\|$  la norme euclidienne. Un tel espace est donc complet.

**Suites extraites**

**10.5.5 Définition:** On appelle *suite extraite (ou sous-suite)* d'une suite  $u = (u_n)_{n \in \mathbb{N}}$  d'éléments de  $E$  toute suite  $v = (v_k)_{k \in \mathbb{N}}$  pour laquelle il existe une fonction  $\varphi : \mathbb{N} \rightarrow \mathbb{N}$  strictement croissante vérifiant :

$$\forall k \in \mathbb{N}, v_k = u_{\varphi(k)}$$

**10.5.6 Proposition:** Si  $w$  est une suite extraite de  $v$  qui est elle-même une suite extraite de  $u$ , alors  $w$  est une suite extraite de  $u$

**Preuve:** Il existe  $\varphi$  strictement croissante et  $\psi$  strictement croissante telles que

$$\forall k \in \mathbb{N}, w_k = v_{\varphi(k)}; v_k = u_{\psi(k)}$$

de sorte que

$$w_k = u_{\psi \circ \varphi(k)}$$

, et  $\phi = \psi \circ \varphi$  est donc strictement croissante par composition. □

**10.5.7 Théorème:** Si  $(u_n)$  converge vers  $l$ , alors toute suite extraite de  $(u_n)$  converge également vers  $l$ .

**Preuve:** Soit  $v_k = u_{\varphi(k)}$  une suite extraite de  $(u_n)$  telle que  $\lim_{n \rightarrow \infty} u_n = l$ . Soit  $\varepsilon > 0$ .  $\exists N \in \mathbb{N}$  tel que  $\forall n \geq N, \|u_n - l\| \leq \varepsilon$ . On peut montrer par récurrence, de par la strice croissance de  $\varphi$ , que  $\varphi_k \geq k$ . Donc, pour  $k \geq N$ ,  $\varphi(k) \geq k \geq N$

$$\Rightarrow \|v_k - l\| = \|u_{\varphi(k)} - l\| \leq \varepsilon$$

□

**10.5.8 Définition:** On appelle *valeur d'adhérence* de la suite  $u = (u_n)_{n \in \mathbb{N}}$  d'éléments de  $E$  toute limite d'une suite convergente extraite de  $u$ . On note  $Adh(u)$  l'ensemble des valeurs d'adhérence de la suite  $u$ .

**10.5.9 Théorème:** (admis) De toute suite réelle bornée, on peut extraire une sous-suite convergente.

Il s'agit d'un résultat classique (théorème de Bolzano-Weierstrass) que l'on voit souvent en première année. La preuve la plus courante utilise la dichotomie.

**Compacité dans les espaces métriques**

On munit  $\mathbb{R}^m$  de l'une des distances usuelles pour en faire un espace métrique, par exemple : la distance  $d((x_j), (y_j)) = \max_{1 \leq j \leq m} |x_j - y_j|$  (max de la soustraction des coordonnées). Comme les distances usuelles sont équivalentes (équivalence des normes en dimension finie), la notion d'ouvert ne dépend pas de la distance choisie.

**10.5.10 Théorème:** Les trois conditions suivantes sont équivalentes :

- (1) L'espace métrique  $(X, d)$  est compact.
- (2) Toute suite dans  $X$  possède au moins une valeur d'adhérence.
- (3) De toute suite dans  $X$  on peut extraire une suite qui converge.

Pour la démonstration, on doit d'abord établir deux lemmes.

**10.5.11 Lemme:** (Lemme 1) Soit  $(X, d)$  un espace métrique tel que toute suite possède au moins une valeur d'adhérence.

Alors pour tout recouvrement  $\{U_i\}_{i \in I}$  de  $X$  par des ouverts, il existe  $\varepsilon > 0$  tel que :

$$\forall x \in X \exists i \in I \quad t.q \quad B(x, \varepsilon) \subset U_i$$

Un nombre tel que  $\varepsilon$  dans l'énoncé précédent est appelé nombre de Lebesgue du recouvrement  $\{U_i\}_{i \in I}$ .



**Preuve:** du Lemme 1 : Si un tel  $\varepsilon$  n'existe pas, une suite  $\{x_n\}_{n \in \mathbb{N}}$  telle que  $B(x_n, \frac{1}{n}) \not\subset U_i, \forall i \in I$  Soit  $a \in X$  une valeur d'adhérence de la suite  $x_n$  ; il existe un ouvert  $U_{i_0}$  qui contient  $a$  ; alors pour  $k$  assez grand, que l'on peut prendre aussi tel que  $n_k > \frac{1}{\varepsilon}, x_{n_k} \in B(a, \varepsilon/2)$ , et alors  $B(x_{n_k}, \frac{1}{n_k}) \subset B(a, \varepsilon) \subset U_{i_0}$ , contradiction.  $\square$

**10.5.12 Lemme:** *Lemme 2 Soit  $(X, d)$  un espace métrique tel que toute suite possède au moins une valeur d'adhérence. Alors pour tout  $\varepsilon > 0$ , il existe un ensemble fini  $\{a_1, \dots, a_N\} \subset X$  tel que :*

$$X = \bigcup_{n=1}^N B(a_n, \varepsilon)$$

**Preuve:** Prenons  $a_1 \in X$  quelconque ; si  $X = B(a_1, \varepsilon)$ , le lemme est démontré. Sinon il existe  $a_2 \in X \setminus B(a_1, \varepsilon)$  ; on recommence avec  $a_1, a_2$  : si  $X = B(a_1, \varepsilon) \cup X = B(a_2, \varepsilon)$ , on a fini, sinon on prends  $a_3 \in X \setminus (B(a_1, \varepsilon) \cup X = B(a_2, \varepsilon))$ , et ainsi de suite.

Si le procédé ne s'arrête pas, on peut construire une suite infinie  $a_1, a_2, \dots, a_n, \dots$  telle que pour  $m \leq n, d(a_n, a_m) \geq \varepsilon$ . Il est clair que l'on ne peut pas extraire une suite qui converge d'une telle suite ; le procédé doit donc s'arrêter après un nombre fini de pas, ce qui démontre le lemme.  $\square$

**Preuve:** du théorème Si  $X$  est compact et  $\{x_n\}_{n \in \mathbb{N}} \subset X$  une suite, les fermés  $F_N = \overline{\{x_n, n \geq N\}}$  sont non vides et on peut appliquer une remarque précédente pour en conclure que  $\bigcap_{N \in \mathbb{N}} F_N$ , qui est l'ensemble des valeurs d'adhérence, est non vide. Ceci montre que (1)  $\Rightarrow$  (2), et il suit d'une proposition précédente que (2) et (3) sont équivalents.

Il reste à voir que (2)  $\Rightarrow$  (1). Soit  $\{U_i\}_{i \in I}$  un recouvrement de  $X$  par les ouverts ; soit  $\varepsilon$  un nombre de Lebesgue de ce recouvrement, fourni par le lemme 1, et soient  $a_1, \dots, a_N \in X$  les points fournis par le lemme 2. Si  $U_{i_n}$  est un ouvert qui contient  $a_{i_n}$ , les ouverts  $U_{i_1}, \dots, U_{i_N}$  recouvrent  $X$ .  $\square$

Comme application, on peut caractériser les compacts de  $\mathbb{R}^n$  en partant du théorème de Cantor qui affirme que si  $I_1 \supset I_2 \dots \supset I_k \supset \dots$  est une suite décroissante d'intervalles fermés et bornés de  $\mathbb{R}$ , alors  $\bigcap I_k \neq \emptyset$ . La preuve du théorème de Cantor est intimement liée à la construction des nombres réels : si  $I_k = [a_k, b_k]$ , on a que  $a_k \leq b_1$  et  $b_k \geq a_1$ , donc il suit des propriétés fondamentales des nombres réels que l'on peut poser

$$a = \sup \{a_k\}, b = \inf \{b_k\};$$

on aura  $a \leq b$  et  $\bigcap_{k \geq 1} I_k = [a, b]$ . Dans le cas où la longueur de  $I_k$  tend vers 0, on aura  $a = b$  et  $\bigcap_{k \geq 1} I_k = \{a\}$ .

Ici, l'espace  $K$  considéré est un espace métrique.

**10.5.13 Théorème:** *Toute suite bornée d'éléments de  $\mathbb{K}$  admet au moins une valeur d'adhérence.*

**10.5.14 Définition:** Une partie  $K$  de  $E$  est dite *compacte* si toute suite d'éléments de  $K$  possède au moins une valeur d'adhérence dans  $K$ , c'est à dire :

$$\forall (u_n) \in K^{\mathbb{N}}, \exists \varphi : \mathbb{N} \longrightarrow \mathbb{N}, tq : \lim_{n \rightarrow \infty} u_{\varphi(n)} = l$$

(avec  $\varphi$  strictement croissante).

**10.5.15 Théorème:** *Toute partie compacte est fermée et bornée.*

**Preuve:** Soit  $K$  une partie compacte. Montrons que  $K$  est fermée : soit  $(x_n)_{n \in \mathbb{N}}$  une suite convergente d'éléments de  $K$  et posons  $l$  sa limite. Puisque  $l$  est l'unique valeur d'adhérence de la suite  $(x_n)_{n \in \mathbb{N}}$  (toute suite extraite converge vers  $l$ ), alors, par définition d'un compact,  $l \in K$ . Montrons maintenant qu'elle est bornée : par l'absurde, si  $K$  n'était pas bornée, il existerait une suite  $(x_n)_{n \in \mathbb{N}}$  d'éléments de  $K$  telle que  $\|x_n\| > n, \forall n \in \mathbb{N}$ . Donc  $\lim_{n \rightarrow \infty} x_n = \infty$ , or, cette suite n'a pas de valeur d'adhérence, car toute suite extraite vérifie (par le fait que  $\varphi(n) > n$  si  $\varphi$  est strictement croissante dans  $\mathbb{N}$ )

$$\|v_n\| = \|u_{\varphi(n)}\| > \varphi(n) > n$$

, donc toute suite extraite diverge également vers  $+\infty$ .  $\square$

**10.5.16 Théorème:** *Toute partie fermée d'une partie compacte est elle-même compacte.*

**Preuve:** Soit  $(x_n)_{n \in \mathbb{N}}$  une suite d'éléments de  $F$ . Comme  $F \subseteq K$ , c'est aussi une suite d'éléments de  $K$ , donc il existe une sous-suite de  $(x_n)_{n \in \mathbb{N}}$  (d'éléments de  $F$ ) convergeant vers  $l$ . Mais,  $F$  étant fermé,  $l \in F$ . □

**10.5.17 Théorème:** *En dimension finie, les parties compactes sont exactement les parties fermées et bornées.*

**Preuve:** En raison d'un des théorème ci-dessus (deuxième au dessus), il n'y a qu'une seule implication à démontrer :

Soit  $K$  une partie fermée et bornée d'un espace vectoriel  $E$  de dimension finie  $p \in \mathbb{N}$ . Si  $p = 0, E = \{0_E\}$  et  $K = \{0\}$  ou  $K = \{0_E\}$ . Dans les 2 cas,  $K$  est une partie compacte.

Sinon, on introduit une base  $e = (e_1, e_2, \dots, e_p)$  de  $E$  et on considère la norme  $\|\cdot\|_{\infty, e}$ . Soit  $u = (u(n))_{n \in \mathbb{N}}$  une suite d'éléments de  $K$ . Notons  $u_1, u_2, \dots, u_p$  les coordonnées de  $u$ . Soit  $v \in (\mathbb{K}^p)^{\mathbb{N}}$  définie par  $v(n) = (u_1(n), \dots, u_p(n))$ . Puisque  $K$  est bornée,  $\exists M \in \mathbb{R}^+$  tel que

$$\forall x \in K, \|x\| \leq M$$

Donc, en particulier

$$\forall n \in \mathbb{N}, \forall x \in K, \|u(n)\| \leq M$$

ce qui signifie :

$$\forall 1 \leq j \leq p, \forall n \in \mathbb{N}, |u_j(n)| \leq M$$

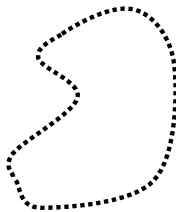
,  $|\cdot|$  étant la valeur absolue ou le module. La suite  $v$  est donc élément du compact  $[-M, M]^p$  si  $\mathbb{K} = \mathbb{R}$  et du compact  $D(0, M)^p$  (disque fermé) si  $\mathbb{K} = \mathbb{C}$ . Elle admet donc une valeur d'adhérence et il existe  $\varphi : \mathbb{N} \rightarrow \mathbb{N}$  strictement croissante telle que  $v((\varphi(n)))_{n \in \mathbb{N}}$  converge, donc les coordonnées  $(u_i(\varphi(n)))_{n \in \mathbb{N}}$  et finalement,  $u((\varphi(n)))_{n \in \mathbb{N}}$  converge. De plus,  $u((\varphi(n)))_{n \in \mathbb{N}} \in K^{\mathbb{N}}$  et  $K$  est fermé donc la limite de cette suite est dans  $K$ . □

**Ouverts, Fermés, bornés**

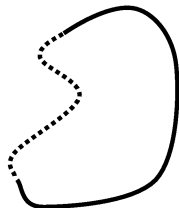
On suppose qu' $E$  est un espace normé, (on pourrait même se restreindre à un espace métrique) on définit les notions d'ouverts, fermés et bornés.

**10.5.18 Définition:** On appelle *voisinage* d'un élément  $a$  de  $E$  toute partie  $V$  vérifiant :  $\exists \alpha > 0$  tel que  $B(a, \alpha) \subseteq V$  ( $B(a, \alpha) = B_o(a, \alpha)$ , boule ouverte).

**10.5.19 Définition:** On dit qu'un ensemble  $U$  est *ouvert* si pour tout  $x \in U$ ,  $x$  est voisinage de chacun de ses points, c'est à dire, si pour tout  $x \in U$ ,  $\exists \alpha > 0$  tel que  $B_o(x, \alpha) \subseteq U$ .

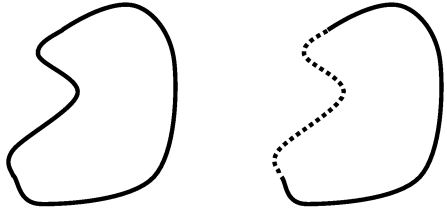


Partie ouverte



Partie non ouverte

**10.5.20 Definition:** Une partie  $F$  de  $E$  est dite *fermée* si son complémentaire dans  $E$  est ouvert.



Partie fermée

Partie non fermée

**10.5.21 Proposition:** -Une réunion (finie ou infinie) de parties ouvertes est une partie ouverte. (1)  
 -Une intersection finie de parties ouvertes est une partie ouverte. (2)  
 -Une intersection (finie ou infinie) de parties fermées est un fermé. (3)  
 -Une union finie de parties fermées est fermé. (4)

**Preuve:** (1). Soit  $(U_i)_{i \in I}$  une famille de parties ouvertes de  $E$  et de  $U = \bigcup_{i \in I} U_i$ .

Soit  $a \in U$ , il existe  $i \in I$  tel que  $a \in U_i$ . Puisque  $U_i$  est un ouvert, il existe  $\alpha > 0$  tel que  $B(a, \alpha) \subset U_i$  et donc  $B(a, \alpha) \subset U$ .

(2). Soit  $(U_i)_{1 \leq i \leq n}$  est une famille de parties ouvertes de  $E$  et  $U = \bigcap_{i=1}^n U_i$ .

Soit  $a \in U$ . Pour tout  $i \in \{1, \dots, n\}$ , il existe  $\alpha_i > 0$  tel que  $a \in U_i$ . Pour  $\alpha = \min \{\alpha_1, \dots, \alpha_n\} > 0$ , on pour tout  $i \in \{1, \dots, n\}$ ,  $B(a, \alpha) \subset B(a, \alpha_i) \subset U_i$  donc  $B(a, \alpha) \subset U$ .

(3) et (4). Par passage au complémentaire d'une union ou d'une intersection d'ouverts.  $\square$

**10.5.22 Proposition:** On a équivalence entre :

(i)  $F$  est fermée

(ii) Pour tout  $x_n \in F^{\mathbb{N}}$ ,  $x_n \rightarrow a \Rightarrow a \in F$

**Preuve:** (i)  $\Rightarrow$  (ii) Par contraposée.

Supposons qu'il existe  $(x_n) \in F^{\mathbb{N}}$  telle que  $x_n \rightarrow a$  et  $a \notin F$ .

Soit  $\alpha > 0$ . Pour  $n$  assez grand,  $\|x_n - a\| < \alpha$  donc  $x_n \in B(a, \alpha)$ , et donc  $B(a, \alpha) \cap F \neq \emptyset$ .

Ainsi  $a \in C_{EF}$  et  $\forall \alpha > 0, B(a, \alpha) \not\subset C_{EF}$

La partie  $C_{EF}$  n'est pas ouverte et donc  $F$  n'est pas fermée.

(i)  $\Rightarrow$  (ii) Par contraposée.

Supposons  $F$  non fermée i.e.  $a \in C_{EF}$  non ouvert.

Il existe  $a \in C_{EF}$  tel que  $\forall \alpha > 0, B(a, \alpha) \cap F \neq \emptyset$

Soit  $n \in \mathbb{N}$ . Pour  $\alpha = 1/(n+1) > 0$  il existe  $x_n \in B(a, 1/(n+1)) \cap F$ .

En faisant varier  $n$ , ceci détermine une suite  $x_n \in F^{\mathbb{N}}$  telle que  $x_n \rightarrow a$  avec  $a \notin F$ .  $\square$

**10.5.23 Definition:** Soit un ensemble  $E$  et  $X$  une partie de  $E$ . On appelle *intérieur* de  $X$  le plus grand ouvert de  $E$  inclus dans  $X$  et *adhérence* de  $X$  le plus petit fermé de  $E$  contenant  $X$ . Pour une définition plus formelle :  $Adh_E(X) = \bigcap_{X \subseteq Y_i, Y_i \text{ ferme}} Y_i$  (intersection de tous les fermés contenant  $X$ ) et

$Int_E(X) = \bigcup_{Y_i \subseteq X, Y_i \text{ ouvert}} Y_i$  (réunion de tous les ouvertures inclus dans  $X$ ).

**10.5.24 Proposition:**  $-a \in Int(X) \Leftrightarrow \exists \alpha > 0$  tel que  $B(a, \alpha) \subseteq X$

$-a \in Adh(X) \Leftrightarrow$  pour tout  $\alpha > 0, B(a, \alpha) \cap X \neq \emptyset$

-(caractérisation séquentielle des points adhérents) : On a équivalence entre : (i)  $a$  est adhérent à  $X$  et (ii)  $\exists x_n \in X^{\mathbb{N}}$  tel que  $x_n \rightarrow a$ .

**Preuve:** Soyons méthodiques : Remarquons que  $\text{Int}(X)$  est un ouvert (réunion infinie de parties ouvertes) et  $\text{Adh}(X)$  est un fermé (intersection infinie de parties fermés). De plus, par définition :  $\text{Int}(X) \subseteq X \subseteq \text{Adh}(X)$ . Ainsi, le sens  $\Rightarrow$  de la première proposition exprime le fait que  $\text{Int}(X)$  est ouvert. L'autre sens se déduit par définition :  $B(a, \alpha) \subseteq X$ . Or  $B(a, \alpha)$  est une boule ouverte, donc un ouvert inclus dans  $X$ . On va maintenant prouver le sens  $\Leftarrow$  de la troisième proposition : puisque  $\text{Adh}(X)$  est fermée et  $X \subseteq \text{Adh}(X)$ ,  $x_n \in \text{Adh}(X)^{\mathbb{N}}$  et donc  $a \in \text{Adh}(X)$  par la caractérisation séquentielle des parties fermées. Prouvons ensuite le sens  $\Rightarrow$  de la deuxième proposition par contraposée : si  $\exists \alpha > 0$  tel que  $B(a, \alpha) \cap X = \emptyset$ ,  $B(a, \alpha) \subseteq C_E(X)$  ( $C_E$  est le complémentaire dans  $E$ ) donc, par passage au complémentaire,  $X \subseteq C_E(B(a, \alpha))$ . Or,  $C_E(B(a, \alpha))$  est un fermé, car  $B(a, \alpha)$  est une boule ouverte, et il contient  $X$ . Donc  $\text{Adh}(X) \subseteq C_E(B(a, \alpha))$ , ce qui permet de déduire :  $a \notin \text{Adh}(X)$  (car  $a \in B(a, \alpha) \Rightarrow a \notin C_E(B(a, \alpha))$ ).

Prouvons maintenant que  $\forall \alpha > 0, B(a, \alpha) \cap X \neq \emptyset \Leftrightarrow \exists x_n \in X^{\mathbb{N}}$ , ce qui nous permettra de démontrer le sens  $\Leftarrow$  de la deuxième proposition et, par suite, une fois ce sens démontré, cela prouve également le sens  $\Rightarrow$  de la troisième proposition (écrivez les implications sur un papier si vous n'êtes pas convaincu de cela). Si  $\forall \alpha > 0, B(a, \alpha) \cap X \neq \emptyset$ , soit  $n \in \mathbb{N}$  tel  $\alpha = \frac{1}{n+1}$ . Par application de la propriété :  $B(a, \frac{1}{n+1}) \cap X \neq \emptyset$ . Soit  $x_n$  un élément de cet ensemble : cela définit une suite telle que  $\lim_{n \rightarrow \infty} \|x_n - a\| \leq \frac{1}{n+1} \rightarrow 0$ , donc  $x_n \rightarrow a$ . Maintenant, si  $\exists x_n \in X^{\mathbb{N}}$  tel que  $x_n \rightarrow a, \forall \alpha > 0, \exists N \in \mathbb{N}$  tel que  $\forall n \geq N, \|x_n - a\| \leq \alpha$ , et donc  $B(a, \alpha) \cap X \neq \emptyset$ .  $\square$

**10.5.25 Proposition:** Soit  $f$  une fonction continue :  $f : E \rightarrow G$ . L'ensemble  $A = \{x \in E, f(x) = k \in F\}$  où  $F$  est un fermé (on peut également considérer, à la place de l'égalité, les signes  $\leq, \geq$ ). Soit  $g$  une fonction continue :  $g : E \rightarrow G$ . L'ensemble  $B = \{x \in E, g(x) < k \in U\}$  où  $U$  est un ouvert (on peut également considérer, à la place le signe  $>$ ).

**Preuve:** (i) Soit une suite  $x_n \in A^{\mathbb{N}} : f(x_n) = k \forall n \in \mathbb{N}$ . On peut passer à la limite car  $f$  est continue, soit : si  $(x_n)_{n \rightarrow \infty} \rightarrow a \in E, (f(x_n))_{n \rightarrow \infty} \rightarrow f(a)$ . Or,  $f(x_n) = k \Rightarrow f(a) = k$ . C'est une conséquence de la continuité de  $f$ . Donc  $A$  est fermé.

(ii) Soit  $y \in B : g(y) < k$ . Or,  $g$  est continue  $\Rightarrow$  pour  $\epsilon = \frac{k - g(y)}{2}, \exists r$  tel que  $\forall x \in B(y, r)$  (boule ouverte),  $|g(x) - g(y)| < \epsilon \Rightarrow g(x) < \frac{g(y) + k}{2} \leq \frac{k + k}{2} = k$ . Donc  $B$  est ouvert.

Remarque : il existe des résultats un peu plus généraux sur le lien entre continuité et topologie, mais non nécessaires dans ce TIPE.  $\square$

**10.5.26 Définition:** On appelle *frontière* d'une partie  $X$  de  $E$  l'ensemble  $Fr(X) = \text{Adh}(X) / \text{Int}(X)$ .

**10.5.27 Définition:** Une partie  $X$  de  $E$  est dite *dense* dans  $E$  si  $\text{Adh}(X) = E$ .

**10.5.28 Proposition:** Si deux fonctions  $f$  et  $g$  sont continues avec

$$f, g : E \rightarrow F$$

et égale sur une partie  $X$  dense dans  $E$ , alors  $f = g$  sur tout l'ensemble  $E$ .

**Preuve:** Soit  $x \in E$ . Il existe  $x_n \in X^{\mathbb{N}}$  telle que  $x_n \rightarrow x$  car  $X$  est dense dans  $E$ .  $f(x_n) = g(x_n) \forall n \in \mathbb{N}$ , donc à la limite, et par continuité de  $f$  et  $g$ ,  $f = g$ .  $\square$

**10.5.29 Définition:** Une partie  $X$  d'un  $\mathbb{K}$ -espace vectoriel  $E$  de dimension finie est dite *bornée* si  $\exists r > 0$  tel que  $X \subseteq B(0, r)$ .

### Espaces normés, Normes

On a déjà défini ce qu'était une norme dans la partie 4.1. On peut également préciser que la norme est une application à valeurs dans  $\mathbb{R}_+$ . Nous allons voir son utilité :

**10.5.30 Définition:** Soit  $\mathbb{K}$  un corps commutatif et  $E$  un  $\mathbb{K}$ -espace vectoriel. Une norme est une application

$$N : E \rightarrow \mathbb{R}^+$$

$$x \mapsto N(x)$$

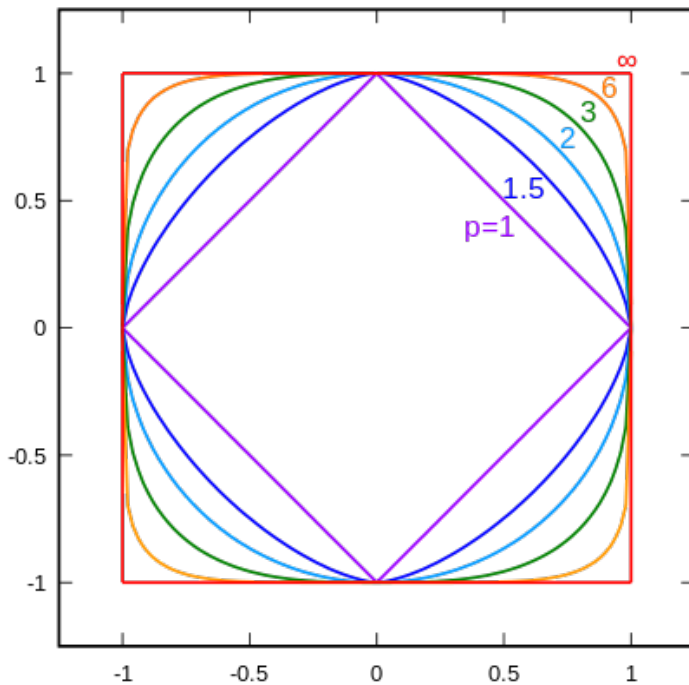
vérifiant les propriétés suivantes :

$$N(x) = 0 \Leftrightarrow x = 0$$

$$N(\lambda x) = |\lambda| N(x)$$

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

Ensemble de vecteurs de normes 1 pour les  $p$ -normes :  $\|x\|_p = (\sum_{i=1}^n x_i^p)^{\frac{1}{p}}$ .



**10.5.31 Définition:** Un espace normé est un espace muni d'une norme.

**10.5.32 Définition:** Soit  $X$  un espace topologique. On appelle *boule de rayon  $a$  et centrée en  $x$*  pour la norme  $\|\cdot\|$  l'ensemble  $\{y \in X, \|y - x\| < a\}$ . Puisu'on a le signe  $<$ , on parle de boule ouverte. Si on avait  $\leq$ , on parlerait de boule fermée. On la note  $B_o(x, a)$  (respectivement  $B_f(x, a)$ ).

**10.5.33 Définition:** Une fonction  $f$  est *continue* en  $y$  si pour tout  $\epsilon > 0, \exists \alpha > 0$  tel que  $\|y - x\| \leq \alpha \Rightarrow \|f(y) - f(x)\| \leq \epsilon$

**10.5.34 Définition:** Convergence : Une suite  $u_n$  *converge* vers une limite  $l$  si  $\lim_{n \rightarrow +\infty} \|u_n - l\| = 0$ .

**10.5.35 Proposition:** Dans un  $\mathbb{K}$ -espace vectoriel de dimension finie, toutes les normes sont équivalentes (2 normes  $\|\cdot\|_1$  et  $\|\cdot\|_2$  sont équivalentes si  $\exists \alpha, \beta$  tels que  $\alpha \|x\|_2 \leq \|x\|_1 \leq \beta \|x\|_2$ , pour tout  $x \in X$ ).

**Preuve:** Soit  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension finie  $p$ .

Si  $p = 0$  alors la seule norme sur  $E$  est l'application nulle et le problème est résolu.

Si  $p > 0$  alors on peut introduire une base  $B = (e^1, \dots, e^p)$  de  $E$  et considérer la norme  $c$  définie par : pour tout  $x = x_1 e_1 + \dots + x_p e_p \in E, \|x\|_\infty = \max\{|x_1|, \dots, |x_p|\}$ .

Soit  $N$  une autre norme de  $E$ , montrons que  $N$  et  $\|\cdot\|_\infty$  sont équivalents ce qui permet de conclure.

Pour  $x \in E$  s'écrivant  $x = x_1 e_1 + \dots + x_p e_p$ , l'inégalité triangulaire et l'homogénéité de la norme  $N$  donnent :

$$N(x) \leq k \|x\|_\infty \text{ en posant le } k = N(e_1) + \dots + N(e_p).$$

Ainsi la norme  $N$  est dominée par la norme  $\|\cdot\|_\infty$ .

Inversement, montrons en raisonnant par l'absurde que la norme  $\|\cdot\|_\infty$  est dominée par  $N$ .

Supposons que  $\|\cdot\|_\infty$  n'est pas dominé par  $N$ . Pour tout  $\alpha \in \mathbb{R}^+$ , il existe donc un vecteur  $x \in E$  vérifiant :  $\|x\|_\infty > \alpha N(x)$ .

En prenant  $\alpha = n$  avec  $n$  parcourant  $\mathbb{N}$ , on peut ainsi définir une suite  $(x(n))_{n \in \mathbb{N}}$  d'éléments de  $E$  vérifiant :  $\forall n \in \mathbb{N}, \|x(n)\|_\infty > nN(x(n))$ .

Quitte à diviser le vecteur  $x(n)$  par le scalaire  $\|x(n)\|_\infty$ , on peut aussi supposer que la suite  $(x(n))_{n \in \mathbb{N}}$  vérifie :  $\forall n \in \mathbb{N}, \|x(n)\|_\infty = 1$

On a alors  $N(x(n)) < \frac{1}{n}$  et donc  $x(n) \xrightarrow[n \rightarrow +\infty]{N} 0$ .

Montrons maintenant qu'il existe une suite extraite  $(x(n))_{n \in \mathbb{N}}$  convergeant pour la norme  $\|\cdot\|_\infty$ .

Introduisons les composantes  $(x_1(n))_{n \in \mathbb{N}}, \dots, (x_p(n))_{n \in \mathbb{N}}$  de la suite  $(x(n))_{n \in \mathbb{N}}$ , celles-ci sont bornées car  $\forall n \in \mathbb{N}, \|x(n)\|_\infty = 1$ .

Puisque la suite scalaire  $(x_1(n))_{n \in \mathbb{N}}$  est bornée, il existe, en vertu du théorème de Bolzano-Weierstrass, une suite extraite  $(x_1(\varphi_1(n)))_{n \in \mathbb{N}}$  convergente.

Puisque la suite scalaire  $(x_2(\varphi_1(n)))_{n \in \mathbb{N}}$  est bornée, car extraite de la suite bornée  $(x_2(n))_{n \in \mathbb{N}}$ , il existe une suite extraite  $(x_2(\varphi_1(\varphi_2(n))))_{n \in \mathbb{N}}$  convergente. De plus la suite extraite  $(x_1(\varphi_1(\varphi_2(n))))_{n \in \mathbb{N}}$  est elle aussi convergente car extraite de  $(x_1(\varphi_1(n)))_{n \in \mathbb{N}}$  déjà convergente.

Ainsi, on a construit une extraction commune aux suites  $(x_1(n))_{n \in \mathbb{N}}$  et  $(x_2(n))_{n \in \mathbb{N}}$  conduisant à des suites convergentes.

En poursuivant ce processus, on peut construire une extractrice  $\varphi$  telles que toutes les suites  $(x_1(\varphi(n)))_{n \in \mathbb{N}}, \dots, (x_p(\varphi(n)))_{n \in \mathbb{N}}$  convergent.

La suite  $(x(\varphi(n)))$  est alors convergente pour la norme  $\|\cdot\|_\infty$  ; posons  $l$  sa limite.

Puisque  $\forall n \in \mathbb{N}, \|x(\varphi(n))\|_\infty = 1$ , à la limite on obtient  $\|l\|_\infty = 1$ .

Or la norme  $N$  est dominée par  $\|\cdot\|_\infty$  donc  $N(x(\varphi(n)) - l) \leq k \|x(\varphi(n)) - l\|_\infty \rightarrow 0$  puis  $x(\varphi(n)) \xrightarrow[n \rightarrow +\infty]{N} l$

mais on a déjà  $x(n) \xrightarrow[n \rightarrow +\infty]{N} 0$  et donc  $x(\varphi(n)) \xrightarrow[n \rightarrow +\infty]{N} 0$ .

Par unicité de la limite, on obtient  $l = 0$ .

Cependant  $\|l\|_\infty = 1$ , c'est absurde! □

**10.5.36 Exemple:** *La norme n'est pas unique. Je vais donner 2 exemples de normes matricielles (puisque c'est le sujet qui nous intéresse) :*

$$\|A\| = \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad (1).$$

$$= \| \| A \| \| = \| \| A \| \|_\infty = \max_{1 \leq i, j \leq n} |a_{ij}| \quad (2).$$

On vérifie que (1) et (2) sont des normes :

**Preuve:**

$$\|A\| = 0 \Rightarrow \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = 0 \Rightarrow a_{ij} = 0 \forall i, j \in \mathbb{N} \Rightarrow A = O_n$$

De même,

$$\| \|A\| \| = 0 \Rightarrow \max_{1 \leq i, j \leq n} |a_{ij}| = 0 \Rightarrow a_{ij} = 0 \forall i, j \in \mathbb{N}$$

On vérifie très facilement les 2 propriétés restantes sur les normes. Pour finir, on va montrer que  $\|\cdot\|$  est une norme d'algèbre : soit  $C=AB$

$$\|AB\| = \max_{1 \leq i \leq n} \sum_{j=1}^n |c_{ij}| = \max_{1 \leq i \leq n} \sum_{j=1}^n \left| \sum_{k=1}^n a_{ik} b_{kj} \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n \sum_{k=1}^n |a_{ik} b_{kj}|$$

Or :

$$\sum_{k=1}^n \sum_{j=1}^n |a_{ik}| |b_{kj}| = \sum_{k=1}^n |a_{ik}| \sum_{j=1}^n |b_{kj}| \leq \sum_{k=1}^n |a_{ik}| \|B\| \leq \|A\| \|B\|.$$

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

□

Dans la suite, on se servira de ces 2 normes, donc ne les oubliez pas !

## 10.6 Connexité

### Espaces topologiques connexes

Dans tout ce chapitre  $(X, O)$  et  $(Y, O')$  désignent des espaces topologiques.

**10.6.1 Proposition:** *On dira que l'espace  $(X, O)$  est connexe s'il vérifie l'une des conditions équivalentes suivantes.*

1. Si  $X$  est réunion de deux ouverts disjoints alors l'un de ces deux ouverts est vide et l'autre est égale à  $X$ .
2. Si  $X$  est réunion de deux fermés disjoints alors l'un de ces deux fermés est vide et l'autre est égale à  $X$  ;
3. Si on considère  $\{0, 1\}$  muni de la topologie discrète et  $f : X \rightarrow \{0, 1\}$  une application continue, alors  $f$  est constante sur  $X$ .
4. Les seuls ensembles à la fois ouverts et fermés de  $X$  sont  $X$  lui même et l'ensemble vide.

**Preuve:**  $1 \Leftrightarrow 2$  est évident par passage au complémentaire.

$1 \Rightarrow 3$  : Soit  $f$  une application continue de  $X$  dans  $\{0, 1\}$ . Alors  $\{f^{-1}(0); f^{-1}(1)\}$  représente une partition de  $E$  en deux ouverts (ou deux fermés) de  $E$ . par conséquent l'un de ces deux ouverts est vide et l'autre égale  $X$  tout entier, ce qui implique bien que  $f$  est constante sur  $X$ .

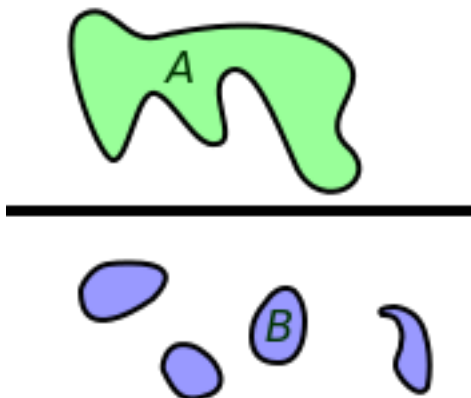
$3 \Rightarrow 1$  : Soient  $U_1$  et  $U_2$  deux ouverts de  $X$  qui définissent une partition de  $X$ . Soit aussi  $f : X \rightarrow \{0, 1\}$  définie par  $f(U_1) = 0$  et  $f(U_2) = 1$ .  $f$  est continue et donc constante sur  $X$ . Donc l'un des deux ouverts est vide et l'autre égale à  $X$  tout entier. Cqfd.

$1 \Rightarrow 4$  : Soit  $U$  un sous ensemble à la fois ouvert et fermé de  $X$ . Alors  $U^c$  est, lui aussi, un sous ensemble ouvert et fermé de  $X$ . Mais  $U$  et  $U^c$  définissent une partition de  $X$  en deux ouverts.  $X$  étant connexe  $U$  est ou vide ou égale à  $X$  tout entier.

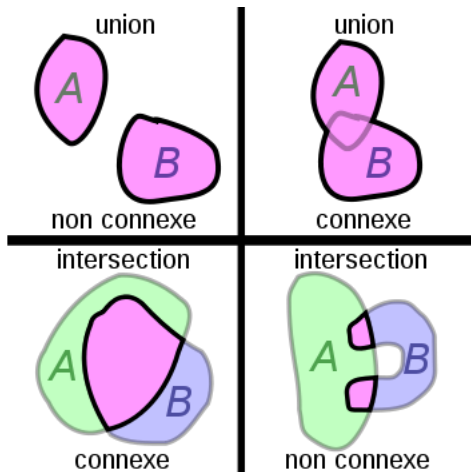
$4 \Rightarrow 1$  : Supposons que  $U$  et  $V$  définissent une partition ouverte de  $X$ . Le complémentaire de  $U$  est alors égale à  $V$  et réciproquement  $V^c = U$ .  $U$  étant ouvert,  $V$  est alors fermé. De même  $U$  est aussi fermé. Mais  $X$  ne possède pas de sous ensemble à la fois ouvert et fermé autre que l'ensemble vide et  $X$ . Donc l'un des deux,  $U$  ou  $V$  est vide l'autre égale à  $X$ , ce qui nous donne le premier point. □

**10.6.2 Définition:** On dira qu'un sous ensemble  $U$  de  $X$  est un *sous espace connexe* de  $X$  (ou un connexe de  $X$ ) si  $U$  est connexe pour la topologie induite de celle de  $X$ .

**10.6.3 Exemple:** Un intervalle de  $\mathbb{R}$  est connexe dans  $\mathbb{R}$  (muni de sa topologie canonique). Les seuls sous ensembles connexes de  $\mathbb{R}$  sont d'ailleurs les intervalles.



Attention, on a pas de résultats vraiment précis sur la réunion et l'intersection de 2 connexes (le seul connu : la réunion de 2 connexes d'intersection non vide est connexe).



**Application continue sur un connexe**

**10.6.4 Théorème:** L'image d'un connexe par une application continue est un sous ensemble connexe de l'espace image de cette application.

**Preuve:** On suppose ici que  $(X, \mathcal{O})$  est un espace topologique connexe et soit  $f : X \rightarrow Y$  une application continue de  $X$  dans  $Y$ . Montrons que  $f(X)$  est un connexe de  $Y$ . Supposons donc qu'il existe une partition  $\{U_1; U_2\}$  de  $f(X)$  en deux ouverts. Posons  $V_1 = f^{-1}(U_1)$  et  $V_2 = f^{-1}(U_2)$ . Les deux ouverts  $U_1$  et  $U_2$  sont des éléments de topologie induite sur  $f(X)$  et sont donc de la forme  $U_1 = f(X) \cap O_1$  et  $U_2 = f(X) \cap O_2$  où  $O_1$  et  $O_2$  sont des ouverts de  $Y$ . De plus, pour  $i = 1, 2$ ,  $f^{-1}(U_i) = f^{-1}(O_i) = V_i$ . Comme  $f$  est continue, on en déduit que les sous ensembles  $V_1$  et  $V_2$  sont des sous ensembles ouverts dans  $X$ . De plus, par construction, leur réunion recouvre  $X$  et leur intersection est vide.  $\{V_1; V_2\}$  est donc une partition de  $X$  en deux ouverts. Comme  $X$  est connexe, l'un de ces deux ouverts est vide et l'autre égale à  $X$  tout entier. Mais ceci implique que l'image par  $f$  de chacun de ces deux ouverts est, respectivement vide et égale à



$f(X)$  tout entier et donc que l'un de nos deux sous ensembles  $U_1; U_2$  de  $f(X)$  est vide et l'autre égale à  $f(X)$ .  $f(X)$  est alors bien connexe.  $\square$

En application de ce théorème :

**10.6.5 Théorème:** (*Theo. des valeurs intermédiaires*) Si une application  $f$  est définie et continue sur un intervalle  $]a, b[$  de  $\mathbb{R}$  (où  $a$  et  $b$  sont des réels quelconques pouvant être égales à respectivement  $-\infty$  et  $+\infty$ ), et si de plus  $a'$  et  $b'$  sont des éléments de  $]a, b[$  tel que  $a' < b'$  alors pour tout  $C \in [f(a'), f(b')]$ , il existe  $c \in [a', b']$  tel que  $f(c) = C$ .

**Preuve:** L'image d'un connexe par une application continue est connexe. Or, les intervalles de  $\mathbb{R}$  sont les sous ensembles connexes de  $\mathbb{R}$ . On en déduit donc que l'image de  $[a', b']$  par  $f$  est un intervalle de  $\mathbb{R}$ . Tout élément de ce dernier possédant un antécédant dans  $[a', b']$ , le théorème est démontré.  $\square$

### Quelques critères de connexité

**10.6.6 Proposition:** Si un sous ensemble  $U$  de  $X$  est connexe, il est de même de son adhérence.

**Preuve:**  $f : \bar{U} \rightarrow \{0, 1\}$  une application continue. (l'adhérence de  $U$  est munie de la topologie induite de celle de  $X$  et  $\{0, 1\}$  est muni de la topologie discrète).  $f$  est donc continue sur  $U$ . Mais  $U$  étant connexe, cela implique que  $f$  est constante sur  $U$ . On peut par exemple supposer que  $f$  vaut 1 sur  $U$ . Soit  $x$  un élément de  $\bar{U}/U$ . Supposons que  $f(x) = 0$ . Comme  $f$  est continue et que  $\{0\}$  est un ouvert de  $\{0, 1\}$  muni de la topologie discrète,  $f^{-1}(0)$  est un ouvert de l'adhérence de  $U$  contenant  $x$ . C'est donc, en particulier, un voisinage de  $x$ . Mais comme  $x$  est adhérent à  $U$ , ce voisinage intersecte nécessairement  $U$ , et donc, par construction de ce voisinage,  $U$  possède des points dont l'image par  $f$  est nulle. Ce qui est absurde, par hypothèse. Donc  $f \equiv 1$  sur l'adhérence de  $U$  et cette adhérence est donc belle et bien connexe.  $\square$

**10.6.7 Proposition:** Soient  $(U_i)_{i \in I}$  une famille de sous ensembles connexes de  $X$  tels  $\bigcap_{i \in I} U_i \neq \emptyset$  alors  $\bigcup_{i \in I} U_i$  est connexe.

**Preuve:** Notons  $U$  la réunion des  $(U_i)_{i \in I}$  et soit une application continue  $f : U \rightarrow \{0, 1\}$ . Soit  $a \in X$  un point de l'intersection des  $(U_i)_{i \in I}$ . La restriction de  $f$  à  $U_i$ ,  $i$  étant fixé dans  $I$ , est encore une application continue à valeurs dans  $\{0, 1\}$ . Comme  $U_i$  est connexe, il s'en suit que  $f$  est constante sur  $U_i$ . On peut supposer, par exemple, que  $f$  vaut 0 sur  $U_i$ . On aura donc  $f(a) = 0$ ,  $i$  étant quelconque dans  $I$ ,  $f$  est alors constante sur chaque  $U_i, \forall i \in I$  et donc que  $f$  est nulle sur  $U$  et donc constante sur  $U$ . Ce qui implique que  $U$  est connexe.  $\square$

**10.6.8 Définition:** Soit  $x$  un élément de  $X$ . On appelle *composante connexe* de  $x$  la réunion des sous ensembles connexes de  $X$  contenant  $x$ .

**10.6.9 Proposition:** Soit  $x$  un élément de  $X$ .

- La composante connexe de  $x$  est le plus grand composant de  $X$  contenant  $x$ .
- La composante connexe de  $x$  est une partie fermée de  $X$ .

**Preuve:** La première partie de la proposition est évidente, par définition de la composante connexe d'un point. La seconde partie s'en déduit aussitôt car, rappelons le, si un ensemble est connexe, il en est de même de son adhérence qui de plus est fermée. Donc si  $U$  est le plus grand connexe de  $X$  contenant  $x$ , il est nécessairement égale à son adhérence qui est aussi connexe et qui contient aussi  $x$ .  $\square$

**10.6.10 Définition:** Soient  $x$  et  $y$  deux éléments de  $X$ . On appelle *chemin d'extrémités  $x$  et  $y$*  (ou chemin joignant  $x$  et  $y$ ) de  $X$  toute application continue  $c : [0, 1] \rightarrow X$  telle que  $c(0) = x$  et  $c(1) = y$ .

**10.6.11 Définition:** On dira que  $(X, O)$  est *connexe par arcs* si tout couple d'éléments de  $X$  peut être joint par un chemin.

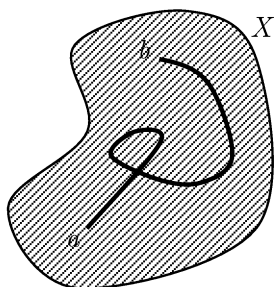
**10.6.12 Proposition:** *Si X est connexe par arc alors X est connexe.*

**Preuve:** Supposons donc que X n'est pas connexe. Soit alors  $\{U, V\}$  une partition de X en deux fermés. Soient aussi x un élément de U et y un élément de V. Comme X est connexe par arc, il existe un chemin  $c : [0, 1] \rightarrow X$  telle que  $c(0) = x$  et  $c(1) = y$ . Notons  $A = \{t \in [0, 1] / c(t) \in U\}$ . Comme A est un sous ensemble majoré de  $\mathbb{R}$ , il possède une borne supérieure que l'on note  $t_0$ . Notons, d'autre part,  $B = \{t \in [0, 1] / c(t) \in V\}$ . B est un sous ensemble minoré de  $\mathbb{R}$  et possède, par conséquent, une borne inférieure que l'on note  $t_1$ . On a nécessairement  $t_0 = t_1$ . Supposons que ce ne soit pas le cas, alors  $t_0 < t_1$ . On peut alors trouver un réel t élément de  $]t_0, t_1[$  mais l'élément  $c(t)$  de X ne peut alors ni être élément de U ni de V. Ce qui est impossible.

Notons  $T = t_0 = t_1$ .

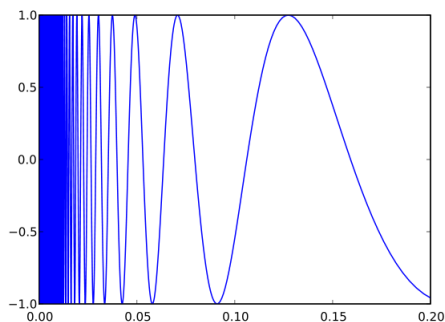
T étant la borne supérieure de A, on peut construire une suite  $(t_n)_{n \in \mathbb{N}}$  d'éléments de A convergeante vers T. Mais U étant fermé et c continu,  $\lim_{n \rightarrow \infty} c(t_n) = c(T)$  est un élément de U. De même, on montrera que  $c(T)$  est un élément de V. Mais U et V ont été supposés disjoints. On aboutit alors à une contradiction et X est bien connexe. □

Voici une petite image :



**Remarque :** Attention, la réciproque est fausse.

L'adhérence C du graphe G de f est connexe, mais pas connexe par arcs



Prouvons le :

Soit  $G = \{(x, \sin(\frac{1}{x})), x \in ]0, 1]\}$ . En tant que graphe d'une fonction continue sur un ensemble connexe, G est connexe, et l'adhérence d'un connexe étant connexe,  $\overline{G}$  est connexe.

Cependant :

**10.6.13 Proposition:** *L'espace topologique G, muni de la topologie induite par celle de  $\mathbb{R}^2$  n'est pas connexe par arcs. Plus précisément, il a deux composantes connexes par arcs :G et  $G' = \{0\} \times [-1, 1]$ .*

**10.6.14 Lemme:** *Les parties connexes de  $\mathbb{R}$  ainsi que celles d'un intervalle de  $I \mathbb{R}$  sont les intervalles de  $\mathbb{R}$  (respectivement, I).*

**Preuve:**

Si  $A$  est une partie connexe de  $\mathbb{R}$  (respectivement,  $I$ ) alors  $A$  est un intervalle, puisque tout réel  $a$  strictement compris entre deux éléments de  $A$  appartient lui aussi à  $A$  : sinon,  $] - \infty, a[ \cap A$  et  $]a, \infty[ \cap A$  (respectivement  $] \inf I, a[ \cap A$  et  $]a, \sup I[ \cap A$ ) formeraient une partition de  $A$  en deux ouverts de  $A$  non vides et disjoints.

Maintenant, montrons qu'un intervalle  $A$  de  $\mathbb{R}$  ou de  $I$  est connexe. En effet,  $A$  puisque toute application continue de  $A$  dans  $\mathbb{R}$  qui ne prend que les valeurs 0 et 1 est constante, d'après le théorème des valeurs intermédiaires.

Attention : cela implique de savoir prouver le théorème des valeurs intermédiaires d'une autre manière que celle que j'ai montré, car c'est précisément ce lemme qu'on a utilisé pour prouver le théorème des valeurs intermédiaires (exemple : par dichotomie ou par la propriété de la borne supérieure).  $\square$

**Preuve:** de la proposition : Montrons déjà que  $\overline{G} = G \cup \{0\} \times [-1, 1] = G \cup G'$ .

$G \subseteq G \cup \{0\} \times [-1, 1]$  : en effet, la fonction  $\sin(\frac{1}{x})$  est bornée par 1 et, pour  $\alpha \in [-1, 1]$ , on pose  $x_n = \frac{1}{\arcsin(\alpha) + 2n\pi}$ . On a donc  $\sin(\frac{1}{x_n}) = \alpha, \forall n \in \mathbb{N}$  et  $\lim_{n \rightarrow \infty} x_n = 0$ . Cette inclusion est donc montrée car  $(x_n, \sin(\frac{1}{x_n})) = (0, \alpha) \in G'$ . La seconde inclusion est évidente, d'où l'égalité des 2 ensembles.

Les espaces  $G$  et  $G'$  sont connexes par arcs : pour  $G$ , on prend  $\gamma(t) = (x + t(y - x), \sin(\frac{1}{x+t(y-x)}))$  qui permet de relier  $(x, \sin(\frac{1}{x}))$  et  $(y, \sin(\frac{1}{y}))$ , tant que  $x$  et  $y \neq 0$ . Pour  $G'$ , c'est évident.

Montrons qu'il n'existe pas de chemin continu reliant  $G$  et  $G'$  : soit  $a = (0, \alpha) \in G'$  et  $b = (\beta, \sin(\frac{1}{\beta}))$ . Supposons qu'il existe  $\gamma : [0, 1] \rightarrow \overline{G}$  tel que  $\gamma(0) = a$  et  $\gamma(1) = b$ . On pose  $\gamma(t) = (u(t), v(t))$  avec  $u, v : [0, 1] \rightarrow \mathbb{R}$  continues, et qui vérifient : soit  $u(t) = 0$  et  $v(t) \in [-1, 1]$ , soit  $u(t) > 0$  et  $v(t) = \sin(\frac{1}{u(t)})$ .

On considère  $\Omega = \{t \in [0, 1], u(t) > 0\}$  : c'est un intervalle ouvert de  $[0, 1]$  car  $u$  est continue. De plus, comme on est dans  $[0, 1]$ , un intervalle de  $\mathbb{R}$ , alors les composantes connexes et ouvertes de  $[0, 1]$  sont les intervalles de la forme :  $]c, d[$ ,  $[0, d[$  ou  $]c, 1]$ , avec  $0 \leq c < d \leq 1$ . On note  $J$  l'une des composantes connexes de  $[0, 1]$ .  $\gamma(0) = (u(0), v(0)) = (0, \alpha)$ ,  $u(0) = 0$  donc, forcément,  $0 \notin \Omega \Rightarrow J = ]c, d[$  ou  $J = ]d, 1] \Rightarrow c \notin J$ . Par conséquent,  $c \notin \Omega$  : par l'absurde, si c'était le cas,  $J \subset J \cup \{c\}$ , qui est également connexe et strictement plus grand que  $J$ , ce qui est impossible par définition de la composante connexe.

Donc  $u(c) = 0$ . L'ensemble  $u(J)$  est un intervalle (car image connexe d'un intervalle), inclus dans  $]0, \infty[$  et comme  $u(c) = 0$ , la borne inférieure de  $u(J)$  est 0. Sur un tel intervalle, la fonction  $x \mapsto \sin(\frac{1}{x})$  prend une infinité de fois chaque valeurs entre  $-1$  et  $1$ , et on déduit que  $\lim_{t \rightarrow c, t > c} \sin(\frac{1}{u(t)})$  n'existe pas alors que  $\lim_{t \rightarrow c, t > c} \sin(\frac{1}{u(t)}) = \lim_{t \rightarrow c, t > c} v(t) = v(c)$  (car  $v(t)$  est continue et, sur  $J, v(t) = \sin(\frac{1}{u(t)})$ ) : c'est une contradiction.  $\square$

**10.6.15 Proposition:** Si  $G/H$  et  $H$  sont connexes,  $G$  est connexe.

**Preuve:** Soit  $f : G \rightarrow \{0, 1\}$  une application continue ; puisque  $H$  est connexe, chaque  $x.H$  l'est aussi et  $f$  est constante sur les classes  $x.H$  ; elle définit donc par passage au quotient une application  $\bar{f} : G/H \rightarrow \{0, 1\}$  continue, donc constante.

Finalement,  $f$  est constante.  $\square$

## 11 Conclusion :

Je remercie Alessandra Frabetti, mon encadrante, pour le choix du sujet, le temps qu'elle a passé à m'expliquer certaines notions et pour l'intérêt qu'elle m'a permis de développer pour la géométrie différentielle.

Je la remercie également pour m'avoir laissé une grande liberté pour la poursuite du TIPE, car j'ai pu approfondir les domaines que je souhaitais, qui m'inspiraient et il est vrai qu'au début de ce TIPE, je ne savais même pas quel allait être le but final, car plusieurs possibilités s'ouvraient à chaque étapes (j'en parle d'ailleurs un peu plus loin).

Je remercie aussi Nguyen Viet Dang, mon professeur de TD de géométrie et calcul différentiel avec qui j'ai pu avoir quelques échanges intéressants sur différents sujets liés à ce TIPE. Je remercie enfin Philippe Caldero pour la confiance qu'il m'a accordé et, plus généralement, l'ensemble des chercheurs en mathématiques de l'Institut Camille Jordan pour la bonne ambiance qui règne au sein du bâtiment braconnier.

Le type de travail que j'ai fait dans ce TIPE est assez varié. Comme on peut le voir dans cette bibliographie, il a consisté à comprendre un certain nombre de notions par moi-même à l'aide de livres ou de pages web, mais également à retrouver certains résultats par moi-même, soit parce que mon encadrante me donnait des résultats en exercices, soit parce que je me posais certains exercices ou enfin, lorsque la documentation que je prenais laissait des preuves en exercices. Enfin, le dernier type de travail est tout simplement par l'enseignement direct : quand je ne comprenais pas quelque chose ou que je bloquais sur une preuve, j'ai pu demander des explications à Alessandra.

## 11.1 Bibliographie

Voici les principales références qui m'ont servi.

1. Matrix groups. An introduction to lie group theory : Andrew baker, chapitres 1,2,3 et 7.
2. Introduction à la théorie des groupes de Lie classiques : Rached Mneimné, Frédéric Testard, chapitres 2 et 3.
3. <http://www.math.ens.fr/viterbo/Cours-Geo-Diff-2012/Poly-Geodiff-2013.pdf>  
Cours de Claude Viterbo sur des notions de géométrie différentielle.
4. Lie groups : J.J Duistermaat et J.A.C Kolk, Chapitre 1.
5. <http://math.univ-lyon1.fr/frabetti/GEO-M2/geometrieM2.pdf>  
Cours d'Alessandra Frabetti pour la partie sur les sections, fibrés et champs de vecteurs.
6. Géométrie différentielle : variétés, courbes et surfaces : Marcel Berger et Bernard Gostiaux, Chapitres 2 et 3.

## 11.2 Index des résultats intéressants prouvés et expliqués dans ce TIPE :

Voici ici un index de résultats intéressants qui ne figurent pas dans la table des matières (aucun intérêt sinon). Malheureusement pour le lecteur, je ne renvoie pas aux pages où l'on retrouve ces résultats (je n'ai pas eu le temps : -) ). La plupart du temps, ces résultats portent sur les groupes ou sur les variétés (car ce sont les chapitres où un part de résultats intéressants ne sont pas dans la table des matières).

- Une isométrie vectorielle réelle est composée de rotations et de symétries.
- Une isométrie vectorielle complexe est une composée de rotations et symétries.
- Le produit cartésien de plusieurs groupes matriciels est un groupe matriciel.
- Le centre de  $GL_n(\mathbb{K})$  est  $\mathbb{K}^*$ .
- Le centre de  $O_n(\mathbb{R})$  est  $\{-I_n, I_n\}$ .
- Le centre de  $U_n(\mathbb{C})$  est  $\{-I_n, I_n\}$ .

- Le centre de  $SL_n(\mathbb{K})$  est  $\{aI_n \in \mathbb{K}t.qa^n = 1\}$ .
- Exemples d'espaces projectifs et caractérisation :  $P^2(\mathbb{R}), PSL_2(\mathbb{R})$  et  $PGL_n(\mathbb{R})$ .
- Tout groupe agit transitivement sur lui même.
- $O_n(\mathbb{R})$  agit transitivement sur  $\mathbb{S}^{n-1}$ .
- $U_n(\mathbb{C})$  agit transitivement sur  $\mathbb{S}^{2n-1}$ .
- $GL_n(\mathbb{K})$  agit transitivement sur  $\mathbb{K}^*$ .
- Surjectivité de l'exponentielle sur  $M_n(\mathbb{C})$  et d'autres ensembles.
- L'exponentielle de  $\mathfrak{sl}_2(\mathbb{R})$  dans  $SL_2(\mathbb{R})$  n'est pas surjective. -L'exponentielle est l'unique sous-groupe à un paramètre.
- $O_n(\mathbb{R})$  est compact, non connexe.
- $SO_n(\mathbb{R})$  est compact, connexe.
- $GL_n(\mathbb{R})$  est non compact, non connexe.
- $GL_n^+(\mathbb{R})$  est non compact, connexe.
- $SL_n(\mathbb{K})$  est connexe, non compact.
- $GL_n(\mathbb{C})$  est non compact connexe.
- $U_n(\mathbb{C})$  est compact, connexe.
- $SU_n(\mathbb{C})$  est compact, connexe.
- $SO_3(\mathbb{R})$  est simple.
- Le cercle  $\mathbb{S}^1$  est une variété.
- La sphère  $\mathbb{S}^2$  est une variété.
- Le tore  $\mathbb{T}^2$  (en tant que tore de révolution) est une variété.
- Le tore  $\mathbb{T}^n$  (en tant que produit cartésien de  $\mathbb{S}^1$ ) est une variété.
- L'intervalle  $[0, 1]$  fermé n'est pas une sous-variété, ni une variété (c'est une variété à bord) tandis que  $]0, 1[$  est une sous-variété et une variété.
- $GL_n(\mathbb{R})$  est une variété.
- $O_n(\mathbb{R})$  et  $SL_n(\mathbb{R})$  sont des sous-variétés réelles de  $GL_n(\mathbb{R})$  (on peut le montrer directement, sans utiliser le théorème de Cartan)
- La spirale logarithmique est un exemple de courbe régulière en tout point mais qui n'est pas une sous-variété.
- Le fibré tangent de  $\mathbb{S}^1$  est parallélisable.
- Le fibré tangent de  $\mathbb{T}^2$  (en tant que tore de révolution) est parallélisable.
- Le fibré tangent de  $\mathbb{T}^n$  (en tant que produit cartésien de  $\mathbb{S}^1$ ) est parallélisable.
- Le fibré tangent de  $\mathbb{S}^2$  n'est pas parallélisable.
- Le cercle  $\mathbb{S}^1$  est un groupe de Lie.
- La sphère  $\mathbb{S}^2$  n'est pas un groupe (donc pas un groupe de Lie).
- La sphère  $\mathbb{S}^3$  est un groupe de Lie.
- Si  $G$  est un groupe connexe,  $\exp(\mathfrak{g})$  engendre  $G$ .
- $\exp(\mathfrak{g})$  dans  $G$  n'est pas forcément surjective même si  $G$  est connexe.
- Si  $\exp(\mathfrak{g})$  dans  $G$  est surjective,  $G$  est connexe.
- L'algèbre de Lie de  $GL_n(\mathbb{K})$  est  $M_n(\mathbb{K})$ .
- L'algèbre de Lie de  $SL_n(\mathbb{K})$  est l'ensemble des matrices de  $M_n(\mathbb{K})$  de trace nulles.
- L'algèbre de Lie de  $O_n(\mathbb{R})$  et  $SO_n(\mathbb{R})$  est l'ensemble des matrices antisymétriques.
- L'algèbre de Lie de  $U_n(\mathbb{C})$  est l'ensemble des matrices antihermitiennes (avec trace nulle pour  $SU_n(\mathbb{C})$ ).
- Le fibré tangent d'un groupe de Lie est parallélisable .
- Il existe un sous groupe ouvert du tore qui n'est pas une sous-variété de celui-ci (ni une variété). -Le groupe de Heisenberg est un groupe de Lie qui n'est pas un groupe matriciel.
- $U(2)/U(1)$  et  $\mathbb{S}^3$  sont homéomorphes
- $SU(2)/U(1)$  et  $\mathbb{S}^2$  sont homéomorphes
- $\mathbb{R}/\mathbb{Z}$  et  $\mathbb{S}^1$  sont homéomorphes

### 11.3 Index de thèmes liés à ce TIPE

Voici maintenant un index de problèmes intéressants qui sont liés à ce TIPE. Le lien n'est pas forcément un lien direct avec le sujet, et le niveau de certains des résultats abordés est très élevé. Le but est de donner une ouverture sur d'autres sujets.

**Autres groupes matriciels :** groupe de Lorentz, groupes symplectiques, groupes spinoriels, corps des quaternions,...

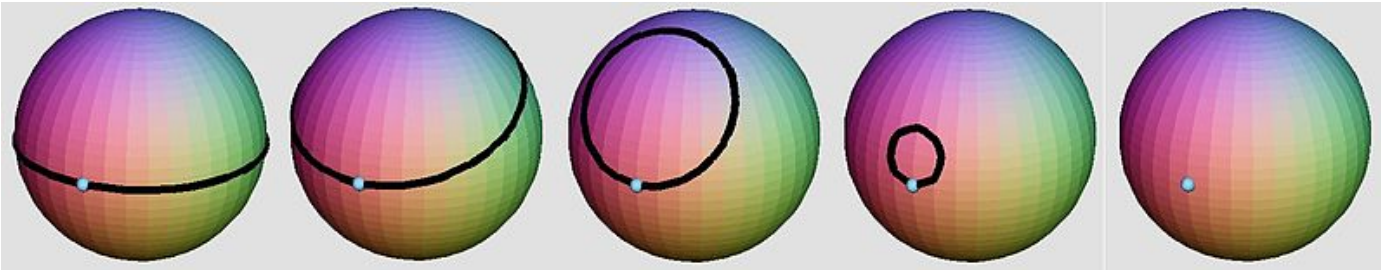
Dans un certains nombre de livres, on retrouve d'autres groupes matriciels tels que le groupe pseudo-orthogonal, les groupes symplectiques, les groupes de Lorentz et qui trouvent des applications en physique. Le livre d'Andrew Baker, *Matrix Groups : An Introduction to Lie Group Theory* en parle en détail aux chapitres 4,5 et 6.

**Espaces homogènes (notion plutôt générale) et projectifs (en particulier : la géométrie projective) :**

Le plan projectif est en fait une compactification du plan  $\mathbb{R}^2$  (c'est à dire, un "plan" fermé et borné), mais qui permet de représenter la direction d'un point vers l'infini. En effet, dans  $\mathbb{R}^2$ , on n'a pas de caractérisation de cette direction, les point qui tendent vers l'infini sont "équivalents". On peut alors introduire de la géométrie sur un plan (ou un autre espace) projectif.

#### Simple connexité

La simple connexité est une notion plus forte que la connexité par arcs : elle raffine la notion de connexité. Un espace simplement connexe est "sans trous" ni "poignées". On formalise cela en disant que tout lacet tracé dans un espace simplement connexe doit pouvoir être réduit continûment (c'est-à-dire par homotopie) à un point. On peut même aller encore plus loin en parlant de  $n$ -connexité. Voici une petite image d'illustration de la définition :



#### Revêtements, groupe fondamental :

Ce sont les notions que la simple connexité nécessite d'introduire.

#### Géométrie différentielle :

Le cours de Claude Viterbo cité dans la bibliographie était très riche à ce sujet. Plus on avance dans les variétés (et donc dans la géométrie différentielle), plus on peut étudier les groupes de Lie avec précision. Il en est de même si on approfondit la notion de groupes. En fait, la combinaison de ces 2 propriétés fait qu'on arrive à de très belles simplifications.

#### Théorème du point fixe de Brouwer :

Voici l'énoncé :

Dans un espace euclidien  $\mathbb{U}^n$  Toute application continue d'une boule fermée d'un espace euclidien dans elle-même admet un point fixe. C'est un résultat topologique très intéressant, qui peut être vu comme une conséquence du théorème de la boule chevelue. On peut trouver la démonstration (qui utilise le théorème de la boule chevelue) ici : <https://www-fourier.ujf-grenoble.fr/~carriere/TerDunias.pdf>

#### Théorème du point fixe de Lefschetz :

Je le mentionne car il est utile lorsque l'on s'intéresse à des groupes de Lie bien particuliers : les groupes de Lie compacts et connexes. Dans ce cas, des notions topologiques permettent de les classer (on utilise la notion de tore maximal).

Le théorème est une généralisation du point fixe de Brouwer. Voici son énoncé :

Soit  $f : X \rightarrow X$  une application continue d'un espace compact triangulable  $X$  dans lui-même. On définit le nombre de Lefschetz  $\Lambda_f$  de  $f$  comme la somme alternée (finie) des traces des endomorphismes induits par  $f$  sur les espaces  $\mathbb{H}_k(X, \mathbb{Q})$  d'homologie singulière de  $X$  à coefficients rationnels :

$\Lambda_f := \sum_{k \geq 0} (-1)^k \text{Tr}(f_* | H_k(X, \mathbb{Q}))$ . Une version simple d'énoncé du théorème de Lefschetz est que si ce nombre  $\Lambda_f$  est non nul, alors il existe au moins un point  $x$  fixe par  $f$ , c'est-à-dire tel que  $f(x) = x$ . Si tout comme moi vous n'avez pas compris tout l'énoncé, retenir qu'il s'agit d'une condition suffisante pour qu'une application continue admette, sous certaines conditions, un point fixe.

### **Théorème de l'invariance du domaine :**

Puisque je parle pas mal de topologie, un grand théorème sur les applications ouvertes est celui-ci :

Soit  $U$  un sous-ensemble ouvert de  $\mathbb{R}^n$  et  $f : U \rightarrow \mathbb{R}^n$  une injection continue, alors  $V = f(U)$  est ouvert et  $f$  est un homéomorphisme entre  $U$  et  $V$ , autrement dit,  $f$  est une application ouverte.

### **Théorème de mesurabilité de Smirnov :**

Toujours dans la topologie, puisque l'on a introduit la notion de paracompacité, voilà un résultat très intéressant, qui utilise des notions locales et parvient à quelque chose de global :

Il affirme qu'un espace est métrisable si et seulement s'il est paracompact (donc séparé) et localement métrisable. En particulier, tout espace métrisable est paracompact (théorème de Arthur Harold Stone) et toute variété topologique paracompacte (même sans base dénombrable) est métrisable.

### **Classification des variétés :**

On aurait pu montrer (mais c'est très difficile) que toute variété compacte et connexe de dimension 1 est difféomorphe à  $\mathbb{R}/\mathbb{Z}$  et que dans le cas contraire, elle est difféomorphe à  $\mathbb{R}$  : c'est la classification des variétés de dimension 1. Cette classification se généralise en dimension 2 et 3. Le mathématicien Grigori Perelman a montré que la seule variété compacte connexe de dimension 3 qui soit simplement connexe est difféomorphe à la sphère.

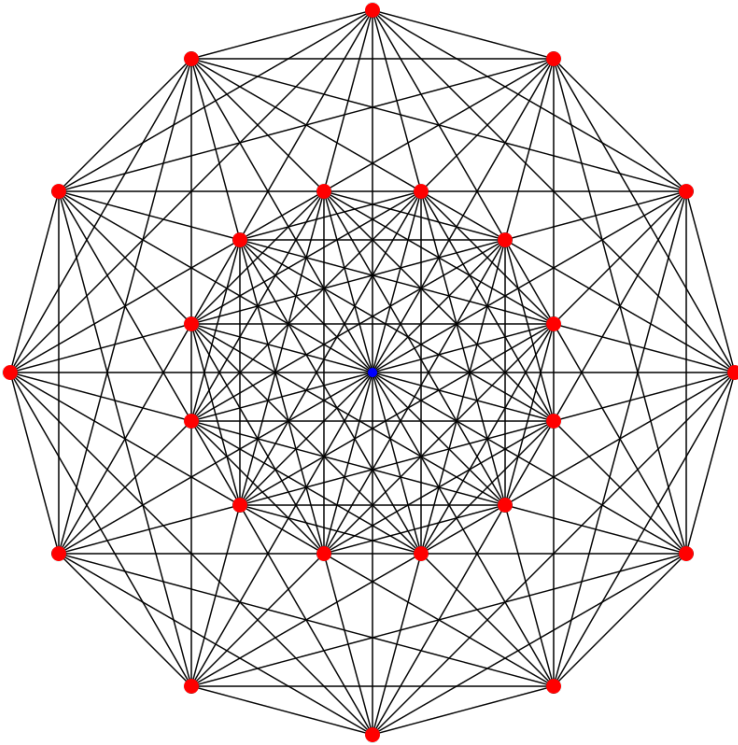
### **Mesure de Haar :**

Elle est essentielle lorsqu'on parle des groupes de Lie. Cependant, je n'ai pour l'instant aucune notion sur cette théorie de la mesure, qu'on voit en L3. Je ne vais donc pas détailler.

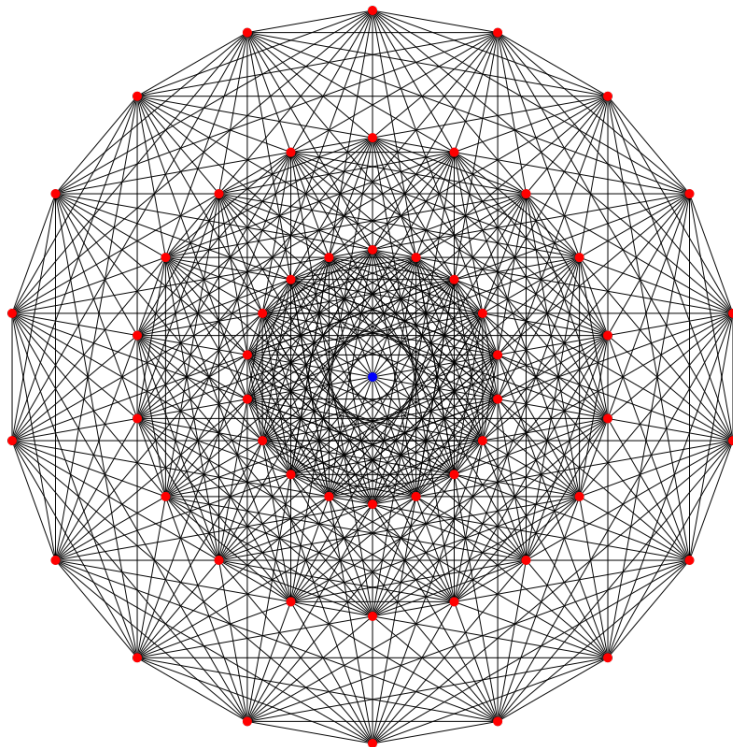
### **Classification des algèbres de Lie et diagrammes de Dynkin :**

Les diagrammes de Dynkin permettent en fait de classer les algèbres de Lie, autrement dit, de déterminer et caractériser l'ensemble des algèbres de Lie qui vérifient certaines conditions (il faut qu'elles soient semi-simples et réductives). Le pdf de Bernhard Keller est très bien expliqué et permet de parvenir à cela. Pour illustrer la beauté de cette théorie, voici quelques exemples de systèmes de racines (en lien avec cette théorie).

Système de racine du groupe de Lie  $E_6$ .



Système de racine du groupe de Lie  $E_7$ .



### Représentations des groupes :

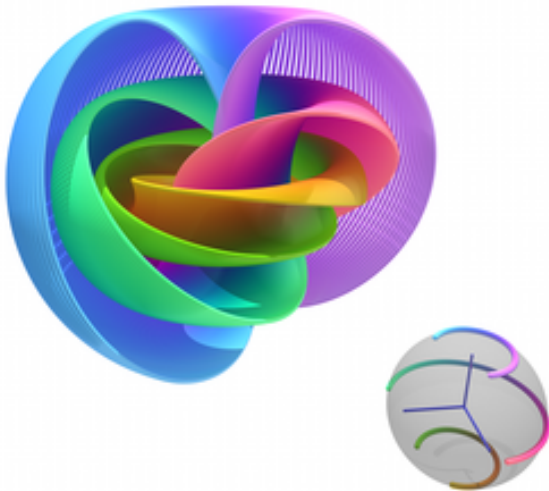
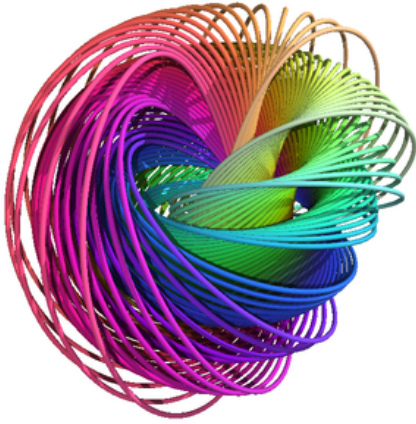
J'ai donné quelques exemples de représentations linéaires, mais il y a une théorie générale utilisant des notions d'algèbre linéaires (tenseurs,...) , qui est beaucoup plus vaste que les quelques notions que



j'ai introduites.

**Fibration de Hopf :**

Le fait que  $SU(2)/U(1) \cong \mathbb{S}^2 \cong \mathbb{S}^3/\mathbb{S}^1$  permet d'étudier la fibration de Hopf : celle-ci donne une partition de la sphère à 3-dimensions  $\mathbb{S}^3$  par des grands cercles. Plus précisément, elle définit une structure fibrée sur  $\mathbb{S}^3$ . Voici 2 images :



Pour plus d'informations sur cette fibration, on en a une très jolie explication sur ce lien (aux chapitres 7 et 8) :

[https://www.youtube.com/embed/cL7BpDrRc4s?list=PLw2BeOjATqrtiLPWvH\\_VeXmmBRmwcEwLz](https://www.youtube.com/embed/cL7BpDrRc4s?list=PLw2BeOjATqrtiLPWvH_VeXmmBRmwcEwLz)

Il s'agit d'un film de vulgarisation très intéressant.

**Groupes de Lie :**

Bien évidemment, il y a encore beaucoup de choses à dire sur les groupes de Lie, et ce TIPE n'est vraiment qu'un début. Une des questions que laisse ce TIPE en suspens est la démonstration du 3ème théorème de Lie de façon globale.